



Norges miljø- og
biovitenskapelige
universitet

Masteroppgave 2021 30 stp

Fakultet for landskap og samfunn

En studie av hvordan avstand til togstasjon påvirker kvadratmeterprisen for ulike leilighetsstørrelser

A study of how distance to the train station affects
the price per square meter for different apartment
sizes

Anders Nicolai Aaby Trætteberg
Simen Hovda Larsen

Master i Eiendomsutvikling

Forord

Denne masteroppgaven er en obligatorisk del av det toårige masterstudiet i Eiendomsutvikling ved Norges Miljø- og Biovitenskapelige Universitet i Ås.

Masteroppgaven har vært en lærerik, spennende og utfordrende prosess som har gitt mye god kunnskap og faglig fordypning.

Vi takker først og fremst Terje Holsen som hovedveileder for denne oppgaven. Gjennom interessante diskusjoner, støtte og tilbakemeldinger har oppgaven falt på plass. Vi retter også en stor takk til Hilde Vinje som veileder og en hjelpende hånd med statistikken. Gyda Grendstad må også takkes for god hjelp som biveileder fra Statens Vegvesen. Til slutt retter vi en stor takk til familie og kjære for god støtte og oppmuntring underveis.

I tillegg har Eiendomsverdi hjulpet med datamateriale, noe vi har satt stor pris på og gjort denne studien mulig.

Sammendrag

I denne masteroppgaven studeres det hvordan avstand til togstasjon påvirker leilighetspriser i byene Lillestrøm, Sandvika og Ski. I tillegg vil det bli gjort grundigere undersøkelser på om ulike leilighetsstørrelser blir påvirket forskjellig av avstanden til togstasjon. Formålet er å gi en ny innsikt i hvordan ulike kjøpegrupper vurderer nærhet til togstasjon som et attributt etter den hedoniske metoden. Her vil studiens resultater sammenlignes med tidligere funn for å se om det er likheter, eller motsetninger når et boligsegment justeres etter størrelse.

Studien baserer seg i hovedsak på teori om tilbud og etterspørsel i et makro- og mikroøkonomisk perspektiv, mono- og polysentrisk byteori samt bid-rent teori for boligmarkedet. De forskjellige teoriene belyser både hvordan avstand til togstasjon kan påvirke leilighetsprisene, samt gi en forståelse for hvordan ulikheter i kjøpegruppene kan påvirke lokasjonsvalg.

Den hedoniske metoden brukes som et grunnlag for å kunne studere enkeltattributter og den implisitte verdien som tilhører attributtet. Metoden bygger på et prinsipp om at en vare består av en rekke slike attributter, og at for en kjøper tilfører attributtene en verdi. Denne verdien kommer normalt sett ikke til syne annet enn gjennom salgsprisen av varen, men gjennom bruken av multippel regresjon kan den implisitte prisen og betalingsviljen for avstand til togstasjon måles. Gjennom bruken av multippel regresjon vil avstandsvariabelen justeres i forhold til andre utvalgte variabler for å gi et så nøyaktig estimat som mulig.

Studiens utvalg omfatter alle blokkleiligheter omsatt i perioden 2007-2020 for byene Lillestrøm, Ski og Sandvika. I tillegg avgrenses observasjonene til å kun være inntil 2 km fra togstasjonen. Observasjonene er hentet fra Eiendomsverdi sin database gjennom særegen avtale.

Resultatene viser at det er moderat til stor forskjell i hvilken grad ulike kjøpegrupper verdsetter avstand til togstasjon. Funnene viser at leilighetsprisene synker med økt avstand i tråd med eksisterende teori, men at det er store forskjeller når leilighetssegmentet deles opp i ulike undersegmenter. I tillegg viser resultatene at det også eksisterer forskjeller mellom de tre byene. Ikke bare er det forskjeller i hvor mye avstand påvirker leilighetsprisene, men det er også forskjeller i hvor mye de ulike segmentene blir påvirket.

Abstract

This master thesis studies how distance to train station affects the prices of apartments in Lillestrøm, Sandvika and Ski. In addition, a more thorough study will be made into whether different apartment sizes are affected differently by the distance to the train station. The purpose is to provide a new insight into how different purchasing groups consider proximity to the train station as an attribute according to the hedonic method. The results of this study will be compared with previous findings to see if there are similarities or contradictions when housing segments are adjusted according to size.

The study is mainly based on macro- and microeconomic theory, mono- and polycentric theory as well as bid-rent theory for the housing market. The various theories shed light on how distance to the train station can affect apartment prices, as well as provide an understanding of how differences in the purchasing groups can affect location choices.

The hedonic method is used as a basis for being able to study individual attributes and the implicit value of the attribute. The method is based on a principle that an item consists of several attributes, and that for a buyer these attributes add value. This value does not normally appear other than through the selling price of the item, but with the use of multiple regression, the implicit price and willingness to pay for the distance to the train station can be measured. Using multiple regression, the distance variable will be adjusted in relation to other selected variables to provide as accurate an estimate as possible.

The study sample includes all apartments in blocks sold in the period 2007-2020 for the cities of Lillestrøm, Ski and Sandvika. In addition, the observations are limited to only being within 2 km from the train station. The observations are taken from Eiendomsverdis database through a special agreement.

The results show that there is a moderate to large difference in the degree to which different purchasing groups value distance to the train station. The findings show that apartment prices fall with increasing distance in line with existing theory, but that there are large differences when the apartment segment is divided into different sub-segments. In addition, the results show that there are also differences between the three cities. Not only are there differences in how much distance affects apartments prices, but there are also differences in how much the different segments are affected.

Innholdsfortegnelse

1. Innledning	1
1.1 Bakgrunn	1
1.2 Problemstilling	3
1.3 Kontekst	5
1.3.1 Regional plan for areal- og transport i Oslo og Akershus	5
1.3.2 Lillestrøm	8
1.3.4 Sandvika	12
1.3.3 Ski	15
1.4 Avgrensninger for oppgaven	19
2 Teori	20
2.1 Tilbud og etterspørsel	21
2.1.1 Etterspørsel	22
2.1.2 Skifte i etterspørselskurven	24
2.1.3 Tilbud	25
2.1.4 Skifte i tilbudskurven	26
2.1.5 Markedslikevekt	26
2.2 Bid-rent	28
2.3 Monosentrisk byteori	31
2.5 Polysentrisk byteori	34
2.6 Hypotese 1	36
2.7 Hypotese 2	36
3 Metode og datamateriale	37
3.1 Oppbygging av oppgaven og valg av forskningsdesign	37
3.2 Hedonisk metode	39
3.3 Attributter	40
3.4 Analyseverktøy	43
3.4.1 Regresjonsanalyse	43
3.5 Rensing av data	49
3.6 Utvalgte attributter	51
3.6.1 Operasjonalisering av variablene og deskriptiv statistikk	52
3.6.2 Observasjoner	60
4 Analyse	65
4.1 Innledning	65
4.2 Introduksjon av begreper	65
4.3 Enkel lineær regresjon	67
4.4 Multipel regresjon	69
4.4.1 Testing av modellen	70
4.5 Lineær multipel regresjon for hver by	73
4.5.1 Lillestrøm	74
4.5.2 Sandvika	78
4.5.3 Ski	81
4.5.4 Oppsummering av Lineær Multipel regresjon	84

4.6 Datarensing.....	84
4.7 Log-Level multippel regresjon for hver by	85
4.7.1 Lillestrøm.....	86
4.7.2 Sandvika	88
4.7.3 Ski.....	90
4.7.4 Oppsummering av log-level multippel regresjon	92
4.8 Oppsummering av resultater og beregninger	92
5 Hypotesetesting.....	95
6 Konklusjon	97
7 Diskusjon	98
8 Kritikk til oppgaven.....	107
9 Forslag til fremtidig forskning.....	109
Referanser.....	111
Vedlegg.....	116

Figurer

Figur 1: Regional plan for areal- og transport i Oslo og Akershus, oversikt over hvilke jernbanelinjer og byer som er pekt ut som utviklingsbyer.....	6
Figur 2: Regional plan for areal- og transport i Oslo og Akershus, viser hvilke avstander som er pekt ut til utbygging Kilde: (Akershus fylkeskommune; Oslo Kommune, 2015, s. 28)	7
Figur 3: Oversiktskart over Osloregionen – Innerstrekninger inn til Oslo. Lillestrøm er markert med svart pil, kilde: (Nesvåg, Handstanger, & Martinsen, 2015, s. 9)	8
Figur 4: Historisk utvikling i befolkningsveksten for ulike aldersgrupper i Lillestrøm kommune, Kilde: (Kommuneprofilen, 2021, a).....	10
Figur 5: Historisk utvikling i Forholdet mellom ulike husholdninger, Kilde: (Kommuneprofilen, 2021, b)	10
Figur 6: Egen figur basert på tall fra SSB. Tabellen viser en oversikt over fordelingen av boligtyper i Lillestrøm kommune i perioden 2010-2020, Kilde: (SSB, 2021, b)	12
Figur 7: Oversiktskart over Osloregionen – Innerstrekninger inn til Oslo. Sandvika er markert med svart pil, kilde: (Nesvåg, Handstanger, & Martinsen, 2015, s. 9).....	13
Figur 8: Innbyggere etter aldersgruppe i prosent i Bærum kommune (Kommuneprofilen, 2021, e)	14
Figur 9: Andel husholdninger etter type i prosent i Bærum kommune (Kommuneprofilen, 2021, f).....	14
Figur 10: Oversiktskart over Osloregionen – Innerstrekninger inn til Oslo. Ski er markert med svart pil, kilde: (Nesvåg, Handstanger, & Martinsen, 2015, s. 9).....	16
Figur 11: Innbyggere etter aldersgruppe i prosent (Kommuneprofilen, 2021, c).....	17
Figur 12: Andel husholdninger etter type i prosent (Kommuneprofilen, 2021, d).....	17
Figur 13: Egen figur basert på tall fra SSB. Viser utvikling i de ulike boligtypene fra 2010 til 2020 i Ski kommune, Kilde: (SSB, 2021, b).	19
Figur 14: Viser grafisk etterspørselskurven (Bang, Demand Curve, 2019, a).....	23
Figur 15: Viser grafisk et skifte i etterspørselskurven, kilde: (Bang, Shift in demand curve, 2019, b)	25
Figur 16: Viser grafisk tilbudskurven (Bang, Supply Curve, 2019, c).....	25
Figur 17: Viser grafisk et skifte i tilbudskurven (Bang, Shift in supply curve, 2019, d).....	26
Figur 18: Viser markedslikevekten hvor S representerer tilbud, og D representerer etterspørsel (Bang, Equilibrium Quantity, 2019, e).....	27
Figur 19: Viser Bid-Rent funksjoner for 3 forskjellige bruksdunsjoner av en eiendom med forskjellig produktivitet og sensitivitet for transportkostnader, Kilde: (Geltner et al., 2014, s.66).....	28
Figur 20: Likevekt for ulike lokasjoner etter ulike inntektsklasser, kilde: (Fujita, 1989, s. 27). Den buede kurven viser prisnivået på boliger i ulik avstand til sentrum. R_i viser krysningspunktet for betalingsviljen for en husstand med lavere inntekt og boligprisen. R_j viser på sin side krysningspunktet for betalingsviljen for en husstand med høyere inntekt og boligpriser.	30
Figur 21: Monosentrisk byteori, kilde:(Geltner et al., 2014, s.85).	32
Figur 22: Leieprisnivåer I en polysentrisk (flerkjernet) region, Kilde:(Geltner et al, 2014, s. 89).	34
Figur 23: Scatterplot som viser enkel lineær regresjon mellom en avhengig variabel og en uavhengig variabel. Den svarte linjen er regresjonslinjen. Kilde: (Brooks, 2008, s. 31).	44
Figur 24: Method of OLS fitting a line to the data by minimising the sum of squared residuals, Kilde (Brooks, 2008, s. 32).....	45

Figur 25: Plottet viser en enkel observasjon, sammen med best tilpassede linje, residualverdien og forventet verdi, Kilde: (Brooks, 2008, s. 32).	45
Figur 26: Regresjonsmodell med et R^2 tall på 1, kilde: (Brooks, 2008, s. 109).....	47
Figur 27: Regresjonsmodell med et R^2 tall lik 1, kilde: (Brooks, 2008, s. 109).	48
Figur 28: a) Histogram over byggeår for Lillestrøm b) Histogram over byggeår for Sandvika, c) Histogram over byggeår for Ski.....	53
Figur 29: a) Histogram over avstand til togstasjon i meter for Lillestrøm b) Histogram over avstand til togstasjon i meter for Sandvika c) Histogram over avstand til togstasjon i meter for Ski.	54
Figur 30: a) Histogram over antall kvadratmeter for Lillestrøm b) Histogram over antall kvadratmeter for Sandvika, c) Histogram over antall kvadratmeter for Ski.....	55
Figur 31: a) Histogram over boligens alder for Lillestrøm b) Histogram over boligens alder for Sandvika, c) Histogram over boligens alder for Ski.	57
Figur 32: a) Histogram over hvilken etasje observasjonen ligger i for Lillestrøm b) Histogram over hvilken etasje observasjonen ligger i for Sandvika, c) Histogram over hvilken etasje observasjonen ligger i for Ski.	58
Figur 33: a) Histogram over kvadratmeterpris for Lillestrøm b) Histogram over kvadratmeterpris for Sandvika, c) Histogram over kvadratmeterpris for Ski.....	59
Figur 34: Oversikt over observasjoner etter at de er kategorisert etter størrelse for Lillestrøm. Rød er små leiligheter, blå er mellomstore leiligheter og gul er store leiligheter. Svart prikk markerer togstasjonen.	60
Figur 35: Oversikt over observasjoner etter at de er kategorisert etter størrelse for Sandvika. Røde er små leiligheter, blå er mellomstore leiligheter og gule er store leiligheter. I de områdene det er grønt ligger blå (mellomstore) og gule (store) over hverandre. Oransje er på sin side røde (små) og gule (store) oppå hverandre. Svart prikk markerer togstasjonen.	61
Figur 36: Oversikt over observasjoner etter at de er kategorisert etter størrelse for Ski. Rød er små leiligheter, blå er mellomstore leiligheter og gul er store leiligheter. I de områdene det er grønt ligger blå (mellomstore) og gule (store) over hverandre. Oransje er på sin side røde (små) og gule (store) oppå hverandre. Svart prikk markerer togstasjonen.	62
Figur 37: Egenprodusert enkel lineær regresjonslinje med kvadratmeterpris (KvmF) på Y-aksen og Meter på X-aksen.....	67
Figur 38: Egenprodusert enkel lineær regresjonsmodell.	68
Figur 39: Egenprodusert lineær regresjonsmodell med alle variabler inkludert.....	69
Figur 40: Egenprodusert Residuals vs fitted plot som eksempel og introduksjon.....	72
Figur 41: Egenprodusert Q-Q plot som eksempel og introduksjon.	73
Figur 42: Egenprodusert lineær regresjonsmodell for Lillestrøm.....	74
Figur 43: De fargede linjene viser regresjonslinjene for kvadratmeterpris for ulike leilighetsstørrelser. Kvadratmeterprisen for ulike leilighetsstørrelser på Y-aksen og avstand til togstasjon på X-aksen for Lillestrøm.	76
Figur 44: Residuals vs fitted plot for lineær regresjonsmodell for Lillestrøm.	77
Figur 45: Q-Q plot for lineær regresjonsmodell for Lillestrøm.	77
Figur 46: lineær regresjonsmodell for Sandvika.....	78
Figur 47: GG plot for kvadratmeterprisen for ulike leilighetsstørrelser på Y-aksen og avstand til togstasjon på X-aksen for Sandvika.	79
Figur 48: Residuals vs fitted plot for lineær regresjonsmodell for Sandvika.	80
Figur 49: Q-Q plot for lineær regresjonsmodell for Sandvika.....	80
Figur 50: Lineær regresjonsmodell for Ski.....	81
Figur 51: GG plot for kvadratmeterprisen for ulike leilighetsstørrelser på Y-aksen og avstand til togstasjon på X-aksen for Ski.	82
Figur 52: Residuals vs fitted plot for lineær regresjonsmodell for Ski.....	83

Figur 53: Q-Q plot for lineær regresjonsmodell for Ski.	83
Figur 54: Log-level regresjonsmodell for Lillestrøm.	86
Figur 55: Residuals vs fitted for Log-level regresjonsmodell for Lillestrøm.	87
Figur 56: Q-Q plot for Log-level regresjonsmodell for Lillestrøm.	87
Figur 57: Log-level regresjonsmodell for Sandvika.	88
Figur 58: Residuals vs fitted plot for Log-level regresjonsmodell for Sandvika.	89
Figur 59: Q-Q plot for Log-level regresjonsmodell for Sandvika.	89
Figur 60: Log-level regresjonsmodell for Ski.	90
Figur 61: Residuals vs fitted plot for Log-level regresjonsmodell for Ski.	91
Figur 62: Q-Q plot for Log-level regresjonsmodell for Ski.	91
Figur 63: Utrekninger basert på resultatene fra log-level regresjonsmodell. Estimaterne for «Meter» er omregnet til en prosentvis endring i kvadratmeterprisen for en leilighet som starter på 80 000kr/kvm ved 0 meter.	93
Figur 64: Viser hvordan avstand til togstasjon påvirker kvadratmeterprisen for de tre leilighetssegmentene i Lillestrøm. På X-aksen er antall meter, og Y-aksen er prosentvis endring i kvadratmeterprisen.	94
Figur 65: Viser hvordan avstand til togstasjon påvirker kvadratmeterprisen for de tre leilighetssegmentene i Sandvika. På X-aksen er antall meter, og Y-aksen er prosentvis endring i kvadratmeterprisen.	94
Figur 66: Viser hvordan avstand til togstasjon påvirker kvadratmeterprisen for de tre leilighetssegmentene i Ski. På X-aksen er antall meter, og Y-aksen er prosentvis endring i kvadratmeterprisen.	95
Figur 67: Oversikt over små leiligheter i Lillestrøm	116
Figur 68: Oversikt over mellomstore leiligheter i Lillestrøm.	116
Figur 69: Oversikt over store leiligheter i Lillestrøm	117
Figur 70: Oversikt over små leiligheter i Sandvika	117
Figur 71: Oversikt over mellomstore leiligheter i Sandvika.	118
Figur 72: Oversikt over store leiligheter i Sandvika.	118
Figur 73: Oversikt over små leiligheter i Ski.	119
Figur 74: Oversikt over mellomstore leiligheter i Ski	119
Figur 75: Oversikt over store leiligheter i Ski	120
Figur 76: Log-Level regresjonsmodell med tre-veis interaksjon mellom Meter: Storrelse og By.	121

Tabeller

Tabell 1: Egen figur tabell på tall fra SSB. Tabellen viser en oversikt over fordelingen av boligtyper i Lillestrøm kommune i 2020, Kilder: (SSB, 2021, b).	11
Tabell 2: Oversikt over antall salg fordelt på boligtyper de siste 5 årene, kilde: (Eiendomsverdi, 2021, a).	15
Tabell 3: Egen tabell basert på tall fra SSB. Tabellen viser en oversikt over fordelingen av boligtyper i Bærum kommune i 2020, Kilde: (SSB, 2021, b).	15
Tabell 4: Egen tabell basert på tall fra SSB. Tabellen viser en oversikt over fordelingen av boligtyper i Ski kommune (Før sammenslåingen) i 2020, Kilde: (SSB, 2021, b).	18
Tabell 5 Deskriptiv tabell over antall leiligheter og avstand til togstasjon.	53
Tabell 6 Deskriptiv tabell over variabelen antall kvadratmeter.	55
Tabell 7 Deskriptiv tabell over variabelen Størrelse.	56
Tabell 8 Deskriptiv tabell over variabelen byggeår.	57
Tabell 9 Deskriptiv tabell over kvadratmeterpriser.	59
Tabell 10 Begrepsoversikt for variablene i regresjonsmodellene.	66
Tabell 11: Multikollinearitetstest av variabler i multippel regresjonsmodell.	70
Tabell 12: Multikollinearitetstest av variabler i multippel regresjonsmodell.	71

1. Innledning

1.1 Bakgrunn

Den regionale planen for areal- og transport for Oslo og Akershus fra 2015 er ment som et overordnet styringsverktøy for byutvikling og sentrumsfortetting i byene i Osloregionen (Akershus fylkeskommune; Oslo Kommune, 2015). Et av hovedmålene i planen er å skape en flerkjernet byutvikling, hvor blant annet Lillestrøm, Sandvika og Ski er pekt ut som vekstbyer. Formålet med en flerkjernet byutvikling er sammensatt. For det første er det en forventet befolkningsvekst frem mot 2030, hvor prognosene tilsier en økning på 60 000 mennesker for Oslo Kommune (Oslo Kommune, 2021). I tillegg peker planen på et eksisterende underskudd på leiligheter i Akershus. Underskuddet er et resultat av en historisk økning i andelene aleneboende og for andelen eldre, og forventningene er at disse andelene vil øke sammen med befolkningsveksten. Et annet delmål er nullvekst i privatbilbruken og at privatbiltransport skal erstattes med gange, sykkel og kollektivtransport, herunder tog. Ved å konsentrere boligutbygging rundt kollektivknutepunkt i regionbyene ønsker man å møte fremtidig boligbehov samtidig som man lokaliserer befolkningen i nærheten av kollektivtransport for å støtte oppunder nullvekstmålet (Akershus fylkeskommune; Oslo Kommune, 2015).

I den forbindelse legger den regionale planen føringer for hvor og hvordan disse byene skal utvikles, med fokus på sentrumsfortetting rundt kollektivknutepunkt. I den regionale planen blir det vist til en 2 km radius som grensen for når mange velger å gå eller sykle til buss og tog. Utenfor denne avstanden forventes det at privatbil er det dominerende transportmiddelet. Med dette som et utgangspunkt har den regionale planen satt et mål om at 80-90% av all utbygging skal skje innenfor 2 km fra kollektivknutepunktet (Akershus fylkeskommune; Oslo Kommune, 2015).

Økende befolkning og byutvikling medfører på sin side en rekke utfordringer, hvorav boligprisvekst er en av dem (Geltner, Miller, Clayton, & Eichholtz, 2014). Boligens pris i forhold til avstand til et bysentrum eller togstasjon er et tema som har blitt studert flere ganger for Oslo-regionen, men også for andre byer og land. Det er spesielt noen studier som vil kunne ha en overføringsverdi og som ligger som inspirasjon for denne studien. Studien til

Ørjan Nybakk Rasch er gjort med hensyn på hvordan prising av boliger varierer med avstand til togstasjon for byer langs øst-, vest og sørgående jernbane ut fra Oslo (Rasch, 2017). Studiens funn viser at det er en monosentrisk effekt der boligene koster mindre med økt avstand fra togstasjonen. De samme funnene finnes i studiene til Joachim Gramm hvor avstand til togstasjon og boligpriser for byer langs Østfoldbanen ble studert (Gramm, 2016) samt Sandhu og Sarna sin studie for hvordan avstand til togstasjon påvirket boligprisene langs jernbanen til Jessheim (Sandhu & Sarna, 2015).

Basert på de nevnte studiene ser det ut til å være en intern monosentrisk priseffekt når byene studeres separat. Også for andre byer i Norge kan slike funn observeres, eksempelvis for Trondheim og Kristiansand (Nguyen, 2012). Samtidig tyder funnene på at det også eksisterer en polysentrisk effekt på boligprisene. Byer som har en større avstand til Oslo, har generelt sett lavere boligpriser enn byer som er nærmere.

Videre er det også gjort interessante funn i Malmø-Lund regionen i Sverige (Bohman & Nilsson, 2016). I denne studien fokuserte Bohman og Nilsson på hvordan prisen på eneboliger blir påvirket av avstand til togstasjon og motorvei. Ved å introdusere et bid-rent teoretisk konsept hvor husholdningens inntekt medregnes som en faktor, ble det gjort funn som viste at husholdninger med lavere inntekt ble påvirket mer av avstand til togstasjon enn de med høyere inntekt. I forbindelse med dette så de også at større boliger falt relativt mindre i pris ved økt avstand til togstasjonen sammenlignet med mindre boliger.

I forbindelse med økt befolkningsvekst og byutvikling vil det være nyttig å få en bedre innsikt i hvordan betalingsviljen kan påvirkes av boligens lokasjon og størrelse. Slik Malmø-Lund studien viste til, eksisterer det forskjeller både for inntektsnivåer og for boligstørrelser, noe som kan tyde på at disse henger sammen. Fujita (1989) og Coulombel (2017) har gjennom sine bid-rent studier av boligmarkedet vist til at det er en slik sammenheng mellom inntektsnivå, husholdningens sammensetning og betalingsvilje for lokasjon, hvorav husholdninger med høyere inntekt har lavere betalingsvilje sammenlignet med husholdninger med lavere inntekt. Ved å bruke ulike boligstørrelser som et mål på ulike kjøpegrupper vil den implisitte betalingsviljen gitt gjennom en hedonisk prismodell gi en bedre forståelse for hvordan kjøpere av ulike leilighetsstørrelser verdsetter avstand til togstasjon. Den implisitte betalingsviljen er et mål på hvor mye kjøperne av for eksempel boliger verdsetter ulike attributter.

Mye av tidligere forskning har vært rettet mot boligmarkedet som en helhet hvor eneboliger, rekkehus, leiligheter og tomannsboliger har blitt studert under ett. Det forskningsfeltet mangler er et dypdykk innenfor et boligsegment, hvor boligsegmentet blir studert grundigere innenfor et geografisk avgrenset område. På bakgrunn av dette og de overordnede føringene fra den regionale areal- og transportplanen, vil denne masteroppgaven ta for seg leilighetssegmentet hvor leilighetene blir kategorisert etter størrelse for å studere undermarkeder.

Det vil være av nytte for både eiendomsutvikler og kommune å få et innblikk i betalingsviljen for nærhet til kollektivtransport, da betalingsviljen er et mål på etterspørsel. Det å utvikle riktige leiligheter, og da også en riktig sammensetning av leiligheter, vil være av betydning for målet om at flere skal benytte seg av kollektivtransport under antagelsen om at folk bruker kollektivtransport i større grad når de bor nærme.

1.2 Problemstilling

For å kunne legge opp til en god leilighetsstruktur hvor det bygges ulike leilighetsstørrelser, vil en forståelse for betalingsviljen, og da etterspørselen for avstand til togstasjon være viktig. Ikke bare handler det om å kunne tilby leiligheter for å møte fremtidig befolkningsvekst eller redusere biltrafikk, men det handler også om å møte markedets etterspørsel etter lokasjon på et detaljnivå. Gjennom å ta hensyn til betalingsviljen kjøperne har for lokasjon, kan man legge opp en byutviklingsstrategi som sikrer at «rett bolig havner på rett sted» sett ut ifra boligkjøperens etterspørsel. For denne masteroppgaven vil målet være å studere sammenhengen mellom leilighetspriser for ulike leilighetsstørrelser og avstand til togstasjon for å kunne gi økt kunnskap om markedets etterspørsel. Gjennom den regionale areal- og transportplanen legges det overordnede føringer for hvordan utvalgte byer skal håndtere sentrumsfortetting, men med lite detaljstyring for de ulike byene. Ved å studere effekten avstand har på leilighetsprisene i Lillestrøm, Ski og Sandvika vil forståelsen for hvordan markedets etterspørsel endres fra by til by. Som følge av dette, vil problemstillingen vi ønsker å besvare gjennom dette studien være:

Hvordan påvirkes kvadratmeterprisen for ulike leilighetsstørrelser av avstand til togstasjon i Lillestrøm, Sandvika og Ski?

For å kunne besvare hovedproblemstillingen er den delt opp i tre underproblemstillinger som vil bli besvart hver for seg.

- 1. Følger kvadratmeterprisen for leiligheter monosentrisk byteori i form av at salgsprisen blir lavere jo større avstanden blir?*
- 2. Er det ulik betalingsvilje for avstand til togstasjon for ulike leilighetsstørrelser?*
- 3. Er betalingsviljen for avstand til togstasjon lik for byene Lillestrøm, Ski og Sandvika?*

Den første underproblemstillingen vil være en studie av hvordan leilighetspriser blir påvirket av avstand til togstasjon på et overordnet plan. Formålet er å gi en innsikt i hvor stor grad betalingsviljen for lokasjon er i leilighetssegmentet.

Den andre underproblemstillingen bygger videre på dette, samtidig som den vinkles mer inn mot hovedproblemstillingen. Effekten som avstand til togstasjon har på kvadratmeterprisen, vil bli målt for ulike leilighetsstørrelser hvor formålet er å studere om betalingsviljen endres i ulike delmarkeder.

For å gi en bedre forståelse for hvordan markedet vurderer pendlerdistansen lokalt for de tre byene, ønsker vi gjennom underproblemstilling tre å sammenligne resultatene fra de to første underproblemstillingene. Lillestrøm, Sandvika og Ski er valgt basert på lokasjon og størrelser. De ligger på hver sin side av Oslo, med Lillestrøm i øst, Ski i sør og Sandvika i vest. De har i tillegg relativt lik reisetid inn til Oslo, og togstasjonen befinner seg i bysentrum. Regional plan trekker frem at fremtidig boligutvikling må vurderes lokalt. En pendler er villig til å gå lengre distanser i tettbygde strøk sammenlignet med spredt bebyggelse (Akershus fylkeskommune; Oslo Kommune, 2015, s. 28) . Gjennom å sammenligne hvordan markedet vurderer de tre byene, vil det kunne gi ny innsikt for hvordan fremtidig byutviklingsstrategi kan møte etterspørselen.

1.3 Kontekst

Formålet med dette kapitlet er å gi kontekst til studien, både i form av geografisk utvalg, overordnede føringer som ligger som inspirasjon og valg av boligsegment. Regional plan for areal- og transport vil først bli presentert etterfulgt av en introduksjon av Lillestrøm, Sandvika og Ski.

1.3.1 Regional plan for areal- og transport i Oslo og Akershus

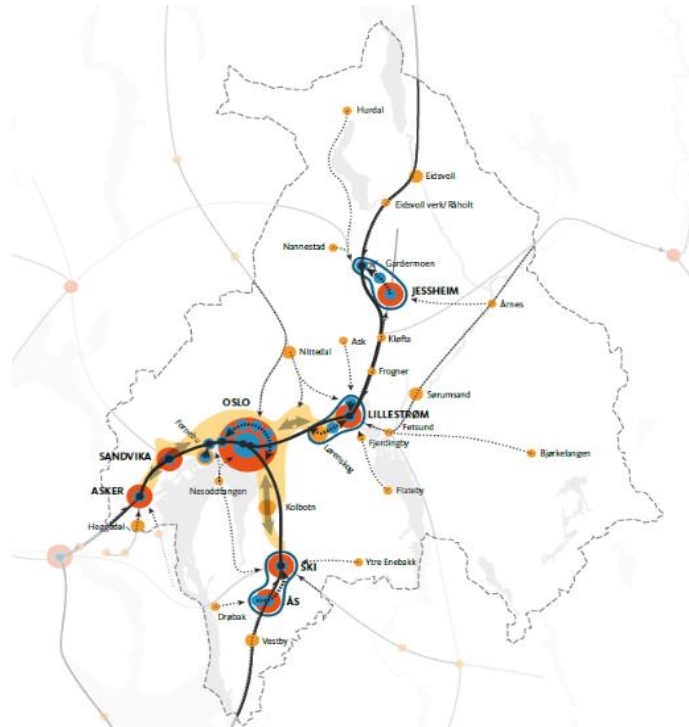
Planen er basert på at dagens by- og tettstedsstruktur videreutvikles, og skal være et overordnet styringsverktøy for stat, fylker og kommune (Akershus fylkeskommune; Oslo Kommune, 2015, s. 9). Det vil pekes på interessante føringer og retningslinjer som følger av planens arealstrategi. Her vil det være enkelte punkter som er av spesiell interesse for denne studien.

Som overordnede mål sier planen blant annet at “Utbyggingsmønsteret skal være arealeffektivt basert på prinsipper om flerkjernet utvikling og bevaring av overordnet grønnstruktur” (Akershus fylkeskommune; Oslo Kommune, 2015, s. 9) . Med dette legger planen opp til at fremtidig regional utvikling skal spres over flere byer, og ikke kun i Oslo. Herunder følger det to underordnede mål av interesse:

- “Det legges til rette for å håndtere en sterk fremtidig vekst i befolkningen” (Akershus fylkeskommune; Oslo Kommune, 2015, s. 9)
- «Planen skal bidra til at de nasjonale og regionale klima- og miljømålene blir oppfylt, inkludert at all trafikkvekst skal skje med sykkel, gange og kollektivt» (Akershus fylkeskommune; Oslo Kommune, 2015, s. 9).

Planen peker på et forventet underskudd av leiligheter frem mot 2030 som et resultat av økt andel eldre og små husholdninger. Ifølge planen er det allerede et underskudd på leiligheter blant annet i Akershus, og med en økende befolkningsvekst, forventes dette underskuddet å øke (Akershus fylkeskommune; Oslo Kommune, 2015, s. 9).

For å skape en flerkjernet utvikling er det pekt ut flere regionale vekstbyer (markert med røde sirkler), herunder Lillestrøm, Ski og Sandvika som vist under i figur 1.



Figur 1 Regional plan for areal- og transport i Oslo og Akershus, oversikt over hvilke jernbanelinjer og byer som er pekt ut som utviklingsbyer, kilde: (Akershus fylkeskommune; Oslo Kommune, 2015).

Disse byene ligger forholdsvis nærme Oslo, og har en direkte tilknytning til hovedstaden via tog og buss. Planens krav til arealutnyttelse legger her føringer for områdeutnyttelsen i disse byene. For de regionale byene, og byer i tilknytning til bybåndet, skal arealutnyttelsen være på 80-100%. Med 80-100% arealutnyttelse er målet å skape “bærekraftige steder der daglige behov kan dekkes innenfor gangavstand” noe som er inspirert av en dansk eksempelstudie. (Akershus fylkeskommune; Oslo Kommune, 2015, s. 26),

Videre viser også planen overordnede føringer for hvor denne utbyggingen skal lokaliseres. Her er det gangavstand som blir pekt på som vurderingsgrunnlag. Planen presiserer følgende: “Akseptabel gangavstand er situasjonsavhengig, og varierer etter omgivelser, hvem som går (bl.a. alder), og hva som er formålet med turen. Vi aksepterer lengre avstander i tettbygde byområder med variert aktivitet og tilrettelegging, og i omgivelser med høy kvalitet” (Akershus fylkeskommune; Oslo Kommune, 2015, s. 28). Det pekes altså på at det burde

gjøres lokale vurderinger med tanke på hvor utbygging burde lokaliseres. I en mindre by er altså folk ikke fullt så villige til å gå langt sammenlignet med en større by.

Når man omtaler gangavstand, er det som oftest 1 km som blir brukt som definisjonen på akseptabel gåavstand (Akershus fylkeskommune; Oslo Kommune, 2015, s. 28). En studie gjort av Nasjonal Gåstrategi kom frem til at dette tilsvarer 5-10 minutters gange. Passerer man 1 km, eller 10 minutters gange, øker andelen som benytter bil betraktelig (Statens vegvesen, 2012).



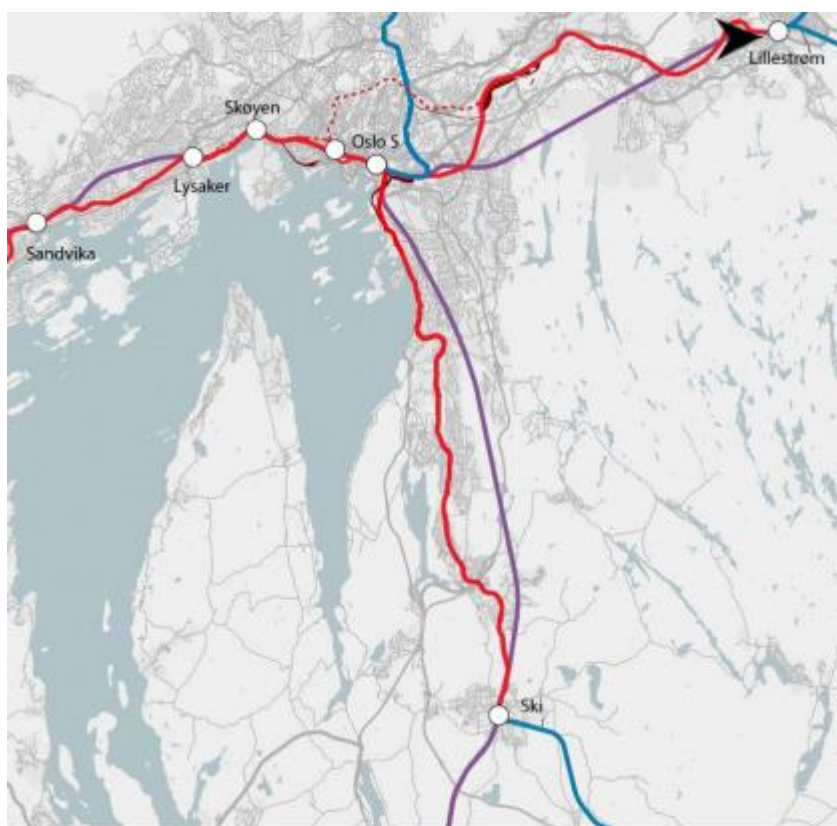
Figur 2: Regional plan for areal- og transport i Oslo og Akershus, viser hvilke avstander som er pekt ut til utbygging, kilde: (Akershus fylkeskommune; Oslo Kommune, 2015, s. 28).

I tilknytning til sentrale kollektivknutepunkt skal det bygges innenfor 1 km, og i de regionale byene er grensen satt til 2 km fra sentrum. For syklende er akseptabel reiseavstand noe lengre, ca. 3 km. Det vises derimot til at å bruke sykkelavstand som mål på akseptabel reiseavstand er noe problematisk. For det første er ikke alle kapable til å sykle grunnet for eksempel alder. Sykkel er også langt mer sesongavhengig sammenlignet med å gå (Akershus fylkeskommune; Oslo Kommune, 2015, s. 28). På bakgrunn av dette vil det for denne studien bli brukt 2 km som områdeavgrensning.

Som nevnt er den regionale planen en overordnet inspirasjon for denne studien. Da planen viser til et eksisterende underskudd på leiligheter, og at dette underskuddet vil øke med tiden grunnet befolkningsvekst, er det naturlig å studere leilighetssegmentet.

1.3.2 Lillestrøm

Lillestrøm ble i 2020 slått sammen til Lillestrøm kommune sammen med Skedsmo, Fet og Sørum, og ble også en del av det nye fylket Viken. Etter sammenslåingen har nå kommunen ca. 87 500 innbyggere og er Norges niende største kommune. Byens navn stammer fra gården den opprinnelig ble bygget ut ifra, Lille Strøm. I dag er Lillestrøm by administrasjonssenter for Lillestrøm kommune (Askheim, Lillestrøm, 2021, a).



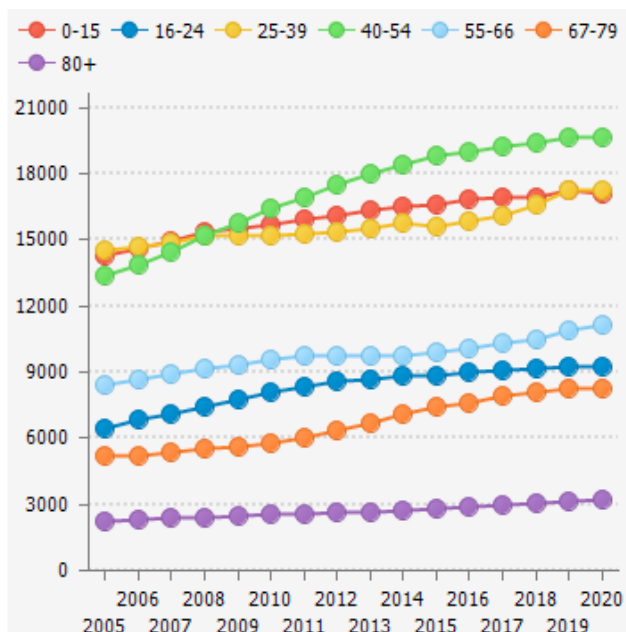
Figur 3: Oversiktskart over Osloregionen – Innerstrekninger inn til Oslo. Lillestrøm er markert med svart pil, kilde: (Nesvåg, Handstanger, & Martinsen, 2015, s. 9).

Lillestrøm var en egen kommune frem mot 1962 og fikk bystatus i 1998 (Askheim, Lillestrøm, 2021, a). Byen ligger på østsiden av Oslo, med et sammenhengende, tettbebygde bånd inn til hovedstaden som gjør byen til en forlengelse av Oslo by. Per 9. februar 2021 hadde byen 12 000 innbyggere og 18 500 innbyggere inkludert Kjeller som ligger en kort avstand unna.

Allerede i 1904 hadde byen jernbane og veiforbindelser, og dette året åpnet også dobbeltsporet til Oslo. Som en del av byggingen av ny hovedflyplass på Gardermoen i 1998 ble det også bygget ut ny jernbane gjennom Lillestrøm hvor flytoget mellom Asker og Gardermoen stopper (Askheim, Lillestrøm, 2021, a). Lillestrøm er et sentralt knutepunkt for nord-østre del av Viken kommune, hvor både buss og tog stopper regelmessig og knytter tidligere Romerike sammen. Med jernbanen i sentrum ligger byen ca. 10 minutters togreise unna Oslo S.

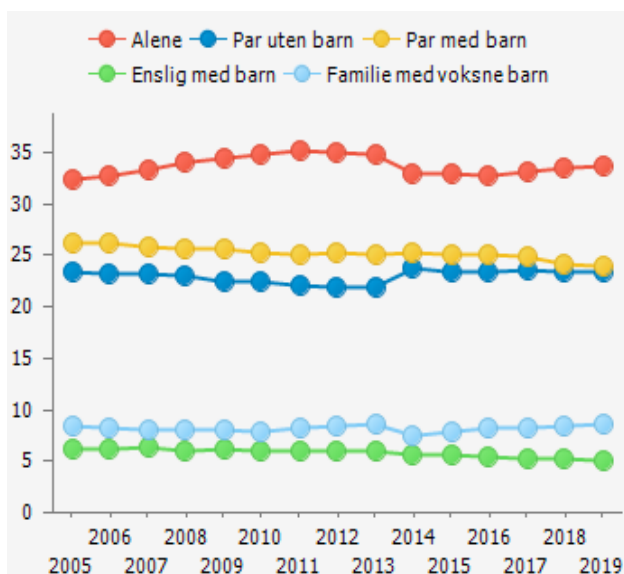
Dagens Lillestrøm er bygd opp rundt togstasjonen, hvor både boligtilbud og næringsarealer er utviklet med utgangspunkt i togstasjonen i sentrum. Fra å være en by med hovedsakelig småhusbebyggelse, ble det etter andre verdenskrig introdusert blokkbebyggelse. I nyere år har det pågått en stor fornyelse av byen, hvor Lillestrøms byplan legger stor vekt på boligtilbud og attraktivitet for beboere og utvidelse av areal for næringsliv (Lillestrøm kommune, 2020). Med en forventet økning i befolkningen til ca. 94 000 i 2030 for kommunen tilsvarer dette en økning på nesten 8% de neste 9 årene. Frem mot 2050 er det forventet at befolkningen øker videre, og med dagens forutsetninger vil kommunens befolkning øke til 104 300, dvs. en økning på ca. 20% (SSB, 2021, a).

Hvis man ser på befolkningsveksten for de ulike aldersgruppene i Lillestrøm (figur 4) ser man at enkelte grupper har økt mer enn andre. Spesielt aldersgruppen 45-54 år har økt betraktelig i perioden 2005-2020. Aldersgruppen 67-79 år har også hatt en betydelig vekst (Kommuneprofilen, 2021, a).



Figur 4 Historisk utvikling i befolkningsveksten for ulike aldersgrupper i Lillestrøm kommune, Kilde: (Kommuneprofilen, 2021, a).

Aleneboere utgjør den største gruppen (figur 5), og utgjør over 30% av husholdningene. Par med barn og par uten barn utgjør henholdsvis ca. 25% hver seg (Kommuneprofilen, 2021, b).



Figur 5 Historisk utvikling i Forholdet mellom ulike husholdninger, Kilde: (Kommuneprofilen, 2021, b).

Det er ikke kun befolkningsvekst som står sentralt i kommunens planer for Lillestrøm by. Det legges også stor vekt på å utvikle byen som et arbeidssenter hvor det skal kunne etableres ny næring (Lillestrøm kommune, 2020). Per i dag domineres næringslivet av engros- og lagervirksomhet, handel og andre servicefunksjoner. I tillegg er det en del statlige

forskningsinstitusjoner på Kjeller. I området i og rundt Lillestrøm er det grafisk industri, verkstedindustri og kjemisk industri. Lillestrøm har med andre ord et variert næringsliv innenfor mange forskjellige bransjer. Til tross for et rikt næringsliv arbeider likevel over halvparten av de yrkesaktive i kommunen utenfor kommunens grenser (Askheim, Lillestrøm, 2021, a).

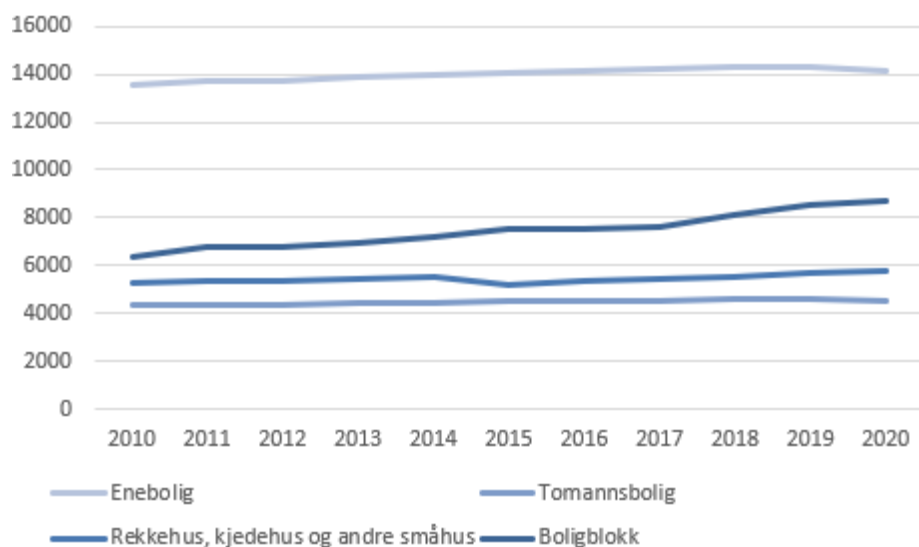
Det er som tidligere nevnt et eksisterende underskudd på leiligheter i Lillestrøm. Det rapporteres at Lillestrøm by kunne hatt den største befolkningsveksten i landet dersom det hadde blitt bygget flere boliger (Kunnskapsbyen, 2019). Ikke bare skal byen bygges opp for egen befolkningsvekst, men det skal i tillegg tas hensyn til den rollen som regional vekstby Lillestrøm har fått i Oslo-regionen. Det må derfor settes et stort fokus på byutvikling og utvikling av gode arealplaner for å skape gode bo- og leveforhold i en framtid med byvekst (Lillestrøm kommune, 2020)

At det er underskudd på leiligheter, kan også ses ut ifra boligfordelingen i kommunen. Eneboliger er det dominerende boligsegmentet og utgjør ca. 43% av den totale boligmassen. Selv med økt utbygging av leiligheter de siste årene, utgjør leilighetssegmentet kun ca. 26%. Siden 2010 har leilighetssegmentet økt med 36,7% (SSB, 2021, b).

Tabell 1: Egen figur tabell på tall fra SSB. Tabellen viser en oversikt over fordelingen av boligtyper i Lillestrøm kommune i 2020, Kilder: (SSB, 2021, b).

	2020 i %	
Enebolig	14136	43 %
Tomannsbolig	4556	14 %
Rekkehus, kjedehus og andre småhus	5732	17 %
Boligblokk	8709	26 %
Totalt	33133	

Med fremtidig vekst ønsker også Lillestrøm kommune at byen skal bli et gå- og sykkelvennlig sted, noe som støtter oppunder den regionale planens mål om nullvekst i privatbilbruk. Byen skal utvikles med hensyn til miljøvennlig pendling, hvor hovedtyngden av destinasjoner er innenfor gå- og sykkelavstand. For å oppnå dette skal det legges vekt på å utvikle sentrum langs eksisterende kollektivakser og kollektivknutepunkter. En utvikling av bysentrum er prioritert, hvor det legges vekt på fortetting (Lillestrøm kommune, 2020)



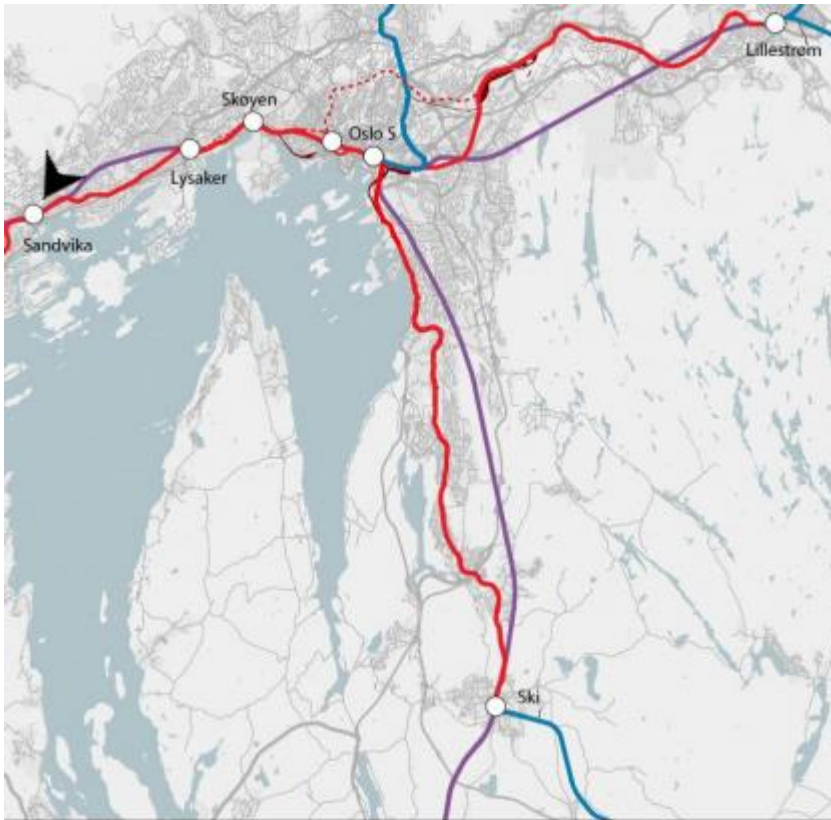
Figur 6: Egen figur basert på tall fra SSB. Tabellen viser en oversikt over fordelingen av boligtyper i Lillestrøm kommune i perioden 2010-2020, Kilde: (SSB, 2021, b).

1.3.4 Sandvika

De som først bodde i Sandvika, var tilknyttet gårder som lå rundt i området. Frem til i dag regnes Malmskrivergården å være den eldste bygningen. Gården ble bygd rundt 1640 tallet og er i dag fredet. Sandstrendene langs Sandvikselva var godt egnet til fiskeplasser, og ble etter hvert en viktig utskipningshavn for kalk, jern og trelast. Som følge av dette begynte fiskere, jakteskipperer og håndverkere å bosette seg i Sandvika (Bjørge, et al., 2012).

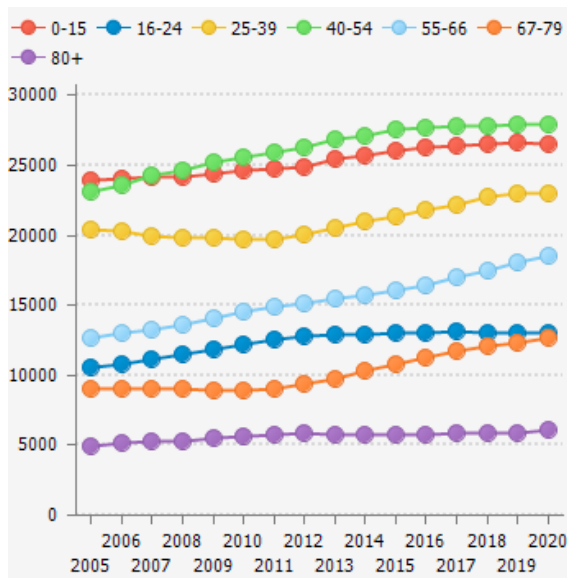
I 1835 ble det etablert en dampskipsforbindelse med Christiania. Omkring slutten av 1800 tallet økte folketallet da middelklassen flyttet til Sandvika. Blant dem var det leger, kjøpmenn, hotellfolk og håndverkere som bosatte seg i området. I 1872 ble Drammensbanen etablert. Jernbanen medførte økt handel og nye etableringer av industriselskaper. Fra 1960-tallet så man et skifte i bygningsstrukturen i Sandvika, hvor trehusbebyggelsen ble erstattet med nyere og moderne bygg (Bjørge, et al., 2012).

Sandvika ligger i dag veldig sentralt i Bærum kommune, hvor mange av de større infrastrukturlinjene, blant annet E18, fører til Sandvika. E18 er en av hovedfartsårene inn til Oslo. Sandvika stasjon har også god kapasitet av tog med blant annet flytoget, inter-city tog og lokaltog. Dette gjør at Sandvika knytter omkringliggende kommuner til Oslo. Sandvika ligger rundt 15 km unna Oslo sentralstasjon. En togtur fra Sandvika stasjon til Oslo sentralstasjon tar ca. 16 minutter.



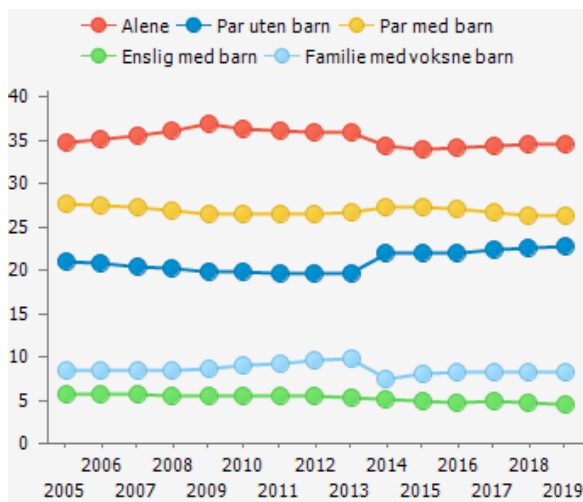
Figur 7: Oversiktskart over Osloregionen – Innerstrekninger inn til Oslo. Sandvika er markert med svart pil, kilde: (Nesvåg, Handstanger, & Martinsen, 2015, s. 9).

Bærum kommune har et innbyggertall på 126 841 (Thorsnærs & Askheim, 2021). Ettersom Sandvika ikke er geografisk avgrenset, var det problematisk å finne innbyggertall for Sandvika, og de presenterte dataene gjelder for Bærum kommune. Aldersfordelingen i Bærum kommune er veldig lik som både Lillestrøm og Ski, (figur 8). Det er flest i aldersgruppen 40-54 år tett etterfulgt av aldersgruppen 0-15 år. Over tidsperioden 2005 til 2020 har det også her vært de samme aldersgruppene som har hatt størst vekst. Aldersgruppene 40-54 år, 55-66 år og 67-79 år har alle hatt en betydelig vekst. Likevel ser det ut til at aldersgruppen 40-54 år har avtatt noe, og at aldersgruppen 25-39 år har hatt en større vekst enn de andre kommunene (Kommuneprofilen, 2021, a).



Figur 8: Innbyggere etter aldersgruppe i prosent i Bærum kommune, kilde: (Kommuneprofilen, 2021, a).

Figur 9 viser befolkningens fordeling etter husstander.



Figur 9: Andel husholdninger etter type i prosent i Bærum kommune, kilde: (Kommuneprofilen, 2021, b).

De som bor alene, er også dominerende for Bærum. Sammenlignet med Lillestrøm er forskjellene mellom alene-beboere, par uten barn og par med barn større. I Bærum ser det ut til at andelen alene-beboere er større, og andelen par med barn mindre. De utgjør likevel totalt sett ca. samme andel som for de andre kommunene. Andelen par uten barn hadde en oppgang fra 2013 til 2014, noe som er likt som Lillestrøm (Kommuneprofilen, 2021, b).

Bane Nor planlegger å utvide den eksisterende jernbanen til et dobbeltspor. (Norconsult, 2020). Dette er noe som kan medføre økt boligetterspørsel i Sandvika. Et godt kollektivtilbud

og lavere boligpriser sammenlignet med Oslo kan være faktorer som gjør at flere bosetter seg i Sandvika.

Som tidligere nevnt er Sandvika et viktig knutepunkt i Bærum kommune, og kommunen sier selv at de ønsker å prioritere utbygging i Sandvika slik at folk bruker kollektivtransporten i større grad (Bærum kommune, 2019). Tall fra Eiendomsverdi (Tabell 3) viser antall salg etter boligtyper de siste 5 årene. I Sandvika er 15% av boligene eneboliger, 2% er rekkehus og 79% er leiligheter av det totale salget (Eiendomsverdi, 2021, a).

Tabell 2: Oversikt over antall salg fordelt på boligtyper de siste 5 årene, kilde: (Eiendomsverdi, 2021, a).

Boligtype	Salg		Boliger		Avvik
	Antall	%	Antall	%	
Enebolig	97	7 %	632	15 %	-8 %
Rekkehus	35	3 %	85	2 %	1 %
Leilighet	1 189	87 %	3 366	79 %	8 %
Tomannsbolig	39	3 %	201	5 %	-2 %
Totalt	1 360	100 %	4 284	100 %	0 %

For resten av Bærum kommune er det en jevnere fordeling av boliger hvor man i 2020 hadde følgende fordeling:

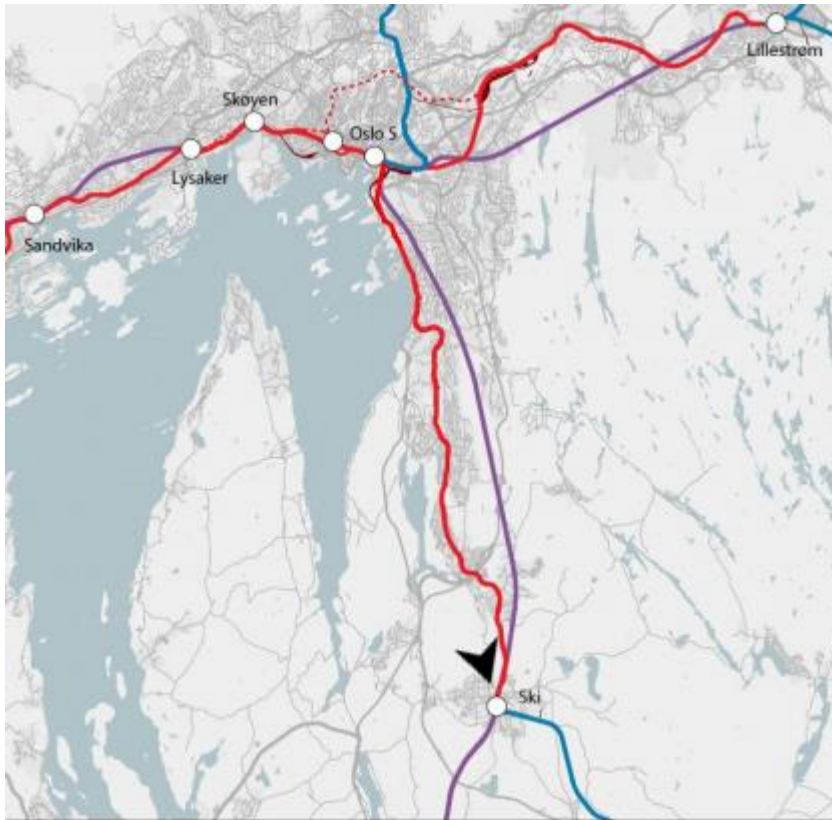
Tabell 3: Egen tabell basert på tall fra SSB. Tabellen viser en oversikt over fordelingen av boligtyper i Bærum kommune i 2020, Kilde: (SSB, 2021, b).

	Antall	%
Enebolig	15715	31 %
Tomannsbolig	7200	14 %
Rekkehus, kjedehus og andre småhus	9104	18 %
Boligblokk	19178	37 %
Totalt antall boliger	51197	

Antallet leiligheter i Bærum kommune har økt med 32% siden 2010, noe som er den høyeste veksten av alle boligtypene (SSB, 2021, b).

1.3.3 Ski

Ski var en egen kommune frem til kommunesammenslåingen i 2020. Da ble Ski kommune og Oppegård kommune slått sammen med Nordre Follo kommune, som det i dag heter. Nordre Follo er en del av Viken fylke, som består av tidligere Akershus, Buskerud og Østfold. Ski er det største tettstedet i Folloregionen og fungerer i dag som administrasjonssenter for Nordre Follo (Thorsnæs & Askheim, 2020).



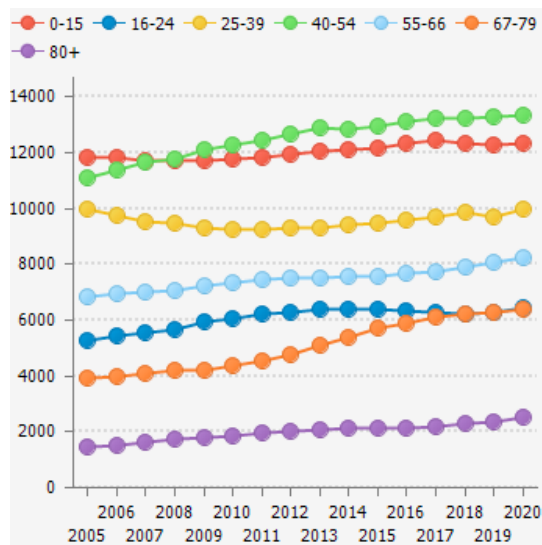
Figur 10: Oversiktskart over Osloregionen – Innerstrekninger inn til Oslo. Ski er markert med svart pil, kilde: (Nesvåg, Handstanger, & Martinsen, 2015, s. 9).

Per 1.januar.2020 var innbyggertallet på 58 237 i Nordre Follo Kommune (Askheim, Nordre Follo, 2021, b). Frem mot 2030 er det forventet at innbyggertallet øker til 70.000, altså en vekst på ca. 20%. Arealstrategien til Nordre Follo kommune er at Ski skal utvikles som en regional by og derfor skal ta den største delen av befolkningsveksten i Nordre Follo (Nordre Follo kommune, 2019, a)

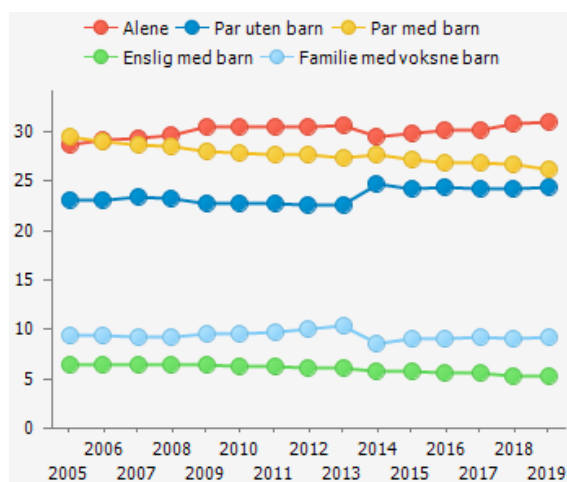
Næringslivet i Ski består hovedsakelig av grafisk industri, verkstedindustri, kjemisk industri, produksjon av gummi- og plastprodukter, samt næringsmiddelindustri. I tillegg til dette er jordbruket, spesielt korndyrking, en viktig del av Ski. Store deler av denne driften skjer i sentrale og søndre deler av gamle Ski kommune. (Thorsnæs & Askheim, 2020).

I gamle Ski kommune var det i 2019 et innbyggertall på 30 843 og kommunen hadde et landareal på 162 km². Store deler av dette landskapet er preget av skogsområder (Thorsnæs & Askheim, 2020). Østfoldbanen åpnet i 1879, og det var etter åpningen av denne banen at innbyggertallet for gamle Ski kommune økte. Spesielt fra 1960 årene opplevde kommunen en kraftig vekst i folketallet. Fra 1996 til 2006 var det en befolkningsvekst på 12,3% (Thorsnæs & Askheim, 2020).

I likhet med Lillestrøm, er det også for Nordre-Follo aldersgruppen 40-54 år som er den dominerende aldersgruppen, tett etterfulgt av 0-15 år (figur 11). Videre ser det ut til at aldersgruppene 40-54, 55-66 og 67-79 år har hatt den største veksten. De yngre befolkningsgruppene har holdt seg forholdsvis stabile (Kommuneprofilen, 2021, a).



Figur 11: Innbyggere etter aldersgruppe i prosent, kilde: (Kommuneprofilen, 2021, a).



Figur 12: Andel husholdninger etter type i prosent, kilde: (Kommuneprofilen, 2021, b).

Alene-beboere og par med barn har lenge vært de dominerende kategoriene, men fra 2013 ble det et skifte hvor også andelen par med barn økte. Disse utgjør til sammen ca. 85% av befolkningen i kommunen. I likhet med Lillestrøm er det enslige med barn, og familie med voksne barn som utgjør de minste gruppene med ca. 15% i Nordre Follo (Kommuneprofilen, 2021, b).

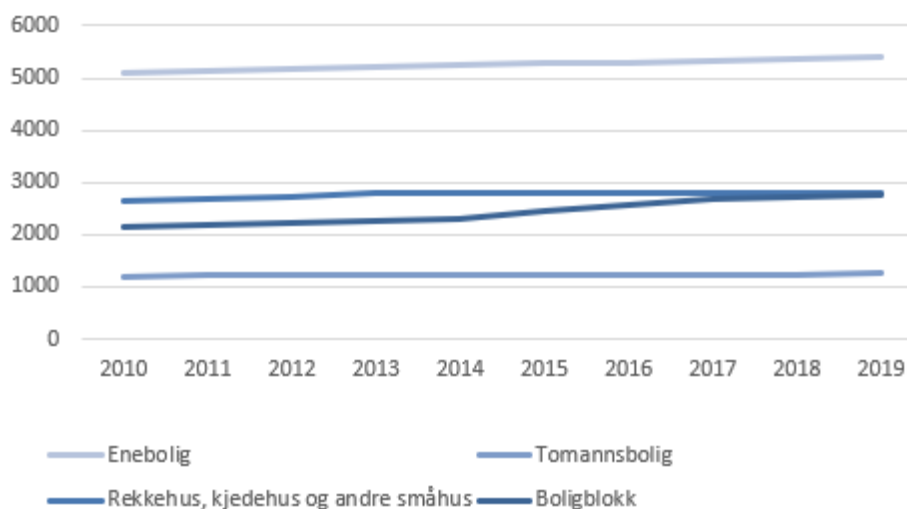
I Regional plan for areal og transport i Oslo og Akershus er Ski utpekt som en regional by og et kollektivknutepunkt (Akershus fylkeskommune; Oslo Kommune, 2015, s. 7). I dag er Follobanens nye dobbeltspor under utvikling, og når denne er ferdig vil reisetiden fra Ski inn til Oslo reduseres til ca. 11 minutter. I tillegg har man den allerede etablerte Østfoldbanen. Når Follobanen er ferdigstilt vil Ski ha et meget godt tilbud av kollektivtransport. Med E18 i nærheten vil Ski være et kollektivknutepunkt for tog og buss som knytter resten av Nordre Follo kommune nærmere Oslo. Nordre Follo kommune har lagt opp til at hoveddelen av byutviklingen i Ski skal skje i nærheten av Ski stasjon for å oppnå en kompakt byutvikling. De prioriterte utbyggingsområdene skal ligge innenfor 10 minutters gangavstand til stasjonen i Ski (Nordre Follo kommune, 2019, b).

I det som tidligere var Ski kommune utgjorde eneboliger 44% av alle boligene i 2019, og var det største boligsegmentet (tabell 4). Tomannsboliger utgjorde kun 10%, og rekkehus 23%. Boligblokker, altså leiligheter, utgjorde også 23% av all boligmasse (SSB, 2021, b).

Tabell 4: Egen tabell basert på tall fra SSB. Tabellen viser en oversikt over fordelingen av boligtyper i Ski kommune (Før sammenslåingen) i 2020, Kilde: (SSB, 2021, b).

	2019	i %
Enebolig	5394	44 %
Tomannsbolig	1259	10 %
Rekkehus, kjedehus og andre småhus	2810	23 %
Boligblokk	2773	23 %
Totalt	12236	

Fra 2010 og frem til 2019 har antallet leiligheter steget med 29%, sammenlignet med eneboliger hvor antallet har steget med 6% i samme periode. Denne utviklingen er i tråd med et skifte i folks preferanser der mange flere ønsker seg en leilighet. Det har også vært et skifte i strategien til kommunen som ønsker boligutvikling i nærhet til kollektivknutepunktet. Nordre Follo kommune har i sin kommuneplan estimert en boligvekst på rundt 500 boliger pr år i perioden 2019 til 2030.



Figur 13: Egen figur basert på tall fra SSB. Viser utvikling i de ulike boligtypene fra 2010 til 2020 i Ski kommune, Kilde: (SSB, 2021, b).

1.4 Avgrensninger for oppgaven

For å kunne besvare problemstilling og underproblemstillinger gjøres det noen overordnede avgrensninger innledningsvis. Disse avgrensningene er til for å forklare det geografiske studieområdet, hvilket boligsegment som studeres, samt hvilket tidsrom dataene vil være basert på.

Som nevnt vil byene Lillestrøm, Ski og Sandvika studeres og sammenlignes. Valget av disse byene er inspirert av den regionale planen i tillegg til at de ligger i omtrentlig lik avstand fra Oslo i form av reisetid med tog. Regional plan trekker riktignok også frem andre byer som vekstbyer, slik som Ås og Asker. Disse er ikke inkludert i denne studien da Ås anses for å være en for liten by, og Asker ligger lengre unna i reisetid. At Ås regnes for å være liten har to årsaker. For det første krever en studie som dette mange observasjoner, noe det ikke er i Ås. For det andre er det satt en radius på 2 km fra togstasjonen som studieområde for hver by. Den geografiske utstrekningen for Ås tettsted er ikke stort nok. Det burde også nevnes at både Lillestrøm og Ski har i etterkant av vedtaket om regional plan vært gjennom kommunesammenslåing, noe som kan påvirke data på kommunalt nivå. Det vil likevel ikke være av stor betydning, da uthentede data hentes ut for en såpass liten radius at endringer i kommunegrensene ikke får betydning.

Radiusen på 2 km er også tatt fra regional plan, hvor denne bruker 2 km som et mål på akseptabel gang- og sykkelavstand i byområder. For å nå målene om flerkjernet byutvikling som både skal ta unna for befolkningsvekst, samt stimulere til nullvekst i biltrafikken, må 80-90% av boligutvikling skje innenfor denne avstanden (Akershus fylkeskommune; Oslo Kommune, 2015, s. 28).

Regional plan viser også til et eksisterende underskudd på leiligheter, spesielt for tidligere Akershus. Underskuddet på leiligheter er også forventet å øke i fremtiden som et resultat av befolkningsveksten. Ikke bare pekes det på en generell befolkningsvekst, men en enda større økning i aleneboere og eldre legger press på leilighetssegmentet. For studien er det derfor valgt å studere leilighetssegmentet. I studiene for Norge (Gramm, 2016) (Sandhu og Sarna, 2015) (Nguyen, 2012) og (Hoen, 2018), har leiligheter vært inkludert en rekke ganger når det kommer til hvilke boligsegmenter som har blitt studert. Det som derimot ikke er studert i samme grad, er et dypdykk innenfor et segment hvor det gjøres en inndeling for å studere ulike undermarkeder.

Som nevnt krever en studie som dette mange observasjoner. Da det geografiske området allerede er avgrenset, er den eneste måten å øke dataomfanget at tidsrommet som studeres økes. I de nevnte studiene har det blitt brukt et tidsrom på 2-4 år, noe som for vårt tilfelle vil være svært begrensende. I denne studien brukes et tidsrom fra 2007 til 2020 for observasjoner. For en statistisk analyse som denne studien utfører er det ikke behov for å ta hensyn til boligprisvekst, da det gjøres like-for-like analyser. Med dette menes det at leiligheter fra 2007 analyseres opp mot andre leiligheter fra 2007, og det vil ikke skape en feilkilde i modellen at boligprisene som brukes er høyere i 2020 enn i 2007.

2 Teori

For å besvare oppgavens problemstillinger må tematikken, altså forholdet mellom byutvikling og det å bygge boliger «på rett sted», settes i kontekst til overordnede teoretiske konsepter. I dette kapitlet skal det presenteres og forklares hvilket teorigrunnlag som ligger til grunn for studien. Studien baserer seg på en strategi om polysentrisk byutvikling basert på den regionale planens mål om flerkjernet byutvikling. Med en avgrensning for Lillestrøm, Ski og Sandvika er formålet å øke kunnskapen om hvordan markedet etterspør nærhet til togstasjon for de tre byene. En sammenligning og vurdering av byene vil kunne gi mer

kunnskap om forholdet mellom bystruktur og markedets ønske om å bo i gå- og sykkelavstand til kollektivknutepunkt. Ikke bare er dette viktig for å møte etterspørselen, men det vil også være med på å nå hovedmålene om nullvekst i bilbruk og forbedret miljø.

Ved å tilby boliger, og da leiligheter på rett sted, vil man i økt grad få beboere som velger å gå eller sykle, gjerne i kombinasjon med kollektivtransport, fremfor å kjøre bil. Monosentrisk byteori vil gi ytterligere forståelse for bystrukturen samt samspillet mellom lokasjon og pendlerkostnad. I tillegg vil teorien kunne gi et perspektiv på hvordan inntekt påvirker lokasjonsmønsteret.

Tilbud og etterspørsel gjennom et makro- og mikroøkonomisk perspektiv vil på sin side gi en bedre forståelse for de økonomiske drivkreftene i markedet. Innledende teori om tilbud og etterspørsel fra både et makroøkonomisk og mikroøkonomisk ståsted vil bli presentert. Gjennom individuelle preferanser fra mikroøkonomisk teori skapes det kunnskap om hvordan markedet består av enkeltpersoner som alle har personlige preferanser, og at disse preferansene blant annet kan endres basert på inntektsnivå. Bid-rent gir en grundigere forklaring på dette, og knytter individuelle preferanser inn mot boligmarkedet og begrepet betalingsvilje. Igjen blir forholdet mellom inntekt og lokasjon belyst, men på et grundigere nivå enn mono- og polysentrisk teori.

2.1 Tilbud og etterspørsel

Tilbud og etterspørsel er en markeds mekanisme som brukes til å forstå hvorfor priser på ulike produkter blir til. For tilbudssiden har man tilbudskurven som representerer prisen tilbyderne er villig til å selge en vare eller tjeneste for. Det samme gjelder for etterspørselssiden, hvor etterspørselskurven representerer prisen den som etterspør varen eller tjenesten er villig til å gi. Det er når det er likevekt mellom tilbudskurven og etterspørselskurven man får prisen på en vare eller tjeneste (Cooper & Andrew, 2011).

Tilbud og etterspørsel kan sees på fra et makroperspektiv og et mikroperspektiv. Fra et makroperspektiv ser man på den samlede effekten av markedsaktørene i økonomien gjennom en tilbud- og etterspørselskurve. Fra et mikroperspektiv er man mer interessert i de

individuelle valgene som utgjør tilbud- og etterspørselskurven (Cooper & Andrew, 2011, s. 80) .

NOU2002:2, Utredning om boligmarkedet gjort av Boligutvalget, beskriver boligmarkedet som et komplisert og uoversiktlig marked, hvor aktørene er ulike på både tilbuds- og etterspørselssiden. I tillegg er de valgene som blir tatt ofte tatt for en lengre tidshorisont (NOU 2002:2, 2002, s. 17).

Kompleksiteten kan komme av at boligmarkedet er delt opp i delmarkeder. Insentiver en kjøper har for investeringen er også med på å gjøre markedet uforutsigbart. En bolig er enten et konsum-, et kapital- eller et formuesgode. Et boligkjøp er dermed avhengig av hvilke formål boligen skal tjene. På den ene siden kjøpes boliger for å tjene et boformål. På den andre siden kan boligkjøp være den største investeringen mange gjør i livet sitt, og det er dermed en investering man gjør i et kapitalobjekt. I boligutvalgets rapport lister de opp fire økonomiske hoveddrivkrefter; boligpriser, bokostnader og boutgifter, tomtepriser, byggekostnader og nyboligpriser samt boligbygging og boligfordeling (NOU 2002:2, 2002, s. 17). Bokostnader og boutgifter vil være et sentralt tema i denne oppgaven, da pendlerkostnader antas å falle innunder denne faktoren.

For å kunne gi en god forklaring på hvordan tilbud og etterspørsel i boligmarkedet fungerer, gis det videre en forenklet fremstilling fra et makroøkonomisk perspektiv og et mikroøkonomisk perspektiv.

2.1.1 Etterspørsel

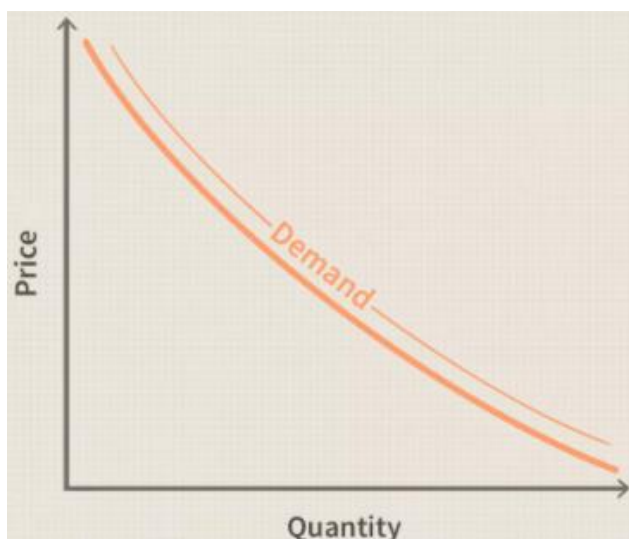
Fra et makroperspektiv i et boligmarked representerer etterspørselskurven mange av husholdningene i markedet (figur 14). Dersom prisen synker på en bolig, vil etterspørselen i markedet øke. Teorien om tilbud og etterspørsel har to effekter. Dersom prisen på en vare eller gode synker, velger flere å kjøpe i stedet for å ikke å eie noe. Ved prisreduksjon på en vare vil flere kjøpe et større kvantum. I boligmarkedet kan man enten leie eller eie. Ved forutsetningene i tilbud og etterspørsel vil flere husholdninger gå fra å leie til å kjøpe seg en bolig dersom prisen på bolig synker. Ettersom prisen på en bolig er høy er det kun de færreste

som har råd til å kjøpe seg flere boliger sammenliknet med andre varer som f.eks. klær (Cooper & Andrew, 2011, s. 80).

Fra et mikroøkonomisk perspektiv er betalingsviljen til hver enkelt husholdning en viktig prisdriver i etterspørselssiden på boligmarkedet. Betalingsviljen er i stor grad avhengig av hva hver enkelt husholdning har i inntekt og formue. Videre avhenger også betalingsviljen av hvilke økonomiske utgifter disse husholdningene har som f.eks. bokostnader og boutgifter. Med bokostnader menes rentekostnad og drifts- og vedlikeholdskostnader, i tillegg må man trekke fra skattefordelen (skattefradrag for gjeldsrenter) ved å eie boligen og verdistigningen på boligen (NOU 2002:2, 2002, s. 20).

Frem til nå er betalingsviljen kun blitt diskutert i lys av inntekt og utgifter hver enkelt husholdning har. Betalingsviljen blir også påvirket av hvordan husholdningene vektlegger andre goder opp mot kjøp av bolig, og hvordan man betrakter bolig opp mot andre investeringer. Forskjellige husholdninger har forskjellige preferanser hvor noen vil prioritere å bruke mer penger på for eksempel reise kontra boligkjøp. Andre husholdninger kan prioriterer andre investering som f.eks. aksjer. Forskjellige husholdninger kan ha samme betalingsevne, men ulike preferanser og dermed ulik betalingsvilje (NOU 2002:2, 2002, s. 20).

I denne delen har vi ikke diskutert hvordan de forskjellige egenskapene til boligen påvirker betalingsviljen til husholdningene. Dette vil bli diskutert i teorien om Bid-Rent og den hedoniske metoden.



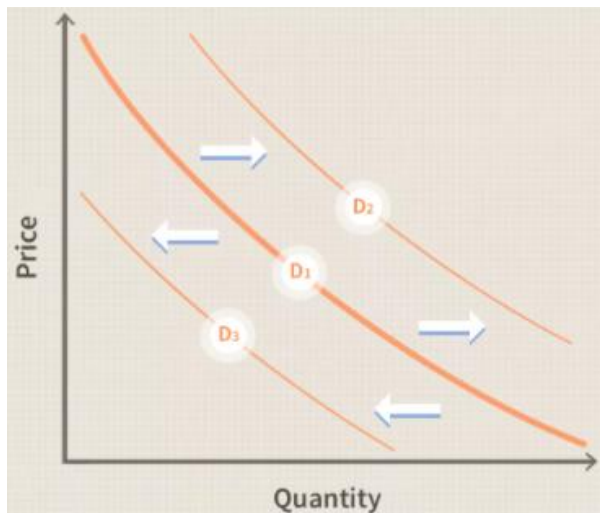
Figur 14: Viser etterspørselskurven, kilde: (Bang, Demand Curve, 2019, a).

2.1.2 Skifte i etterspørselskurven

Etterspørselen forutsetter at prisene varierer, men at alle andre faktorer holdes konstant. Dersom disse faktorene endres, vil man få et skifte i etterspørselskurven (Figur 15). I boligmarkedet er det en rekke faktorer som kan påvirke etterspørselskurven (Cooper & Andrew, 2011, ss. 80-81). I et makroøkonomisk perspektiv vil boliglånsforskriften som ble innført i 2017 være et godt eksempel på en slik faktor. Forskriften satte blant annet begrensninger på hvor mye hver person kan ta opp i lån. Etter at denne forskriften ble innført kan man ikke låne mer enn 5 ganger brutto årsinntekt (Regjeringen, 2016). Med denne begrensningen på låneopptak er det nærliggende å tro at dette har hatt en effekt på kjøpekraften til husholdningene innenfor boligmarkedet. Andre faktorer som kan redusere etterspørselen er rentehevinger, billigere leiepriser for bolig og dårlig økonomiske utsikter. Dersom noen av disse faktorene slår inn vil etterspørselen synke. Redusert etterspørsel fører til et mindre behov for boliger på et hvert prisnivå som igjen vil stimulere til at etterspørselskurven flytter seg til venstre (Cooper & Andrew, 2011, s. 82). Denne effekten er vist i figur 15.

Tidligere nevnt er husholdningenes preferanser en viktig driver i etterspørselen i et mikroøkonomisk perspektiv, og et skifte i husholdningenes preferanser kan føre til et skifte i etterspørselskurven. Helt siden Covid-19 brøt ut har mange sittet på hjemmekontor og i samme periode har boligprisene steget kraftig, med hele 12 % i Oslo (Christensen, 2021). Noe av grunnen til denne boligprisveksten kan være at folk har innsett at de trenger mer plass, som kan ha ført til et skifte i preferansen mot større boliger. Som et resultat av dette har betalingsviljen økt og presset prisene oppover.

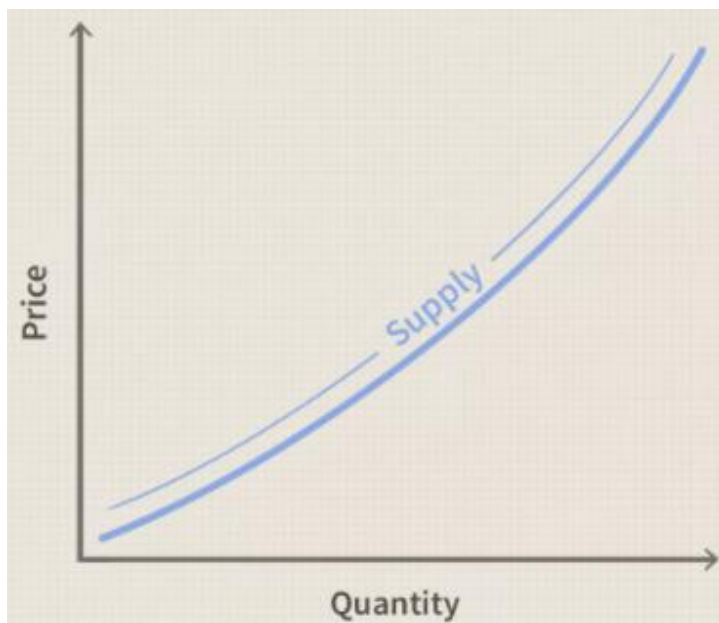
Som tidligere nevnt er husholdningenes preferanser en viktig driver i etterspørselen i et mikroøkonomisk perspektiv, og et skifte i husholdningenes preferanser kan føre til et skifte i etterspørselskurven. Helt siden Covid-19 brøt ut har mange sittet på hjemmekontor og i samme periode har boligprisene steget kraftig, med hele 12 % i Oslo (Christensen, 2021). Noe av grunnen til denne boligprisveksten kan være at folk har innsett at de trenger mer plass, som kan ha ført til et skifte i preferansen mot større boliger. Som et resultat av dette har betalingsviljen økt og presset prisene oppover.



Figur 15: Viser et skifte i etterspørselskurven, kilde: (Bang, Shift in demand curve, 2019, b).

2.1.3 Tilbud

Tilbudskurven er en motvekt til etterspørselskurven (figur 16). I et makroperspektiv vil tilbudskurven representere alle selgere i økonomien. Dersom prisen på en vare eller tjeneste stiger så stiger tilbudet av denne varen eller tjenesten (Cooper & Andrew, 2011, ss. 83-84). Tilbudet økes ved at det bygges nye boliger, og tilbudet synker ved blant annet at bygninger brennes ned eller at bygninger får endret bruk. Av det totale antallet boliger så utgjør bygging av nye boliger kun en liten del (NOU 2002:2, 2002, ss. 20-21).

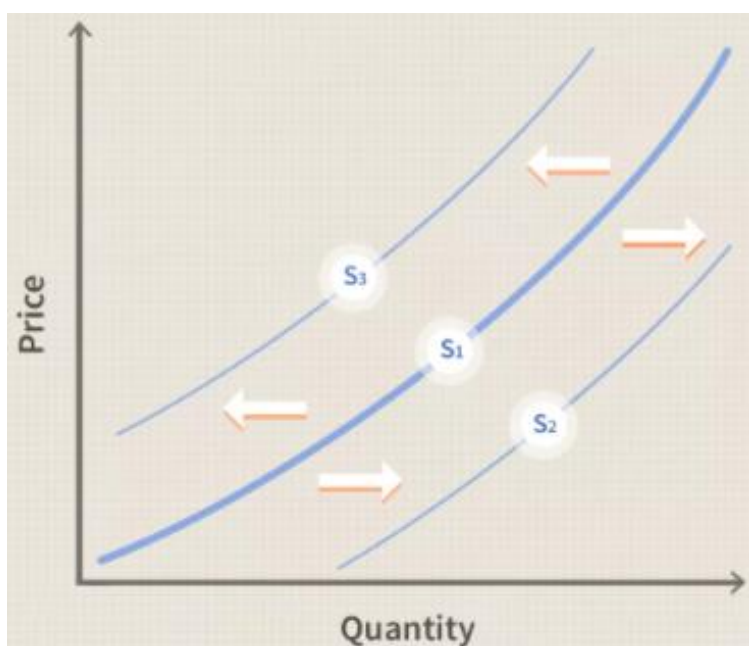


Figur 16: Viser tilbudskurven, kilde: (Bang, Supply Curve, 2019, c).

2.1.4 Skifte i tilbudskurven

Prisøkning på en vare eller tjeneste vil ha to effekter på tilbudssiden. Når prisen stiger vil flere selskaper komme inn på markedet som igjen fører til økt etterspørsel på disse varene og tjenestene. I tillegg vil alle eksisterende selskaper i markedet øke produksjonen som igjen fører til et høyere tilbud (Cooper & Andrew, 2011, ss. 83-84).

Følgende tilbudskurve (figur 17) forutsetter at når prisen på en vare eller tjeneste endrer seg vil alle de andre faktorer holdes konstant. Tilbudskurven vil i likhet med etterspørselskurven flytte seg dersom noen av disse faktorene endrer seg. Dersom tilbudskurven flytter seg til venstre vil man ha et lavere tilbud på et hvert prisnivå (Cooper & Andrew, 2011, ss. 84-85).



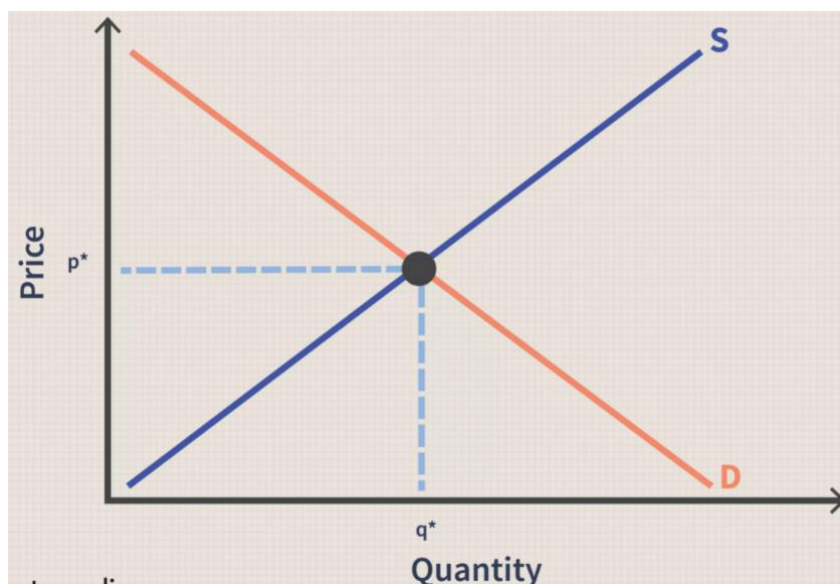
Figur 17: Viser et skifte i tilbudskurven, kilde (Bang, Shift in supply curve, 2019, d).

For at boligutviklere skal kunne bygge flere boliger er de nødt til å regulere tomtene som er tenkt til boligutvikling. Dersom denne prosessen tar lang tid, eller at man ikke får vedtatt en reguleringsplan, vil dette påvirke tilbudet av boliger, som igjen kan føre til et skifte i tilbudskurven. Andre faktorer som kan føre til et slikt skifte i tilbudskurven er økte byggekostnader eller dårlig tilgang på kapital.

2.1.5 Markedslikevekt

Det er i punktet hvor etterspørselskurven og tilbudskurven møtes at man finner markedslikevekten. I dette punktet er det like mange kjøpere og selgere som er villig til å

gjennomføre handelen til den prisen som er angitt av det punktet som er markedsliekevekten (Cooper & Andrew, 2011, ss. 85-87). Figur 18 viser punktet for markedsliekevkt.



Figur 18: Viser markedsliekevkten hvor S representerer tilbud, og D representerer etterspørsel, kilde (Bang, *Equilibrium Quantity*, 2019, e).

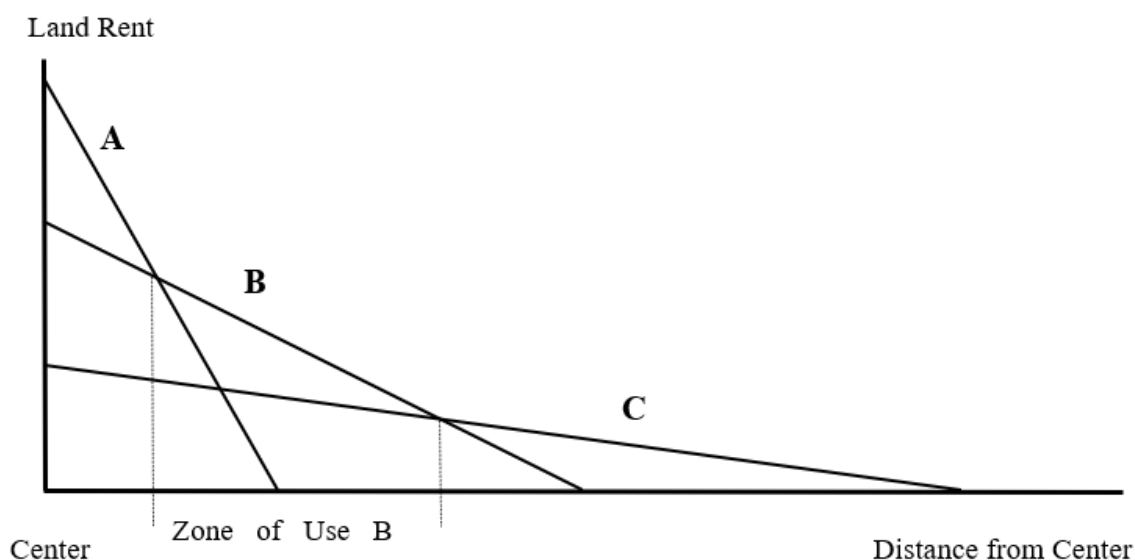
Tilbud- og etterspørselskurven vil krysses kun en gang, vist med det svarte punktet i figur 18. Dersom prisen på en vare eller tjeneste faller utenfor prisen for markedsliekevkt, vil markedsmekanismen fra teorien om tilbud og etterspørsel sørge for å gjenopprette balansen mellom tilbud og etterspørsel. La oss si at selskap A har en enebolig ute for salg til et gitt prisnivå. Dersom etterspørselen er lav på dette prisnivået er selskap A nødt til å senke prisen slik at flere blir interessert i å kjøpe denne eneboligen. Ved å gjøre dette gjenoppretter man balansen mellom kjøper og selger (Cooper & Andrew, 2011, ss. 85-87).

Som tidligere nevnt vil alle husholdninger som har høyere betalingsvilje enn krysningspunktet mellom tilbud- og etterspørselskurven har mulighet til å kjøpe seg en bolig. Over tid vil husholdningene som tilhører etterspørselssiden oppleve inntektsvekst som vil øke betalingsviljen. Det vil også være befolkningsvekst som gjør at behovet for boliger øker. Dermed vil dette flytte på etterspørselskurven. Dersom det bygges flere boliger enn det som tas vekk vil også tilbudskurven flytte seg (NOU 2002:2, 2002, ss. 19-20).

2.2 Bid-rent

Bid-rent teorien ble presentert av William Alonso i 1960 gjennom en videreutvikling av Von Thunen (1826) sin modell for utnyttelse av areal. I Alonso sin modell ble populasjon, hvor mange som jobber og leienivå introdusert som faktorer som påvirker kjøperens villighet til å bo nærme bysentrum. Alonso sin modell kunne derfor brukes for boligmarkedet (Alonso, 1960).

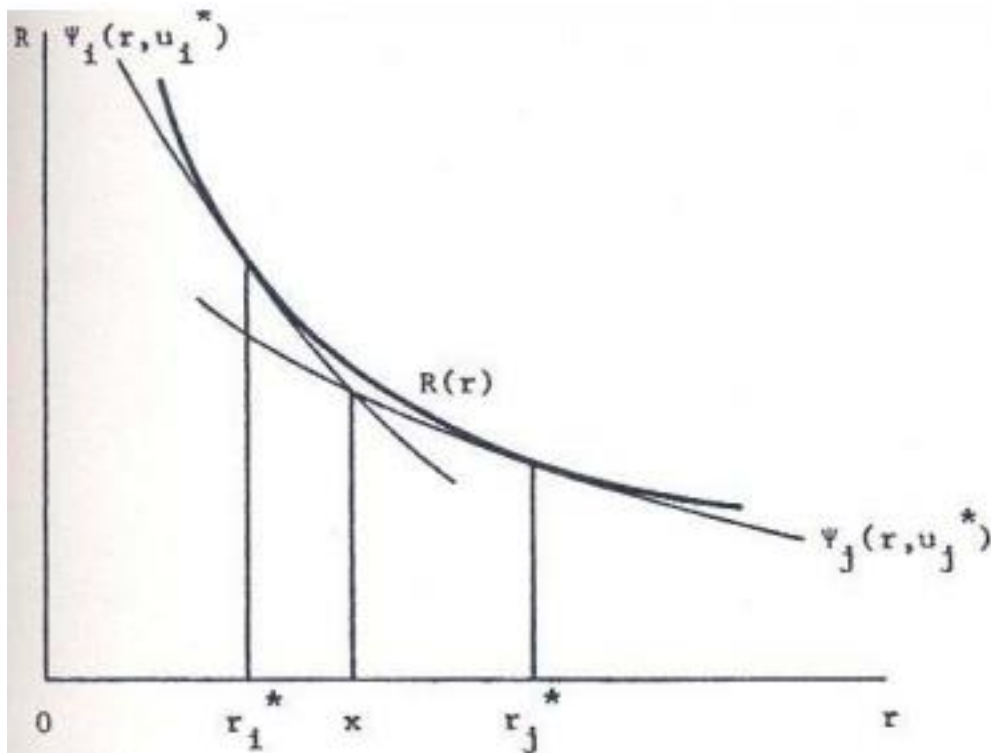
Bid-rent presentert av Geltner og Miller omhandler i hovedsak ulike markedsaktører. Teorien forklarer sammenhengen mellom transportkostnader og leiekostnader etter prinsippet om HBU (Highest Best Use) i et konkurranseutsatt marked, forstått som at markedene konkurrerer om de samme arealene. Bid-rent kurven representerer den maksimale betalingsviljen for et område. Hver aktør i markedet vil ha sin egen bid-rent funksjon. I vurderingen av bid-rent teorien vil det som nevnt være transportkostnadene som er avgjørende, hvor lavere transportkostnader vil medføre høyere betalingsvilje (Geltner et al, 2014, s.66). Eksempelvis vil en bedrift som baserer seg på vareleveringer ha en høy betalingsvilje for en lokasjon i nærheten av hovedveinettet. Geltner og Miller sin bid-rent teori kan illustreres slik:



Figur 19: Viser Bid-Rent funksjoner for 3 forskjellige bruksdunsjoner av en eiendom med forskjellig produktivitet og sensitivitet for transportkostnader, kilde: (Geltner et al, 2014, s.66).

Slik figuren viser, vil ulike næringsmarkeder ha ulike bid-rent funksjoner ettersom transportkostnadene endres. Kontor-, handel- og servicemarkedet er i stor grad avhengig av kunder for å lykkes, og må derfor plasseres i områder hvor det er størst konsentrasjon av kunder (kurve A). Dette vil i de fleste tilfeller være i Central Business District (CBD). I den forbindelse ser man en bratt kurve, da betalingsviljen drastisk endres når avstanden til CBD øker. Samtidig vil ikke virksomheter slik som lager og distribusjon være like avhengige av kunder for å lykkes, og kan plassere seg lenger unna CBD hvor transportkostnadene er lavere (kurve B). Kurven er dermed flatere sammenlignet med forrige marked ettersom transportkostnadene er mindre følsomme for avstand. På den andre siden vil ikke dette markedet konkurrere om de mest sentrale lokasjonene, da heller ikke betalingsviljen er stor nok. Til slutt finner vi «Residential», eller boligmarkedet (kurve C). I likhet med næringsmarkedene blir også boligmarkedet påvirket av avstanden til CBD. Jo nærmere lokasjonen er, jo mer stiger verdien. Boligmarkedet vil likevel ikke konkurrere om de mest sentrumsnære arealene, da betalingsviljen ikke er høy nok hos kjøperne. Med andre ord er boligkjøpere minst følsomme for transportkostnader (Geltner et al., 2014, s.66).

I senere tid har modellen til Alonso blitt utviklet en rekke ganger, blant annet av Masahisa Fujita (Fujita, 1989). Fujita viser til at lokasjonen for en husholdning i hovedsak baserer seg på en avveining mellom avstand til sentrum (pendlerkostnader) og prisen for tomten. Med andre ord blir det en konkret vurdering husstanden gjør sett opp mot inntektsnivået husstanden har, og hvor mange arbeidende husholdningen består av (Fujita, 1989). Resultatene i studien viser at single og par uten barn gjerne trekker inn mot sentrum uavhengig av inntektsnivå. Utenfor den indre sirkelen vil man finne større husholdninger, som for eksempel par med ett eller flere barn, med middels inntekt. I ytterkanten av byen vil de rikere og større husholdningene befinne seg. Fujita viser i sin modell at husholdningens plassering påvirkes av flere faktorer, og ved å sortere befolkningen etter inntektsklasser, kan man se skiller i betalingsviljen til de ulike klassene ved ulik lokasjon. Figur 20 viser hvordan bid-rent kurven for lav- og høy inntektsklasse er forskjellig, hvor høy inntektsklasse har lavere betalingsvilje for lokasjon i sentrum enn lav inntektsklasse. Malmø-Lund studien viser noe av de samme funnene presentert av Fujita. I studien gjort for eneboliger så man det at store boliger faller mindre i pris når avstanden økte sammenlignet med små boliger (Bohman & Nilsson, 2016).



Figur 20: Likevekt for ulike lokasjoner etter ulike inntektsklasser, kilde: (Fujita, 1989, s. 27). Den buede kurven viser prisnivået på boliger i ulik avstand til sentrum. R_i viser krysningpunktet for betalingsviljen for en husstand med lavere inntekt og boligprisen. R_j viser på sin side krysningpunktet for betalingsviljen for en husstand med høyere inntekt og boligpriser.

Blant annet Nicolas Coulombel har forsket videre med utgangspunkt i Fujita sin studie, og har gjennom studien «Residential choice and household behavior: State of the Art» vist til liknende funn som Fujita, altså at ulike inntektsnivåer skaper ulike markeder (Coulombel, 2010). Selv om funnene viser at inntektsnivået for husholdningen er en betydelig faktor i vurderingen om hvor husholdningen skal bosette seg, viser Coulombel at husholdningens sammensetning er av større betydning enn først antatt av Fujita (Coulombel, 2010).

Utredningen av boligmarkedene og boligpolitikken gjort av Boligutvalget i 2002 peker også på at både husholdningens sammensetning og husholdningens inntekt er faktorer som påvirker betalingsviljen for å bo sentralt i en by. I utredningen vises det til en lineær sammenheng med antall arbeidende, og betalingsviljen for å bo sentralt. Ved en dobling av personer i husholdningen med arbeid, dobles også betalingsviljen for en sentral bolig. Når det kommer til forholdet mellom pendlerkostnad, og boligpris, peker utredningen på at et slikt forhold tilsier at prisen på små boliger blir påvirket mindre av avstand, enn hva store boliger gjør. Forholdet mellom reisetid og boligpris kommer av at lavtlønnede husholdninger har lavere tidskostnader sammenlignet med de som har høy lønn (NOU 2002:2, 2002, ss. 30-31).

Forklart på en annen måte betyr det at tiden lavtlønnede bruker på å pendle koster dem mindre penger i lønn sammenlignet med de med høy inntekt. Denne tolkningen er en motsetning av hva studiene (Fujita, 1989), (Coulombel, 2010) og (Bohman & Nilsson, 2016) har, da boligutvalgets utredning tilsier at store boliger er mer sensitive for avstand enn hva små boliger er.

NOUen tolker resultatene om at prisen på små boliger blir påvirket mer av avstand til sentrum enn prisen på store boliger som at lavtlønnede har større behov for å bo sentralt fordi reisekostnadene er relativt sett høyere i forhold til lønnsnivået. På den andre siden viser NOU2002:2 at det er en motsatt effekt med utgangspunkt i at reisetiden er mindre kostbar for lavtlønnede enn den er for høytlønnede (NOU 2002:2, 2002).

I en studie gjort av Lemoy, Raux og Jensen med utgangspunkt i Alonso-Mills-Muth metoden (monosentrisk modell) hvor datasettet var fiktivt og randomisert, ble det vist til at en større by gjerne har ring-effekter med tanke på hvor personer lokaliserer seg basert på inntekt (Lemoy, Raux, & Jensen, 2016). Studien viste at bysentrum i hovedsak bestod av husholdninger med høyere inntekt, etterfulgt av en ring med husholdninger som har lavere inntekt. I ytterkant av byen fant man en ring hvor det igjen var rikere husholdninger. Det at husholdninger med høyere inntekt bor i sentrum støttes av en annen studie gjort for en rekke Britiske storbyer gjort av David Cuberes og Jenniffer Roberts (Cuberes & Roberts, 2015). I denne studien viser funnene at de som flytter eller bor i bysentrum, er gjerne de som har opplevd en stor økning i inntekten over kort tid.

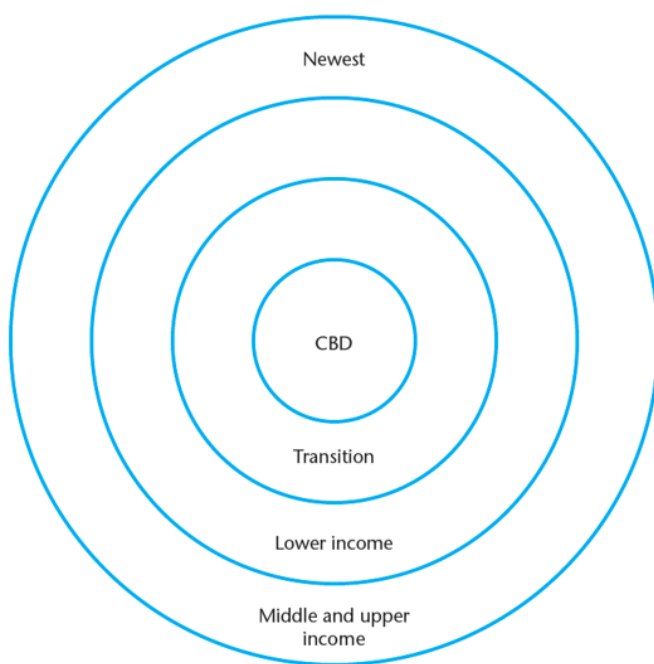
2.3 Monosentrisk byteori

Det er også hensiktsmessig å beskrive hvordan prisutvikling er en faktor av boligens beliggenhet sett i et større bilde. Den monosentriske byteorien omhandler nettopp dette, og tar hensyn til forholdet mellom boligens plassering i forhold til bykjernen og gir en forklaring på prisdannelse i boligmarkedet.

Den monosentriske byteorien har stått sentralt i forståelsen for hvordan og hvorfor en by vokser størrelsesmessig. William Alonso med sin studie «The Historic and the Structural Theories of Urban Form: Their Implications for Urban Renewal» (Alonso, 1964), Edwin S. Mills med sin studie “Studies in the Structure of the Urban Economy” (Mills, 1972) og Muth

(1969) har alle vært sentrale i utviklingen av teorien. I senere tid har teorien blitt gjengitt og utviklet en rekke ganger, blant annet av Geltner og Miller (Geltner et al, 2014, s.69). Det er først og fremst deres tolkning av teorien vi baserer oss på.

Geltner og Millers monosentriske byteori forklarer hvordan byen vokser fysisk og økonomisk sett. I likhet med de første variantene av teorien, baserer Geltner og Miller sin teori seg på en by med kun ett sentrum, også kalt CBD (Central Business District). CBD er midtpunktet i den sirkulære byen hvor alle jobber, og omringes av boligdelen av byen som igjen stopper i ytterkant/randsonen (Geltner et al, 2014, s.67).



Figur 21: Monosentrisk byteori, kilde:(Geltner et al. 2014, s.85).

I Geltner og Millers tolkning av teorien presenteres verden svært forenklet. For å på en enkel måte forklare prinsippet, må antagelsene om at alle husholdninger er like, og har en inntekt (Y) legges til grunn. Disse husholdningene bruker inntekten kun på bolig, reise og andre konsumvarer, representert med (X). Reisekostnaden avhenger av to faktorer; selve tiden det tar å komme seg til CBD, og de faktiske kostnadene slik som billettpriser. I modellen anses begge som en kostnad, noe som betyr at jo lengre unna CBD en person kommer, desto høyere er kostnaden i form av tid. I denne modellen antas det også at alle hus og tomter er identiske. Den eneste variabelen er avstanden til CBD.

Kostnadene tilknyttet bolig blir videre fordelt mellom leie av tomt og leie av bolig. Leieprisen for tomt på sin side styres av alternativkostnaden, og blir av Geltner og Miller målt basert på avkastning av å dyrke mark (jordbruk) (Geltner og Miller 2014, s. 68). Dette fører til at leie av bolig er billigst i ytterkanten/randsonen av byen, hvor boligprisene er lik byggekostnadene. Boligprisene vil som følge av denne funksjonen øke jo nærmere CBD man beveger seg i byen, da reisekostnadene blir mindre.

I Geltner og Miller sin teori blir byen større som et resultat av tre faktorer; befolkningsvekst, reduksjon i transportkostnader eller økt inntekt. Faktorene gjør at etterspørselen etter boliger øker og at boligprisene øker. En reduksjon i transportkostnadene vil også kunne oppnå samme effekt, da selve kostnaden for å komme seg fra bolig til CBD reduseres, og marginen for å bygge nye boliger blir noe større. Det blir også pekt på en fjerde faktor som vil påvirke utviklingen av byen, og dette er en endring i alternativkostnaden. Som nevnt, er leiekostnadene styrt etter avkastningen av å dyrke mark. Dersom avkastningen i jordbruket reduseres, vil også leieprisene på tomten gå ned fordi jordbrukseieren krever mindre for tomten.

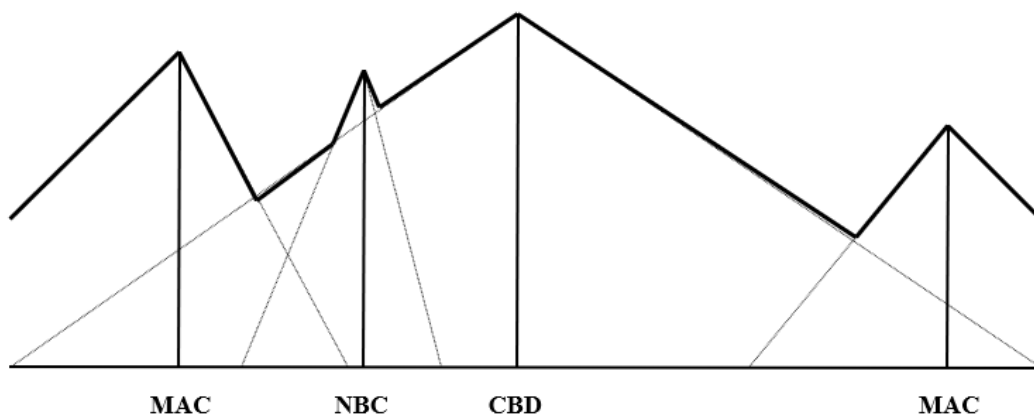
Dersom betalingsviljen til kjøperne er høyere enn summen av bygge- og tomtekostnadene, vil en by utvikle seg ettersom det er profitt å hente for utbyggere. Det antas også at ny utbygging i noen tilfeller kun skjer i randsonen som et resultat av at det ikke er tilgjengelige arealer for utbygging innenfor. Byutvikling vil til tross for dette skje på flere måter, hvorav de kanskje mest åpenbare er gjennom transformasjon eller byfortetting. Ved transformasjon endrer man bygningsfunksjonen til å romme mennesker, og ved byfortetting forsøker man å bygge byen tettere og tettere.

Som en effekt av denne teorien priser man boligene etter lokasjonspremien boligen medfører i tillegg til selve verdien av boligen. Med andre ord medregnes reisekostnadene, eller reduksjonen i reisekostnadene, ut fra boligens lokasjon. Som nevnt er teorien forenklet på flere måter. Ingen mennesker er like, og boligkjøp regnes for å være svært subjektivt. En bolig dekker mange behov, og disse behovene verdsettes ulikt fra person til person. Videre vil det svært sjeldent kun være ett CBD, men flere CBD spredt over store områder. Overordnet sett ser man at storbyer blir utviklet fra ett sentrum, og boligpriser påvirkes av avstanden til sentrum. For denne studien er det derfor interessant å se om teorien kan overføres til et annet type sentrum, altså togstasjon. Det vil også være av interesse å se om teorien er lik på tvers av

leilighetsstørrelser og priser, eller om det er forskjeller innenfor et boligsegment. Tidligere studier viser til funn som antyder at det kan være en forskjell innenfor et segment (Bohman & Nilsson, 2016) (Gramm, 2016).

2.5 Polysentrisk byteori

En annen teori som også vil være av interesse er teorien om polysentriske byer. Med polysentrisk teori menes det en by eller en region med flere sentrum, og er sånn sett en utvidelse av den monosentriske teorien (Geltner et al, 2014, s. 89). Teorien vil ikke være like overførbart som den monosentriske teorien, da studien ikke tar for seg Oslo eller nærhet til Oslo som en variabel. Likevel forutsetter vår studie at Oslo regnes for å være CBD for den utvalgte regionen. Teorien vil derimot ytterligere belyse forholdet mellom inntekt og lokasjon i tillegg til å legge et grunnlag for å sammenligne Lillestrøm, Ski og Sandvika.



Figur 22: Leieprisnivåer i en polysentrisk (flerkjernet) region, kilde:(Geltner et al, 2014, s. 89).

Geltner og Miller sier at i den virkelige verdenen er det byer og regioner som består av flere mindre steder som alle har sitt eget lokale sentrum, slik som figur 22 viser. Disse vil gjerne være i forbindelse med havneområder, flyplasser, sportssentre, medisinske sentre eller kollektive knutepunkt. Figuren over illustrerer hvordan den monosentriske teorien blir ført et steg videre. CBD ligger fortsatt som det store sentrumet, men hvor det også eksisterer mindre, mer lokale sentrum som tilfører en funksjon, også kalt NBC (neighborhood business center) (Geltner et al, 2014, s. 89). Lillestrøm, Ski og Sandvika vil alle på hver sin måte bli påvirket av tilknytningen til Oslo, hvor denne tilknytningen kan være noe av forklaringen for de høye boligprisene. Samtidig oppfyller hvert av disse stedene en egen funksjon, da alle har kjøpesentre, sykehus, skoler, godt etablert infrastruktur og ikke minst, har et kollektivt

knutepunkt direkte inn til Oslo. På mange måter kan Lillestrøm, Ski og Sandvika anses som NBC-er, da de har en tilknytning til Oslo, men ikke er store nok til å være en MAC. Med MAC menes det «major activity center», og regnes for å være andre storbyer (Geltner et al, 2014, s. 89).

Den polysentriske teorien baserer seg også på at ikke alle personer har lik inntekt. De med lavest lønn vil måtte trekke ut mot randsonen som et resultat av at boligprisene er lavere, og de med høyere lønn vil på sin side trekke inn mot sentrum. Til forskjell fra monosentrisk teori, tilfører den polysentriske teorien et konsept som tilsier at det også vil være en tiltrekningskraft for å etablere seg i randsonen. Grunnen for denne tanken er at etableringskostnadene er langt lavere. Med etablering av bedrifter og funksjoner i randsonen, vil det igjen tiltrekke beboere, da disse bedriftene og funksjonene tilbyr arbeidsplasser og oppfyller hverdagsbehov. Det vil dermed etter hvert skapes nye og mindre sentrum i ytterkanten av den eksisterende storbyen.

Funn fra tidligere studier støtter oppunder både mono- og polysentrisk teori, og hvordan avstand til sentrum påvirker boligprisene. For Oslo-regionen har det blitt gjort studier for blant annet togstrekningen Oslo-Jessheim, hvor det ble pekt på at økt avstand til togstasjon jevnt over har en negativ effekt på boligprisene (Sandhu & Sarna, 2015). I tillegg ble det gjort funn som tilsier at denne effekten er ulik desto lenger unna byen ligger Oslo. Funn gjort i en studie av strekningen Oslo-Moss støtter oppunder dette (Gramm, 2016). Hver by har sin egen monosentriske effekt på boligprisene, men det foreligger også en ekstra polysentrisk effekt når man ser på Oslo som CBD.

Funn fra tidligere studier støtter oppunder både mono- og polysentrisk teori, og hvordan avstand til sentrum påvirker boligprisene. For Oslo-regionen har det blitt gjort studier for blant annet togstrekningen Oslo-Jessheim, hvor det ble pekt på at økt avstand til togstasjon jevnt over har en negativ effekt på boligprisene (Sandhu & Sarna, 2015). I tillegg ble det gjort funn som tilsier at denne effekten er ulik desto lenger unna byen ligger Oslo. Funn gjort i en studie av strekningen Oslo-Moss støtter oppunder dette (Gramm, 2016). Hver by har sin egen monosentriske effekt på boligprisene, men det foreligger også en ekstra polysentrisk effekt når man ser på Oslo som CBD.

2.6 Hypotese 1

Med en forståelse for den monosentriske byteorien er hypotese 1 utarbeidet for å teste hvilken effekt avstand til togstasjonen har på leilighetspriser. Gjennom denne hypotesen vil det kunne besvares om leilighetssegmentet blir påvirket av avstand til togstasjon på et overordnet plan, samtidig som det vil gi en bedre innsikt i hovedproblemstillingen for denne masteroppgaven.

Monosentrisk byteori:

Hypotese 1

- H1: Leilighetspriser blir påvirket av avstand til togstasjon
- H0: Leilighetspriser blir ikke påvirket av avstand til togstasjon

2.7 Hypotese 2

I den presenterte mikroøkonomiske teorien, bid-rent teorien og polysentrisk teorien er det en underliggende enighet om at etterspørselen etter bolig endres i takt med inntekt. På bakgrunn av dette ønsker vi å teste om monosentrisk teori påvirker leilighetssegmentet i forskjellig grad når vi kategoriserer leilighetene etter størrelse. Boligmarkedet er et sammensatt heterogent marked og det vil derfor være interessant å studere forskjeller i et tilsynelatende homogent leilighetsmarked. I studien for Malmø-Lund regionen i Sverige samt Boligutvalgets utredning ble det gjort funn som antydte at de ulike leilighetsstørrelsene representerte ulike inntektsnivåer og kjøpegrupper (Bohman & Nilsson, 2016) (NOU 2002:2, 2002, ss. 30-31). Med dette som et utgangspunkt, burde teorien om at avstand påvirker boligprisen være varierende for de ulike leilighetsstørrelsene, da det basert på inntektsnivåer er ulik etterspørsel etter ulike leilighetsstørrelser. Hypotesen vil bli testet for hver by separat.

Hypotese 2 er dermed:

- H1: Kvadratmeterprisen for ulike leilighetsstørrelser blir påvirket i ulik grad av avstand til togstasjon
- H0: Kvadratmeterprisen for ulike leilighetsstørrelser blir ikke påvirket i ulik grad av avstand til togstasjon

3 Metode og datamateriale

3.1 Oppbygging av oppgaven og valg av forskningsdesign

I denne studien vil det være hensiktsmessig å benytte seg av en kvantitativ metode med et kausalt design (Sukamolsen, 2007). Dette er ønskelig ettersom målet for studien er å påvise, eller avkrefte en eventuell statistisk sammenheng for et fenomen. Med kausalitet menes det årsakssammenheng, og et slikt design brukes ofte når en ønsker å påvise sammenheng mellom to variabler, som i dette tilfellet er sammenhengen mellom avstand til togstasjon og kvadratmeterprisen for leiligheter (Sander, 2019). Mye av teorien rundt kausalitet kan spores tilbake til David Hume sin analyse av kausalitet. Han mente at kausalitet hadde 3 kjennetegn. Det første kjennetegnet er «Asymmetri» og menes med at årsaken kommer før virkningen (Skog, 1998, s. 16). For eksempel, leiligheter ligger x antall meter fra togstasjonen og har en størrelse på x antall kvm. Både nærhet til togstasjonen og størrelse vil påvirke et salg. Årsaken i denne sammenheng, er dermed at leiligheten blir bygd med ulike attributter som fører til salg, hvor selve salget er virkningen. Det andre kjennetegnet er «lokalitet», som betyr at virkningen skal skje nær årsaken i både tid og rom. Dette kjennetegnet er omdiskutert, men Hume mente at det må foreligge en ubrutt kausalkjede fra «fortidig forhold og frem til nåtid» (Skog, 1998, s. 16). For denne studien skjer virkningen svært nærme årsaken hvor årsaken er at noen selger en leilighet som fører til virkningen at noen kjøper leiligheten. Det siste kjennetegnet er «konstant konjunksjon» som vil si at for hver gang det forekommer en årsak påvirker det virkningen. Dette kjennetegnet er også omdiskutert. Grunnen til dette er fordi virkningen skal komme hver gang det foreligger en årsak. For at dette skal skje må omstendighetene være de samme hver gang. Dersom omstendighetene endres fra gang til gang kan man risikere at virkningen uteblir (Skog, 1998, s. 16). Når en leilighet er solgt (årsak), så må det være en som har kjøpt denne leiligheten (virkning). I denne sammenheng vil årsak alltid utløse en virkning.

Valget av kvantitativ metode underbygges av datamaterialets størrelse. For å kunne studere en slik sammenheng er man avhengig av større mengder data, noe som vil være lite hensiktsmessig å tilegne seg gjennom en kvalitativ metode.

For å kunne tilegne seg store mengder med data, er det normalt å bruke sekundærdata innhentet av en annen institusjon. De er gjerne hentet for et formål av institusjonen, og brukes

til et annet formål av undersøkeren. Syed Muhammad Sajjad Kabir nevner derimot at gjennom bruken av sekundærdata medføres det en økt risiko lavere validitet, til tross for at de er hentet av en anerkjent institusjon. Validitetsrisikoen er igjen knyttet opp mot at dataene i utgangspunktet ikke ble innhentet for samme formål, og må vurderes i forbindelse med databearbeiding (Kabir, 2016, ss. 204-205). For denne studien har Eiendomsverdi stått sentrale i datainnhenting. Gjennom en særegen avtale har vi fått tilgang til et uttrekk fra deres databaser.

I forbindelse med denne studien hvor målet er å kunne påvise sammenhenger og trender, vil det måtte vurderes validitet og pålitelighet/reliabilitet av både data og funn. I boken «Discovering Statistics Using SPSS» av Andy Field beskrives validitet som at metoden og dataene faktisk måler det som skal måles (Field, 2009, s. 11). Validitet kan igjen deles inn i to deler; ytre og indre validitet (Skog, 1998, ss. 85-87). Den ytre validiteten omhandler i hvilken grad studiens resultater kan generaliseres. Eksempelvis vil den ytre validiteten for denne studien omhandle i hvilken grad funnene fra studiens begrensede område, både i form av geografisk lokasjon, tidsperspektiv og tematikk kan generaliseres for et større område. Den indre validiteten omhandler i hvilken grad et forsøk eller en studie kan forklares gjennom utvalgte hypoteser. Et eksempel på indre validitet er det som heter definisjonsvaliditet, og går ut på at en valgt indikator eller variabel faktisk måler det den skal måle. Med andre ord er definisjonsvaliditet et mål på samsvaret mellom en indikator og den teoretiske definisjonen. For denne studien vil det bety at variablene faktisk måler det som skal måles. For å oppnå høy indre validitet må blant annet attributter eller variabler tilknyttet leiligheter i tydelig grad operasjonaliseres og defineres slik at de måler det de faktisk skal måle.

Field sier videre at validitet ikke er et godt nok vurderingsgrunnlag alene, og må også ses i sammenheng med reliabiliteten til metoden. Med reliabilitet menes det at metoden eller den statistiske analysen produserer de samme resultatene flere ganger, og at tolkningen av resultatene er den samme (Field, 2009, s. 12). En høy reliabilitet oppnås når den samme testen viser de samme resultatene om igjen, eksempelvis at det er en sammenheng mellom avstand til togstasjon og kvadratmeterpris på leiligheter. I statistiske analyser omtales gjerne signifikans som et mål på reliabilitet, hvor høy signifikans betyr at testen er pålitelig.

3.2 Hedonisk metode

Teoriene om tilbud og etterspørsel samt mono- og polysentriske byer har gitt et grunnlag for hvordan prisdannelsen i byer oppstår og utvikler seg. I tillegg har det gitt en forståelse for de overordnede økonomiske kreftene som påvirker boligmarkedet. Det er derimot viktig å trekke inn subjektive preferanser, da kjøpergruppen i boligmarkedet varierer mye, slik teorien om mikro-økonomi og bid-rent poengterer.

Hedonisk metode ble først presentert av Sherwin Rosen som en måte å måle verdien av goder tilknyttet en bolig på (Rosen, 1974). Metoden til Rosen stammer fra tanken om at en bolig, som en hvilken som helst vare, består av et sett med attributter. Attributtene kan være alt fra størrelse og etasje, utsikt, nærhet til sentrum eller alder. Med andre ord er det alle direkte og indirekte goder som følger med boligen. Verdien av attributtene vil ikke være synlige i seg selv (implisitt), men uttrykkes gjennom salgsprisen på boligen. Gjennom å bryte opp en bolig ned til de ulike attributtene kan man estimere den implisitte prisen hvert attributt har (Herath & Maier, 2010, s. 1). Denne måten å estimere attributtverdien på har blitt brukt i en rekke studier med ulike formål. Den kan anvendes for å studere boligprisutvikling, etterspørsel etter boliger eller forskjeller i boligsammensetningen på nabolagsnivå. Nabolagkarakteristikker er studert i en rekke studier, men interne attributter ved selve boligen er derimot et tema som har blitt studert i mindre grad (Herath & Maier, 2010, s. 1).

I Liv Oslands studie ble interne attributter for boliger i Haugesund kommune studert (Osland, 2001). Formålet med studien var å estimere den implisitte prisen som er en forklaring på betalingsviljen for en rekke interne boligattributter. I studien ble blant annet noen av disse attributtene inkludert: garasje, eneboligens størrelse, alder, antall rom og utsikt (Osland, 2001). Attributtene som inngår i en bolig, kan i hovedsak deles opp i to grupper: attributter til selve boligen som f.eks. boligareal og innredning (interne), og attributter som er eksterne som f.eks. avstand til bysentrum, avstand til skole, utsikt, miljø eller klima. Eksterne attributter er blant annet omtalt i artiklene «Hedonic Estimation of Housing Market Prices: A Further Comment» i tidsskriftet *The Review of Economics and Statistics* (Straszheim, 1974), og i studien «Segmentation in urban housing markets» (Schnare & Struyk, 1976).

SSB har brukt den hedoniske metoden siden 1992 for å studere og predikere historisk og fremtidig prisutvikling på boliger (Takle, 2012, s. 14). Det Europeiske statistikkbyrået, også

kalt Eurostat, skriver i sin rapport fra 2013 at den hedoniske metoden mest sannsynlig er den beste metoden for å ta hensyn til de ulike attributtene som følger en eiendom (Eurostat, 2013, s. 57).

Det er en generell enighet om at den hedoniske metoden burde brukes for å forstå hvordan ulike variabler kan påvirke boligpriser. For denne studien vil metoden være sentral i forståelsen for hvordan og hvorfor nærhet til togstasjon kan være et attributt som etterspørres. Gjennom å bryte opp en bolig i et sett med attributter kan de studeres hver for seg, og på denne måten kan også betalingsviljen for nærhet til togstasjon som en ekstern variabel estimeres. En utfordring vil være å inkludere alle variabler som kan være relevante. For det første er leiligheter forskjellige, og kan ha ulike attributter. For det andre er også kjøpegruppen ulik, hvor attributter av verdi for en person, ikke nødvendigvis har en betydning for noen andre. En detaljert begrunnelse for valg av attributter brukt i denne studien vil gis i neste delkapittel.

3.3 Attributter

Den hedoniske prisfunksjonen sier at prisen på en bolig er totalsummen av alle variabler som inngår i en bolig (Rosen, 1974). Som vi har vært inne på, er et boligkjøp svært subjektivt, og hva som regnes for å være et attributt av betydning for en kjøper, ikke vil være like viktig for en annen. Det har blitt gjort få empiriske studier på hva som regnes for å være viktige attributter, og det er derfor også vanskelig å si noe om hvilke attributter som er riktige å inkludere i denne studien. En av de få studiene som har blitt gjennomført, er Hårsman sin studie «Housing demand models and housing market models for regional and local planning» hvor det ble gjennomført en intervjuundersøkelse for Stockholm. Analysen konkluderte med at boligareal, standarden på boligen, avstand i tid til sentrum og boligtype er attributter som verdsettes (Hårsmann, 1981). Seppo Laakso har i sin studie «Urban housing prices and the demand for housing characteristics», laget en oversikt over attributter inkludert i tidligere empiriske studier, hvorav disse er relatert til selve boligen (Laakso, 1997):

- Boligens størrelse (antall rom, eller m²)
- Tomtestørrelse
- Boligens/bygningens alder

De valgte attributtene er svært like med de som ble brukt i en Nederlandsk studie av Eric Pels. Studien så på hvordan avstand til togstasjon påvirket boligpriser. I studien ble det benyttet disse boligrelaterte attributtene (Pels, 2005):

- Boligens størrelse (m²)
- Bygningens alder
- Antall rom
- Antall bad

Det må også tas stilling til hvor mange attributter som burde inkluderes i analysene. Laakso viser til at det er hensiktsmessig å begrense antall attributter. I Laakso sin studie fra Helsinki ble det brukt 4-5 attributter. Ved å begrense antall variabler vil man i større grad kunne spesifisere modellen og dermed oppnå reliable estimater og resultater som enklere kan tolkes (Laakso, 1997, s. 47). På den andre siden vil for få variabler resultere i mangelfulle modeller, da det ikke er nok variabler etter det teoretiske grunnlaget. Studien sier derimot lite om hvor mange variabler som er mange nok (Laakso, 1997). Studien viser også til i hvor stor grad de ulike attributtene påvirker prisen. Naturlig nok påvirker boligens størrelse prisen, der større leiligheter er relativt sett billigere enn små leiligheter målt i pris pr m². Studien konkluderte også med at boliger med en større, ubebygd tomt rundt relativt sett var større enn en bolig hvor tomten var helt bygd ut. (Laakso, 1997).

Boligens alder påvirker også prisen på boligen, men ikke kun i negativ forstand. Frem til et visst punkt påvirkes prisen negativt, men vil igjen stige når bygget overstiger en viss alder. Studien fra Helsinki viser også til en positiv sammenheng mellom avstand til togstasjon og boligprisen. Jo nærmere togstasjonen boligen befinner seg, jo høyere er prisen (Laakso, 1997). Christopher Hewitt og W.E Hewitt viser derimot gjennom studien «The Effect of Proximity to Urban Rail on Housing Prices in Ottawa» at det kan være en negativ prispremie i forbindelse med avstanden. At boligen ligger for nærme togstasjonen har ikke alltid positive virkninger, da det blant annet medfører økt støy, sjenerende utsikt og i enkelte tilfeller økt kriminalitet (Hewitt & Hewitt.W.E, 2012, s. 6)

Det må også vurderes på hvilken måte leilighetene skal segmenteres. Laakso sin studie viser videre til amerikanske studier som har tatt for seg nettopp denne problemstillingen. I disse studiene har antall rom, eller antall soverom vært et mål på boligens størrelse, istedenfor antall m² (Laakso, 1997, s. 220). Laakso retter derimot noe kritikk til å avgjøre boligens

størrelse basert på antall rom eller soverom, da det gjør tolkning av resultatene noe vanskeligere. Til tross for dette viste studiene fortsatt at prisen økte med størrelsen (Laakso, 1997). Disse funnene stemmer godt overens med tidligere studier for Norge og Sverige, hvor boligens størrelse var målt ut fra hvor mange kvadratmeter den er på (Gramm, 2016) (Sandhu & Sarna, 2015) (Bohman & Nilsson, 2016) (Osland, 2001).

Studien gjort av Laakso peker derimot på en mulig feilkilde i den økonometriske analysen ved å bruke slike variabler. Mange av variablene henger sammen, og vil derfor kunne påvirke hverandre (Laakso, 1997). Et eksempel her kommer av den monosentriske teorien, hvor byer blir utviklet innenfra og ut. Det følger naturlig nok at de byggene eller boligene som er nærmest sentrum er eldst, og når alder er en variabel som har innvirkning på prisen, vil altså lokasjon og alder påvirke hverandre og kunne gi et multikollineært utslag på prisen. Med multikollinearitet menes det at to uavhengige variabler, for eksempel avstand til sentrum og boligalder, påvirker hverandre samtidig som de påvirker boligprisen. Et slikt forhold forsterker påvirkningseffekten på boligprisen og er lite ønskelig i statistiske analyser. Laakso påpeker videre at et større datasett vil redusere problemet med multikollinearitet (Laakso, 1997, s. 97). Hvor stort datasett en studie må ha for å redusere dette problemet blir ikke spesifisert.

I forbindelse med den hedoniske pristeorien, så Liv Osland i sin studie på bruken av attributter. I studien brukte hun følgende attributter: Alder, garasje, avstand til sentrum, utsikt, peis, antall WC, hybel, enebolig, terrassehus, kjedebolig, soverom, kjeller og parkett. (Osland, 2001).

Det er tydelig at antallet attributter som inkluderes i slike studier varierer mye, noe Laakso også har påpekt. (Laakso, 1997, s. 48). Selv om antall attributter brukt i tidligere studier varierer vil attributtene som har blitt brukt tidligere være et godt grunnlag for valg av attributter som skal tas med videre. Det er verdt å merke seg at dette fungerer kun som et vurderingsgrunnlag som vi vurderer vår studie opp mot. Valg av attributter i denne studien vil først og fremst være styrt av de dataene som har vært tilgjengelige fra Eiendomsverdi.

3.4 Analyseverktøy

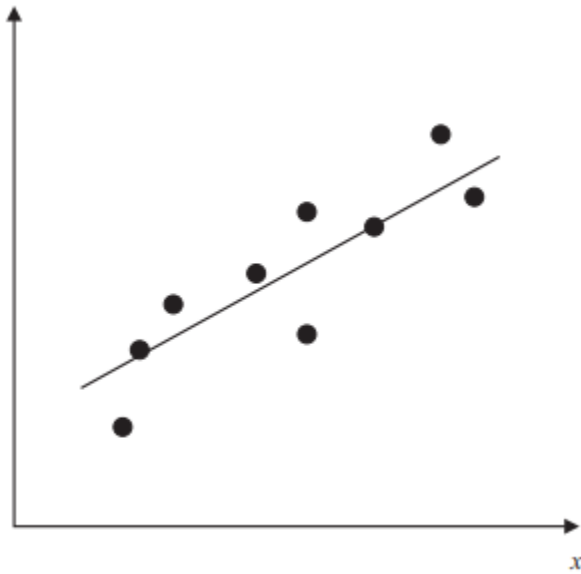
Statistikk er et sentralt verktøy i alle studier hvor kvantitative data brukes. Det er «vitenskapen for planlegging av undersøkelser, innsamling og presentasjon av tallmateriale, og analyse og beslutninger ut fra innsamlede data» (Bjørnstad, 2018). Den hedoniske metoden vil ligge som et grunnlag for valg av fremgangsmåten for studien, og sier i tråd med teorien om etterspørsel og tilbud at boligprisen er et resultat av en rekke attributter tilknyttet boligen og ulike preferanser hos kjøperne. For å kunne analysere eventuelle sammenhenger mellom de attributtene som inkluderes i denne studien, vil regresjonsanalyse være det mest aktuelle. Problemstillingen er som tidligere nevnt rettet mot sammenhengen mellom avstand fra togstasjon og kvadratmeterprisen. I denne delen presenterer vi en overordnet teoretisk framstilling av hvordan en regresjonsanalyse er bygd opp. Senere vil studiens spesifikke modell tydeliggjøres og presenteres.

3.4.1 Regresjonsanalyse

En regresjonsanalyse prøver å beskrive en sammenheng mellom en bestemt variabel og en eller flere andre variabler (Braut & Dahlum, 2018). Rent matematisk har man en variabel y og variabel x_1 , hvor man ser på endringer i y ut ifra endringer i x_1 . Man skiller mellom enkel regresjonsanalyse og multipl regressjonsanalyse (Thrane, 2017, ss. 23-27).

Enkel lineær regresjon

I enkel lineær regresjonsanalyse har man en variabel y som er den avhengige variabelen og kun en variabel x som er den uavhengige variabelen. La oss si at man skal se på sammenhengen mellom boligpris og boligareal, hvor boligpris er den avhengige variabelen y og boligareal er den uavhengige variabelen x . For at man skal kunne se på sammenhengen mellom disse variablene er man nødt til å samle inn data som sier hvor mye forskjellige boliger har blitt solgt for, og hvor store disse er. Etter at dataene er innhentet legger man dette inn i en modell og finner ut hvor mye boligprisen endrer seg når boligareal endrer seg 1 m^2 (Thrane, 2017, ss. 23-27). Figur 23 viser dette i et scatterplot.



Figur 23 Scatterplot som viser enkel lineær regresjon mellom en avhengig variabel og en uavhengig variabel. Den svarte linjen er regresjonslinjen, kilde: (Brooks, 2008, s. 31).

Den svarte linjen er regresjonslinjen, altså den best tilpassede rette linjen mellom de svarte observasjonspunktene. Ligningen for regresjonslinjen, som også viser sammenhengen mellom variablene y og x er gitt ved:

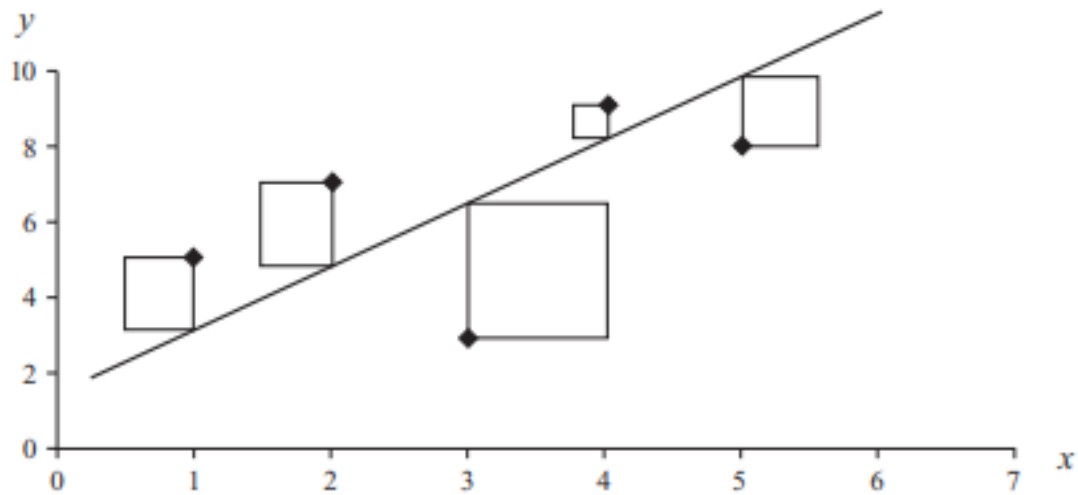
$$y = \alpha + \beta x$$

α representerer det punktet hvor regresjonslinjen krysser y -aksen, og β representerer helningen eller stigningen på regresjonslinjen. Eksemplet som tidligere er blitt introdusert kan matematisk beskrives slik: Boligpris = $\alpha + \beta$ boligareal, hvor modellen sier at all bevegelse i boligpris skjer med endring i boligareal (Thrane, 2017, ss. 28-29). I den virkelige verden vet man at andre variabler også vil påvirke boligprisen. Måten man har løst dette problemet på er å legge inn en feilterm (Residualen) u . Feiltermen tar hensyn til alle variabler som ikke tas med i modellen. Dette gjør at man får en ny ligning som er gitt ved (Thrane, 2017, ss. 43-45):

$$y = \alpha + \beta x + u$$

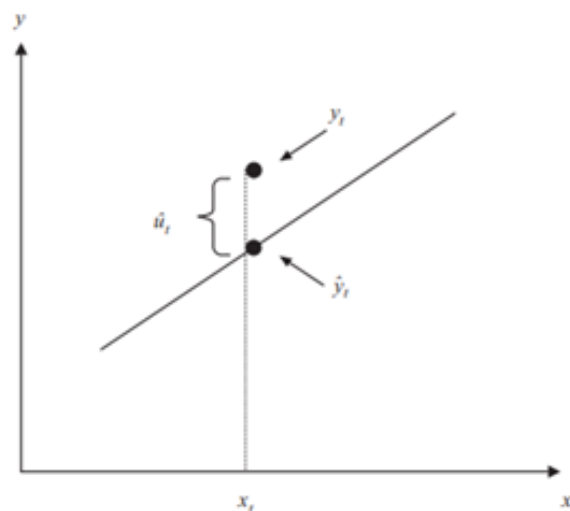
Posisjonen til regresjonslinjen er bestemt av α og β , så alt man trenger å vite er verdiene til α og β som er beregnet ut ifra observasjonene man har hentet inn. Det er forskjellige måter å regne ut disse verdiene på, men den mest kjente er minste kvadratsums metode. Denne metoden går ut på at man tegner inn en regresjonslinje som best passer punktene på alle

observasjonene i grafen. Naturligvis vil man ikke treffe alle med en rett line, derfor blir man nødt til å beregne avstanden mellom punktene og regresjonslinjen. Siden det vil være noen punkter over og under denne regresjonslinjen som regnes som positive og negative verdier, vil disse sørge for at totalsummen blir null. Minste kvadratsums metode går ut på å måle avstanden mellom punktene og regresjonslinjen, kvadrere dette og minimere den totale summen av alle kvadratene (Thrane, 2017, ss. 28-29)(figur 24).



Figur 24: Illustrerer hvordan minste kvadratsum metode fungerer, kilde (Brooks, 2008, s. 32).

La y_t være datapunktet for selve observasjonen og \hat{y}_t være punktet på regresjonslinjen (Figur 25). Avstanden mellom disse punktene er gitt ved \hat{u}_t . t representerer antall observasjoner. Har man for eksempel 10 observasjoner er $t = 10$.



Figur 25: Plottet viser en enkel observasjon, sammen med best tilpassede linje, residualverdien og forventet verdi, kilde: (Brooks, 2008, s. 32).

Man vil derfor få følgende utregning:

$$\hat{u}_1^2 + \hat{u}_2^2 + \hat{u}_3^2 + \dots + \hat{u}_t^2$$

Summen av alle residualene kalles «Residual sum of squares» (RSS). $\hat{\alpha}$ og $\hat{\beta}$ erstatter verdiene for α og β og ligningen for regresjonslinjen blir $\hat{y} = \hat{\alpha} + \hat{\beta}x_t$. L representerer verdien for RSS og summerer for alle observasjoner fra $t = 1$ til T, hvor T er antall observasjoner (Brooks, 2008, ss. 29-33).

$$L = \sum_{t=1}^T (y_t - \hat{y}_t)^2 = \sum_{t=1}^T (y_t - \hat{\alpha} - \hat{\beta}x_t)^2$$

L er minimert med hensyn til $\hat{\alpha}$ og $\hat{\beta}$, slik at man kan finne verdiene for α og β som minimerer RSS som gir linjen som er nærmest dataene. Deretter differensierer man L med hensyn til $\hat{\alpha}$ og $\hat{\beta}$ ved å sette den første deriverte til null. Ligningen for å regne ut $\hat{\beta}$ som er koeffisientestimatoren for brattheten på regresjonslinjen:

$$\hat{\beta} = \frac{\sum x_t y_t - T \bar{x} \bar{y}}{\sum x_t^2 - T \bar{x}^2}$$

Ligningen for $\hat{\alpha}$, som er verdien der regresjonslinja krysser y-aksen, er gitt ved:

$$\hat{\alpha} = \hat{y} + \hat{\beta} \bar{x}$$

Når man har regnet ut $\hat{\alpha}$ og $\hat{\beta}$ kan man estimerer hvor mye y endrer seg i forhold til endringer i x (Brooks, 2008, s. 33).

Multipel regresjon

I eksemplet om forholdet mellom boligpris og boligareal, hvor boligpris var variabelen y og boligareal var variabelen x tar man utgangspunkt i at det er kun variabelen boligareal som påvirker boligprisen. I den virkelige verden er dette som tidligere nevnt ikke realiteten.

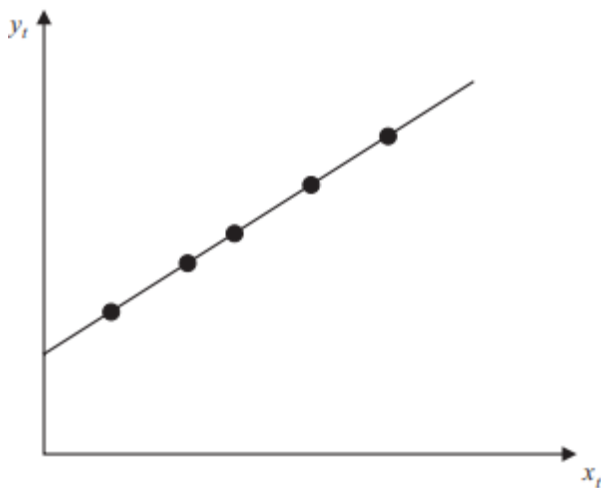
Variabler som lokasjon, alder på bygningen og boligens størrelse vil også være drivende for boligprisen. Dette er variabler som må være med i en regresjonsmodell, og man må derfor endre på ligningen:

$$y_t = \beta_1 + \beta_2 x_{2t} + \beta_3 x_{3t} + \dots + \beta_k x_{kt} + u_t, \quad t = 1, 2, \dots, T$$

Variablene $x_{2t}, x_{3t}, \dots, x_{kt}$, er forklarende variabler som kan påvirke variabelen y . Koeffisientene $\beta_1, \beta_2, \beta_3, \dots, \beta_k$ er parametere som måler effekten til alle de forklarende variablene som er inkludert i modellen. Teoretisk sett gjør man en regresjon for hver enkelt koeffisient, hvor man holder alle andre variabler konstante (Thrane, 2017, ss. 49-51). Tar man for eksempel β_2 som måler effekten av lokasjon (x_2) på boligpris, og samtidig holder effekten av alder på bygning (x_3) og størrelse (x_4) konstant.

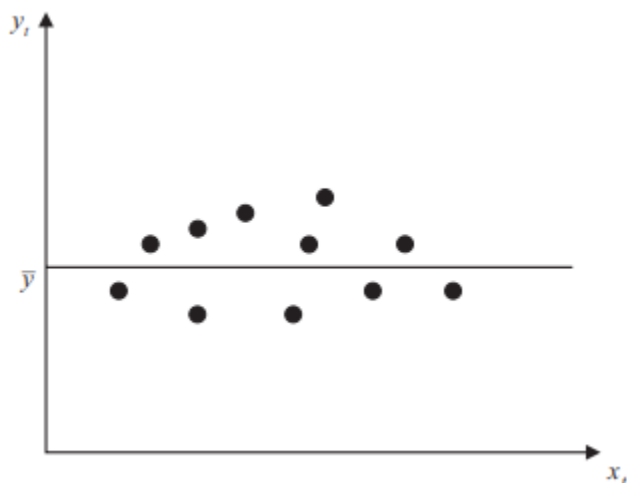
Forklaringskraften til modellen

For å teste hvor god modellen er, ser man på hvor mye av variasjonen variablene x forklarer. Totalvariasjonen i den avhengige variabelen y , er gitt ved TSS. Variasjonen som forklares av variablene x er gitt ved MSS. For å se på forklaringskraften til variablene x regner man ut MSS/TSS , som gir oss R -squared/ R^2 . R^2 gir oss et tall mellom 0% og 100%, hvor 100% betyr at variablene x som er tatt med i modellen forklarer totalvariasjonen i y . Figur 26 viser grafisk et R^2 på 100%, hvor observasjonene ligger helt på regresjonslinjen (Brooks, 2008, s. 109).



Figur 26: Regresjonsmodell med et R^2 tall på 1, kilde: (Brooks, 2008, s. 109).

Figur 27 viser et eksempel på en regresjonsmodell hvor $R^2 = 0\%$. Her ser man at regresjonslinjen er helt flat, og observasjonene er mer tilfeldig plassert



Figur 27: Regresjonsmodell med et R^2 tall lik 1, kilde: (Brooks, 2008, s. 109).

Validitet og reliabilitet for regresjonsmodellen

Ved å bruke regresjonsmodell for å analysere og besvare problemstillingen for denne studien, er det enkelte verdier som kan brukes for å si noe om validiteten og reliabiliteten.

Signifikansnivået og p-verdien er begge mål på hvor reliable estimatene er. Likevel må både studiens validitet og reliabilitet også vurderes etter størrelsen på datamaterialene (Field, 2009, ss. 11-12). Signifikansnivået brukes for å forkaste eller beholde 0-hypotesen man tester, og testes som regel på 95% nivå (Di Leo & Sardanelli, 2020). 0-hypotesen forkastes når signifikansnivået er over 95%, som vil si at man er mer enn 95% sikker på at estimatet er riktig. Dette er viktig med tanke på reliabiliteten, da et signifikansnivå på 95% betyr at man i 95 av 100 forsøk vil få samme resultater, noe som anses for å være reliabelt.

Videre brukes R^2 for å vurdere modellens forklaringskraft, altså validiteten (Snee, 1977). R^2 brukes til å si noe om hvor mye av kvadratmeterprisen de valgte variablene forklarer. R^2 har en verdi mellom 0 og 1, hvor en verdi nærmere 0 betyr at modellen forklarer lite av variasjonene i kvadratmeterprisen. Dersom verdien er høyere, betyr det at modellen forklarer mer av kvadratmeterprisen og at validiteten er høy. I tillegg må R^2 -adjusted også vurderes. Denne verdien tar hensyn til at modellen kan ha flere variabler, noe som er tilfellet i en multipel regresjonsmodell. En liten differanse mellom de to R^2 -verdiene indikerer at modellen forklarer mye av variasjonen i kvadratmeterpris, og at forklaringskraften ikke er kunstig høy fordi det er inkludert flere variabler (Snee, 1977).

Gjennom disse verdiene kan modellen vurderes ut ifra de overordnede kriteriene for forskning, altså reliabilitet og validitet. Ved å vurdere både variabler og modellen basert på

signifikansnivå og forklaringskraft (R^2) kan man si noe om hvor godt modellen svarer på problemstillingen, og hvor sannsynlig det er at funnene vil kunne gjenskapes i senere studier.

Logaritmisk form

Med utgangspunkt i datasettet som er benyttet for denne studien, vil det være aktuelt å vurdere spesifikke grep for å tilpasse regresjonsmodellen. Som nevnt forutsetter en regresjonsmodell et lineært forhold mellom de variablene som blir inkludert, altså en forutsetning som ikke alltid er virkelighetsnær. Ved tilfeller hvor forholdet ikke er lineært, vil det medføre muligheter for feil eller skjevhet i modellen. Grep for å tilpasse modellen for å redusere disse skjevhetene, samtidig som man beholder antagelsen om et lineært forhold, er å log-transformere de gjeldende variablene (Lane, 2021).

3.5 Rensing av data

Rensing, eller vasking av data gjøres for å sikre validiteten og reliabiliteten til studien så vel som å fjerne manglende eller ekstreme verdier. Da kvantitative studier ofte baserer seg på sekundærkilder, vil det være muligheter for feilinformasjon, manglende informasjon, eller informasjon som ikke umiddelbart kan tolkes (Osborne, 2009). Det kan også forekomme ekstremverdier, enten på grunn av ekstreme observasjoner, men også gjennom særegne innføringsmetoder fra utgiver av datasettet. Eksempelvis kan manglende informasjon bli registrert med «-999», noe som vil kunne medføre en stor påvirkning i analysene. Gjennom en rensing, og en tilpasning av datamaterialet, vil det sikre mer reliable og valide resultater så vel som å redusere sjansen for feil.

For å sikre validitet og reliabilitet måtte dataene fra Eiendomsverdi renses. Gjennom et uttrekk fra deres databaser, fikk vi ca. 33000 observasjoner for kommunene Lillestrøm, Nordre-Follo og Bærum over perioden 2007-2020. Hver observasjon ble oppgitt med følgende informasjon:

- Bolig ID
- Registreringsdato
- Salgsdato
- Tinglysningsdato
- Kommune
- Postnummer
- Pris
- Fellesgjeld
- Prisantydning
- Eierform

- Byggeår
- P-rom (Antall kvadratmeter)
- BRA (Bruksareal)
- Etasje
- Antall etasjer i blokka
- Soverom
- Balkong
- Parkering
- Koordinater

Da denne studien ser på et område med en radius på 2 km fra togstasjonen, måtte datamaterialet tilpasses ved hjelp av en formel for å omregne koordinater for hver enkelt observasjon til avstand fra togstasjonen. Dette ble gjort gjennom følgende formel i Excel:

$$=ARCCOS(COS(RADIANER(90-S2))*COS(RADIANER(90-U2))+SIN(RADIANER(90-S2))*SIN(RADIANER(90-U2))*COS(RADIANER(T2-V2)))*6371$$

I en noe forenklet forstand beregner formelen avstand i luftlinje fra ett sett med koordinater til ett annet. I tillegg tar formelen hensyn til jordens kurve for å være mest mulig nøyaktig. Koordinatene er representert ved S2, U2, T2 og V2 hvorav S2 er latitude for observasjonen og U2 er longitude. T2 og V2 er på sin side latitude og longitude for den gjeldende togstasjon. Avsluttende verdi "6371" er jordens omkrets i km. Formelen omgjør altså to sett med koordinater, eller to sett med latitude- og longitudeverdier om til en avstand i km. Ved å gjøre dette kunne vi omgjøre alle koordinater i datasettet til en faktisk avstand fra kommunens respektive togstasjon til hver enkelt observasjon. Slik kunne vi også filtrere ut de observasjonene som var utenfor studieområdet. Som nevnt er den beregnede avstanden i luftlinje, og den faktiske avstanden vil i de fleste tilfeller være lengre da veier gjerne ikke er korteste vei. Dette er derimot ikke mulig å beregne på en effektiv måte, og en manuell oppmåling vil være ressurskrevende utover hva denne studien kan gjøre.

«Antall etasjer i blokka» som en variabel var i utgangspunktet inkludert under en antagelse om at det forelå en prispremie for å bo i toppetasjen. Det viste seg derimot at denne variabelen hadde mangelfull informasjon, da tilhørende verdier hadde «NULL», «2» eller «3». Dette gjaldt også for «Balkong», hvor verdiene var enten «NULL» eller «1». Vi kontaktet Eiendomsverdi angående dette for å få en oppklaring i verdiene. «NULL» kunne representere mangel på informasjon, eller ingen balkong, og med denne usikkerheten i datagrunnlaget valgte vi å fjerne variablene i sin helhet. Også «Parkering» måtte fjernes av

denne årsaken. For «Etasje» forelå det også noen observasjoner med mangelfulle verdier. Vi fjernet disse observasjonene, og beholdt variabelen hvor det var korrekte verdier.

I vurderingen om det skulle brukes P-rom (antall kvadratmeter) eller BRA som et mål på leilighetsstørrelsen, så vi at dataene over BRA hadde store hull. Siden vi hadde begge variablene, falt valget på P-rom hvor dataene var mer fullstendige. Det samme gjaldt for vurderingen mellom tinglysningsdato og salgsdato, hvor salgsdato ble foretrukket. I tillegg anså vi det som mer hensiktsmessig å bruke salgsdato for å få et så presist mål på når leiligheten ble omsatt som mulig. Med tinglysningsdato anså vi det som en viss fare for forsinkelser, noe som kunne flyttet observasjonen til et annet år. «Postnummer», «Eierform» og «Registreringsdato» var tilhørende informasjon som var lite relevante for studien. Disse ble ikke inkludert videre.

Vi så også at det forelå store sprik i noen av verdiene. Salgsprisen bar preg av at mye av boligens verdi var flyttet til fellesgjelden, noe som mest sannsynlig er har vært et salgstriks for å gjøre boligen mer attraktiv gjennom å redusere inngangsprisen. Da dette gjaldt en god del observasjoner for alle byene, valgte vi å justere salgsprisen ved å inkludere fellesgjelden, noe som igjen ga oss en «Totalpris». Etter denne justeringen, fjernet vi øverste og nederste percentil, da det forelå ekstremverdier i begge ender av skalaen. Noen observasjoner hadde fortsatt svært lav pris, og noen observasjoner hadde svært høy. Leilighetsstørrelsene var også svært varierende, alt fra 10 kvm, til 320 m². For de minste leilighetene, valgte vi å fjerne alt under 14 m². For de store leilighetene, gjorde vi en mer tilpasset vurdering avhengig av hvilken by det var. Vi ekskluderte de observasjonene som klart lå utenfor «normalen» for den respektive byen. Eksempelvis fjernet vi en observasjon på 320 m², og en observasjon på 290 m² for Sandvika. Utover dette har vi også gjort en generell rensning på de resterende variablene, da det forekom enkelte observasjoner hvor en eller flere variabler hadde manglende informasjon.

3.6 Utvalgte attributter

For denne studien har vi som tidligere nevnt benyttet oss av Eiendomsverdi for å innhente data. Både banker, eiendomsmeglere og takstbransjen benytter seg av denne kilden og sekundærdataene innhentet fra Eiendomsverdi må derfor kunne anses for å være sikre

(Eiendomsverdi, 2021, b). Det vil likevel være nødvendig med en operasjonalisering av variablene for å ytterligere sikre validitet og reliabilitet.

3.6.1 Operasjonalisering av variablene og deskriptiv statistikk

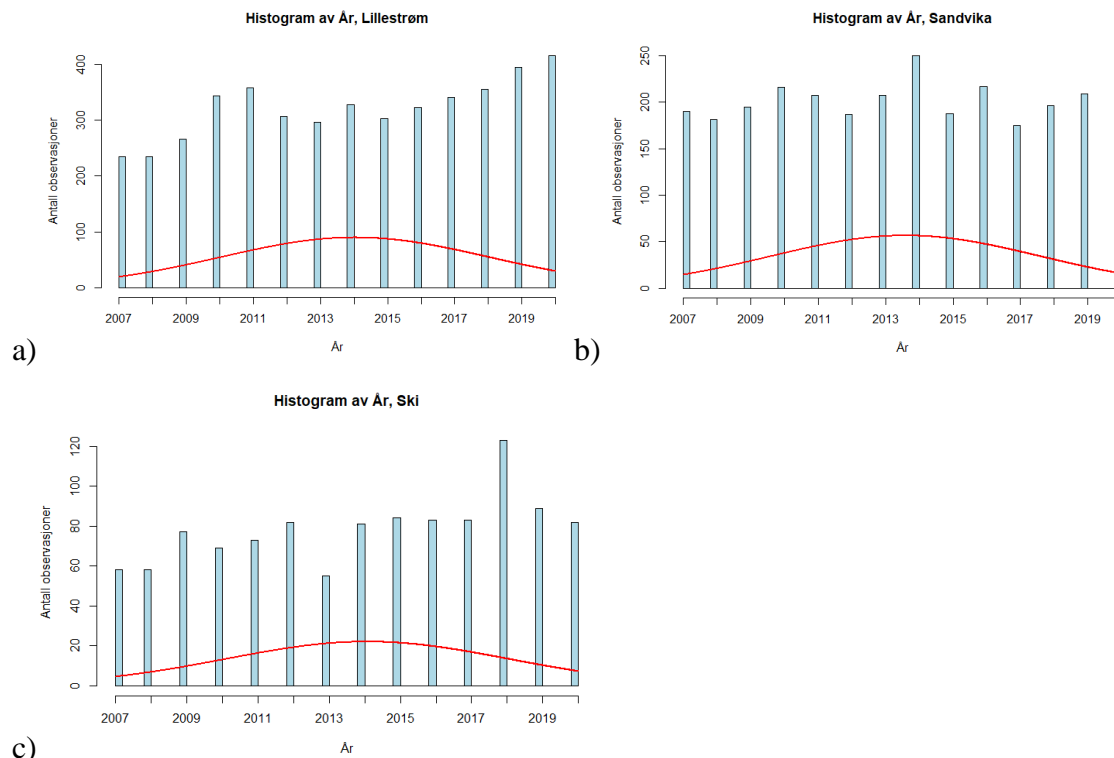
Med operasjonalisering menes det å formulere klare kriterier for hvordan attributtene skal presenteres i studien. Dette innebærer å klargjøre hvordan målingen av variablene skal gjennomføres for å ytterligere sikre validitet og reliabilitet (Dahlum & Grønmo, 2020). I forbindelse med operasjonaliseringen velger vi å presentere deskriptiv statistikk til hver variabel. Formålet er å tydeliggjøre hvordan variablene ser ut.

Boligsegmentet

Det vil være hensiktsmessig å operasjonalisere boligsegmentet i denne studien. Da vi primært tar utgangspunkt i den regionale planen for areal- og transport for Oslo og Akershus vil det være aktuelt å studere blokkleiligheter. I rapporten er det blitt understreket at for å kunne oppnå en høy utnyttelse rundt de kollektive knutepunktene er leiligheter noe det må bygges mer av. Ved å spesifisere boligsegmentet til å kun gjelde blokkleiligheter, vil vi minimere variasjoner i boligprisene samt andre uforutsette variabler. Hva som i utgangspunktet regnes for å være leiligheter er ikke ensidig blokkleiligheter, og ved å filtrere ut andre former for leiligheter, forsøker vi å skape et så homogent segment som mulig. Ved å gjøre dette minimerer vi risikoen for kategoriske feil i datamaterialet.

Tidsperspektiv

Studien tar for seg blokkleiligheter i tidsrommet 2007-2020. Et så stort tidsrom benyttes ettersom en studie som dette avhengig av mange observasjoner (Laakso, 1997). Ved å benytte et tidsrom på 13 år vil antall observasjoner som inkluderes i modellen øke. Etter samtaler med Eiendomsverdi i forbindelse med innhenting av dataene, ble det pekt på at 2007 var grensen for når de anså dataene for å være gode nok til å kunne brukes i en slik studie. I målet om å øke antall observasjoner ville det vært hensiktsmessig å gå lengre tilbake i tid, men det ble her en avveining mellom kvaliteten på dataene og antall observasjoner. Som allerede nevnt vil ikke tidsrommet påvirke modellen i form av generell boligprisvekst. Regresjonsmodellen gjør like-for-like analyser, noe som betyr at den sammenligner observasjoner fra 2007 med andre observasjoner for 2007 og observasjoner fra 2020 med de fra 2020.



Figur 28: a) Histogram over byggeår for Lillestrøm b) Histogram over byggeår for Sandvika, c) Histogram over byggeår for Ski.

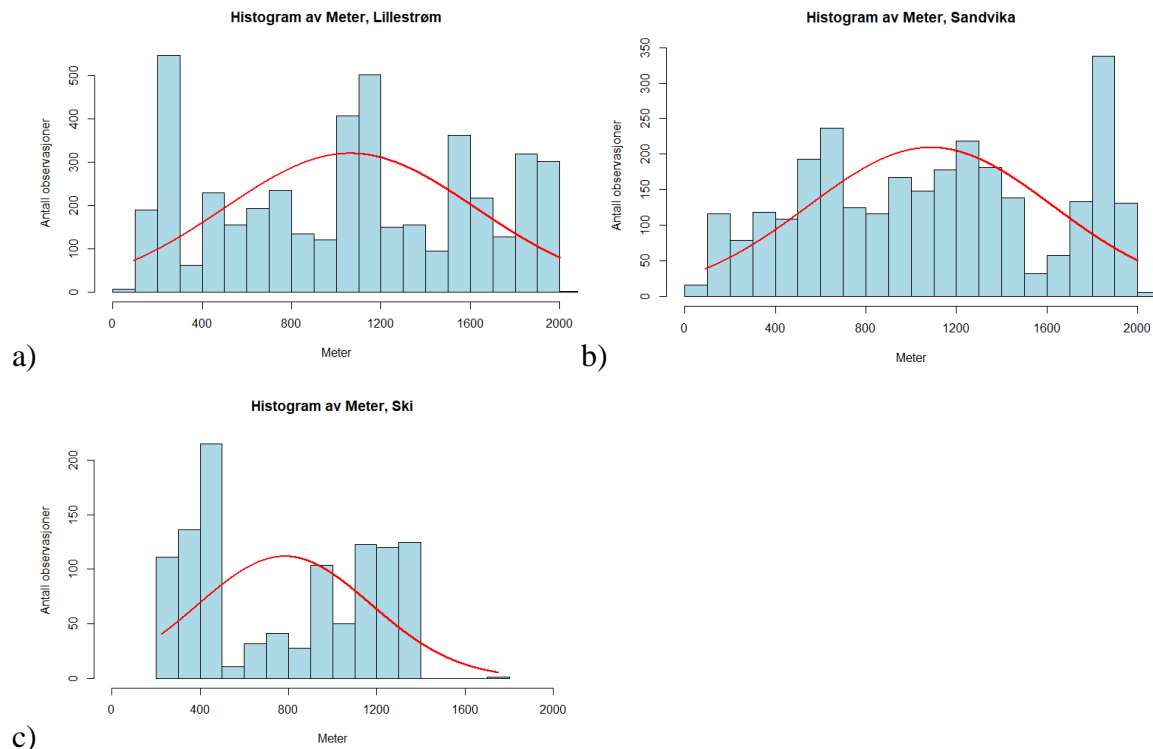
Det er stor variasjon i antall observasjoner for de ulike byene. For Lillestrøm er det registrert 4505 observasjoner, Sandvika har 2836 og Ski har 1097. Videre ser vi at det er små variasjoner i når leilighetene er solgt. Det er lokale forskjeller med tanke på hvor mange boliger som har blitt solgt per år. Sandvika har hatt et forholdsvis jevnt salg hvert år, og Lillestrøm har et noe økende salg fra 2007 til 2020, trolig på grunn av en økning av boligutbyggingen i sentrum. Ski har på sin side hatt svingninger, med noen år med færre salg, og noen år med flere salg.

Avstand til togstasjonen

Avstanden til togstasjonen er den uavhengige variabelen vi ønsker å teste i denne studien. Som tidligere nevnt er avstanden beregnet i luftlinjen gjennom to sett med koordinater.

Tabell 5 Deskriptiv tabell over antall leiligheter og avstand til togstasjon.

Avstand i meter	N	Min	Max	Gjennomsnitt	Standardavvik
Lillestrøm	4505	92	2003	1062	561,95
Sandvika	2836	90	2002	1086,5	540,92
Ski	1097	228	1750	783	391,62



Figur 29: a) Histogram over avstand til togstasjon i meter for Lillestrøm b) Histogram over avstand til togstasjon i meter for Sandvika c) Histogram over avstand til togstasjon i meter for Ski.

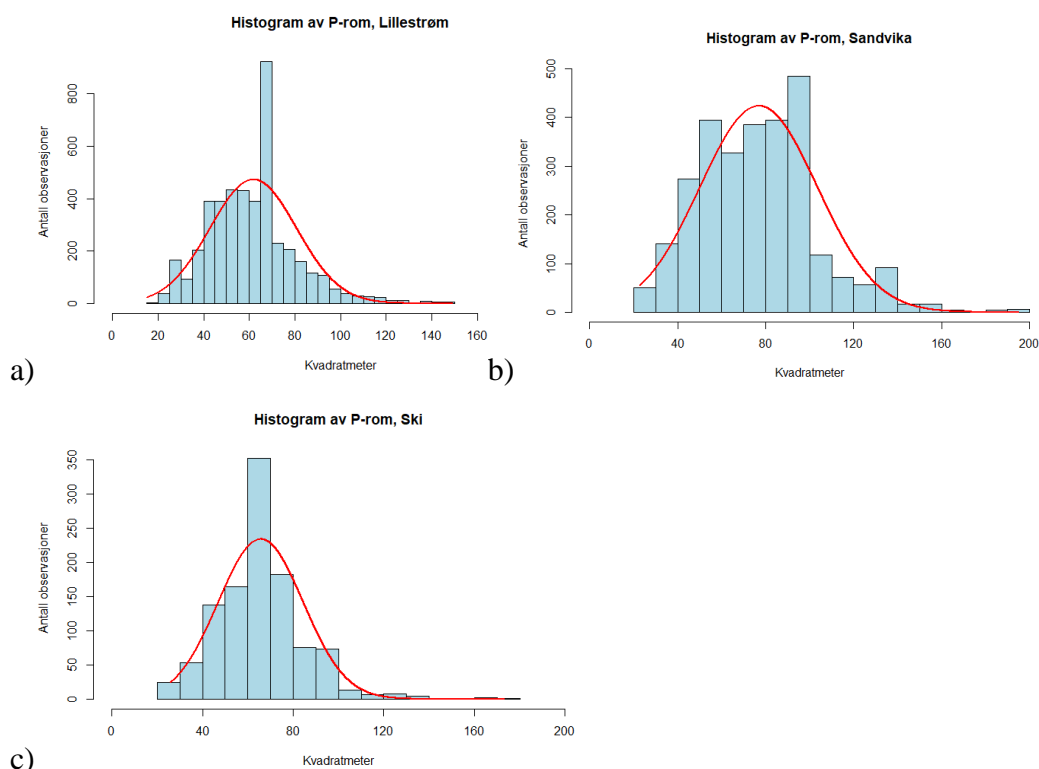
Vi ser at Ski skiller seg ut når det kommer til hvor langt unna togstasjonen leilighetene befinner seg. Ski er sammenlignet med Lillestrøm og Sandvika en mindre by, noe som også vises i tallene over. I Lillestrøm og Sandvika ligger de nærmeste leilighetene ca. 90 m fra togstasjonen. I Ski derimot ligger nærmeste leilighet hele 228 m unna. I tillegg ser vi at det ikke er noen observasjoner utenfor 1750 m i Ski. Da grensen er satt til 2 km for studien, vil det ikke være mulig å fylle ut denne for Ski. For Ski er den gjennomsnittlige boligen også ca. 300 m nærmere togstasjonen.

Antall kvadratmeter

For denne studien vil antall kvadratmeter indirekte være med som en variabel i form av den kategoriske variabelen "Størrelse" som blir forklart under. Årsaken til dette er for å minimere risikoen for multikollinearitet. Vi har testet hvorvidt antall kvadratmeter kan stå som en separat variabel. Testene ga ingen ytterligere økning i forklaringskraften til modellene, og vi antar derfor at antall kvadratmeter blir inkludert på en tilfredsstillende måte gjennom kategoriseringen av leilighetsstørrelsene.

Tabell 6 Deskriptiv tabell over variabelen antall kvadratmeter.

Antall kvadratmeter	N	Min	Max	Gjennomsnitt	Standardavvik
Lillestrøm	4505	14	149	61,8	19,00
Sandvika	2836	23	195	77	26,71
Ski	1097	26	173	65,7	18,72



Figur 30: a) Histogram over antall kvadratmeter for Lillestrøm b) Histogram over antall kvadratmeter for Sandvika, c) Histogram over antall kvadratmeter for Ski.

Sandvika ser ut til å ha en stor andel store leiligheter med et gjennomsnitt på 77 m² sammenlignet med 61,8 m² for Lillestrøm og 65,7 m² for Ski. Sandvika har også den største leiligheten på hele 195 m². Maksverdien for hver by representerer den øvre grensen for hvor det befant seg tre eller flere observasjoner med samme antall kvadratmeter. Denne filtreringen er gjort av samme årsak som kvadratmeterprisen, da ekstremverdier påvirket den statistiske analysen i betydelig grad.

Boligens størrelse

Det har som nevnt vært ulike metoder for hvordan størrelsen på boliger skal måles. For denne studien tar vi utgangspunkt i Laakso sin fremgangsmåte, og velger derfor å bruke antall kvadratmeter som grunnlag for kategorisering. For å kunne påvise en eventuell forskjell i lokasjonspremien for ulike størrelser på leiligheter, må leilighetssegmentet ytterligere deles opp. Studien vil se på tre størrelser; små, mellomstore og store leiligheter. Hvordan

leilighetene skal kategoriseres etter størrelse baseres på plan- og bygningsetatens fordelingsnorm for leiligheter (Plan- og bygningsetaten, Oslo Kommune, 2016, s. 7): Ekstra leiligheter $>35 \text{ m}^2$, små leiligheter $35\text{-}49 \text{ m}^2$, mellomstore leiligheter $50\text{-}79 \text{ m}^2$ og store leiligheter $80+ \text{ m}^2$. For denne studien vil ekstra leiligheter medregnes i kategorien «Liten leilighet». Grunnen for dette valget er for å skape et bredere segment, da $35\text{-}49 \text{ m}^2$ vil inkludere få observasjoner. Kategoriseringen blir dermed «Liten leilighet» = $>49 \text{ m}^2$, «Mellomstor leilighet» = $50\text{-}79 \text{ m}^2$ og «Stor leilighet» er $80+\text{m}^2$.

Når man ser på intervallene for hver kategori, ser man at store leiligheter omfatter alle leiligheter over 80 m^2 . For datasettet vi benytter, er den største leiligheten på ca. 320 m^2 , altså 250 m^2 større enn minimumsgrensen for kategorien. Dette skaper et stort sprik og store variasjoner innad i kategorien. Som tidligere nevnt er det derfor gjort en rensing på disse verdiene.

Tabell 7 Deskriptiv tabell over variabelen Størrelse.

Størrelsekategori	Liten	Mellomstor	Stor	Totalt
Lillestrøm	896	2968	641	4505
Sandvika	317	1229	1290	2836
Ski	116	786	195	1097

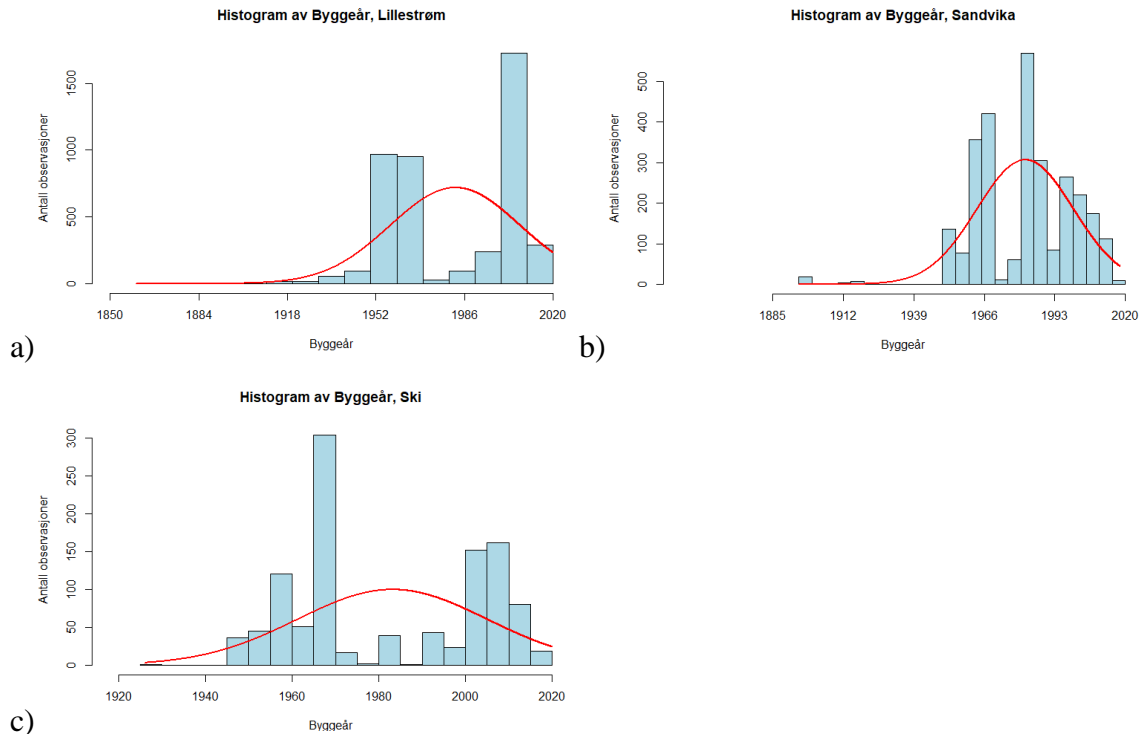
Når det kommer til boligens størrelse, er dette en kategorisk variabel. Vi ser her at det er store variasjoner i hvor mange observasjoner det er for hver kategori for de ulike byene. I Lillestrøm er majoriteten av leilighetene innenfor den mellomstore kategorien, med en ganske lik fordeling på «Liten» og «Stor». I Sandvika er det langt færre små leiligheter, og langt flere store. Ski har forholdsvis lik fordeling mellom de ulike leilighetsstørrelsene som i Lillestrøm, men har færre observasjoner totalt.

Boligens alder

Boligens alder vil som tidligere nevnt også være en faktor som påvirker prisen på boligen. Slik Laakso påpekte, vil alder og lokasjon kunne medføre en multikollinear sammenheng. Dette følger naturlig nok av den monosentriske byteorien, da byutviklingen starter innenfra og ut. Som et resultat av dette, vil boliger i sentrum være eldre, noe som jevner ut prisforskjellen fra sentrumsnære boliger og nyere boliger bygget lengre unna.

Tabell 8 Deskriptiv tabell over variabelen byggeår.

Byggeaar	N	Min	Max	Gjennomsnitt	Standardavvik
Lillestrøm	4505	1860	2020	1982	24,97
Sandvika	2836	1895	2018	1982	18,45
Ski	1097	1926	2020	1983	21,93

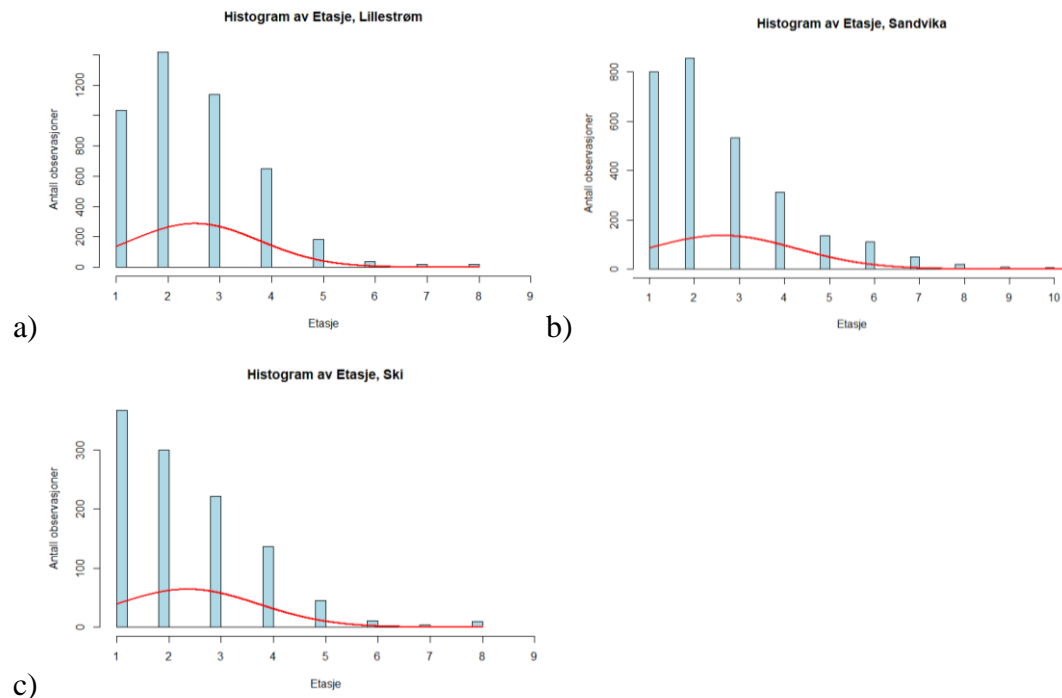


Figur 31: a) Histogram over boligens alder for Lillestrøm b) Histogram over boligens alder for Sandvika, c) Histogram over boligens alder for Ski.

Lillestrøm ser ut til å være den byen hvor boligutviklingen startet tidligst, med observasjoner helt tilbake til 1860. Sandvika har første observasjon fra 1895, og Ski i 1926.

Etasje

Vi har også valgt å bruke etasje som en uavhengig variabel. Det ble som tidligere nevnt diskutert hvorvidt vi skulle inkludere en variabel for å kunne ta hensyn til om leiligheten lå i toppetasjen. Da variabelen som tilsa om observasjonen lå i toppetasjen utgikk, forkastet vi denne ideen. Variabelen vil derfor kun være et mål på hvilken etasje leiligheten ligger i, uten en merverdi for om den ligger i toppetasjen.



Figur 32: a) Histogram over hvilken etasje observasjonen ligger i for Lillestrøm b) Histogram over hvilken etasje observasjonen ligger i for Sandvika, c) Histogram over hvilken etasje observasjonen ligger i for Ski.

Som man kan se, er det ikke store variasjoner eller sprik i denne variabelen. En mulig utslagsgivende faktor er at Lillestrøm og Ski har maksimumsverdi på 8 etasjer, hvorav Sandvika ligger på 10 etasjer.

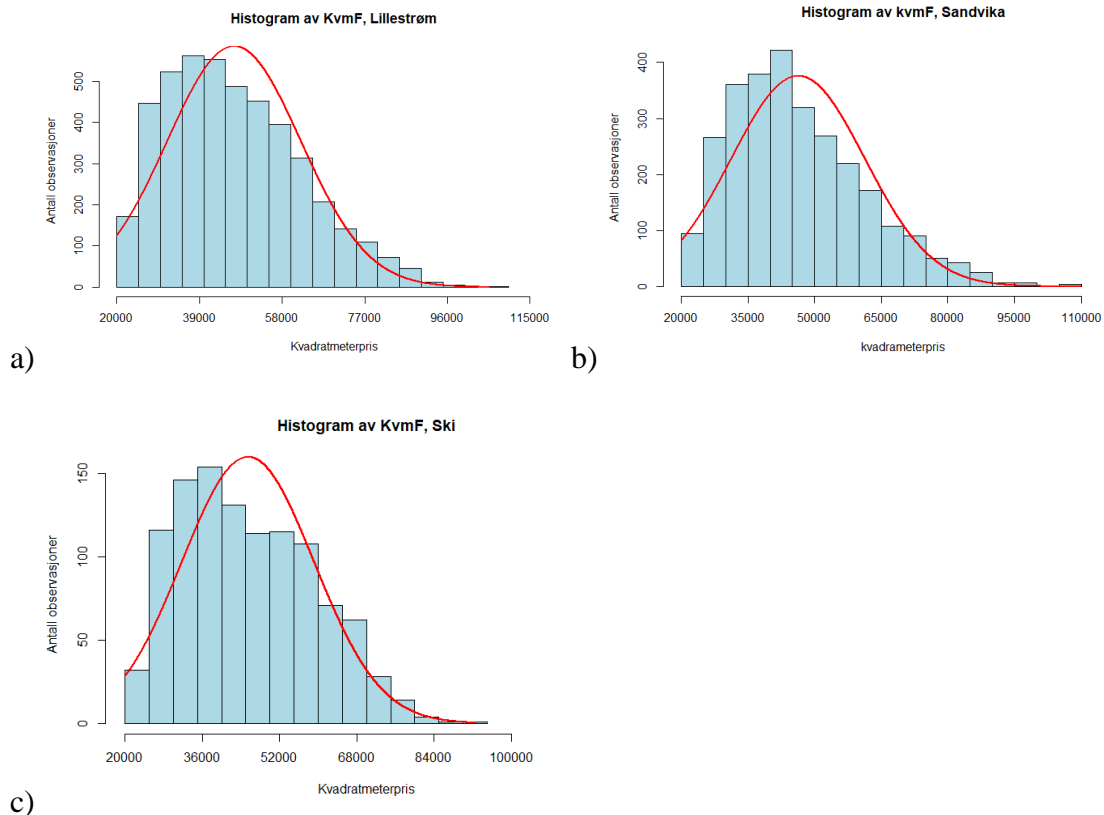
Kvadratmeterpris

Kvadratmeterprisen er den avhengige variabelen i denne studien, altså den variabelen vi forsøker å se om blir påvirket av avstand til togstasjonen. Forholdet mellom boligpris, antall kvadratmeter og de kategoriske størrelsene er intrikat. I en analyse hvor formålet er å studere boligpriser, er antall kvadratmeter (prom) en viktig forklaringsvariabel. Samtidig har vi, gjennom valget å bruke kategoriske størrelser, påvist moderat til høy multikollinearitet mellom prom og leilighetens størrelse (kategoriske variabel) (tabell 11). For å redusere behovet av antall kvadratmeter (prom) som en forklaringsvariabel, valgte vi å bruke kvadratmeterprisen istedenfor salgspris eller totalpris da kvadratmeterprisen er mindre følsom for hvor mange kvadratmeter leiligheten er.

Det har ikke vært behov for å justere boligprisene med hensyn til hvilket år leilighetene har blitt solgt, da en regresjonsanalyse gjennomført i R-studio tar hensyn til dette automatisk gjennom en like-for-like analyse, som beskrevet i kapittel 1.4. Regresjonsanalysene vil med andre ord sammenligne observasjoner som har like variabler f.eks. salgsår.

Tabell 9 Deskriptiv tabell over kvadratmeterpriser.

Kvadratmeterpris	N	Min	Max	Gjennomsnitt	Standardavvik
Lillestrøm	4504	20125	118735	47043	15405,80
Sandvika	2836	20164	109885	46324	15054,11
Ski	1097	20258	92464	45512	13680,85



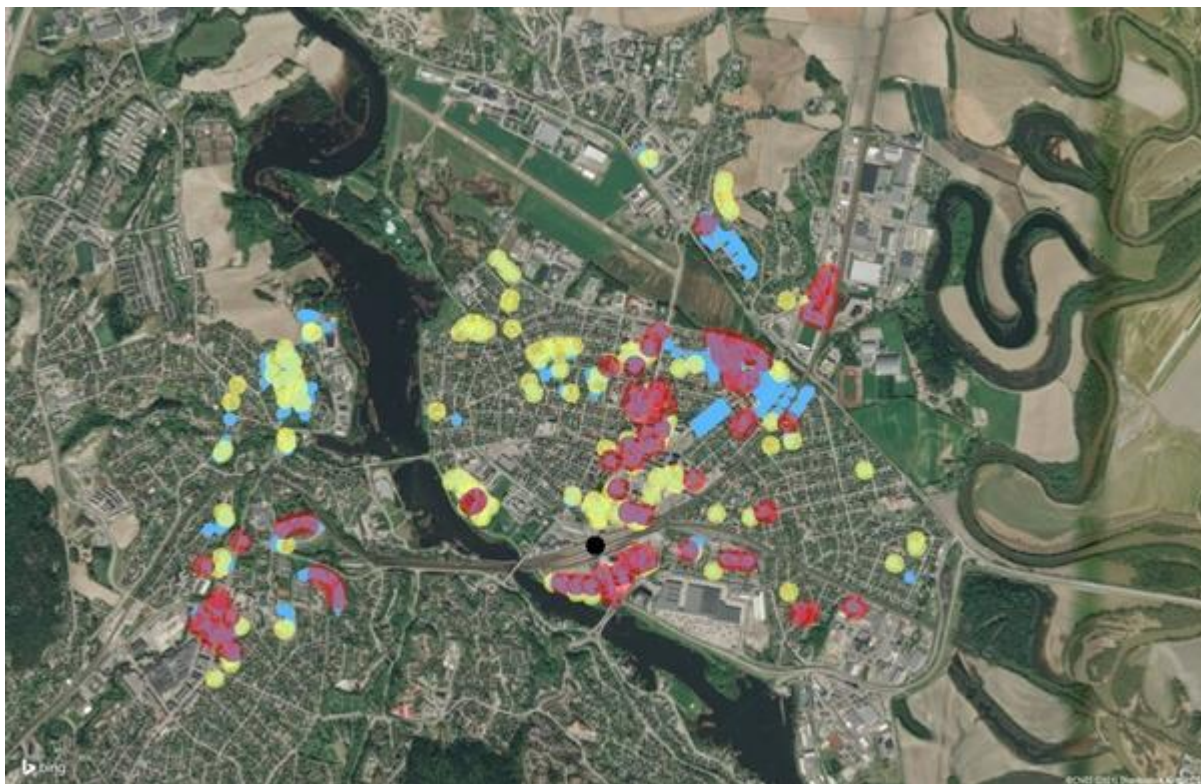
Figur 33: a) Histogram over kvadratmeterpris for Lillestrøm b) Histogram over kvadratmeterpris for Sandvika, c) Histogram over kvadratmeterpris for Ski.

Minimumsverdien ligger på ca. 20 000kr/m² da det er satt en grense for å utelukke ekstremverdier. Datasettet inneholdt kvadratmeterpriser på under 10 000kr/m², noe vi anså som lite representativt da dette mest sannsynlig gjaldt oppussingsobjekter eller objekter som ikke hadde blitt justert for fellesgjeld. I tillegg påvirket disse ekstremverdiene de statistiske resultatene i forholdsvis stor grad. Vi har også fjernet den øverste percentilen i datasettet, nettopp av samme årsak.

Som et resultat av disse tilpasningene, ser vi at minimumsverdien er så og si lik for de tre byene. På den andre siden ser vi at Lillestrøm har den høyeste verdien, og Ski den laveste. Likevel ser vi at den gjennomsnittlige kvadratmeterprisen er forholdsvis lik for de tre byene.

3.6.2 Observasjoner

I forbindelse med den deskriptive statistikken ser man at det er store forskjeller i antall observasjoner for hver by. Det er også er forskjeller i hvor mange observasjoner det er innenfor hver størrelseskategori. For å få en bedre forståelse av hvordan størrelseskategoriene og spredningen av observasjonene er dette fremstilt i kart (Figur 34, 35 og 36 under).



Figur 34: Oversikt over observasjoner etter at de er kategorisert etter størrelse for Lillestrøm. Rød er små leiligheter, blå er mellomstore leiligheter og gul er store leiligheter. Svart prikk markerer togstasjonen.

Leilighetene i Lillestrøm er spredt omtrentlig jevnt utover avstanden på 2 km. En annen faktor som burde påpekes er at flere av observasjonene på vestre side av Glomma ligger nærmere andre knutepunkt, slik som Strømmen. Ikke bare er det egen togstasjon her, men også Norges største kjøpesenter i antall butikker (Strømmen Storsenter, 2020). Vedlegg 1 viser de tre størrelseskategoriene i egne kart for Lillestrøm.

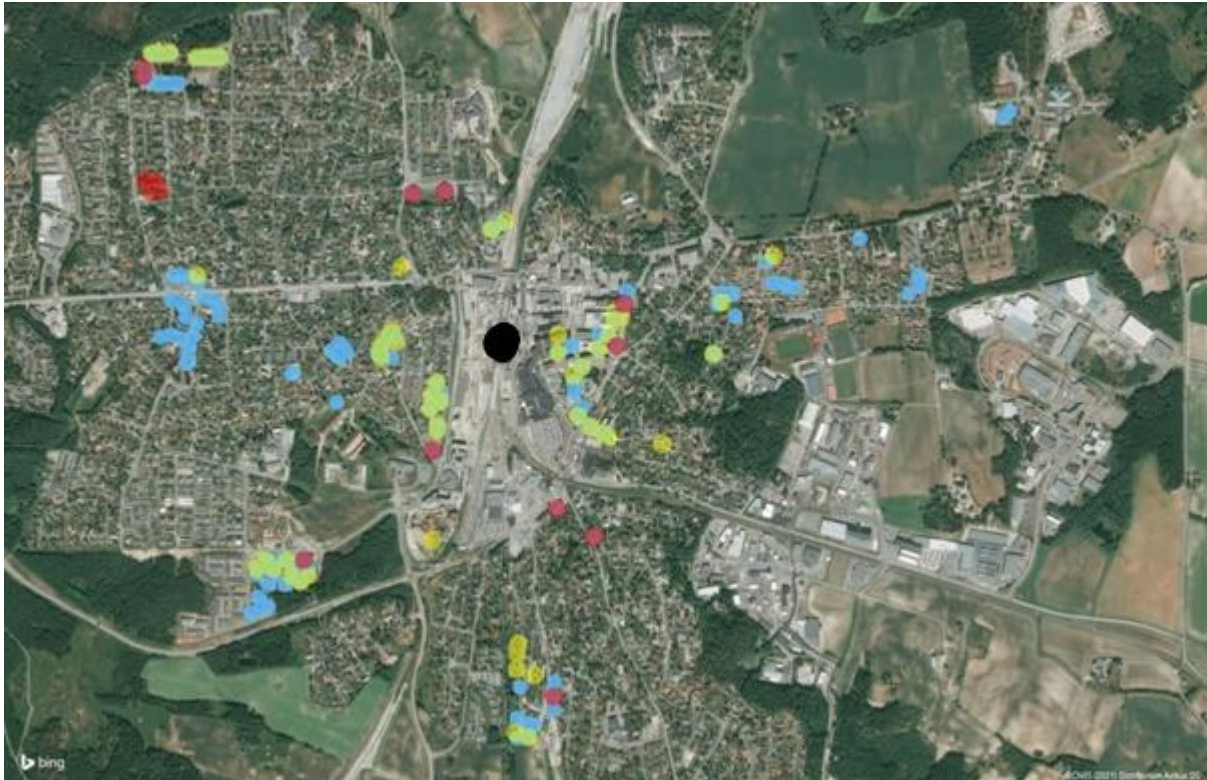
Under vises figur 35. Figuren viser oversikt over observasjonene i Sandvika med fargekode for størrelse.



Figur 35: Oversikt over observasjoner etter at de er kategorisert etter størrelse for Sandvika. Røde er små leiligheter, blå er mellomstore leiligheter og gule er store leiligheter. I de områdene det er grønt ligger blå (mellomstore) og gule (store) over hverandre. Oransje er på sin side røde (små) og gule (store) oppå hverandre. Svart prikk markerer togstasjonen.

For Sandvika er det også en god spredning på observasjonene. I Sandvika er det en stor andel av store og mellomstore leiligheter i kort avstand til togstasjonen, og de fleste små ligger litt utenfor. I tillegg ser man at E18 skaper en naturlig barriere i sør. Dette kan påvirke boligprisene noe, da større deler av området rundt Sandvika stasjon er i nærhet til hovedveier og tilhørende anlegg. Videre skaper også Oslofjorden en ytterligere barriere i sør-øst. Antagelig medfører ikke dette noen store endringer i boligprisene som kan påvirke resultatene, men det gjør at observasjonene ligger i en «vifteform». Vedlegg 2 viser de tre størrelseskategoriene i egne kart for Sandvika.

Under viser figur 36 en oversikt over observasjonene i Ski. Observasjonene er fargekodet etter størrelse.



Figur 36: Oversikt over observasjoner etter at de er kategorisert etter størrelse for Ski. Rød er små leiligheter, blå er mellomstore leiligheter og gul er store leiligheter. I de områdene det er grønt ligger blå (mellomstore) og gule (store) over hverandre. Oransje er på sin side røde (små) og gule (store) oppå hverandre. Svart prikk markerer togstasjonen.

I motsetning til Lillestrøm og Sandvika ser man at observasjonene ikke har en like god spredning i Ski. Mange av observasjonene ligger i klynger, noe som fører til at det blir mange observasjoner innenfor en bestemt avstand fra togstasjonen. I en slik studie som dette hadde det vært ønskelig at det var jevnere fordeling av observasjoner over avstanden på 2 km. I tillegg er det en rekke observasjoner som befinner seg i strøk med eneboliger, hvor det kan antas at kollektivdekningen inn til sentrum ikke er særlig god. Etersom det ikke er en jevn spredning av leiligheter utover radiusen på 2 km kan det være at dette påvirker resultatene. Vedlegg 3 viser de tre størrelseskategoriene i egne kart for Ski.

4 Analyse

4.1 Innledning

I dette kapitlet presenterer vi studiens analyser som brukes for å besvare hypotesene og problemstillingen. I første del vil en oversikt med forklaring av begreper og forkortelser brukt i analysene bli presentert. Videre vil analysene som er utført presenteres. Det vil gis en introduksjon til utskrift fra R-studio og tolkning av disse. Videre vil det gis en introduksjon til hvilke tester som er utført innledningsvis, og hvilke tester som skal utføres i forbindelse med hovedmodellen. Hovedmodellen presenteres med utgangspunkt i en lineær multippel regresjon etterfulgt av log-level multippel regresjon.

For analysene har vi benyttet oss av R-studio (versjon 4.0.4), hvor vi har muligheten til å gjennomføre de statistiske modellene samt residual- og normalfordelingstester av regresjonsmodellene. I tillegg vil vi kunne uthente illustrative grafer som et hjelpemiddel for å tolke resultatene. Hypotesene vi ønsker å teste er følgende:

Hypotese 1

- H1: Kvadratmeterprisen for leiligheter blir påvirket av avstand til togstasjon
- H0: Kvadratmeterprisen for leiligheter blir ikke påvirket av avstand til togstasjon

Hypotese 2

- H1: Kvadratmeterprisen for ulike leilighetsstørrelser blir påvirket i ulik grad av avstand til togstasjon
- H0: Kvadratmeterprisen for ulike leilighetsstørrelser blir ikke påvirket i ulik grad av avstand til togstasjon

4.2 Introduksjon av begreper

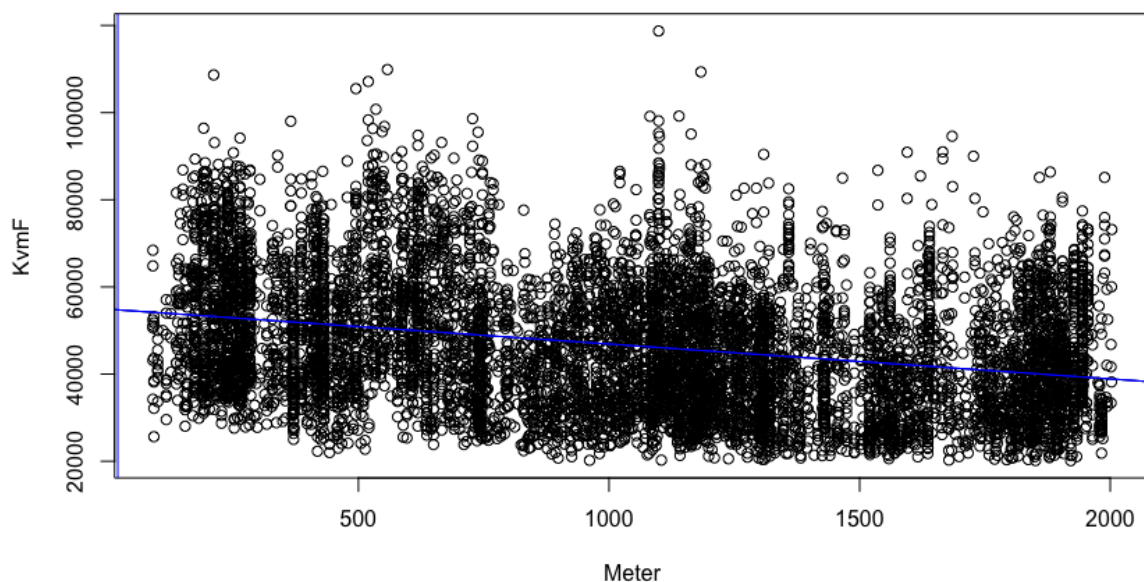
For at R-studio skal kunne lese kolonner og rader i importerte Excel-filer, har vi måttet gjøre noen endringer på navn og begreper i analysen. Tabell 10 viser en oversikt over de ulike begrepene med tilhørende forklaring.

Tabell 10 Begrepsoversikt for variablene i regresjonsmodellene

Begrep	Forklaring
Intercept	Angir på hvilket nivå kvadratmeterprisen krysser Y-aksen. I dette punktet er alle X-variabler 0
KvmF	Kvadratmeterprisen for observasjonene. Denne er justert for fellesgjeld. Avhengig variabel som studeres sammen med meter
Meter	Avstand til togstasjon i meter, maks verdi på 2000 m. Hovedeffekt som studeres
Prom	Antall kvadratmeter. Brukes kun i analyser for introduksjon. Fjernes på grunn av multikollinearitet
Byggeaar	Hvilket år leilighet ble bygd. Kontrollvariabel
Aar	Hvilket år leiligheten ble solgt mellom 2007-2020. Kontrollvariabel
Etasje	Hvilken etasje observasjonen ligger i. Kontrollvariabel
KommuneLillestrøm	Alle observasjoner lokalisert i Lillestrøm by
KommuneBarum	Alle observasjoner lokalisert i Sandvika by
KommuneNordreFollo	Alle observasjoner lokalisert i Ski by
StorrelseLiten	Alle observasjoner innenfor kategorien «liten leilighet». Vil i de separate analysene for hver by ligge i Intercept. Dette forklares nærmere.
StorrelseMellomstor	Alle observasjoner innenfor kategorien «mellomstor leilighet»
StorrelseStor	Alle observasjoner innenfor kategorien «stor leilighet»
Interaksjonsledd	Statistisk metode for å sammenligne to kategoriske variabler, og estimere forskjellen mellom dem.
Meter:StorrelseMellomstor	Interaksjon mellom meter og mellomstore leiligheter. Viser estimert endring ved å gå 1 m ut fra togstasjonen, og øke leilighetsstørrelsen fra liten til mellomstor.
Meter:StorrelseStor	Interaksjon mellom meter og store leiligheter. Viser estimert endring ved å gå 1 m ut fra togstasjonen, og øke leilighetsstørrelsen fra liten til stor.

4.3 Enkel lineær regresjon

Innledningsvis presenteres lineær regresjon på sin enkleste form, altså med kun en avhengig og en uavhengig variabel. I dette eksempelet er kvadratmeterprisen på Y-aksen og «Meter» på X-aksen. Figur 37 viser en enkel lineær modell som er gjort med utgangspunkt i datasettet for alle observasjoner innenfor 2 km avstand fra togstasjonen og er uavhengig av by. Figuren viser at når avstanden til togstasjonen øker, synker kvadratmeterprisen. De svarte punktene er observasjonene, og den blå linjen er regresjonslinjen.



Figur 37: Egenprodusert enkel lineær regresjonslinje med kvadratmeterpris (KvmF) på Y-aksen og Meter på X-aksen.

Under vises utskrift fra R-studio. Utskriften viser regresjonsmodellen hvor kvadratmeterprisen er på Y-aksen og «Meter» er på X-aksen (Figur 38).

```

Call:
lm(formula = KvMF ~ Meter, data = dta4)

Residuals:
    Min       1Q   Median       3Q      Max
-29392 -11532  -2019   9361  80310

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 54946.32    338.94  162.11  <2e-16 ***
Meter        -8.04      0.29  -27.73  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 14540 on 8483 degrees of freedom
Multiple R-squared:  0.08311, Adjusted R-squared:  0.083
F-statistic: 768.9 on 1 and 8483 DF, p-value: < 2.2e-16

```

Figur 38: Egenprodusert enkel lineær regresjonsmodell.

Utskriften av regresjonsmodellen viser en rekke verdier, hvorav noen av verdiene vil være av større interesse enn andre. Den første verdien av interesse er «Estimate», altså koeffisientestimatene for de valgte variablene. Det er koeffisientestimatet som beskriver om regresjonslinjen stiger eller synker. I dette tilfellet viser modellen at for hver meter avstanden øker med, synker kvadratmeterprisen med -8,04 kr. Intercept og tilhørende estimat er en variabel som angir hva kvadratmeterprisen for en leilighet er dersom de tilhørende X-variablene er 0. I dette eksemplet betyr det at kvadratmeterprisen er på 54946.32kr når avstanden til togstasjonen er 0 meter.

Neste verdi av interesse er P-verdien (Pr(>t)). Signifikansnivået er av betydning i vurderingen om hvor gode variablene er i å forklare kvadratmeterprisen, altså validiteten. Dersom signifikansnivået faller under 95% sikkerhet, vil det måtte gjøres en vurdering for om variabelen burde utelates. I denne modellen ser vi at begge estimatene er på laveste nivå, og godt innenfor grensen for 95% signifikans.

«Multiple R-squared» og «Adjusted R-squared» er også verdier av interesse. Den førstnevnte viser forklaringskraften til modellen, altså hvor mye av kvadratmeterprisen som avstand til togstasjon i meter forklarer. I denne modellen er verdien svært lav, kun på 8,3%. Variabelen «Meter» forklarer altså kun 8,3% av kvadratmeterprisen. «Adjusted» R-squared er et mål på hvor stor forklaringskraften modellen har, tatt i hensyn hvor mange variabler som legges inn (Investopedia, 2021). Adjusted R-squared vil derfor lavere ettersom den trekker ifra for hver

variabel som legges inn. Forholdet mellom disse må vurderes, da en stor differanse tilsier at det kan være svakheter i modellen, slik som multikollinearitet.

4.4 Multippel regresjon

Etter å ha presentert et eksempel på en enkel, lineær regresjonsmodell med kun avstand til togstasjon som variabel, vil følgende modell ta utgangspunkt i den hedoniske teorien, som sier at flere variabler påvirker boligprisene. Variablene «Etasje», «Byggeaar» og «Aar» vil brukes i analysen som kontrollvariabler. Formålet med kontrollvariablene er å justere modellen med hensyn på hovedvariabelen som er avstand til togstasjon. Det vil gis en introduksjon av modellen for å tydeliggjøre variablene og for å vurdere hvor gode variablene som legges inn er basert på signifikansnivået. I følgende modell blir også de kategoriske variablene for by og leilighetsstørrelse lagt inn. Regresjonsmodellen som presenteres under er kjørt basert på alle observasjoner for Lillestrøm, Ski og Sandvika (figur 39)

```
Call:
lm(formula = KvmF ~ Meter + Storrelse + Etasje + Byggeaar + Aar +
    prom + Kommune, data = dta4)

Residuals:
    Min       1Q   Median       3Q      Max
-27280  -5321   -695    4685  105507

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)   -5.609e+06  4.467e+04 -125.551 < 2e-16 ***
Meter          -8.153e+00  1.697e-01  -48.057 < 2e-16 ***
StorrelseMellomstor -7.233e+03  3.140e+02  -23.037 < 2e-16 ***
StorrelseStor   -6.377e+03  5.374e+02  -11.865 < 2e-16 ***
Etasje         4.989e+02  6.289e+01   7.934  2.4e-15 ***
Byggeaar       4.940e+01  2.381e+00  20.746 < 2e-16 ***
Aar            2.770e+03  2.220e+01  124.732 < 2e-16 ***
prom          -7.924e+01  7.442e+00  -10.647 < 2e-16 ***
KommuneLILLESTROM  -2.375e+03  2.071e+02  -11.466 < 2e-16 ***
KommuneNORDRE FOLLO -5.205e+03  3.011e+02  -17.290 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 8156 on 8475 degrees of freedom
Multiple R-squared:  0.7117,    Adjusted R-squared:  0.7114
F-statistic: 2325 on 9 and 8475 DF,  p-value: < 2.2e-16
```

Figur 39: Egenprodusert lineær regresjonsmodell med alle variabler inkludert.

Signifikansnivåene for alle variabler er innenfor 99% og det er ingen av variablene som må utelates i senere modeller basert på dette. Det som derimot må vurderes, er som tidligere

nevnt samspillet mellom variabelen «prom» og de kategoriske variablene for «Storrelse». Årsaken ligger i at for å avgjøre hvilken størrelseskategori hver leilighet ligger i, har antall kvadratmeter blitt brukt, og det vil derfor være en fare for multikollinearitet.

Forklaringskraften til modellen sammenliknet med modellen for enkel lineær regresjon har gått betraktelig opp. Modellen forklarer 71% av kvadratmeterprisen sammenliknet med 8% for modellen fra enkel lineær regresjon. En forklaringskraft på 71% anses for å være bra tatt i betraktning at den gjøres for tre ulike byer samtidig.

4.4.1 Testing av modellen

Modelltestene vil være av betydning for å vurdere hvor god modellen er. Som nevnt er det fare for at variablene «prom» og «Storrelse» kan ha et multikollineært forhold. For å teste om det eksiterer et multikollineært forhold, gjøres dette gjennom VIF (Variance inflation factor). En VIF-test tester hvor mye de valgte variablene påvirker hverandre.

Tabell 11: Multikollinearitetstest av variabler i multipel regresjonsmodell.

GVIF	
<i>Meter</i>	1,156944
<i>Storrelse</i>	4,127530
<i>Aar</i>	1,010722
<i>Byggeaar</i>	1,135184
<i>Etasje</i>	1,048653
<i>Kommune</i>	1,189489
<i>prom</i>	3,970308

I tilfeller hvor de kategoriske variablene består av mer enn to verdier (i dette tilfellet tre; liten, stor og mellomstor samt Lillestrøm, Ski og Sandvika) må man bruke GVIF (generalized variance inflation factor) (University of Cambridge, 2015). GVIF-verdien viser her at variablene «Storrelse» og «prom» har en påvirkning på hverandre, da verdiene er henholdsvis 4,12 og 3,97. Det er ingen fasit på hva som er en høy verdi og når man burde fjerne en variabel. Vi har tatt utgangspunkt i studien til Robinson og Schumacker som regner en verdi på 4-5 for å være moderat til høy (Robinson & Schumacker, 2009). Dersom variablene ikke

påvirker hverandre (multikollinearitet) vil GVIF-verdien være rundt 1. At «prom» og «Storrelse» påvirker hverandre er som tidligere nevnt naturlig, da leilighetsstørrelsen «Liten», «Mellomstor» eller «Stor» avhenger av hvor mange kvadratmeter leiligheten er.

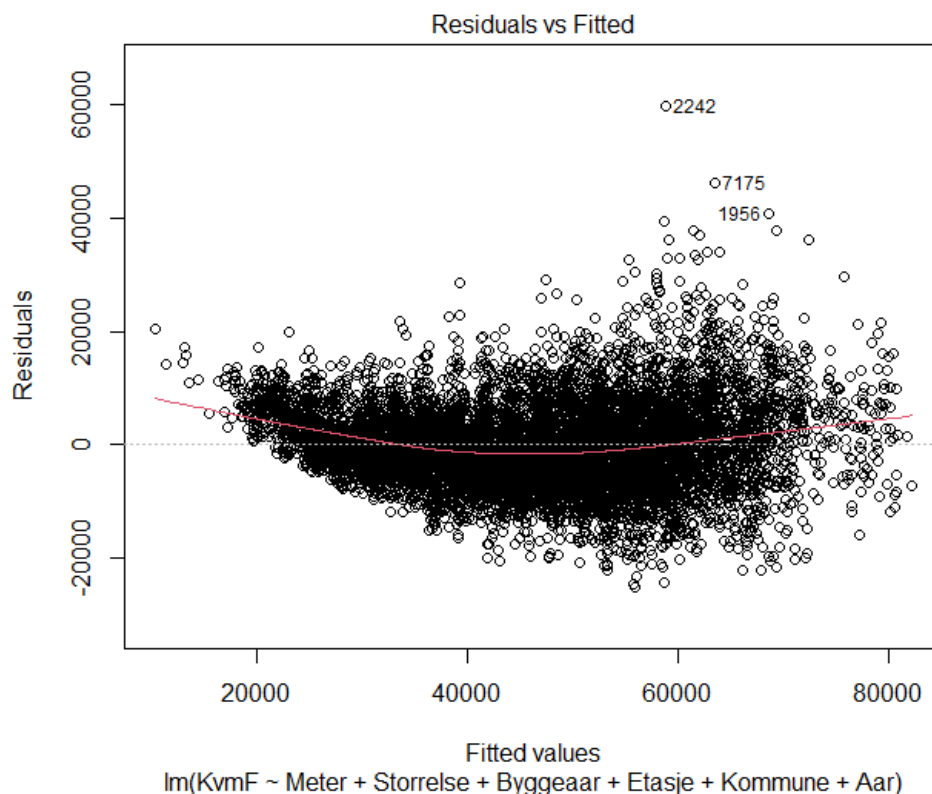
Basert på testen, vil «prom» fjernes som er variabel fra regresjonsmodellen da det antas at forholdet mellom «prom» og «Storrelse» er såpass multikollineært at det vil påvirke resultatene. I vurderingen av hvilken variabel som skal fjernes, må det tas utgangspunkt i hva formålet med studien er. For å kunne undersøke om avstand til togstasjon påvirker leilighetspriser for ulike størrelseskategorier, må variabelen «Storrelse» beholdes, og «prom» fjernes. Etter å ha fjernet «prom» var det ingen tegn på multikollinearitet mellom variablene da GVIF-verdiene var rundt 1 (tabell 12).

Tabell 12: Multikollinearitetstest av variabler i multippel regresjonsmodell.

	GVIF
<i>Meter</i>	1,156196
<i>Storrelse</i>	1,147383
<i>Aar</i>	1,010643
<i>Byggeaar</i>	1,134086
<i>Etasje</i>	1,044845
<i>Kommune</i>	1,182491

De siste testene som vil gjøres, og som vil repeteres for regresjonsmodellen for hver by, er Residuals vs Fitted (figur 40) og Q-Q plot (figur 41). Residuals vs fitted plottet i figur 42 viser residualverdiene på Y-aksen og «fitted» verdier på X-aksen. Med «fitted» verdier menes det estimerte verdier, eller respons. Modelltesten brukes for å se om det er linearitet og peke ut avvikende observasjoner. For å se om det er linearitet er det observasjonenes (de svarte punktene) plassering som ses på. Dersom de ligger tilfeldig spredt rundt den stiplede linjen, er dette et tegn på linearitet. I tilfeller hvor de ligger systematisk eller at noen av observasjonene har ekstremverdier som avviker fra gjennomsnittet, vil det føre til at den røde linjen (som er gjennomsnittet) for eksempel har en U-form. Dette er et tegn på at modellen ikke er fullstendig lineær (Som er en forutsetning er lineære regresjon). Det er likevel viktig å notere seg at tolkningen av disse plottene kan være veldig subjektive og man kan ikke forvente at den røde linjen er helt rett i alle regresjonsanalyser (Simon & Young, 2021). Q-Q

plottet (figur 41) er et grafisk verktøy som viser sannsynligheten for at datasettet som er brukt kommer fra en fordeling som f.eks. en normalfordeling. Det er en forutsetning for lineær regresjonsanalyse at den avhengige variabelen er normalfordelt. Med andre ord kan man si at Q-Q plot gjør en sammenligning av estimert verdi for, i vårt tilfelle, kvadratmeterpris opp mot en teoretisk verdi og angir normalfordelingen i kvadratmeterprisen som avhengig variabel. Dersom det er lite avvik i estimatene, vil de svarte punktene ligge på den stiplede linjen. Dersom det er større avvik, vil de svarte punktene ikke ligge på linjen noe som betyr at kvadratmeterprisen for disse observasjonene ikke er normalfordelte. I likhet med Residuals vs Fitted plottet er dette kun en grafisk representasjon og må tolkes subjektivt. Plottet er derfor ingen fasit, og må heller brukes som en indikator på om forutsetningen om normalfordeling holder (Ford, 2015).

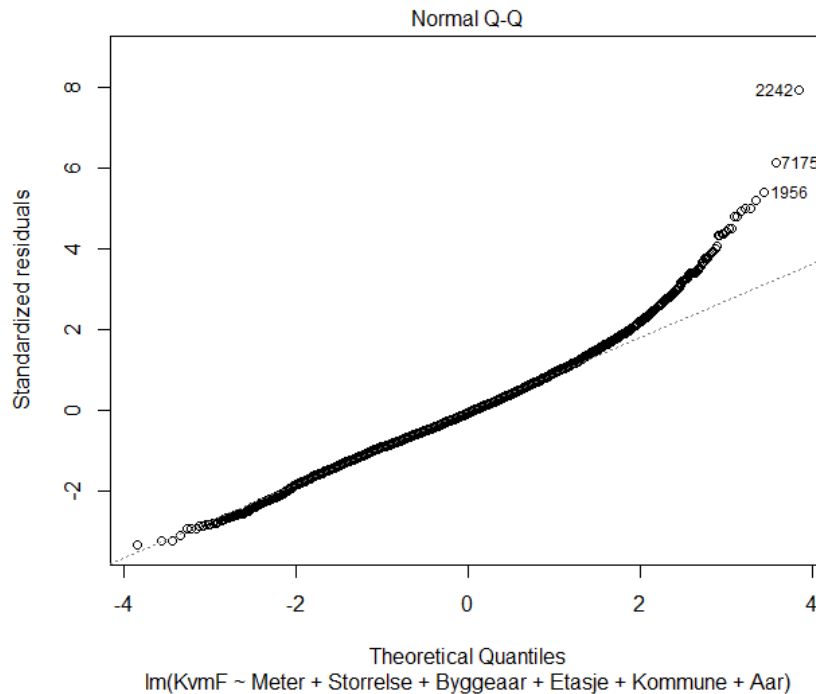


Figur 40: Egenprodusert Residuals vs fitted plot som eksempel og introduksjon.

I figur 40 kan man observere «Residuals vs Fitted» plottet til regresjonsmodellen som er vist i figur 39. Den røde linjen er som nevnt et gjennomsnitt av alle residualverdiene. For denne modellen kan man observere at den røde linjen har en U-form. Dette kan som nevnt over tyde på at det ikke er det fullstendig lineært forhold mellom kvadratmeterpris (den avhengige variabelen) og de øvrige uavhengige variablene. I dette tilfellet ser det ut til at det er noen

ekstremverdier som påvirker resultatet, og noen av de mest ekstreme er oppgitt med et observasjonsnummer.

Q-Q plottet (normal probability plot) (figur 41) brukes som nevnt for å se om det er avvik i den estimerte kvadratmeterprisen og teste om kvadratmeterprisen er normalfordelt.



Figur 41: Egenprodusert Q-Q plot som eksempel og introduksjon.

Figur 41 viser tegn på at det er avvik på høyre side. Dette kan som nevnt indikere at kvadratmeterprisen ikke er normalfordelt basert på modellen i figur 39.

Ettersom dette kun er en introduksjon til hvordan modellene testes, vil det ikke foretas ytterligere justeringer foreløpig. Datasettet som er brukt for denne presentasjonen omfatter alle byene, og hva som regnes for å være avvik her, er ikke nødvendigvis et avvik når modellene for Lillestrøm, Ski og Sandvika kjøres separat. For å unngå variasjoner tilknyttet hvilken by observasjonen befinner seg i, vil det gjøres spesifikke tester for regresjonsmodellen for hver separat by.

4.5 Lineær multippel regresjon for hver by

Som nevnt er en multippel regresjon den mest hensiktsmessige modellen å bruke for å ta hensyn til den hedoniske metoden. Basert på testene over, er det grunnlag for å anta at analysene burde kjøres separat for de tre byene. Valget om separate modeller støttes av

Robinson og Schumacker i deres artikkel «Interaction Effects: Centering, Variance Inflation Factor, and Interpretation Issues" (Robinson & Schumacker, 2009). I deres studie viste resultatene at for datasett hvor det er variasjon i variabelverdiene, eller hvor det er nødvendig med flere interaksjonsledd, er det hensiktsmessig å bryte opp modellene i separate tester. Ikke bare vil en stor modell muligens vise mangel på signifikansnivå, men det vil også være lettere å tolke estimatene fra separate tester. I tillegg vil separate tester kunne bidra til å redusere variasjoner, slik som at Lillestrøm, Ski og Sandvika har ulike prisnivåer for leiligheter. Det blir presentert eksempler på disse effektene i vedlegg 4 for å gi et innblikk i hvordan en større modell byr på disse utfordringene.

4.5.1 Lillestrøm

Figur 42 viser regresjonsmodellen som ble gjort med datasettet for Lillestrøm. For å besvare hypotesene, er det introdusert et interaksjonsledd i modellen. Leddet viser det estimerte forholdet mellom avstand til togstasjon og leilighetsstørrelsen. Variabelen «Meter» representerer effekten for små leiligheter, og de to siste leddene «Meter:StorrelseMellomstor» og «Meter:StorrelseStor» er et forholdstall som må legges til hovedvariabelen «Meter».

```
Call:
lm(formula = KvmF ~ Meter * Storrelse + Etasje + Byggeaar + Aar,
    data = dta)

Residuals:
    Min       1Q   Median       3Q      Max
-24431  -4509   -425    4255   39463

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)   -5.859e+06  5.424e+04 -108.019 < 2e-16 ***
Meter          -5.099e+00  3.867e-01  -13.186 < 2e-16 ***
StorrelseMellomstor -5.822e+03  5.417e+02  -10.746 < 2e-16 ***
StorrelseStor   -6.269e+03  6.797e+02   -9.224 < 2e-16 ***
Etasje         6.578e+02  8.857e+01    7.427 1.33e-13 ***
Byggeaar       1.380e+02  4.618e+00   29.887 < 2e-16 ***
Aar            2.802e+03  2.690e+01  104.164 < 2e-16 ***
Meter:StorrelseMellomstor -2.273e+00  4.602e-01  -4.939 8.14e-07 ***
Meter:StorrelseStor  -2.776e+00  6.111e-01  -4.542 5.72e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 7169 on 4494 degrees of freedom
Multiple R-squared:  0.7828,    Adjusted R-squared:  0.7824
F-statistic: 2025 on 8 and 4494 DF,  p-value: < 2.2e-16
```

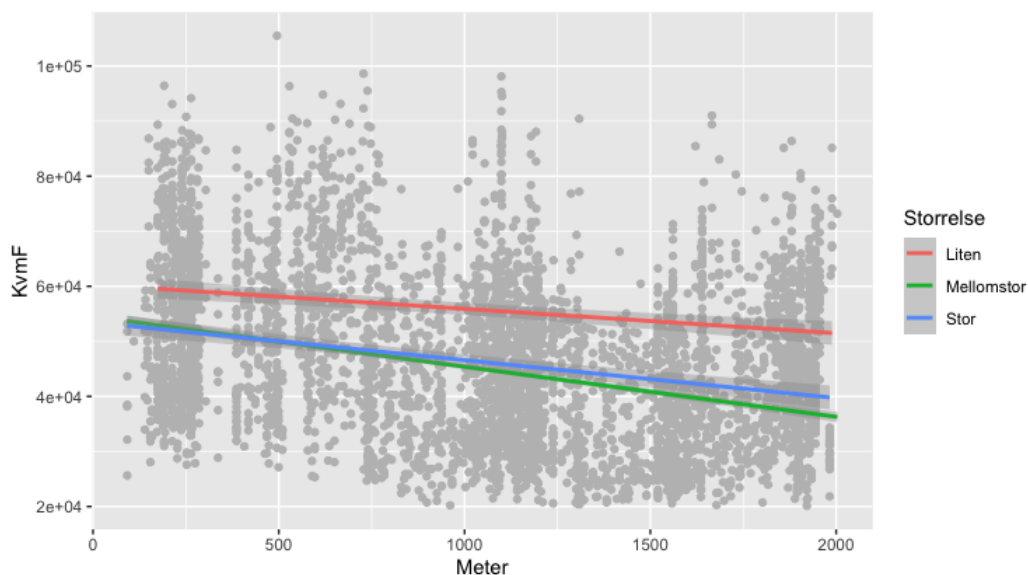
Figur 42: Egenprodusert lineær regresjonsmodell for Lillestrøm.

Alle variablene i regresjonsmodellen har signifikansnivå på 99%. I tillegg ligger modellens forklaringskraft på 78,28%, noe vi anser for å være høyt med tanke på hvor mange variabler som kan forklare boligprisen.

Som man kan se, har det blitt introdusert et interaksjons ledd i regresjonsmodellen, eksempelvis «Meter:StorrelseMellomstor». Interaksjonsleddet kan forstås som at ved en økning på 1 meter i avstanden til togstasjon i tillegg til å gå fra liten til mellomstor leilighet, synker kvadratmeterprisen med -2.273kr. Det er verdt å merke seg at regresjonsmodellen her, og for senere i analysen, ikke vil synliggjøre små leiligheter. Årsaken til dette ligger i at R-studio bruker små leiligheter (som en kategorisk variabel) som sammenligningsgrunnlag for de andre størrelsene. Resultatet av dette er at små leiligheter ligger i «Intercept». Samtidig vil «Meter» svare til hvor mye avstand til togstasjon påvirker kvadratmeterprisen for små leiligheter.

Når det kommer til selve estimatene, ser vi umiddelbart at det er en forskjell mellom de ulike størrelsene. «Meter» som hovedeffekt for små leiligheter har en verdi på -5,099, altså ved en økning i avstanden fra togstasjonen på 1 meter, synker kvadratmeterprisen på små leiligheter med 5,099kr. Tilsvarende synker kvadratmeterprisen for mellomstore leiligheter med: $-5,099 + -2,273 = 7,372$ kr og for store: $-5,099 + -2,778 = -7,875$ kr. Små leiligheter ser ut til å bli påvirket minst av avstanden til togstasjon, og forholdet mellom store og mellomstore leiligheter ser ikke ut til å være stort.

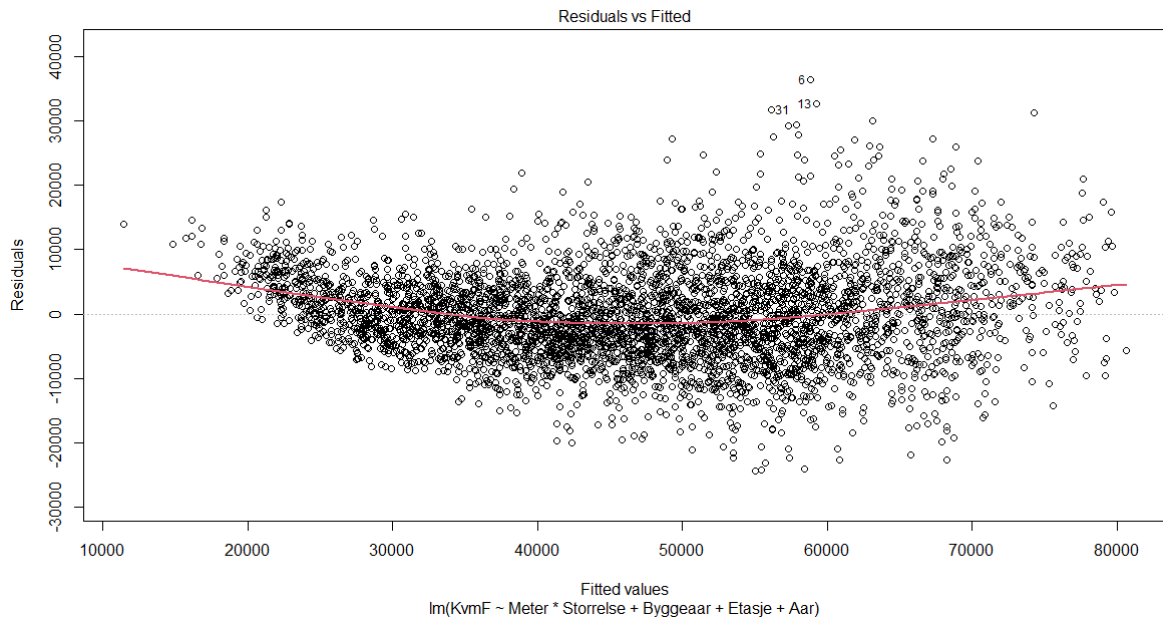
Figur 43 viser den gjennomsnittlige nedgangen i kvadratmeterprisen for de ulike leilighetsstørrelsene. Rød linje er små leiligheter, grønn linje er mellomstore og blå linje er store.



Figur 43: De fargede linjene viser regresjonslinjene for kvadratmeterpris for ulike leilighetsstørrelser. Kvadratmeterprisen for ulike leilighetsstørrelser på Y-aksen og avstand til togstasjon på X-aksen for Lillestrøm.

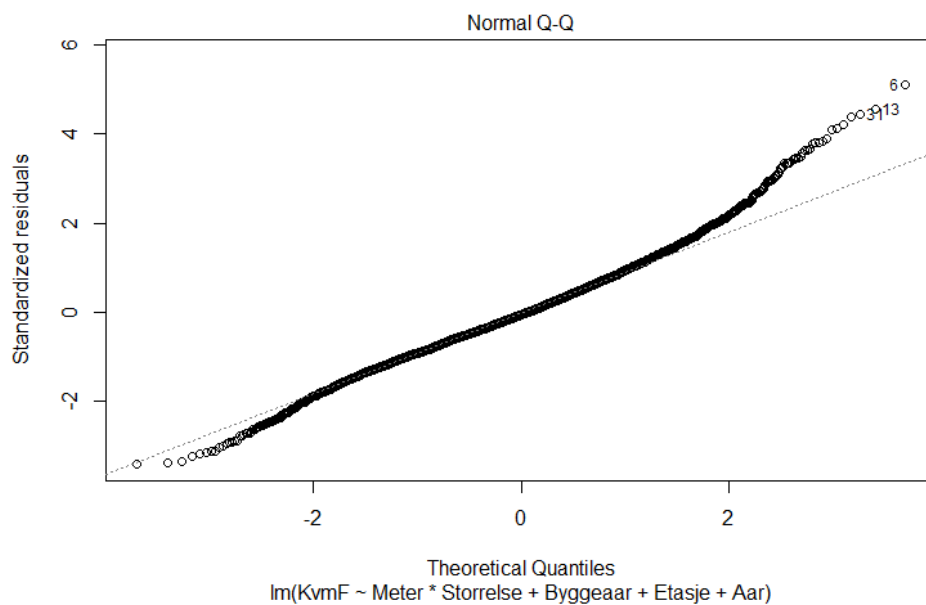
Figuren gir en god forståelse for hvordan kvadratmeterprisen synker med avstanden for de ulike leilighetsstørrelsene. Figuren representerer virkeligheten, og viser den faktiske nedgangen i kvadratmeterprisen for de ulike leilighetsstørrelsene når avstanden til togstasjonen øker. Figuren vil være et godt sammenligningsgrunnlag for regresjonsmodellen, da den viser hvordan virkeligheten er. Som man kan se, er det faktisk kvadratmeterprisen for mellomstore leiligheter som synker mest. Årsaken er trolig samspillet mellom de andre variablene. Et eksempel på hvorfor man ser dette, kan være at det er flere eldre mellomstore leiligheter lengre unna togstasjonen sammenlignet med store leiligheter. Ved sammenligning med regresjonsmodellen ser man at den isolerte avstandseffekten ikke er synlig her. Derfor vil figuren over brukes som et sammenligningsgrunnlag, mens estimatene fra regresjonsmodellen egner seg bedre til å si noe om hvordan avstand til togstasjon påvirker kvadratmeterprisen isolert sett.

Under blir modellens Residuals vs fitted-plot vist (figur 44). I dette plottet, som i likhet med plottet i figur 40, kan man observere at den røde linjen har en U-form. Det vil si at det også i denne modellen er indikasjoner på at det ikke foreligger et fullstendig lineært forhold. I denne modellen kan man også se at det foreligger noen ekstremverdier som observasjon nr. 6, 13 og 31. Disse vil det være aktuelt å se nærmere på, og det kan være grunnlag for en ytterligere rensing av data for å oppfylle forutsetningen om et lineært forhold.



Figur 44: Residuals vs fitted plot for lineær regresjonsmodell for Lillestrøm.

Q-Q plottet på sin side (figur 45) viser liknende avvik. De samme observasjonene blir trukket frem, og ligger langt ute på høyre side av figuren. Det ser ut til at det er spesielt de høyere kvadratmeterprisene som er avvikende, da linjen for observasjonene stiger opp fra normallinjen. Dette kan være en indikasjon på at datasettet for Lillestrøm ikke er normalfordelt.



Figur 45: Q-Q plot for lineær regresjonsmodell for Lillestrøm.

4.5.2 Sandvika

Modellen for Sandvika har andre estimater og signifikansnivåer sammenlignet med Lillestrøm (figur 46).

```
Call:
lm(formula = KvmF ~ Meter * Storrelse + Etasje + Byggeaar + Aar,
    data = dta2)

Residuals:
    Min       1Q   Median       3Q      Max
-23513  -5206   -800    4402   43559

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)   -5.723e+06  7.746e+04 -73.882 < 2e-16 ***
Meter         -8.627e+00  9.036e-01  -9.547 < 2e-16 ***
StorrelseMellomstor -1.767e+04  1.385e+03 -12.753 < 2e-16 ***
StorrelseStor   -1.337e+04  1.380e+03  -9.692 < 2e-16 ***
Etasje        -1.051e+02  9.429e+01  -1.114  0.265
Byggeaar       2.335e+02  9.402e+00  24.833 < 2e-16 ***
Aar           2.646e+03  3.753e+01  70.502 < 2e-16 ***
Meter:StorrelseMellomstor  5.836e+00  1.009e+00  5.785 8.06e-09 ***
Meter:StorrelseStor   -7.372e-01  9.863e-01  -0.747  0.455
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 7975 on 2827 degrees of freedom
Multiple R-squared:  0.7201,    Adjusted R-squared:  0.7193
F-statistic: 909.3 on 8 and 2827 DF,  p-value: < 2.2e-16
```

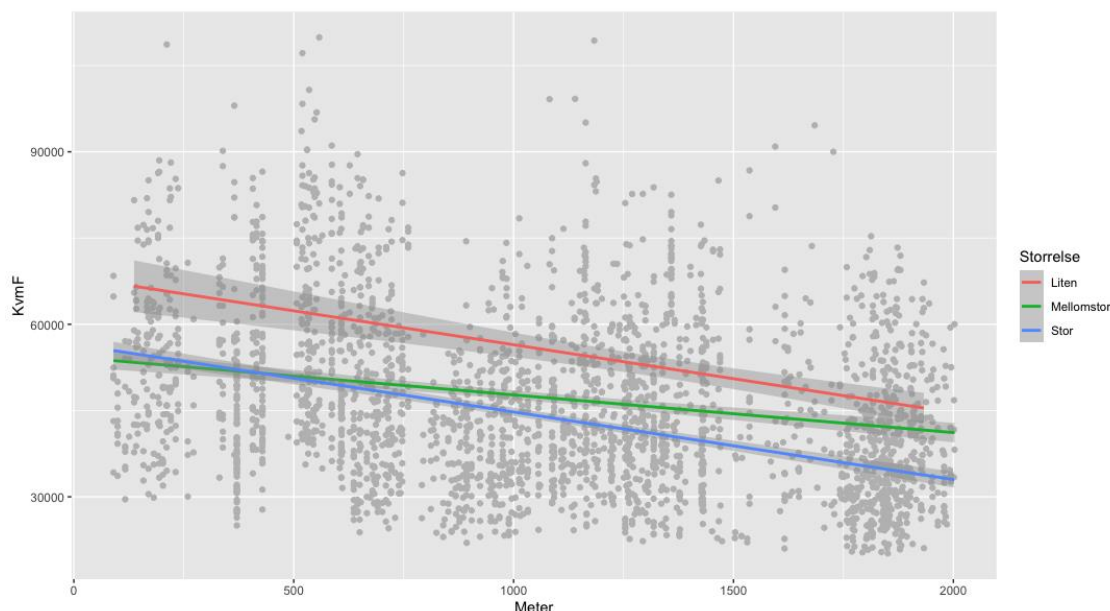
Figur 46: lineær regresjonsmodell for Sandvika.

For det første er ikke lenger estimatet for «Etasje» lenger innenfor målet om 95% signifikansnivå, og har også en negativ koeffisient, noe som betyr at når man går fra for eksempel 1. etasje til 2. etasje, synker kvadratmeterprisen med ca. 105kr. At estimatet er negativt har en sammenheng med signifikansnivået, og vi gir det ingen ytterligere tolkning. Interaksjonsleddet for avstand til togstasjon og store leiligheter er heller ikke signifikant, da det kun er innenfor 55%, noe som er overraskende lavt. Det kan være flere årsaker til hvorfor interaksjonsleddet ikke er signifikant. For det første kan det komme av at ikke eksisterer noen sammenheng mellom de to variablene. For det andre kan også årsaken være at det er store variasjoner i datamaterialet og da kvadratmeterprisen for store leiligheter. Testingen av modellen vil gi ytterligere innsikt i hva som forårsaker at interaksjonsleddet ikke lenger er signifikant.

De resterende variablene er signifikante (99% signifikansnivå). Små leiligheter har et estimat på -8,627. For hver meter avstanden fra togstasjonen øker vil kvadratmeterprisen reduseres med 8,627 kr. Mellomstore leiligheter har et positivt estimat, og kvadratmeterprisen blir altså

påvirket mindre enn små leiligheter. Estimaten tilsier at kvadratmeterprisen for mellomstore leiligheter synker med $-8,627 + 5,836 = -2,791$ kr. Modellens forklaringskraft er til tross for bortfall av to variabler forholdsvis høy på 72%.

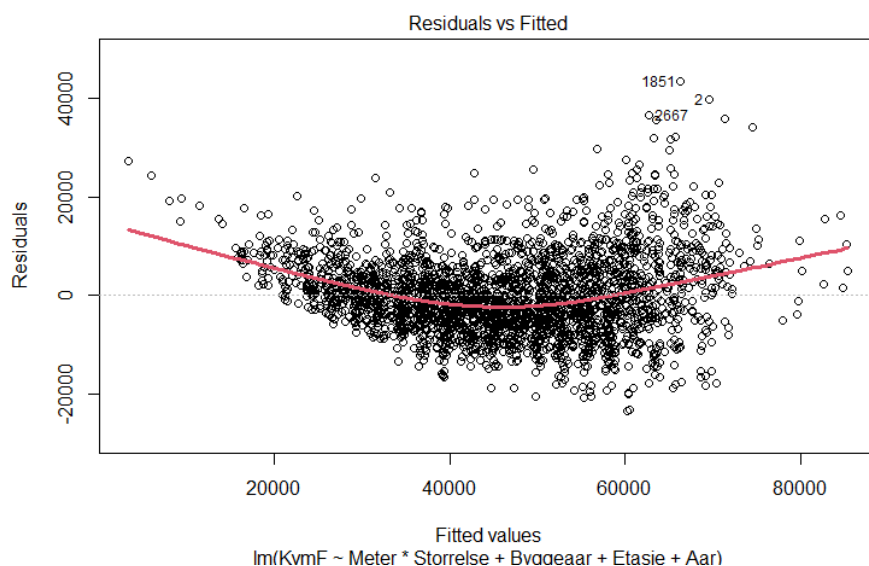
Det at mellomstore leiligheter blir påvirket i mindre grad av avstanden enn små leiligheter ser vi godt ut ifra figuren under (figur 47).



Figur 47: GG plot for kvadratmeterprisen for ulike leilighetsstørrelser på Y-aksen og avstand til togstasjon på X-aksen for Sandvika.

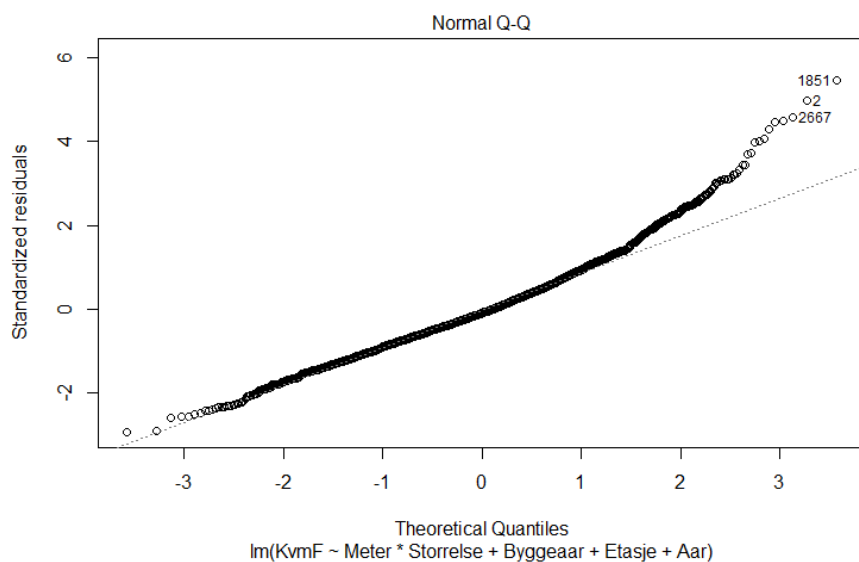
I likhet med analysen for Lillestrøm, viser figur 47 den virkelige nedgangen i kvadratmeterpris når avstanden blir større. Selv om signifikansnivået for interaksjonsleddet til store leiligheter ikke lenger var innenfor 95%, ser det ut til at kvadratmeterprisen for små og store leiligheter synker omtrent like mye basert på figur 49. Figuren vil ikke i seg selv gi grunnlag for å konkludere at det er en signifikant effekt, men den antyder at det kan være variasjoner i datamaterialet som er årsaken til manglende signifikansnivå.

Av Residuals vs fitted i figur 48 ser det ut til at det er langt større avvik i begge ender sammenlignet med plottet for Lillestrøm som fører til at U-formen på den røde linjen kommer mye tydeligere frem. Dette indikerer at forutsetningen om et lineært forhold står svakere i datasettet for Sandvika, sammenlignet med datasettet for Lillestrøm. Fra figur 48 kan man også observere at observasjonene er mer systematiske plassert og flere tydeligere ekstremverdier sammenlignet med Lillestrøm



Figur 48: Residuals vs fitted plot for lineær regresjonsmodell for Sandvika.

I Q-Q plottet ser det ut til at det spesielt er i de høyere prisnivåene at avvikene forekommer (figur 49). Det ser også ut til at det er noen observasjoner som skiller seg spesielt mye ut, slik som observasjon 1851, 2 og 2667. Basert på Q-Q plottet kan det se ut til datasettet for Sandvika ikke er normalfordelt i likhet med datasettet for Lillestrøm.



Figur 49: Q-Q plot for lineær regresjonsmodell for Sandvika.

4.5.3 Ski

I figur 52 nedenfor vil resultater fra lineær multippel regresjonsmodellen for Ski presenteres

```
Call:
lm(formula = KvmF ~ Meter * Storrelse + Etasje + Byggeaar + Aar,
    data = dta3)

Residuals:
    Min       1Q   Median       3Q      Max
-14869  -4186   -331    3823   24881

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)   -5.510e+06  9.280e+04 -59.372 < 2e-16 ***
Meter          6.819e+00  1.905e+00  3.580 0.000358 ***
StorrelseMellomstor -4.433e+03  1.595e+03 -2.778 0.005558 **
StorrelseStor   -3.630e+03  1.724e+03 -2.106 0.035430 *
Etasje         5.206e+02  1.396e+02  3.728 0.000203 ***
Byggeaar       1.746e+02  9.584e+00  18.214 < 2e-16 ***
Aar            2.588e+03  4.700e+01  55.076 < 2e-16 ***
Meter:StorrelseMellomstor -8.979e+00  1.951e+00 -4.603 4.66e-06 ***
Meter:StorrelseStor   -1.586e+01  2.126e+00 -7.458 1.79e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6042 on 1088 degrees of freedom
Multiple R-squared:  0.8064,    Adjusted R-squared:  0.805
F-statistic: 566.5 on 8 and 1088 DF,  p-value: < 2.2e-16
```

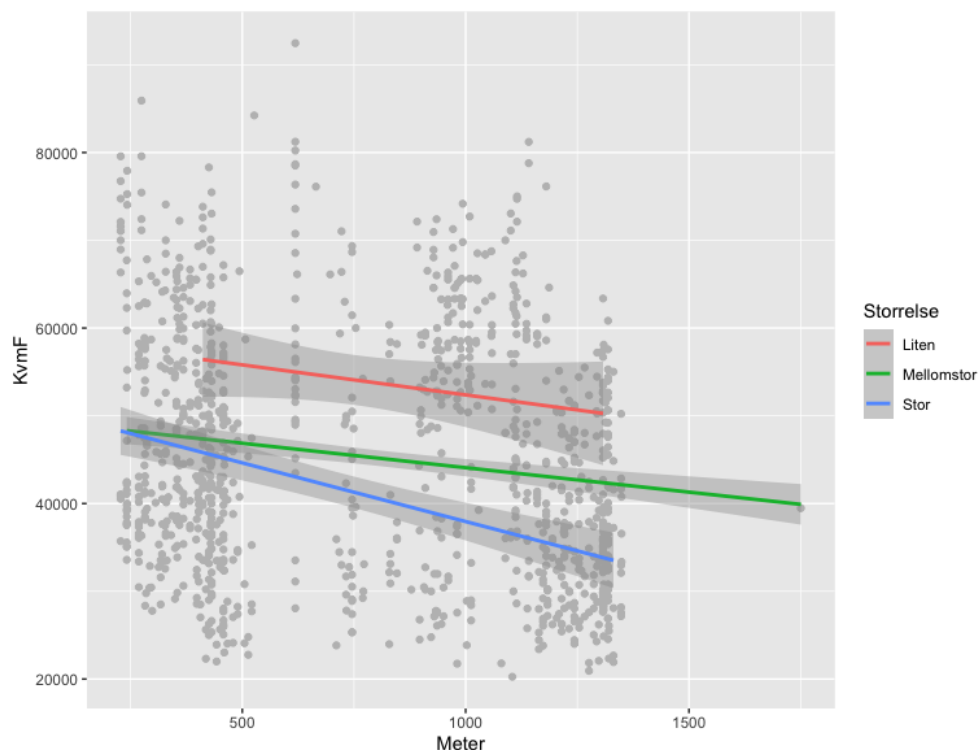
Figur 50: Lineær regresjonsmodell for Ski.

Regresjonsmodellen for Ski viser signifikante effekter på 99% nivå for alle variabler og interaksjonsledd, utenom «StorrelseStor» og «StorrelseMellomstor». Disse er likevel signifikante innenfor 95% nivå. Det som derimot ser ut til å skille seg ut, er estimatet for avstandseffekten på små leiligheter (Meter) som har et positivt fortegn. Kvadratmeterprisen for små leiligheter øker med 6,819kr når avstanden til togstasjonen øker med 1 meter. I likhet med Sandvika kan det være flere årsaker, og estimatet må tolkes ut ifra datasettet som har blitt brukt. Mest sannsynlig er dette et resultat av svært få observasjoner for kategorien små leiligheter. Selv om Ski har ca. 1100 observasjoner, er det kun 116 av disse i kategorien liten slik som vi så i den deskriptive statistikken.

For mellomstore leiligheter ser vi at det igjen er et negativt forhold. Prisnedgangen per meter er her på $6,819 + -8,979 = -2,16$ kr. Denne nedgangen er forholdsvis lik som for Sandvika. Store leiligheter i Ski ser ut til å ha et prisfall per meter på $6,819 + -15,86 = -9,041$ kr, noe

som er den klart største effekten av de tre byene. Modellen for Ski anses for å ha en svært høy forklaringskraft med 80,64%.

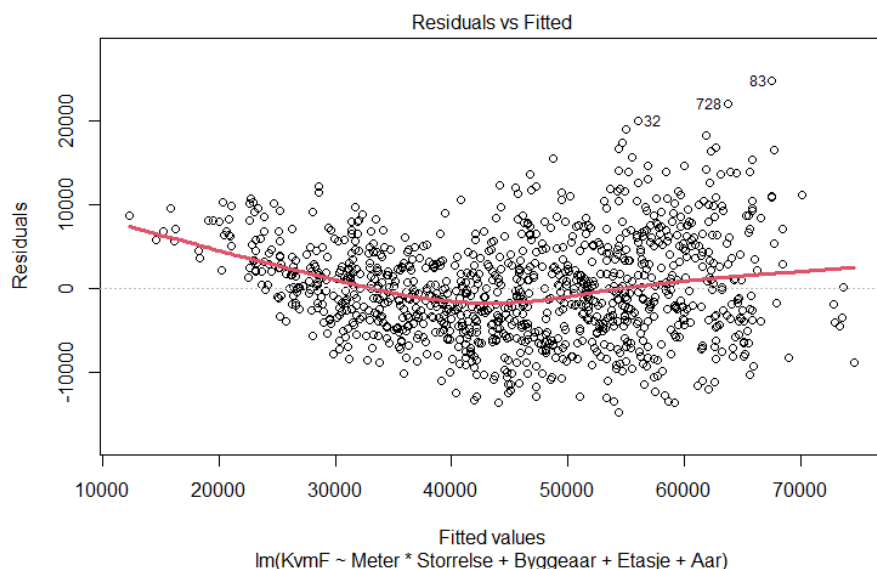
Antagelsen om at antall observasjoner påvirker estimatene styrkes av observasjonene vist i figur 51.



Figur 51: GG plot for kvadratmeterprisen for ulike leilighetsstørrelser på Y-aksen og avstand til togstasjon på X-aksen for Ski.

Som figur 51 viser er det ingen observasjoner utenfor ca. 1350 meter fra togstasjonen (med unntak av to). Det er tydelige tegn på at observasjonene ligger i klynger. Ved for eksempel 600 meter ser man en rekke observasjoner som tilhører samme sted, da de ligger i en rett linje over hverandre. I likhet med figuren for Sandvika (figur 47), ser man også her at estimatene fra regresjonsmodellen er forskjellige fra hva virkeligheten tilsier. Figuren her illustrerer at kvadratmeterprisen for små leiligheter ikke øker når avstanden blir større, og at det muligens er andre variabler som heller reduserer prisen når avstanden blir større.

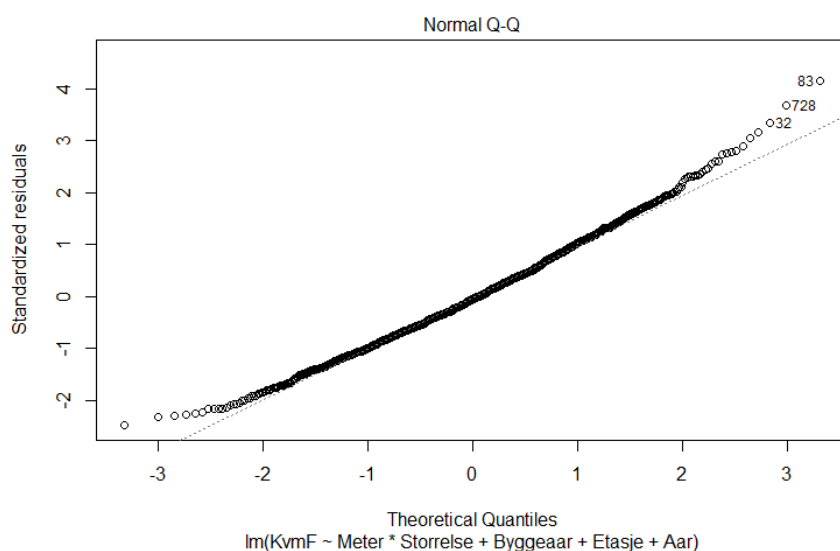
For modelltestingen ser plottet for Residuals vs fitted ut som en kombinasjon av Lillestrøm og Sandvika (figur 52).



Figur 52: Residuals vs fitted plot for lineær regresjonsmodell for Ski.

På venstre side er det en bratt synkende kurve og på høyresiden ser man en svak stigende kurve som også her danner en slags U-form. Igjen ser det ut til at det er noen ekstremverdier som forårsaker store deler av denne U-formen. I likhet med datasettet for Lillestrøm og Sandvika, står forutsetningen om et lineært forhold noe svakt for datasettet for Ski.

Q-Q plottet har derimot en forholdsvis jevn linje, sammenlignet med hva Residuals vs Fitted skulle tilsi (figur 53). Den beveger seg noe opp nederst til venstre. Samtidig ser det også ut til at det er enkeltobservasjoner som skiller seg ut i begge ender. Her blir observasjon 83, 728 og 32 pekt ut som de største avvikene. Det ser altså ut til at forutsetningen om normalfordeling står sterkere i datasettet for Ski.



Figur 53: Q-Q plot for lineær regresjonsmodell for Ski.

4.5.4 Oppsummering av Lineær Multippel regresjon

Totalt sett viser modellene at det varierer hvor mye hvert segment blir påvirket av avstand til togstasjon. For Lillestrøm er små leiligheter det segmentet hvor kvadratmeterprisen synker minst når avstanden øker. Kvadratmeterprisen for mellomstore- og store leiligheter blir påvirket i veldig lik grad. For Sandvika viser modellen andre estimater. Det mellomstore segmentet er det som blir påvirket minst, og avstand til togstasjon ser ikke ut til å ha en sterk effekt på kvadratmeterprisen. For Sandvika viste også modellen en ikke-signifikant sammenheng mellom kvadratmeterprisen på store leiligheter og avstand til sentrum. Ski på sin side har de desidert største forskjellene mellom leilighetssegmentene. Estimaten tilsier at kvadratmeterprisen øker når avstanden til togstasjonen blir større. Mellomstore leiligheter er i likhet med Sandvika det segmentet som blir påvirket minst av avstanden, og kvadratmeterprisen på store leiligheter synker i betraktelig grad. Store leiligheter er det segmentet som blir påvirket mest av avstand for alle byene.

4.6 Datarensing

Ut ifra testingen av regresjonsmodellene er det tegn på at det er avvik i datagrunnlaget, og at det burde gjennomføres en ny datarensing. Formålet er å fjerne ekstremverdier som påvirker datasettet, slik som enkelte observasjoner som har skilt seg ut.

For de avvikende observasjonene så det ut til at kvadratmeterprisen var uforholdsmessig høy (eller lav i noen tilfeller) sett ut ifra de uavhengige variablene. Ved en grundigere gjennomgang så vi at flere av observasjonene hadde unormalt høy kvadratmeterpris. Et eksempel på dette var en observasjon med en kvadratmeterpris på 84 000kr, bygd i 1926, lå i 1. etasje 1800m unna togstasjonen. I enkelte observasjoner var det også verdimangel, hvorav to av observasjonene hadde byggeår satt til 0.

Detaljrensingen ble gjennomført fire ganger for hver by, hvor de avvikende observasjonene ble kontrollert og fjernet. Rensingen resulterte til at 12 observasjoner ble fjernet for hver by, altså 36 totalt. Vi satte en grense på 12 for å ikke fjerne for mange observasjoner og på denne måten manipulere dataene i for stor grad. Selve rensingen vil ikke være synlig i modelltestingen i form av at den ser mye bedre ut, da det alltid vil være relative avvik. Rensingen fjerner derimot de mest ekstreme observasjonene, og kan gi utslag i videre

analyser. Basert på hvor avvikende modelltestene var, vil det måtte gjøres ytterligere tilpasning, gjennom en såkalt modelltransformasjon.

Når det kommer til testingen av regresjonsmodellene er det tegn på at lineær regresjon ikke nødvendigvis er beste tilnærming. For det første så vi ut ifra Residuals vs fitted for både Sandvika og Ski at residualverdiene økte når kvadratmeterprisen ble høy eller lav, altså at den røde linjen fikk en U-form. Også for Q-Q plottet så vi at normalfordelingslinjen gikk opp fra den stiplede linjen når kvadratmeterprisen økte for Lillestrøm og Sandvika. Basert på modelltestene er det tegn på at lineariteten er noe svak, og at dataene ikke er normalfordelte for byene, og krever en logaritmisk transformasjon. Ved å gjøre denne endringen vil kvadratmeterpris endres til logaritmisk form for å jevne ut variasjoner, samtidig som modellen fortsatt vil ha en lineær form. En grundigere forklaring av logaritmisk form vil komme i kapittel 4.7. Utover modelltransformasjon vil det også være aktuelt å gjennomføre en ny rensing av datamaterialet da det er enkeltobservasjoner som er svært avvikende.

4.7 Log-Level multippel regresjon for hver by

En måte å tilpasse modellen på er gjennom en logaritmisk transformasjon av en eller flere variabler. I vurderingen om hvilke variabler som skal transformeres må det tas utgangspunkt i modelltestene som ble presentert i forrige delkapittel. Kenneth Benoit skriver i artikkelen «Linear Regression Models with Logarithmic Transformations» at det er aktuelt å log-transformere en eller flere variabler i tilfeller hvor det er en skjevfordeling i datagrunnlaget. Ved å log-transformere en variabel vil den omgjøres til en ikke-lineær variabel, samtidig som modellen holdes på lineær form (Benoit, 2011).

Da modelltestingen ga utslag på at kvadratmeterprisen ikke har et lineært forhold, noe som også stemmer overens med at kvadratmeterprisen synker når leiligheten blir større, er det i hovedsak denne variabelen som burde transformeres. Skjevfordelingen kan også sees i den deskriptive statistikken (figur 35), hvor histogrammet tydelig viser at majoriteten av datapunktene ligger i den nedre enden av skalaen samt at det er noen leiligheter med svært høye kvadratmeterpriser. En log-transformasjon av boligpris har tidligere vist seg å gi gode resultater, slik som i studien gjort av Hoen hvor man så at gjennom log-transformasjon ble boligprisene langt mer normalfordelte (Hoen, 2018).

Endringen medfører også at estimatene må tolkes på nytt, da forholdet mellom Y og X endres. En endring i avstanden til togstasjon vil nå medføre en prosentvis endring i kvadratmeterprisen. Siden estimatene også er på logaritmisk form, vil en omregning av estimatene fra log-level modellene komme i neste delkapittel.

Grafene som viser den virkelige prisnedgangen når avstanden til togstasjonen øker vil ikke bli presentert på nytt, da en tilpasning av regresjonsmodellen ikke vil ha noen påvirkning på disse (figur 43 og 47 og 51).

4.7.1 Lillestrøm

Figur 54 viser Log-level regresjonsmodell for Lillestrøm

```
Call:
lm(formula = log(KvmF) ~ Meter * Storrelse + Etasje + Byggeaar +
    Aar, data = dta)

Residuals:
    Min       1Q   Median       3Q      Max
-0.58637 -0.09151  0.00037  0.09412  0.54685

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    -1.190e+02  1.109e+00 -107.331 < 2e-16 ***
Meter           -9.909e-05  7.904e-06  -12.536 < 2e-16 ***
StorrelseMellomstor -9.388e-02  1.107e-02  -8.478 < 2e-16 ***
StorrelseStor    -9.807e-02  1.389e-02  -7.059 1.93e-12 ***
Etasje          1.106e-02  1.810e-03   6.111 1.08e-09 ***
Byggeaar        3.005e-03  9.439e-05  31.835 < 2e-16 ***
Aar             6.155e-02  5.499e-04 111.940 < 2e-16 ***
Meter:StorrelseMellomstor -6.328e-05  9.407e-06  -6.727 1.95e-11 ***
Meter:StorrelseStor   -7.652e-05  1.249e-05  -6.126 9.77e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1465 on 4494 degrees of freedom
Multiple R-squared:  0.802,    Adjusted R-squared:  0.8017
F-statistic: 2275 on 8 and 4494 DF,  p-value: < 2.2e-16
```

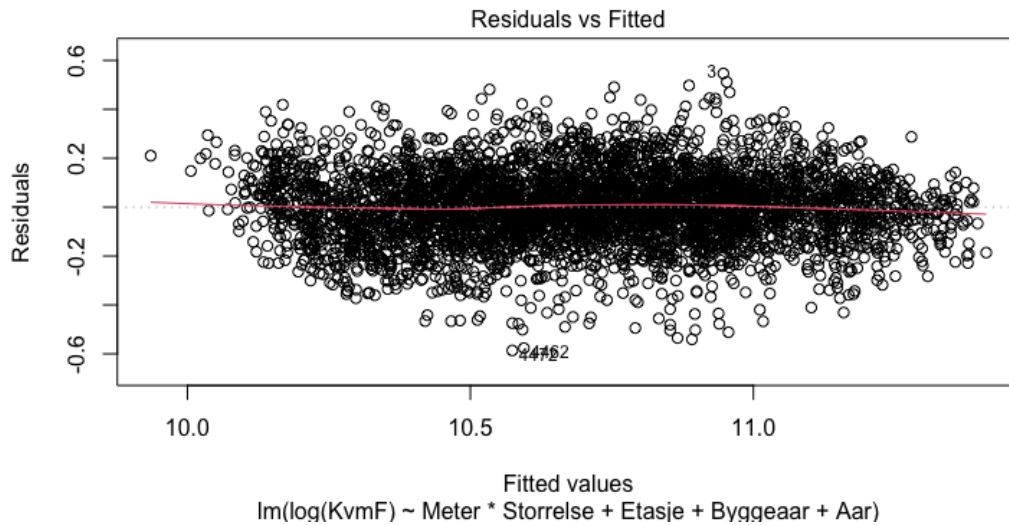
Figur 54: Log-level regresjonsmodell for Lillestrøm.

Log-level modellen viser høye signifikansnivåer, hvor alle variabler har et signifikansnivå på 99%. Modellens forklaringskraft har også økt fra å være på 78,38% til 80,2%.

Residuals vs fitted (figur 55) viser også tegn på at modellen er bedre tilpasset.

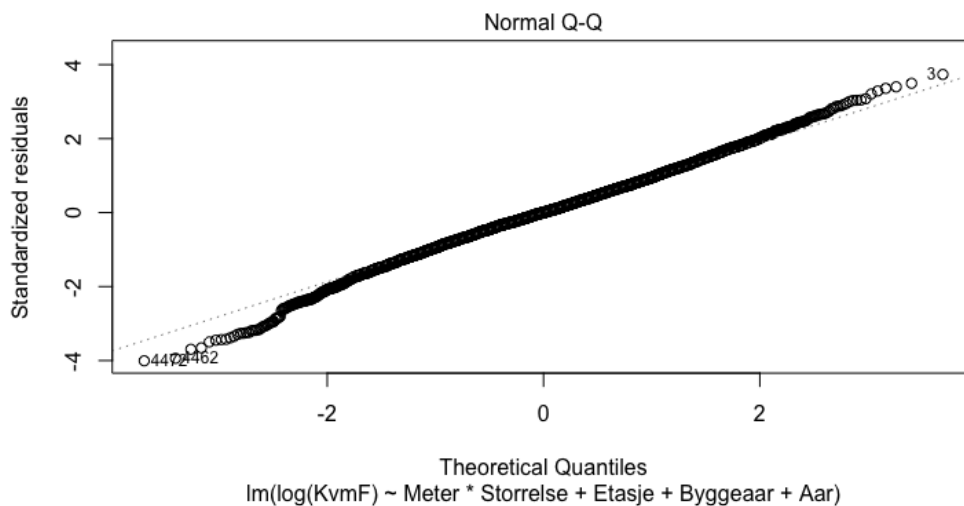
Observasjonene ser ut til å være mer eller mindre tilfeldig spredt. Den røde linjen støtter oppunder at modellen er bedre tilpasset, hvor gjennomsnittlig residualverdi nå er svært lik

null (stiplet linje). Det betyr at med denne tilpasningen er forutsetningen om et lineært forhold i større grad ivarettatt.



Figur 55: Residuals vs fitted for Log-level regresjonsmodell for Lillestrøm.

Q-Q plottet viser samme trend, hvor store deler av observasjonene ligger på rett linje oppå den stiplede linjen (figur 56).



Figur 56: Q-Q plot for Log-level regresjonsmodell for Lillestrøm.

I figur 56 ser det likevel ut til å være noe avvik for de laveste kvadratmeterprisene nederst til venstre. Sammenliknet med modelltesten for lineær regresjon, ser det likevel ut til at modellen er langt bedre tilpasset observasjonene for Lillestrøm. Flere av observasjonene ligger på den stiplede linjen, og avvikene er mindre.

4.7.2 Sandvika

Figur 57 viser log-level regresjonsmodell for Sandvika

```
Call:
lm(formula = log(KvmF) ~ Meter * Storrelse + Etasje + Byggeaar +
    Aar, data = dta2)

Residuals:
    Min       1Q   Median       3Q      Max
-0.51212 -0.10046 -0.00412  0.10070  0.55019

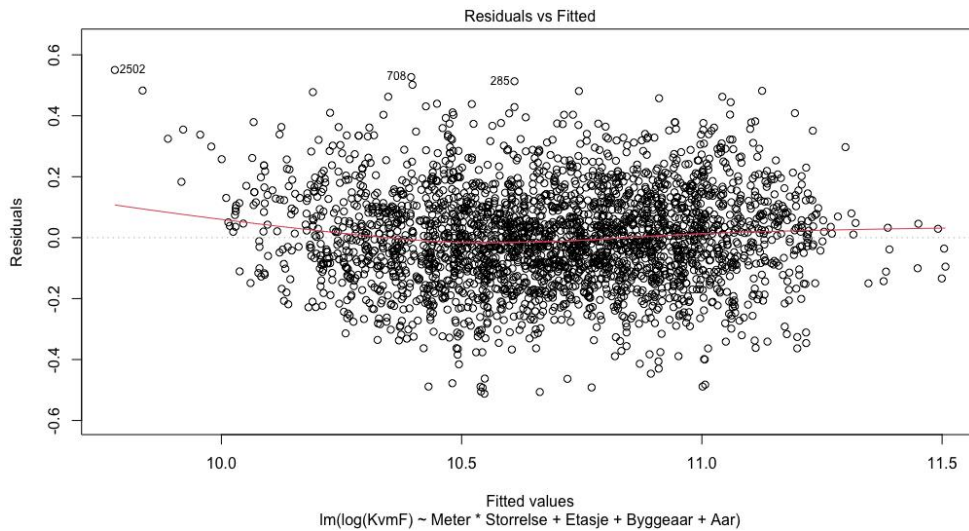
Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      -1.147e+02  1.527e+00 -75.106 < 2e-16 ***
Meter             -1.650e-04  1.781e-05  -9.261 < 2e-16 ***
StorrelseMellomstor -3.280e-01  2.731e-02 -12.012 < 2e-16 ***
StorrelseStor     -2.377e-01  2.720e-02  -8.740 < 2e-16 ***
Etasje           -4.349e-03  1.859e-03  -2.340  0.01936 *
Byggeaar          4.771e-03  1.853e-04  25.744 < 2e-16 ***
Aar               5.777e-02  7.397e-04  78.102 < 2e-16 ***
Meter:StorrelseMellomstor  9.765e-05  1.989e-05  4.910 9.61e-07 ***
Meter:StorrelseStor   -5.121e-05  1.944e-05  -2.634 0.00848 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1572 on 2827 degrees of freedom
Multiple R-squared:  0.758,    Adjusted R-squared:  0.7573
F-statistic: 1107 on 8 and 2827 DF,  p-value: < 2.2e-16
```

Figur 57: Log-level regresjonsmodell for Sandvika.

Etter log-level transformasjon har variablene som var ikke-signifikante i den lineære modellen blitt signifikante. Interaksjonsleddet «Meter:StorrelseStor» har nå et 95% signifikansnivå. Modellen har også fått en økt forklaringskraft. Den lineære modellen hadde 72,01% forklaringskraft, og log-level har her 75,8%.

Residuals vs fitted viser også tegn på forbedring, selv om det heller ikke her er et helt lineært forhold (figur 58).

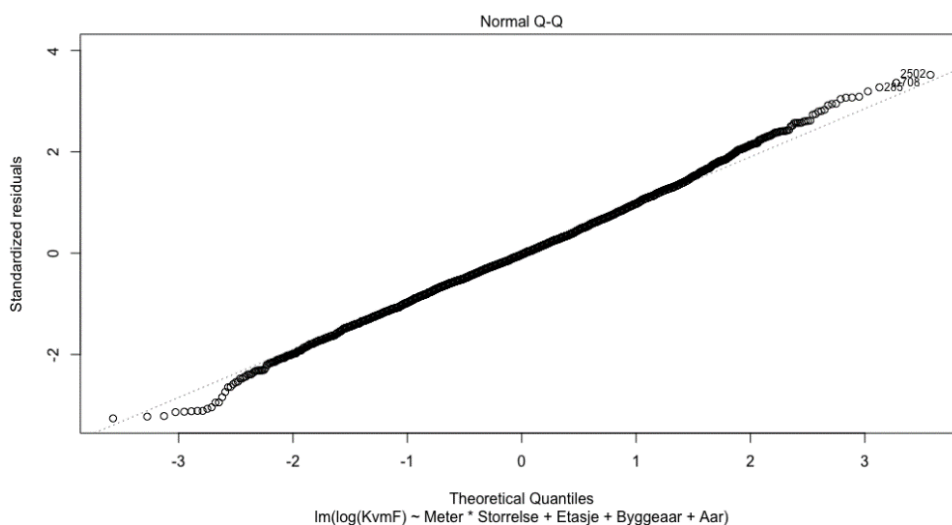


Figur 58: Residuals vs fitted plot for Log-level regresjonsmodell for Sandvika.

Det ser ut til at de største avvikene ligger på venstre side av residual vs fitted plottet, hvor det fortsatt ligger noen observasjoner med ekstremverdier, slik som observasjon 2502.

Observasjonene er generelt sett mer tilfeldig spredt rundt 0-linjen, og man kan også her se at den logaritmiske transformasjonen medfører en bedre tilpasset modell

Q-Q plottet for Sandvika viser også tegn på forbedring (figur 59).



Figur 59: Q-Q plot for Log-level regresjonsmodell for Sandvika.

Spesielt de høye kvadratmeterprisene har fått en god tilpasning, og ligger på en langt rettere linje sammenlignet med den lineære modellen. Det ser derimot ut til at det er noen

observasjoner på andre siden av plottet som er avvikende.. Til tross for dette, anses tilpasningen for å være god, og forutsetningen om normalfordeling ser ut til å være langt bedre for denne modellen.

4.7.3 Ski

Figur 60 under viser log-level regresjonsmodell for Ski.

```
Call:
lm(formula = log(KvmF) ~ Meter * Storrelse + Etasje + Byggeaar +
    Aar, data = dta3)

Residuals:
    Min       1Q   Median       3Q      Max
-0.33480 -0.08817  0.00052  0.08566  0.32385

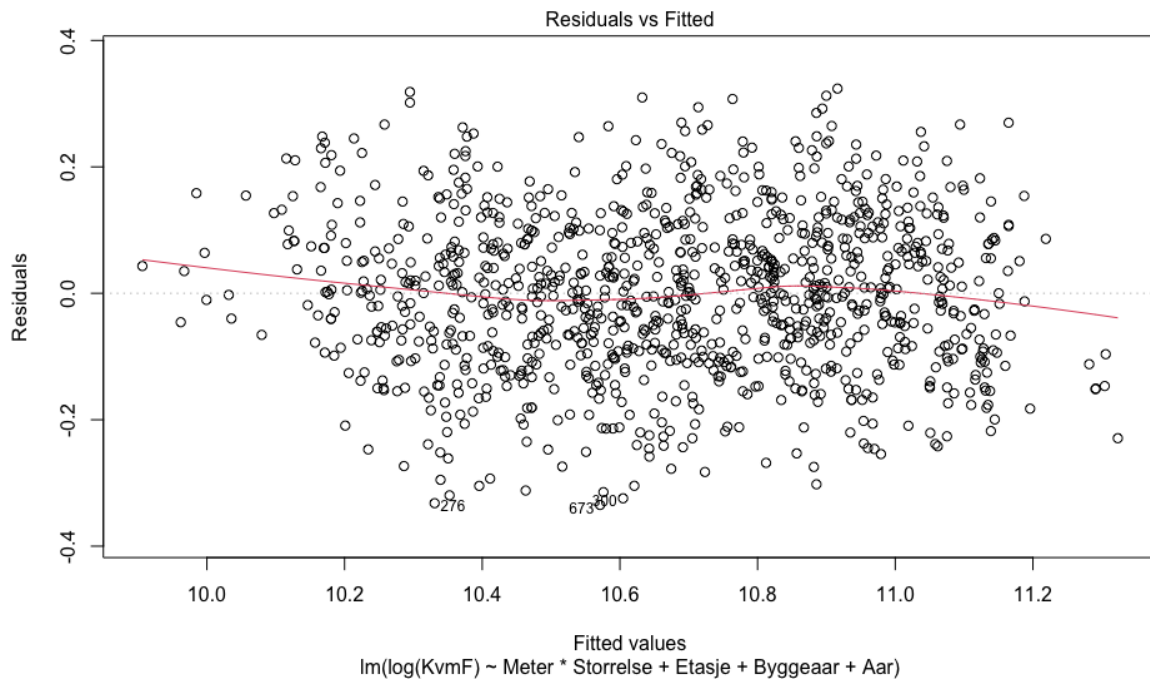
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.157e+02  1.911e+00 -60.526 < 2e-16 ***
Meter        1.399e-04  3.923e-05   3.565 0.000379 ***
StorrelseMellomstor -8.606e-02  3.286e-02  -2.619 0.008940 **
StorrelseStor    -6.098e-02  3.550e-02  -1.718 0.086163 .
Etasje         9.523e-03  2.876e-03   3.311 0.000959 ***
Byggeaar       3.902e-03  1.974e-04  19.770 < 2e-16 ***
Aar            5.895e-02  9.680e-04  60.896 < 2e-16 ***
Meter:StorrelseMellomstor -1.916e-04  4.018e-05  -4.769 2.11e-06 ***
Meter:StorrelseStor    -3.736e-04  4.380e-05  -8.529 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1244 on 1088 degrees of freedom
Multiple R-squared:  0.8351,    Adjusted R-squared:  0.8338
F-statistic: 688.5 on 8 and 1088 DF,  p-value: < 2.2e-16
```

Figur 60: Log-level regresjonsmodell for Ski.

Log-level transformasjonen av modellen for Ski ga noen små utslag. Variabelen «StorrelseStor» er nå kun signifikant på 90%, og «StorrelseMellomstor» er signifikant på 95%. Endringen anses ikke som svært betydelig, da de kun er en bi-variabel som et resultat av interaksjonsleddet. Med dette menes det at R-studio lager egne variabler kun for størrelsene som et resultat av å ha interaksjonsleddet. Variablene viser gjennomsnittlig differanse i kvadratmeterprisen, og er ikke interessante for denne studien. De andre variablene er signifikante på 99%. Modellens forklaringskraft har også her økt. Med en Multiple R-squared på 80,64% for den lineære modellen, ligger den nå på 83,51%.

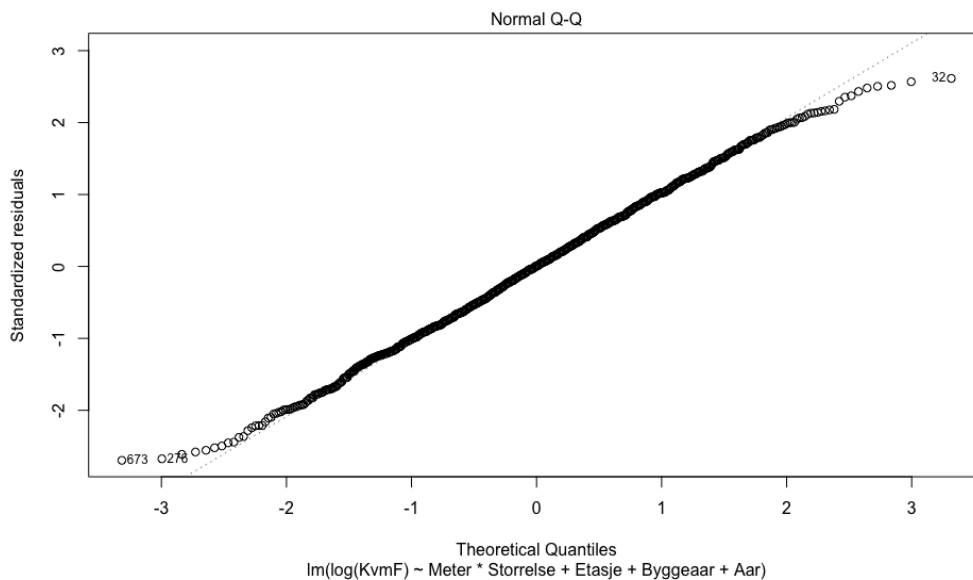
Residuals vs fitted (figur 61) viser et noe mer lineært forhold sammenlignet med modelltesten for lineær regresjon,



Figur 61: Residuals vs fitted plot for Log-level regresjonsmodell for Ski.

Observasjonene er mer tilfeldig spredt rundt den stiplede linjen (figur 63). Etter log-level transformasjonen hadde residualverdiene (rød linje) en noe slyngende form. Sammenlignet med modelltestingen på lineær form, anses dette likevel som en forbedring og forutsetningen om et lineært forhold står noe sterkere.

Q-Q plottet (figur 62) viser noe av de samme svingningene som Residuals vs fitted gjennom avvik i motsatt retning på hver side.



Figur 62: Q-Q plot for Log-level regresjonsmodell for Ski.

I Q-Q plottet er det mindre avvik sammenlignet med Q-Q plottet for lineær regresjon og majoriteten av observasjonene ligger på en rett linje. Basert på modelltestingene ser det ut til at log-transformasjonen har medført at forutsetningen om et normalfordelt datasett er innenfor det man kan forvente for et slikt datasett.

4.7.4 Oppsummering av log-level multippel regresjon

Etter å ha tilpasset modellen har blant annet signifikansnivået for interaksjonsleddet for avstand til togstasjonen og store leiligheter i Sandvika blitt styrket. I tillegg er det en generell økning i modellens forklaringskraft for de tre byene. Når det kommer til testingen av modellene, anser vi alle for å være forbedret i form av at variablene jevnt over forklarer mer av kvadratmeterprisen, og log-transformasjonen har medført et mer lineært forhold. Den samlede vurderingen må derfor være at log-level regresjon er en enda bedre modell, og det vil videre være estimatene fra disse som vil brukes i utregningene og hypotesetestingen. Som nevnt vil det derimot være vanskeligere å beskrive den faktiske endringen i kvadratmeterprisen uten en omregning av estimatene fra logaritmisk form til prosentvis endring. Utregninger vil bli presentert i neste kapittel.

4.8 Oppsummering av resultater og beregninger

Som nevnt medfører en log-transformasjon av den uavhengige variabelen at estimatene må tolkes ulikt sammenlignet med en lineær regresjon. Når den uavhengige variabelen log-transformeres, må den tolkes som en prosentvis endring når X øker med 1. At den uavhengige variabelen log-transformeres betyr at den må omregnes med hensyn på e . I figur 63 blir det presentert en utregning hvor estimatene for avstand og størrelse blir omregnet til å vise en prosentvis endring i kvadratmeterprisen. Avstandseffekten er isolert for å synliggjøre variabelen uten å ta hensyn til de andre variablene. Utregningen er gjort for intervaller på 250 m.

<u>Lillestrøm</u>								
Størrelse	250	500	750	1000	1250	1500	1750	2000
Liten	-2,48 %	-4,95 %	-7,43 %	-9,91 %	-12,39 %	-14,86 %	-17,34 %	-19,82 %
Mellomstor	-4,06 %	-8,12 %	-12,18 %	-16,24 %	-20,30 %	-24,35 %	-28,41 %	-32,47 %
Stor	-4,39 %	-8,78 %	-13,17 %	-17,56 %	-21,95 %	-26,34 %	-30,73 %	-35,12 %

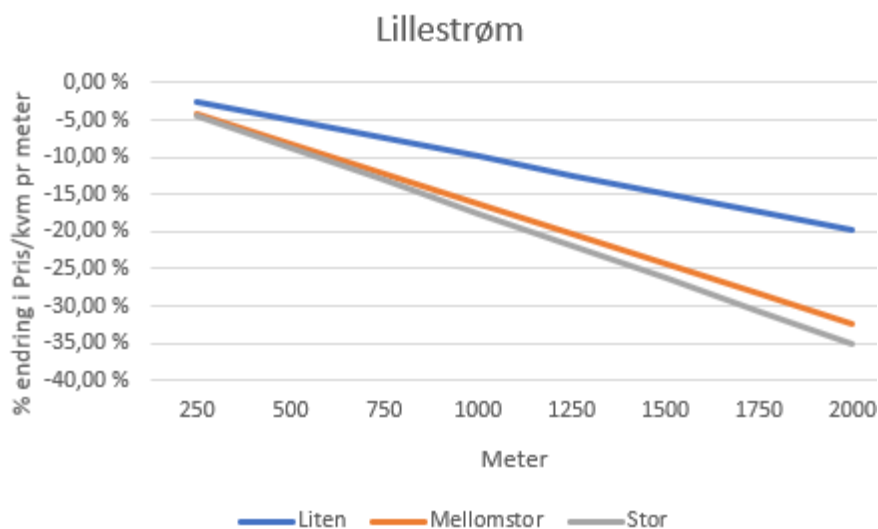
<u>Sandvika</u>								
Størrelse	250	500	750	1000	1250	1500	1750	2000
Liten	-4,12 %	-8,25 %	-12,37 %	-16,50 %	-20,62 %	-24,75 %	-28,87 %	-33,00 %
Mellomstor	-1,68 %	-3,37 %	-5,05 %	-6,73 %	-8,42 %	-10,10 %	-11,78 %	-13,47 %
Stor	-5,40 %	-10,81 %	-16,21 %	-21,62 %	-27,02 %	-32,43 %	-37,83 %	-43,24 %

<u>Ski</u>								
Størrelse	250	500	750	1000	1250	1500	1750	2000
Liten	3,50 %	7,00 %	10,49 %	13,99 %	17,49 %	20,99 %	24,48 %	27,98 %
Mellomstor	-1,29 %	-2,58 %	-3,88 %	-5,17 %	-6,46 %	-7,75 %	-9,04 %	-10,33 %
Stor	-5,84 %	-11,68 %	-17,52 %	-23,36 %	-29,20 %	-35,04 %	-40,88 %	-46,72 %

Figur 63: Utregninger basert på resultatene fra log-level regresjonsmodell. Estimaten for «Meter» er omregnet til en prosentvis endring i kvadratmeterprisen for en leilighet som starter på 80 000kr/kvm ved 0 meter.

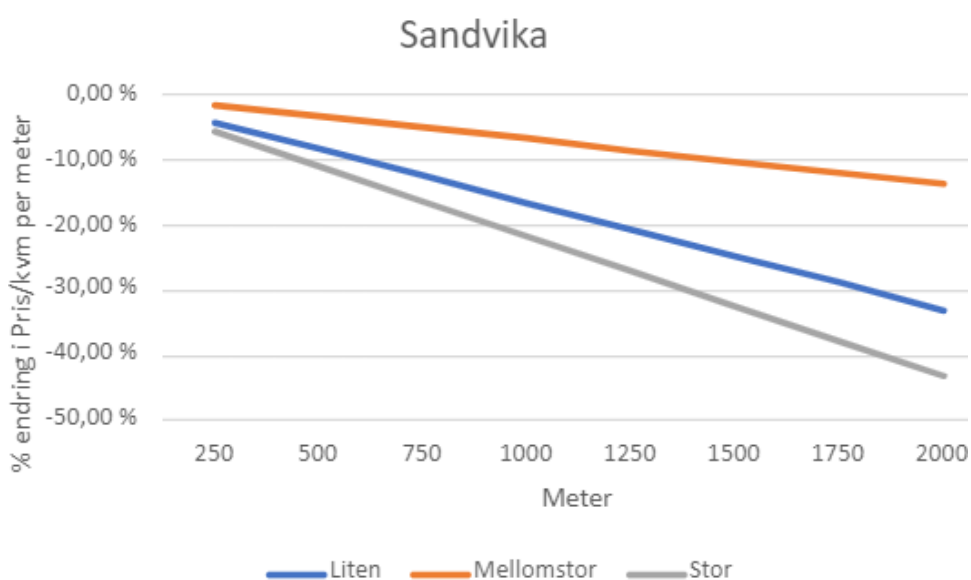
Ved å bruke den isolerte avstandseffekten (estimatene for hvordan avstand påvirker kvadratmeterprisen) for hver leilighetsstørrelse, justeres ikke endringen i kvadratmeterprisen for de andre variablene. Formålet med studien er å se på avstand som en variabel, og hvordan den påvirker leilighetsprisene. Ved å bruke hele modellen som en predikasjon for kvadratmeterprisen synliggjøres ikke avstand til togstasjon som en variabel da de andre variablene vil overskygge avstandseffekten. Predikasjon av kvadratmeterpris er heller ikke poenget med denne studien, og de andre variablene er kun inkludert for å gi et så nøyaktig estimat for avstand som mulig.

Som tidligere nevnt vil det være en viss sammenheng i variablene, slik som at det gjerne er nyere bygg lenger unna togstasjonen. Nyere bygg medfører en økning i boligprisene som justerer for avstanden. Det er likevel interessant å se på den isolerte effekten som avstand har på kvadratmeterprisene. Det er store forskjeller mellom leilighetsstørrelsene for hver av byene, men også byene imellom. For Lillestrøm ser man at små leiligheter i mye mindre grad blir påvirket av avstanden enn hva mellomstore og store leiligheter gjør (figur 64). Ved randsonen av studieområdet, altså på 2 km, utgjør dette mange tusen kroner per kvadratmeter.



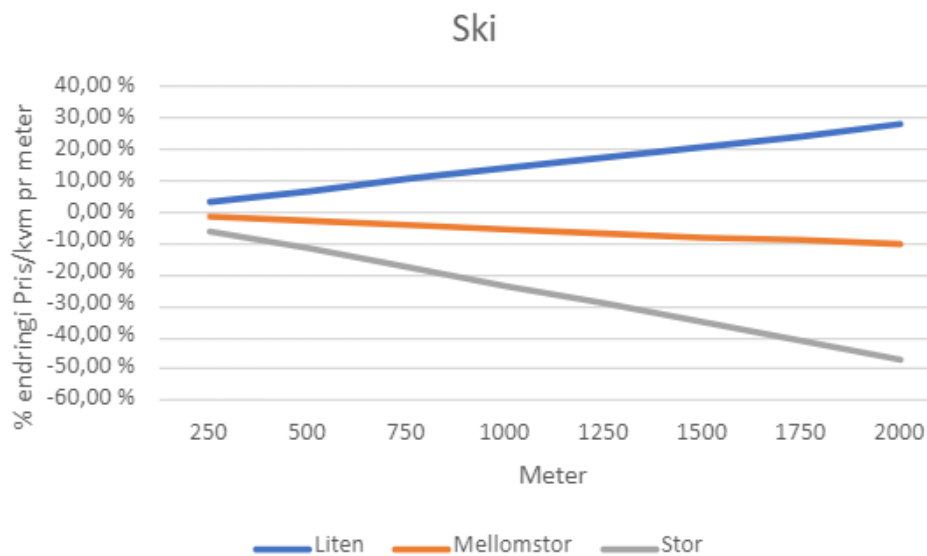
Figur 64: Viser hvordan avstand til togstasjon påvirker kvadratmeterprisen for de tre leilighetssegmentene i Lillestrøm. På X-aksen er antall meter, og Y-aksen er prosentvis endring i kvadratmeterprisen.

For Sandvika (figur 65) er det mellomstore leiligheter som er det segmentet der prisen faller minst. Sammenliknet med Lillestrøm er differansen mellom de ulike segmentene enda større, hvor avstanden isolert sett medfører et prisfall ved 2000m på 43,24% for store leiligheter.



Figur 65: Viser hvordan avstand til togstasjon påvirker kvadratmeterprisen for de tre leilighetssegmentene i Sandvika. På X-aksen er antall meter, og Y-aksen er prosentvis endring i kvadratmeterprisen.

At avstand er en betydelig faktor, ser vi også for Ski (figur 66). Det er viktig å påpeke at det mest sannsynlig foreligger uforholdsmessigheter i datagrunnlaget, da avstandseffekten er positiv for små leiligheter. Dette henger mest sannsynlig sammen med at mange små leiligheter som har blitt bygd i nyere tid, har vært forholdsvis langt unna togstasjonen.



Figur 66: Viser hvordan avstand til togstasjon påvirker kvadratmeterprisen for de tre leilighetssegmentene i Ski. På X-aksen er antall meter, og Y-aksen er prosentvis endring i kvadratmeterprisen.

5 Hypotesetesting

For å kunne besvare problemstillingen for studien, vil vi i første omgang teste underhypotesene vi har satt opp. Hypotesen vil testes ut ifra normalstandarden på 95% signifikansnivå. Med andre ord forkastes H_0 dersom vi med 95% sikkerhet vet at H_1 gjelder.

Hovedproblemstilling

- *Hvordan påvirkes kvadratmeterprisen for ulike leilighetsstørrelser av avstand til togstasjon i Lillestrøm, Sandvika og Ski?*

Underproblemstilling

1. *Følger kvadratmeterprisen for leiligheter monosentrisk byteori i form av at salgsprisen blir lavere jo større avstanden blir?*
2. *Er det ulik betalingsvilje for avstand til togstasjon for ulike leilighetsstørrelser?*
3. *Er betalingsviljen for avstand til togstasjon lik for byene Lillestrøm, Sandvika og Ski?*

Underproblemstilling 1:

Følger kvadratmeterprisen for leiligheter teorien om monosentrisk byteori i form av at prisen blir lavere jo større avstanden til togstasjonen blir?

Hypotese 1

- H0: Kvadratmeterprisen for leiligheter blir ikke påvirket av avstand til togstasjon
- H1: Kvadratmeterprisen for leiligheter blir påvirket av avstand til togstasjon

Med et utgangspunkt i de oppdelte regresjonsmodellene, viser alle tre at avstand som en hovedvariabel er signifikant på 99% signifikansnivå med negative koeffisienter. H0 forkastes og H1 beholdes på 99% signifikansnivå. Basert på dette kan man konkludere med at den monosentriske effekten er gjeldende.

Underproblemstilling 2:

Hvordan blir kvadratmeterprisen for de ulike leilighetsstørrelsene påvirket av avstand til togstasjon for byene Lillestrøm, Sandvika og Ski?

Hypotese 2

- H0: Kvadratmeterprisen for ulike leilighetsstørrelser blir ikke påvirket i ulik grad av avstand til togstasjon
- H1: Kvadratmeterprisen for ulike leilighetsstørrelser blir påvirket i ulik grad av avstand til togstasjon

Lillestrøm:

Basert på log-level transformasjonen av Lillestrøm viste estimatene en forskjell i hvilken grad avstand påvirker leilighetene (figur 54). Forskjellene var på 99% signifikans. H0 forkastes og H1 beholdes på 99% signifikansnivå. Vi kan derfor si med 99% sikkerhet at det er en forskjell på hvor mye avstand til togstasjon påvirker de ulike leilighetsstørrelsene i Lillestrøm by.

Sandvika:

Basert på log-level transformasjonen av Sandvika) viste estimatene en forskjell i hvilken grad avstand påvirker leilighetene (figur 57). Forskjellene var på 95% signifikans. H0 forkastes og

H1 beholdes på 95% signifikansnivå. Vi kan derfor si med 95% sikkerhet at det er en forskjell på hvor mye avstand til togstasjon påvirker de ulike leilighetsstørrelsene i Sandvika by.

Ski:

Basert på log-level transformasjonen av Ski viste estimatene en forskjell i hvilken grad avstand påvirker leilighetene (figur 60). Forskjellene var på 95% signifikans. H0 forkastes og H1 beholdes på 95% signifikansnivå. Vi kan derfor si med 95% sikkerhet at det er en forskjell på hvor mye avstand til togstasjon påvirker de ulike leilighetsstørrelsene i Ski by.

Underproblemstilling 3:

Er betalingsviljen for avstand til togstasjon lik for byene Lillestrøm, Sandvika og Ski?

Underproblemstilling 3 besvares ved å sammenlikne resultatene fra hypotese 2. Jf. Figur 63,64,65 og 66 er det tydelige tegn på at de ulike leilighetsstørrelsene blir påvirket i forskjellig grad avhengig av hvilken by det gjelder. Denne underproblemstillingen kunne blitt testet gjennom en hypotese, men på grunn av kompleksiteten til modellen som besvarer den eventuelle hypotesen, er den svært utfordrende å tolke. Modellen presenteres i vedlegg 4.

6 Konklusjon

Resultatene for underproblemstillingene viser klare tegn på at den monosentriske effekten er gjeldende internt for de tre byene med et signifikansnivå på 99%. Videre er det også tydelige tegn på at de ulike leilighetsstørrelsene blir påvirket i ulik grad av den monosentriske effekten etter en bid-rent teoretisk tolkning av boligmarkedet som sier at ulike kjøpegrupper har ulike preferanser for avstand til togstasjonen. Betalingsviljen er estimert som et mål på å kunne si noe om hvordan ulike kjøpegrupper verdsetter avstand til togstasjon, og resultatene viser her at det er forskjeller både mellom leilighetssegmentene samt mellom byene.

Hovedproblemstillingen for denne oppgaven baserer seg på en sammenligning av resultatene fra underproblemstillingene. Ved å se på utregningene for estimatene av de isolerte variablene for avstand, er det tegn på at det betalingsviljen for avstand til togstasjon er forskjellig når man ser på ulike leilighetssegmenter. Denne forskjellen gjør seg gjeldende på to måter. For det første er den monosentriske effekten i noe mindre grad framtreende for

Lillestrøm enn for de andre byene. Samtidig er det også forskjeller innad i størrelseskategoriene. Store leiligheter er det segmentet som generelt blir påvirket mest av avstand, men blir påvirket i mindre grad i Lillestrøm enn i Sandvika og Ski. Sandvika og Ski ser ut til å være relativt like, der mellomstore og store leiligheter blir påvirket omtrent like mye. Små leiligheter blir påvirket i relativt liten grad av avstand til togstasjon i Lillestrøm sammenlignet med de to andre størrelsene. For Sandvika ser man derimot at små leiligheter blir påvirket langt mer. For Ski viser resultatene at betalingsviljen øker når avstanden øker, noe vi anser for å være et lite troverdig resultat da observasjonsmengden er veldig liten.

Som en konklusjon på hovedproblemstillingen, er det klare tegn på at monosentrisk effekt er til stede, og at denne effekten påvirker kvadratmeterprisen for de ulike leilighetsstørrelsene i ulik grad. Effekten som avstand til togstasjon har på kvadratmeterprisen er også ulik for de tre byene, der det er tegn på at betalingsviljen for leilighetssegmentene er ulik avhengig av hvilken by det gjelder. Gjennom bruken av hedonisk metode har vi påvist at det eksisterer forskjeller i betalingsviljen når det kommer til ulike leilighetsstørrelser. Dette kan tyde på at det er ulike kjøpegrupper for de ulike leilighetsstørrelsene. Ikke bare er forskjellene tydelige mellom leilighetsstørrelsene i en by, men det er også tydelige forskjeller mellom Lillestrøm, Ski og Sandvika.

Funnene er av nytte både for eiendomsutviklere, kommuner og statlige virksomhet. Ved å få innsikt i hvordan ulike kjøpegrupper verdsetter avstand til togstasjon, vil man i ytterligere grad kunne legge opp fremtidige planer rundt dette. Innsikten vil kunne bidra til en enda mer tilpasset arealutvikling hvor formålet er redusert privatbilbruk til fordel for kollektivbruk.

7 Diskusjon

Formålet med masteroppgaven var å studere om kvadratmeterprisen på leiligheter ble påvirket av avstand til togstasjon. Leiligheter ble delt opp i tre kategorier etter normen for kategorisering av leiligheter, hvor liten, mellomstor og stor leilighet ble studert og analysert opp mot hverandre (Plan- og bygningsetaten, Oslo Kommune, 2016, s. 7). Studieområdet tok utgangspunkt i den regionale areal- og transportplanen for Oslo og Akershus, hvor blant annet Lillestrøm, Sandvika og Ski har blitt pekt ut som regionale vekstbyer som skal utvikles for å møte den fremtidige befolkningsveksten (Akershus fylkeskommune; Oslo Kommune, 2015, s. 7). I tillegg er det også et mål at oppfølging av planen skal bidra til nullvekst i

biltrafikken, og at alle skole- og arbeidsreiser i størst mulig grad skal skje gjennom å gå, sykle eller kollektivtransport (Akershus fylkeskommune; Oslo Kommune, 2015, s. 10). For å nå nullvekst i privatbilbruk som et ledd i målet om forbedret miljø, er det behov for kunnskap og forståelse for hvordan markedet etterspør nærhet til togstasjon.

Gjennom bruken av hedonisk metode har den implisitte betalingsviljen for ulike variabler blitt målt. Hovedvariabelen har vært avstand til togstasjon, hvor den estimerte betalingsviljen representerer markedets etterspørsel etter lokasjon. Gjennom bruken av bid-rent teori for boligmarkedet har de tre leilighetskategoriene representert ulike kjøpegrupper. Bid-rent teorien som har blitt presentert viser til at ulike kjøpegrupper (fordelt etter inntekt) har ulik betalingsvilje for lokasjon, og verdsetter pendlerkostnadene i ulik grad (Fujita, 1989) (Coulombel, 2010). Resultatene av regresjonsanalysen viser at avstand til togstasjon har en generell påvirkningseffekt på kvadratmeterprisen for leiligheter. Resultatene viser i de fleste tilfeller et negativt forhold, hvor kvadratmeterprisen er høyest i nærhet til togstasjonen, og synker med økende avstand. Estimaten svarer altså positivt på hypotesen om den interne monosentriske effekten, og er i tråd med gjeldende teori (Geltner et al, 2014, s.67).

I tilknytning den andre hypotesen viser resultatene at den interne monosentriske effekten påvirker de forskjellige leilighetsstørrelsene i ulik grad. Regresjonsanalysene svarer i hovedsak positivt på hypotesene om at kvadratmeterprisen for ulike leilighetsstørrelser blir påvirket i ulik grad av avstand til togstasjon både i Lillestrøm, Sandvika og Ski. Den presenterte bid-rent tolkningen av etterspørselssiden i boligmarkedet kommer tydelig frem i de resultatene for denne studien, hvor det er et skille mellom små, mellomstore og store leiligheter og at det er ulike kjøpegrupper med ulike attributtpreferanser. På den andre siden er resultatene motstridende fra tidligere forskning. Både (Fujita, 1989) og (Coulombel, 2010) sin forskning viser at husholdninger med høyere inntekt har en lavere betalingsvilje for å bo nærme togstasjonen sammenlignet med husholdninger hvor inntekten er lavere. De samme funnene er også presentert for Malmø-Lund regionen i Sverige hvor store boliger hadde en mindre prisnedgang når avstanden økte enn mindre boliger (Bohman & Nilsson, 2016). Funnene i denne studien viser det motsatte, hvor det tilsynelatende er større boliger som blir påvirket mest av avstand.

Det kan være mange årsaker til at funnene er motsatte. For det første må boligsegmentet som studeres sammenlignes. Eneboliger ble studert både av (Fujita, 1989), (Coulombel, 2010) og

(Bohman & Nilsson, 2016), altså et annet boligsegment enn hva som er studert her. Basert på funnene kan det tyde på at eneboliger og leiligheter oppfyller ulike roller for en husholdning. Med utgangspunkt i den hedoniske metoden, er det mulig at andre attributter verdsettes for de som kjøper enebolig, slik som større tomt, mindre støy eller økt avstand fra sentrum (Fujita, 1989), (Coulombel, 2010). Ved en sammenligning av funnene for denne studien med de nevnte studiene, kan det tyde på at dette er tilfelle. Coulombel har pekt på at i avgjørelsen for hvor en kjøper skal bosette seg er det en lang rekke kriterier, og kriteriene kan være ulike for de som velger enebolig fremfor leilighet. Resultatene for denne studien indikerer at for leiligheter er avstand til togstasjon en langt viktigere faktor for husholdninger med høyere inntekt, og at de med lavere inntekt ikke verdsetter nærheten i like stor grad. Med andre ord ser det ut til at leilighetssegmentet i langt større grad operer under monosentrisk teori enn bid-rent, og er sånn sett likere funnene beskrevet i Boligutvalgets utredning (NOU 2002:2, 2002). I diskusjonen om husholdningens inntektsnivå påvirker hvor husholdningen plasserer seg, må også husholdningens sammensetning trekkes frem. Boligutvalgets utredning sier at det er et lineært forhold mellom antall arbeidende og betalingsviljen, hvor en dobling i hvor mange som jobber, også vil doble betalingsviljen (NOU 2002:2, 2002). Fujita fant ut at spesielt aleneboere og par uten barn utgjorde majoriteten av befolkningen i bysentrum, noe som igjen er motsatt av funn fra norsk forskning (Fujita, 1989).

Det geografiske området må også nevnes. I de nevnte studiene har avstanden vært langt større enn 2 km. I en by kan det observeres ringeffekter når man studerer sammensetningen av husholdninger og husholdningens inntekt (Lemoy, Raux, & Jensen, 2016). Med andre ord vil bysentrum bestå av husholdninger med høyere inntekt, etterfulgt av de med lavere inntekt. I randsonen av byen kan det igjen observeres at husholdningens inntekt øker. Funnene gjort av (Lemoy et al, 2016) er like de som er presentert av Fujita (Fujita, 1989). Med disse funnene kan det virke som at 2 km ikke fanger opp den ytterste ringen, og at det i denne studien kun er inkludert de to innerste ringene, noe som også kan forklare at resultatene er ulike. Dersom radiusen fra togstasjonene hadde blitt økt og inkludert denne ytre ringen, altså omlandet, er det mulig at resultatene hadde vist en mindre effekt på store leiligheter. Et annet element i denne tolkningen er sammensetningen av boliger i de ulike ringene. For Lillestrøm, Sandvika og Ski består sentrum i hovedsak av leiligheter, etterfulgt av eneboliger og rekkehus nærmere randsonen. At husholdningens inntekt øker i ytterkant kan også være et resultat av at man fanger opp andre boligtyper, noe som ikke er inkludert i dette studien. Ved å øke radiusen, og inkludere andre boligtyper, kan det være at resultatene i denne studien hadde vært mer like

studien gjort for Malmø-Lund (Bohman & Nilsson, 2016). De nevnte studiene er gjort for ulike land, noe som også kan være et element som påvirker resultatene. En Amerikansk storby fra for eksempel Fujita sin studie, vil mest sannsynlig ha andre funn enn en mindre by i Norge.

At estimatene viser at store leiligheter blir påvirket mest av avstand til togstasjon kan også være et resultat av at husholdninger som har opplevd en brå økning i inntektene flytter til bykjernen (Cuberes & Roberts, 2015). Cuberes og Roberts sin forskning ble gjort for en rekke britiske byer, så resultatene vil ikke være direkte overførbare til det norske boligmarkedet. Det kreves derimot mer forskning på om en brå økning i inntekt påvirker lokasjonsvalget i det norske markedet, og om inntektsendringen også medfører en brå endring i attributtpreferansen.

Det er med andre ord forskjeller i resultatene basert på hvilket boligsegment og geografiske område som blir studert. Med utgangspunkt i den hedoniske metoden, antas forskjellene å komme av at ulike kjøpegrupper har ulike preferanser for attributter som medfølger boligen (Osland, 2001). Det er viktig å merke seg at dette kun er antagelser basert på sammenligning av resultater og datagrunnlag. Regresjonsmodellen i denne studien svarer kun på at det eksisterer ulike kjøpegrupper uttrykt gjennom ulike leilighetsstørrelser. Den vil ikke kunne svare direkte på sammenhengen mellom inntekt og lokasjonsvalg eller valg av leilighetsstørrelse, annet enn at de ulike leilighetsstørrelsene har ulike kjøpegrupper med ulike preferanser. Den vil heller ikke kunne svare på forskjellene i preferansene for kjøpegrupper av leiligheter opp mot andre boligsegmenter. Det vil derfor ikke kunne trekkes en konklusjon om dette, men introduserer en rekke interessante problemstillinger for fremtidig forskning. Noen av disse vil trekkes frem i kapittel 9.

Resultatene må også tolkes ut ifra datamaterialet de er basert på (Laakso, 1997).

Datagrunnlaget som er brukt i denne studien er hentet ut i samarbeid med Eiendomsverdi, en god sekundærkilde for boliginformasjon i Norge. Det antas at Eiendomsverdi er den beste kilden på slike data, og kildegrunnlaget anses som godt. Etter første datarensing var det ca. 8000 observasjoner som gikk inn i regresjonsanalysen. Likevel var det store forskjeller mellom Lillestrøm, Sandvika og Ski, hvor Lillestrøm hadde over halvparten av disse observasjonene med ca. 4500, Sandvika med ca. 2200 og Ski med kun 1000. Videre ble observasjonene fordelt på de tre størrelseskategoriene, noe som resulterte i 896 små, 2968

mellomstore og 641 store for Lillestrøm. Sandvika hadde 317 små, 1229 mellomstore og 1290 store, og Ski endte opp med 116 små, 786 mellomstore og 195 store. For å kunne gjennomføre en regresjonsanalyse, er man som nevnt avhengige av store datasett, spesielt hvis man skal kunne gjøre opp for kategoriske feil (Laakso, 1997). Kategoriske feil, eller effekter, kan for eksempel være observasjoner i klynger eller i form av en naturlig barriere slik som en elv, stor vei eller flyplass.

For Lillestrøm har modellene, både lineær regresjon og log-level transformasjon, vist at datasettet er stort nok til å gi signifikante sammenhenger, og at det ikke tyder på at det foreligger kategoriske feil. I kartet som viser spredningen på observasjonene i Lillestrøm ser det ut til å være en relativt god fordeling over både distanse og himmelretning (figur 36). Det er likevel flere observasjoner som befinner seg på vestsiden av Glomma, og da i kort avstand fra Strømmen. Strømmen har et stort kjøpesenter og egen togstasjon, noe som kan være med på å påvirke resultatene for Lillestrøm. Samtidig ligger disse observasjonene på vestsiden av Glomma, og har en naturlig barriere inn mot Lillestrøm stasjon som kan forsterke effekten.

Sandvika ser også ut til å ha en relativt god spredning over distansen på 2 km. Likevel kan antall observasjoner vise seg å påvirke modellen i noe grad, noe som var tydelig i den lineære modellen hvor interaksjonsleddet for avstand og store leiligheter ikke var signifikant. I kartet som viser observasjonene for Sandvika kan det tyde på at det er en viss trend i datamaterialet (figur 37), noe som kan påvirke resultatene (Benoit, 2011) (Laakso, 1997). Det er forholdsvis få leiligheter i kategorien liten i sentrum, og flest i kategorien stor. Til tross for denne trenden, gir det ingen ensidig god forklaring på hvorfor det er nettopp store leiligheter som blir påvirket. Heller ikke antall observasjoner antas å være utslagsgivende, da det er 1290 i kategorien stor noe som anses for å være nok observasjoner ved sammenligning med Lillestrøm som har ca. 900.

For å få en bedre forståelse for hva som kan ha ført til at estimatene og signifikansnivået var såpass lave for den lineære regresjonsmodellen må modelltransformasjonen trekkes inn. I tilfeller hvor det forekommer skjevheter i en variabel som kan resultere i falske resultater, burde variabelen transformeres til en logaritmisk verdi (Benoit, 2011). Gjennom en logaritmisk transformasjon av kvadratmeterprisen ble modellen langt bedre, noe som kunne observeres av modelltestene. Det at en transformasjon løste dette problemet kan tyde på at skjevheter i kvadratmeterprisen for store leiligheter i Sandvika var utslagsgivende.

Resultatene for Ski kan tyde på skjevheter i datamaterialet. Estimater sier at avstand til togstasjonen medfører en økning i kvadratmeterprisen for små leiligheter både ved lineær og log-level regresjon, noe som stemmer dårlig med hva som kan observeres gjennom gg-plottet hvor den virkelige kvadratmeterprisen blir vist (figur 53). At estimatet er positivt, har mest sannsynlig en sammensatt forklaring. For det første er det kun 116 observasjoner i kategorien liten leilighet, noe som kan antas å være for få. I tillegg er disse observasjonene konsentrert i klynger, noe som kan ses på kart over observasjonene i Ski (se figur 38). Det er med andre ord en dårlig spredning over avstanden på 2 km, noe som kan anses å være en kategorisk feil. I tillegg kan også estimatet ytterligere bli påvirket av utbyggingen i Ski, hvor det i de siste årene har blitt bygget en rekke leilighetsblokker. Salgsdato og byggeår påvirker prisen i stor grad, og når dette ses i sammenheng med at flere av observasjonene har en forholdsvis lang avstand til togstasjonen, kan det gi en bedre forståelse for hvorfor avstandsestimatet er positivt.

I vurderingen av datamaterialet og operasjonaliseringen av variablene ble Plan- og bygningsetatens norm for leilighetskategorisering brukt (Plan- og bygningsetaten, Oslo Kommune, 2016, s. 7). En inndeling av leiligheter etter kvadratmeter har hatt sine fordeler, og gjort det mulig å studere størrelsessegmentene opp mot hverandre. Samtidig vil denne metoden også kunne gi noe av forklaringen på hvorfor store leiligheter i Sandvika har vist tegn på lavere signifikansnivå. Med utgangspunkt i normen fra plan og bygningsetaten i Oslo Kommune er kategorien «store leiligheter» alle leiligheter over 80 kvadratmeter (Plan- og bygningsetaten, Oslo Kommune, 2016, s. 7). Samtidig viser datamaterialet at det eksisterer mange leiligheter som er langt større enn dette, helt opp til 320 kvm. Disse ble riktignok fjernet i datarensingen, men det tegner et bilde på variasjonene som ligger i kategorien «stor». Med økende variasjon innenfor kategorien, følger det trolig med større variasjoner i de andre variablene også, og det vil være vanskeligere å vise til et lineært forhold (Benoit, 2011). Dersom senere studier skal bruke denne metoden, anbefales det å innføre en 4. kategori «ekstra stor». Ikke bare vil det redusere spennet i kategorien «stor», men det vil også kunne gi interessante funn som ikke blir vist her. Det antas at de aller største observasjonene i langt mindre grad blir påvirket av avstand til togstasjon, da disse leilighetene appellerer til en helt annen kjøpegruppe. Som et forslag til videre studier hadde det derfor vært interessant å se hvilke attributter som er av betydning for de aller største boligene. Ikke bare vil det være

nyttig i fremtidig utforming av disse boligene, men det gir også en ytterligere innsikt i ulikhetene i kjøpergruppene.

En annen utfordring med å bruke 2 km radius for Ski er tilknyttet byens størrelse.

Kommunegrensen til Ås krysser gjennom byens vestre side. Da datagrunnlaget ble hentet på kommunenivå, er det enkelte observasjoner som faller utenfor da de befinner seg i Ås kommune selv om de er innenfor radiusen på 2 km. Ikke bare utelater det observasjoner som kunne vært inkludert, men det begrenser også observasjonsområdet til å være mindre enn 2 km. I tillegg kan det se ut til at de resterende observasjonene heller ikke når ut til 2 km, da det kun er registrert to leilighetssalg utenfor 1350 m. Denne utfordringen har også vært gjeldende for Sandvika der kommunegrensen til Asker går innenfor 2 km fra Sandvika togstasjon i vest. Observasjoner i Asker har ikke blitt inkludert av samme årsak som nevnt over. I seg selv vil ikke dette påvirke modellens resultater nevneverdig, da modellen ikke er avhengig av at observasjonene fyller ut radiusen. Det viser derimot at å bruke en fastsatt radius på 2 km for fremtidig boligutvikling, slik den regionale planen legger opp til, kan by på enkelte utfordringer.

For den andre hypotesen har studien påvist at det eksisterer en forskjell i betalingsviljen for nærhet til togstasjon når man justerer for ulike leilighetsstørrelser. Noe av forskjellen mellom leilighetsstørrelsene må antas å komme av ulikheter i datamaterialet, slik som forskjeller i antall observasjoner, skjevheter i fordelingen mellom små, mellomstore og store leiligheter og observasjonenes plassering.

I forbindelse med den tredje underproblemstillingen som legger opp til en sammenligning av de tre byene, viser resultatene at det også eksisterer forskjeller her. Ved sammenligning av de tre byene er effekten minst for Lillestrøm, og størst i Ski. Store leiligheter (det segmentet som blir påvirket i størst grad) har en nedgang på -35,12% (isolert sett) ved 2 km i Lillestrøm. Sandvika har -43,24% for store leiligheter og Ski har -46,72%. Det er riktignok ikke bare forskjeller i hvor mye avstanden påvirker kvadratmeterprisene som er forskjellige, men også hvilke segmenter som blir påvirket mest eller minst. I Lillestrøm er det små leiligheter som blir påvirket minst, med en nedgang i kvadratmeterprisen på -19,82% ved 2 km. Mellomstore har en nedgang på -32,47%. Samtidig observeres det en annen rekkefølge i Sandvika, hvor det er mellomstore leiligheter som blir påvirket minst med -13,34% nedgang i kvadratmeterprisen ved 2 km. Med andre ord blir mellomstore leiligheter i Sandvika påvirket

i mindre grad enn det små leiligheter gjør i Lillestrøm. Små leiligheter i Sandvika har på sin side en nedgang på -33% i kvadratmeterprisen ved 2 km.

Resultatene for Ski er mer utfordrende å gi en god tolkning av, da de trolig blir påvirket av skjevheter og uforholdsmessigheter i datamaterialet som diskutert over. Kvadratmeterprisen for små leiligheter har en positiv koeffisient, og øker med 27,98% ved 2 km. Mellomstore leiligheter i Ski synker med -10,33%, noe som er omtrentlig like mye som for Sandvika.

Det er med andre ord forskjeller både i hvor mye avstand til togstasjon påvirker kvadratmeterprisene for leiligheter på et overordnet plan, og også hvilke segmenter som blir påvirket mest eller minst. Ut ifra en hedonisk-metodisk tolkning kan det tyde på at avstand til togstasjon som et attributt verdsettes ulikt fra by til by, i tillegg til at det er forskjeller innad i de ulike kjøpegruppene. Sett ut ifra tilhørende teori, burde de ulike undersegmentene bli påvirket likt uavhengig av by (Fujita 1989) (Coulombel 2017) (Geltner et al, 2014). Det kan derimot se ut til at den interne monosentriske effekten påvirker undersegmentene ulikt. Ikke bare er attributtene som verdsettes ulikt for de ulike kjøpegruppene når man ser på leilighetsstørrelsene, men kjøpegruppene ser ut til å foretrekke attributtene ulikt avhengig av hvilken by man studerer.

Videre kan også resultatene tyde på at avstandseffekten er sterkere i mindre byer hvor den tettbygde byen er mindre. Ved å sammenligne estimatene for store leiligheter, finner man den største effekten i Ski, etterfulgt av Sandvika. Differansen mellom Lillestrøm og Ski er på ca. 11,5%, altså at kvadratmeterprisen for store leiligheter i Ski synker i snitt med 11,5% mer enn i Lillestrøm. Denne tolkningen støttes derimot kun delvis av resultatene, da mellomstore leiligheter i Ski på sin side har et svært lavt estimat med kun 10,33%. Regresjonsmodellen tilsier at det kun er enkelte kjøpegrupper som blir påvirket av byens størrelse, og at forholdet mellom gang- og sykkelavstand og tettbygde områder og mindre tettbygde områder kun påvirker enkelte kjøpegrupper.

Et annet aspekt i vurderingen av dette forholdet er forklaringskraften til modellen. For Ski har log-level-modellen en forklaringskraft på 83,51%. En mulig årsak til at forklaringskraften er såpass høy kan være antall observasjoner, altså at færre observasjoner medfører høyere forklaringskraft. Samtidig kan det også tolkes dithen at variablene inkludert i modellen forklarer mer av kvadratmeterprisen i Ski, enn den gjør for Lillestrøm og Sandvika som har

henholdsvis 80,2% og 75,8%. Med andre ord er avstand en mer betydelig faktor i Ski enn den er for Lillestrøm og Sandvika. Sandvika har den laveste forklaringskraften, noe som kan tyde på at det er andre variabler eller attributter som er av større betydning. Hvilke variabler som mangler er vanskelig å si, og det er stor variasjon i hva tidligere forskning har brukt (Laakso, 1997), (Osland, 2001).

Datarensningen som er gjennomført kan være et grep som endrer resultatene i en viss grad. Kvadratmeterprisen for leilighetene ble i første omgang justert for fellesgjelden da det viste seg å være mange leiligheter i de tre byene som hadde svært lave kvadratmeterpriser fordi en stor andel av kjøpesummen var flyttet til fellesgjeld. I forbindelse med justeringen av fellesgjeld så det fortsatt ut til å være ekstremverdier i form av enten svært lave, eller svært høye kvadratmeterpriser. I tillegg var det også observasjoner som skilte seg ut i form av størrelse, hvor noen var på over 300 m². Som et grep for å minimere variasjoner i datamateriale, ble også disse fjernet. Selv om dette var et nødvendig grep for å jevne ut variasjoner i datamaterialet, medfører det at datamaterialet ikke er like virkelighetsnært. Det samme gjelder det for de observasjonene som ble fjernet da det var manglende informasjon. Første datarensningen førte til at datamaterialet ble langt mer egnet for en statistisk analyse, men representerer ikke virkeligheten i like stor grad som før rensingen, noe som kan ha påvirket resultatene og forklaringskraften til modellen i noe grad.

Som beskrevet tidligere viste den lineære regresjonen at det fortsatt var variasjoner som påvirket resultatene, og at kvadratmeterprisen (til tross for den første rensingen) fortsatt ikke være lineær. I tillegg viste modelltestene at det var enkelte observasjoner som skilte seg ut, og disse ble fjernet. Log-transformasjonen i seg selv antas å ikke påvirke modellen i særlig stor grad, da det kun medfører en endring i hvordan X-variablene påvirker kvadratmeterprisen. Det som derimot kan ha påvirket resultatene i noen grad er fjerningen av de 36 observasjonene som skilte seg mest ut. Selv om dette medførte et mer lineært forhold til regresjonsmodellen, kan det argumenteres for at ved å utelate enkelte observasjoner, blir datagrunnlaget mindre virkelighetsnært.

Hovedproblemstillingen i denne studien ble utformet for å kunne studere om det eksisterer en forskjell i betalingsviljen for nærhet til togstasjon når det tas hensyn til ulike leilighetsstørrelser. Hypotesene som ble utarbeidet blir alle bekreftet innenfor et signifikansnivå på 95%. Den interne monosentriske effekten observeres i alle byene, og

påvirker leilighetsprisene i negativ forstand. I forskningsspørsmålet om det eksisterer forskjeller i betalingsviljen for ulike leilighetsstørrelser, viser også her modellene at det er tilfelle. Når resultatene fra de separate modellene sammenlignes, er det også tegn på at den interne monosentriske effekten er ulik fra by til by. Med utgangspunkt i forskjellene i resultatene kan det tyde på at byene ikke kan anses som like, og at for eksempel Ski krever en arealstrategi hvor en større andel av utbyggingen skjer tettere inn mot sentrum og togstasjonen. Resultatene viser at betalingsviljen synker bratt for kjøpegruppen i Ski, noe som tyder på at nærhet til togstasjonen er viktig. Når målet om økt kollektivbruk, nullvekst i bilbruken og en mer miljøbærekraftig region står sentralt, må kjøpegruppens etterspørsel også tas hensyn til. De samme trendene kan også observeres både for Sandvika og Lillestrøm. I fremtidig utbygging burde det også tas hensyn til differanser i kjøpegruppene. For noen kjøpegrupper er nærhet til togstasjonen viktigere enn andre, slik resultatene viser.

8 Kritikk til oppgaven

Å estimere boligprisene under en forutsetning om at prisen er et resultat av verdien til alle variabler som inngår i en bolig, byr på utfordringer. For det første er det svært vanskelig å avgjøre hvilke variabler som er gjeldende. Først og fremst er konsumentene forskjellige, og hva som er et viktig attributt for en konsument, er ikke nødvendigvis viktig for en annen. I tillegg er boligmarkedet i seg selv svært forskjellig, hvor lokale variasjoner kan ha mye å si for boligprisene. Med dette følger det at en modell som vi har fremstilt, ikke kan fange opp alle variabler som er aktuelle for å estimere boligpriser.

Regresjonsmodellen tar ikke hensyn til at det er andre lokasjonsrelaterte faktorer enn togstasjonen. Dette er naturligvis gjort for å presisere hvilken effekt avstand til togstasjonen har på boligprisene, men det utelater samtidig den effekten andre faktorer som matbutikk, cafeer, apoteker osv. har. Sånn sett egner ikke modellen seg til å predikere leilighetspriser, da den utelater veldig mange variabler.

Videre er modellen utformet under en forutsetning om en lineær sammenheng mellom avstand og boligprisene. I den virkelige verden er ikke dette tilfelle, og med samme utgangspunkt som i argumentet over, vil det være andre lokale variasjoner som enten trekker prisene opp eller ned. Ikke bare vil lokale variasjoner påvirke prisen, men det er trolig et

ikke-lineært forhold mellom avstand og boligpriser generelt sett. Et eksempel er at man vil trolig se et prisfall innenfor svært kort avstand til togstasjonen grunnet blant annet støy.

I tillegg byr lokale variasjoner mellom byene på utfordringer. Modellen i seg selv er best egnet for en by hvor togstasjonen ligger midt i, uten noen naturlige barrierer. Lillestrøm har Glomma, Ski er i hovedsak utviklet på den ene siden av jernbanen, og Sandvika bærer preg av langt mer spredt bebyggelse. Alle disse variasjonene gjør det vanskelig å utforme en regresjonsmodell som sammenligner byene, da avstandseffekten og det generelle prisnivået er såpass ulik for de tre byene. I tillegg medfører slike ulikheter at det blir vanskelig å bruke en regresjonsmodell hvor alle byene sammenlignes. Ikke bare medfører variasjoner at signifikansnivå og estimater kan påvirkes i såpass stor grad at de ikke lenger er innenfor 95%, men det gjør også modellen vanskelig å tolke.

En annen ting som kan anses for å være en svakhet, er at grunnlaget for å studere de ulike leilighetsstørrelsene stammer fra en antagelse om at husholdninger med høyere inntekt i de fleste tilfeller kjøper større leilighet. Dette er en antagelse som ikke blir sjekket i modellen, og det kan naturligvis være slik at sammenhengen mellom inntekt og leilighetsstørrelse ikke er like tydelig som forutsatt i denne studien. En annen forutsetning som er gjort for studien, men som har blitt lite diskutert, er at det er et likt forhold mellom tilbud og etterspørsel for de ulike leilighetsstørrelsene. Samtidig er trolig forholdet mellom tilbud og etterspørsel i kategorien liten leilighet skjevfordelt sammenlignet med de andre kategoriene – altså at det er større etterspørsel enn tilbud for de små leilighetene sammenlignet med mellomstore og store leiligheter. Dette kan observeres i Oslo, og medfører et stort prispress på små leiligheter. For Lillestrøm, Sandvika og Ski vil for stor etterspørsel i forhold til tilbud kunne medføre generelt høyere priser på små leiligheter lenger unna togstasjonen, noe som også kan påvirke resultatene for denne studien.

Det siste vi ønsker å trekke frem er inndelingen av størrelseskategoriene. Små og mellomstore leiligheter har tilnærmet likt spenn, mens store leiligheter omfatter alle observasjoner som er over 80 m². Dette utgjør, selv etter datarensingen, et spenn på nesten 120 m². I dette spennet kan det være store variasjoner som påvirker resultatene. Det kunne vært hensiktsmessig å lage en 4. Kategori «Ekstra store leiligheter» som tar for seg for eksempel leiligheter over 120 m²+. Ikke bare vil det kunne redusere variasjoner i kategorien «stor leilighet», men det vil også kunne gi interessant innsikt i hvilke variabler som påvirker

kvadratmeterprisen mest for de aller største leilighetene. Å inkludere en slik kategori kan riktignok by på utfordringer i seg selv, da det er få observasjoner som er så store. I datasettet vi har brukt (etter rensing), var det kun 260 leiligheter i denne størrelsen (totalt sett for de tre byene).

9 Forslag til fremtidig forskning

Gjennom denne studien har det kommet opp en rekke interessante temaer som vil være spennende å studere videre. Det første temaet som kom opp, var hvordan pendlerandelen i en kommune eller by påvirker boligprisene. Dette studieforslaget kom i forbindelse med en undersøkelse gjort av Eiendomsverdi på oppdrag av Eiendom Norge (2019), hvor det ble kort drøftet om en høyere pendlerandel medfører større prispress og økte boligpriser i nærheten av togstasjonen (Eiendomsverdi, 2019).

I forbindelse med at vi ønsker å fremme forskning som bygger videre på funn gjort i denne studien, samt øke forståelsen for pendlervaner, boligpriser og byutvikling, viser vi til en studie gjort for Amsterdam i Nederland (Ghebreegiabiher, Pels, & Rietveld, 2005). Studien viste funn som sa at en pendler gjerne reiser litt lenger for å komme seg til en togstasjon med flere avganger enn en togstasjon med færre avganger. I tillegg hadde avstanden mellom disse to togstasjonene også en merkbar effekt på boligprisen som tilsa at togstasjonen med flere avganger (og avstanden til denne) var av større betydning for husholdningen.

I tillegg anbefaler vi å gjøre et grundigere studie rundt bid-rent teorien, og da spesielt forholdet mellom inntekt og valg av leilighetsstørrelse. Det ville vært interessant å se hvordan lønnsnivået påvirker valg av leilighetsstørrelse. Med utgangspunkt i denne studien som viser at kvadratmeterprisen for store leiligheter synker mest med avstanden, vil det være nyttig innsikt dersom det kan bekreftes eller avkreftes at rikere husholdninger faktisk kjøper større leiligheter. Teorien tilsier også at med økt inntekt medfører en endring i betalingsviljen for ulike goder, og spørsmålet som burde stilles er hvorvidt kollektiv transport, og da spesielt togreiser, regnes for å være et normalgode eller luksusgode. Dette vil kunne gi en økt forståelse for kjøpemønsteret for bolig, og hvilke faktorer som inngår i valg av bolig. Dette er igjen et viktig område når det kommer til fremtidig byutvikling.

Et siste forslag som vi hadde likt å se, er en studie av hvordan betalingsviljen og preferansen for avstand til togstasjon er forskjellig fra leiligheter til for eksempel eneboliger. Det har blitt gjort en rekke studier på dette hvor man har påvist små forskjeller. Det forskningsfeltet mangler er en forklaring på hvorfor betalingsviljen er forskjellig for de ulike boligsegmentene. Hypotesen vår er at husholdninger som bosetter seg i enebolig, eller i områder som er mer typiske for eneboliger, i langt mindre grad avhengig av togstasjonen som reisemåte da de besitter en eller flere biler.

Referanser

- Akershus fylkeskommune; Oslo Kommune. (2015). *Regional plan for areal og transport i Oslo og Akershus*. Hentet fra https://viken.no/_f/p1/ib1949fd3-e553-4dd7-9505-4a2519ba6d00/regional-plan-for-areal-og-transport-i-oslo-og-akershus.pdf
- Alonso, W. (1960, Januar). *A theory of the urban land market*. Hentet fra http://lib.cufe.edu.cn/upload_files/other/4_20140526030430_59_A%20theory%20of%20the%20urban%20land%20market.pdf
- Alonso, W. (1964, Mai). *The Historic and the Structural Theories of Urban Form: Their Implications for Urban Renewal*. Hentet fra <https://www.jstor.org/stable/3144355?origin=crossref&seq=1>.
- Askheim, S. (2021, Januar 8, a). *Lillestrøm*. Hentet fra [snl.no: https://snl.no/Lillestr%C3%B8m](https://snl.no/Lillestr%C3%B8m)
- Askheim, S. (2021, Januar 22, b). *Nordre Follo*. Hentet fra [snl.no: https://snl.no/Nordre_Follo](https://snl.no/Nordre_Follo)
- Bang, J. (2019, a). *Demand Curve*. Hentet fra Investopedia: <https://www.investopedia.com/terms/d/demand-curve.asp>
- Bang, J. (2019, b). *Shift in demand curve*. Hentet fra Investopedia: <https://www.investopedia.com/terms/d/demand-curve.asp>
- Bang, J. (2019, c). *Supply Curve*. Hentet fra Investopedia: <https://www.investopedia.com/terms/s/supply-curve.asp>
- Bang, J. (2019, d). *Shift in supply curve*. Hentet fra Investopedia: <https://www.investopedia.com/terms/s/supply-curve.asp>
- Bang, J. (2019, e). *Equilibrium Quantity*. Hentet fra Investopedia: <https://www.investopedia.com/terms/e/equilibrium-quantity.asp>
- Benoit, K. (2011, Mars 17). *Linear Regression Models with Logarithmic Transformations, Artikkel*. Hentet fra <https://kenbenoit.net/assets/courses/ME104/logmodels2.pdf>
- Bjørge, A. S., Groseth, T., Janbu, I. H., Elin, H., Magnesen, G., & Moe, F. L. (2012). *Rik på historie, Hefte utgitt av Bærum kommune*. Hentet fra <https://www.baerum.kommune.no/globalassets/om-baerum-kommune/organisasjon/historisk-arkiv/kulturhistoriskriss.pdf>
- Bjørnstad, J. (2018, Juni 26). *statistikk*. Hentet fra [snl.no: https://snl.no/statistikk](https://snl.no/statistikk)
- Bohman, H., & Nilsson, D. (2016, Oktober). *The impact of regional commuter trains on property values: Price segments and income, Artikkel*. Hentet fra <https://www.sciencedirect.com/science/article/pii/S0966692316300151?via%3Dihub>
- Braut, G. S., & Dahlum, S. (2018, Mai 24). *Regresjonsanalyse*. Hentet fra [snl.no: https://snl.no/regresjonsanalyse](https://snl.no/regresjonsanalyse)
- Brooks, C. (2008). *Introductory Econometrics for Finance, Second edition*. Cambridge University Press.
- Bærum kommune. (2019). *Befolkningsutvikling: Årsrapport*. Bærum: Bærum kommune.
- Christensen, A. R. (2021, Januar 21). *Boligprisene opp 12 prosent i 2020: – For Oslo er det grunn til bekymring. Nordre Aker Budstikke*. Hentet fra <https://nab.no/nyheter/for-oslo-er-det-grunn-til-bekymring/19.22825>.
- Cooper, R., & Andrew, J. (2011). *Microeconomics: Theory Through Applications*. Saylor Foundation.
- Coulombel, N. (2010, Juni 24). *Residential choice and household behavior : State of the art*. Hentet fra

- file:///C:/Users/simen/Downloads/wp_2.2a_residential_choice_and_household_behavior%20(2).pdf.
- Cuberes, D., & Roberts, J. (2015, Oktober). *Household location and income: a spatial analysis for*. Hentet fra https://www.sheffield.ac.uk/polopoly_fs/1.514271!/file/paper_2015022.pdf.
- Dahlum, S., & Grønmo, S. (2020, August 26). *Operasjonalisering*. Hentet fra snl.no: <https://snl.no/operasjonalisering>
- Di Leo, G., & Sardanelli, F. (2020, Mars). *Statistical significance: p value, 0.05 threshold, and applications to radiomics—reasons for a conservative approach*. Hentet fra <https://eurradiolexp.springeropen.com/articles/10.1186/s41747-020-0145-y>.
- Eiendomsverdi. (2019, November 7). *Hvordan påvirker avstand til stasjonen boligprisene i pendlerområder? [Blogginnlegg]*. Hentet fra <https://eiendomnorge.no/aktuelt/blogg/hvordan-pavirker-avstand-til-stasjonen-boligprisene-i-pendleromrader>.
- Eiendomsverdi. (2021, Mai 23, a). *Områderapport av Sandvika*. Hentet fra Eiendomsverdi.no: <https://eiendomsverdi.no/app/appAreaSelection.aspx>
- Eiendomsverdi. (2021, Februar 2, b). *Eiendomsinformasjon satt i system*. Hentet fra <https://eiendomsverdi.no/>: <https://eiendomsverdi.no/>
- Eurostat. (2013). *Handbook on Residential Property Prices Indices (RPPIs)*. Hentet fra <https://ec.europa.eu/eurostat/documents/3859598/5925925/KS-RA-12-022-EN.PDF>.
- Field, A. (2009). *DISCOVERING STATISTICS USING Spss THIRD EDITION*. London: SAGE Publication Ltd.
- Ford, C. (2015, August 26). *Understanding Q-Q Plots*. Hentet fra <https://data.library.virginia.edu/understanding-q-q-plots/?fbclid=IwAR39I844NTnPkEaUIrwSpwMrbCHDhQVxnImrAZ4XHIOYalqzWunlu2PI7FM>.
- Fujita, M. (1989). *Urban economic theory*. Cambridge University press.
- Geltner, D. M., Miller, N. G., Clayton, J., & Eichholtz, P. (2014). *Commercial real estate analysis and investment 3 edition*. Oncourse Learning.
- Ghebreegiabiher, D., Pels, E., & Rietveld, P. (2005, August 27). *Impact of railway station on Dutch residential housing*. Hentet fra https://www.econstor.eu/bitstream/10419/117836/1/ERSA2005_748.pdf.
- Gramm, J. (2016). *En analyse av boligpriser langs Østfoldbanen (Masteroppgave, Norges miljø- og biovitenskapelige iniversitet)*. Hentet fra https://nmbu.brage.unit.no/nmbu-xmlui/bitstream/handle/11250/2401897/Masteroppgaven_Joachim_Gramm.pdf?sequence=1&isAllowed=y.
- Hansen, T. (2015, Mars 19). *Sekundærdata, Artikkel*. Hentet fra [analysen.no: http://www.analysen.no/latest-news/item/sekundaerdata](http://www.analysen.no/latest-news/item/sekundaerdata)
- Herath, S., & Maier, G. (2010). *The hedonic price method in real estate and housing market research: a review of the literature*. Hentet fra <https://ro.uow.edu.au/cgi/viewcontent.cgi?article=1977&context=buspapers>.
- Hewitt, C., & Hewitt, W.E. (2012, Desember). The Effect of Proximity to Urban Rail on Housing Prices in Ottawa. *Journal of Public Transportation* 15(4), ss. 43-65. Hentet fra https://www.researchgate.net/publication/277574318_The_Effect_of_Proximity_to_Urban_Rail_on_Housing_Prices_in_Ottawa.
- Hoen, B. (2018). *Dyre forstadsbyer- en studie av boligpriser, boligtyper og boligstørrelser i Osloområdets regionbyer (Masteroppgave, Universitet i Oslo)*. Hentet fra https://www.duo.uio.no/bitstream/handle/10852/66827/1/Hoen_master.pdf.

- Hårsmann, B. (1981). *Housing demand models and housing market models for regional and local planning*. Hentet fra <https://pascal-francis.inist.fr/vibad/index.php?action=getRecordDetail&lang=en&idt=PASCALBTP8130493678>.
- Investopedia. (2021, April 30). *R-Squared vs. Adjusted R-Squared: What's the Difference?* Hentet fra [investopedia.com: https://www.investopedia.com/ask/answers/012615/whats-difference-between-rsquared-and-adjusted-rsquared.asp](https://www.investopedia.com/ask/answers/012615/whats-difference-between-rsquared-and-adjusted-rsquared.asp)
- Kabir, S. M. (2016). *Basic Guidelines for Research: An Introductory Approach for All Disciplines*. Book Zone Publication.
- Kommuneprofilen. (2021, Mai 10, a). *Antall innbyggere etter aldersgruppe på kommunenivå*. Hentet fra https://www.kommuneprofilen.no/Profil/Kommunefakta/Befolkning_Alder_kommune.aspx.
- Kommuneprofilen. (2021, Mai 10, b). *Andel husholdninger etter type i prosent på kommunenivå*. Hentet fra https://www.kommuneprofilen.no/Profil/Kommunefakta/Hushold_kommune.aspx.
- Kunnskapsbyen. (2019, Januar 8). *Lillestrøm bygger for lite*. Kunnskapsbyen Lillestrøm. Hentet fra <http://kunnskapsbyen.no/byutvikling/lillestrom-bygger-for-lite/>.
- Lane, D. M. (2021, Mai 25). *Online Statistics Education: A Multimedia Course of Study*. Hentet fra <https://onlinestatbook.com/2/index.html>.
- Lemoy, R., Raux, C., & Jensen, P. (2016, August 5). *Where in cities do "rich" and "poor" people live? The*. Hentet fra <https://hal.archives-ouvertes.fr/hal-00805116v3/document>.
- Lillestrøm kommune. (2020). *Byutviklingsplan for Lillestrøm by Fremtidens Lillestrøm del II*. Hentet fra <https://lillestrom.sharepoint.com/sites/Eksterndeling/Delte%20dokumenter/Forms/AllItems.aspx?id=%2Fsites%2FEksterndeling%2FDelte%20dokumenter%2FByutviklingsplan%20Lillestr%C3%B8m%20by%2FByutviklingsplan%5F20201123%5FDEL%202%2Epdf&parent=%2Fsites%2FEkstern>.
- Laakso, S. (1997). *URBAN HOUSING PRICES AND THE DEMAND FOR HOUSING CHARACTERISTICS, A study on housing prices and the willingness to pay for housing characteristics and local public goods in the Helsinki Metropolitan Area*. Hentet fra <https://www.etla.fi/wp-content/uploads/A27.pdf>.
- Mills, E. S. (1972). *Studies in the Structure of the Urban Economy*. Hentet fra <https://eric.ed.gov/?id=ED065300>.
- Nesvåg, L.-P., Handstanger, A. C., & Martinsen, O. J. (2015). *Innerstrekningene av Hovedbanen, Østfoldbanen og*. Oslo: Jernbaneverket, Statens vegvesen, Ruter AS.
- Nguyen, M. E. (2012). *Hvordan påvirker avstand til sentrum boligprisene i Kristiansand, (Masteroppgave, Universitet i Agder)*. Hentet fra <https://uia.brage.unit.no/uia-xmlui/bitstream/handle/11250/135716/Oppgave%20Mai-Elin%20Nguyen.pdf?sequence=1&isAllowed=y>.
- Norconsult. (2020, Oktober 6). *Mulighetsstudie - Sandvika sentrum nord med kollektivknutepunkt*. Hentet fra <https://www.baerum.kommune.no/innsyn/byggesak/wfdocument.ashx?journalpostid=2020229077&dokid=5146750&versjon=1&variant=A&>.
- Nordheim, E., & Bachmann, T. (2013). *Hvilken betydning har utvalgte attributter for omsetningsprisen i fritidsboligmarkedet i Vinje, i perioden 2008-2012 (Masteroppgave, Universitetet i Agder)*. Hentet fra <https://core.ac.uk/download/pdf/225886081.pdf>.

- Nordre Follo kommune. (2019, Juni 12, a). *Kommuneplan 2019-2013 - Samfunnsdel*. Hentet fra http://webhotel3.gisline.no/GisLinePlanarkiv/3020/0213KPLAN2019/Dokumenter/KPLAN2019_samfunnsdel.pdf.
- Nordre Follo kommune. (2019, Juni 12, b). *Kommuneplan 2019-2030 - Arealdel Ski*. Hentet fra <https://www.nordrefollo.kommune.no/globalassets/nordre-follo/kommuneplan/kommuneplan-for-nordre-follo-2019-2030-arealdel-og-planbestemmelser-ski.pdf>.
- NOU 2002:2. (2002, Februar 15). *Boligmarkedene og boligpolitikken - NOU*. Hentet fra <https://www.regjeringen.no/contentassets/80899d9e55ef499c86359694e816207f/no/pdfa/nou200220020002000dddpdfa.pdf>.
- Osborne, J. W. (2009, Desember). *Data Cleaning Basics: Best Practices in Dealing with Extreme Scores*. Hentet fra https://www.researchgate.net/publication/244872484_Data_Cleaning_Basics_Best_Practices_in_Dealing_with_Extreme_Scores.
- Osland, L. (2001, Januar). *Den hedonistiske metoden og estimering av attributtpriser, Artikkel*. Hentet fra https://www.researchgate.net/publication/258092733_Den_hedonistiske_metoden_og_estimering_av_attributtpriser.
- Oslo Kommune. (2021, Mai 10). *Statistikkbank - Befolkningsfremskrivninger*. Hentet fra <https://statistikkbanken.oslo.kommune.no/webview/>.
- Pels, E. (2005, September). *Impact of railway station on Dutch residential housing market*. Hentet fra https://www.researchgate.net/publication/23731849_Impact_of_railway_station_on_Dutch_residential_housing_market.
- Plan- og bygningsetaten, Oslo Kommune. (2016). *Evaluering av bruk og effekt*. Oslo: Oslo Kommune.
- Rasch, Ø. N. (2017). *Analyse av hvordan prising av boliger varierer i avstand fra togstasjon og stasjonens avstand (i tid) til Oslo (Mastergradoppgave, Norges miljø- og biovitenskapelige universitet)*. Hentet fra https://nmbu.brage.unit.no/nmbu-xmlui/bitstream/handle/11250/2465235/M30-EUTV_Master_Rasch.pdf?sequence=1&isAllowed=y.
- Regjeringen. (2016, Desember 14). *Boliglånsforskriften*. Hentet fra <https://www.regjeringen.no/no/tema/okonomi-og-budsjett/finansmarkedene/boliglansforskriften-1.-januar-201730.-juni-2018/id2523977/>.
- Robinson, C., & Schumacker, R. E. (2009, Januar). *Interaction Effects: Centering, Variance Inflation Factor, and Interpretation Issues*. Hentet fra <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.620.5853&rep=rep1&type=pdf>.
- Rosen, S. (1974, Februar). *Hedonic Prices and Implicit Markets: Product Differentiation in pure Competition, Paper*. Hentet fra http://neconomides.stern.nyu.edu/networks/phdcourse/Rosen_Hedonic_prices.pdf?fbclid=IwAR08OwBm25SPY3OAF7rS-rfU-PBwvJS9_Jogmf84SFO6NewalReLf6xcZA8.
- Sander, K. (2019, September 15). *Kausalt design. Studie.no*. Hentet fra <https://studie.no/kausalt-design/>.
- Sandhu, V. A., & Sarna, V. K. (2015). *Effekten av total transporttid på boligpriser i Jessheim (Masteroppgave, Norges miljø- og biovitenskapelige universitet)*. Hentet fra <http://hdl.handle.net/11250/294787>.

- Schnare, A. B., & Struyk, R. J. (1976, April). *Segmentation in urban housing markets*. Hentet fra sciencedirect.com:
<https://www.sciencedirect.com/science/article/abs/pii/0094119076900504?via%3Dihub>
- Simon, L., & Young, D. (2021, Mai 26). *Residuals vs. Fits Plot*. Hentet fra
<https://online.stat.psu.edu/stat462/node/117/>.
- Skog, O. (1998). *Å forklare sosiale fenomener : en regresjonsbasert tilnærming*. Oslo: Gyldendal.
- Snee, R. D. (1977). *Validation of Regression Models: Methods and Examples, Artikkel*. Hentet fra <https://www.tandfonline.com/doi/abs/10.1080/00401706.1977.10489581>.
- SSB. (2021, Mai 5, a). *Kommunefakta om Lillestrøm*. Hentet fra <https://www.ssb.no/kommunefakta/lillestrom>): 2021
- SSB. (2021, Mai 23, b). *06265: Boliger, etter bygningstype (K) 2006 - 2021*. Hentet fra SSB.no: <https://www.ssb.no/statbank/table/06265/tableViewLayout1/>
- Statens vegvesen. (2012). *Nasjonal gåstrategi*. Hentet fra https://www.vegvesen.no/_attachment/528658/binary/851145?fast_title=Brosjyre+g%C3%A5strategi.pdf.
- Store Norske Leksikon. (2021, Mars 27). *Teori*. Hentet fra <https://snl.no/teori>.
- Straszheim, M. (1974, Augst). Hedonic Estimation of Housing Market Prices: A Further Comment. *The Review of Economics and Statistics*, ss. 404-406,
<https://www.jstor.org/stable/1923985>.
- Strømmen Storsenter. (2020). *Praktisk informasjon*. Hentet fra [strommenstorsenter.no: https://strommenstorsenter.no/praktisk-informasjon/media/](https://strommenstorsenter.no/praktisk-informasjon/media/)
- Sukamolsen, S. (2007, Januar). *Fundamentals of quantitative research*. Hentet fra <https://pdf4pro.com/view/fundamentals-of-quantitative-research-chula-20b41d.html>.
- Takle, M. (2012, November). *Boligprisindeksen, Dokumentasjon av metode*. Hentet fra SSB.no: https://www.ssb.no/a/publikasjoner/pdf/notat_201210/notat_201210.pdf
- Thorsnæs, G., & Askheim, S. (2021, April 26). *Bærum*. Hentet fra snø:
<https://snl.no/B%C3%A6rum>
- Thorsnæs, G., & Askheim, S. (2020, Januar 22). *Ski*. Hentet fra snl.no: https://snl.no/Ski-_tidligere_kommune
- Thrane, C. (2017). *Regresjonsanalyse - En praktisk tilnærming*. Oslo: Cappelen Damm Akademisk.
- University of Cambridge. (2015, Januar 22). *Collinearity*. Hentet fra <https://imaging.mrc-cbu.cam.ac.uk/statswiki/FAQ/Collinearity>.

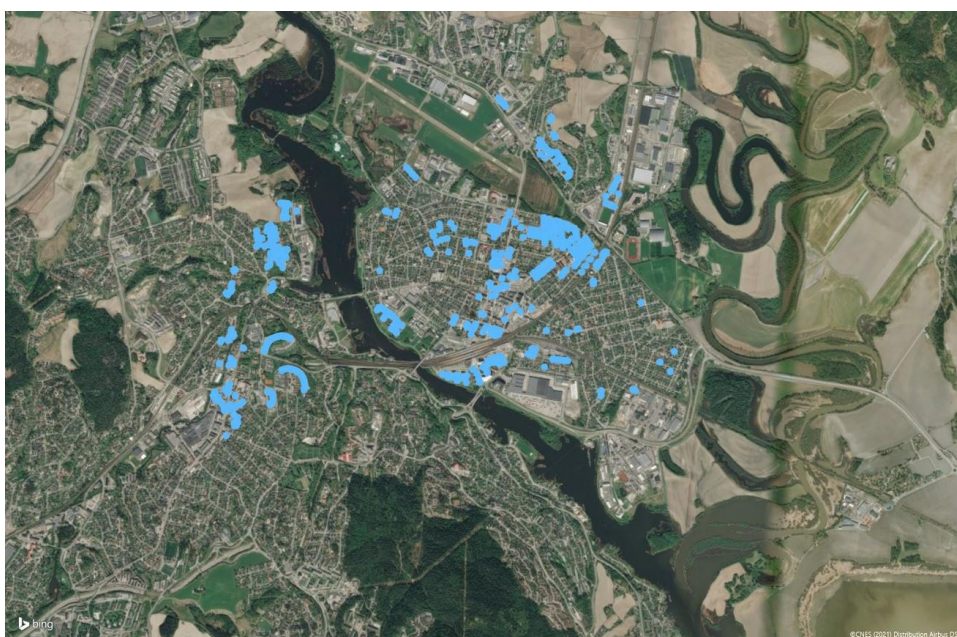
Vedlegg

Vedlegg 1

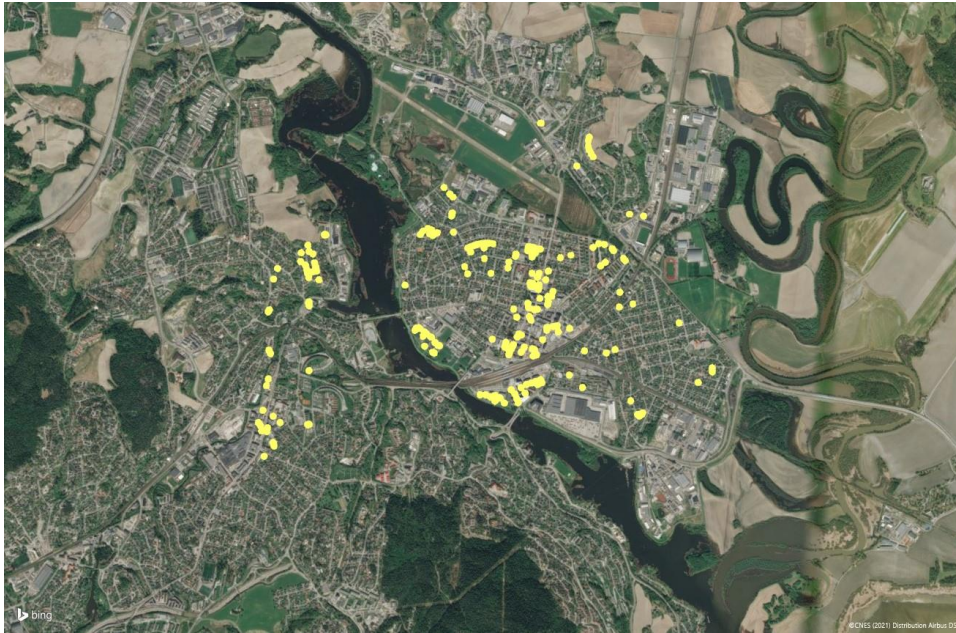
Observasjoner for Lillestrøm fordelt etter kategorier. Rød er liten leilighet, blå er mellomstor leilighet og gul er stor leilighet.



Figur 67 Oversikt over små leiligheter i Lillestrøm



Figur 68 Oversikt over mellomstore leiligheter i Lillestrøm



Figur 69 Oversikt over store leiligheter i Lillestrøm

Som man kan se i kartene for de ulike kategoriene er små leiligheter i hovedsak konsentrert rundt sentrum, og sprer seg i hovedsak fra Sør til Nord, samt en liten klynge i Strømmen. Mellomstore og store leiligheter er i større grad spredt mer tilfeldig rundt i byen.

Vedlegg 2

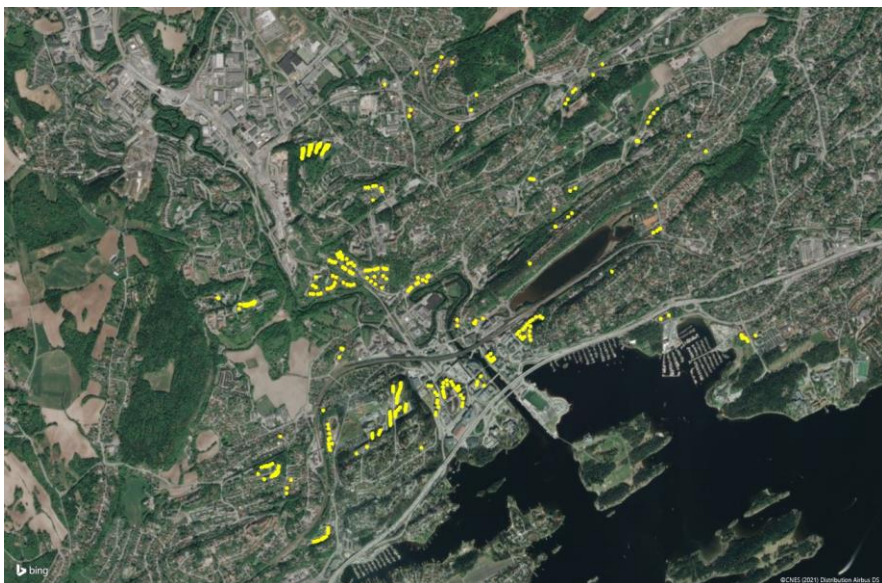
Observasjoner for Sandvika fordelt etter kategorier. Rød er liten leilighet, blå er mellomstor leilighet og gul er stor leilighet.



Figur 70 Oversikt over små leiligheter i Sandvika



Figur 71 Oversikt over mellomstore leiligheter i Sandvika

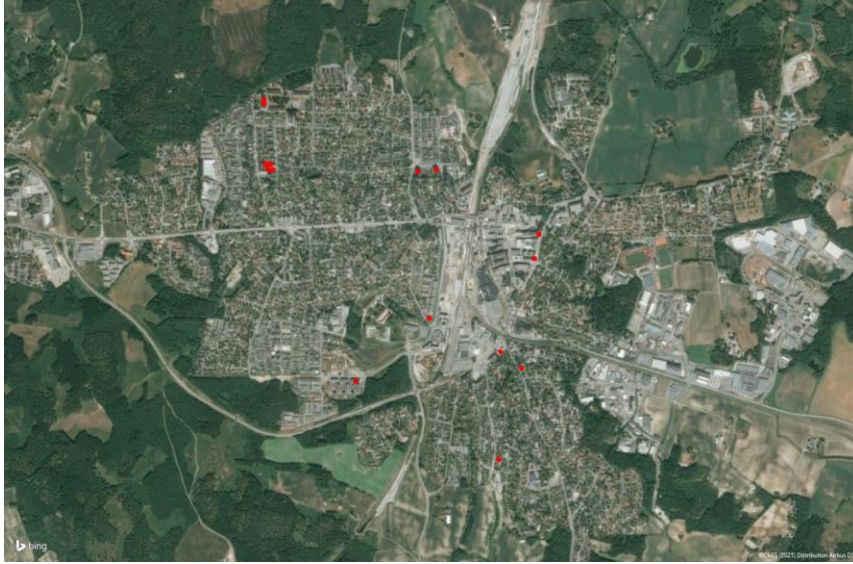


Figur 72 Oversikt over store leiligheter i Sandvika

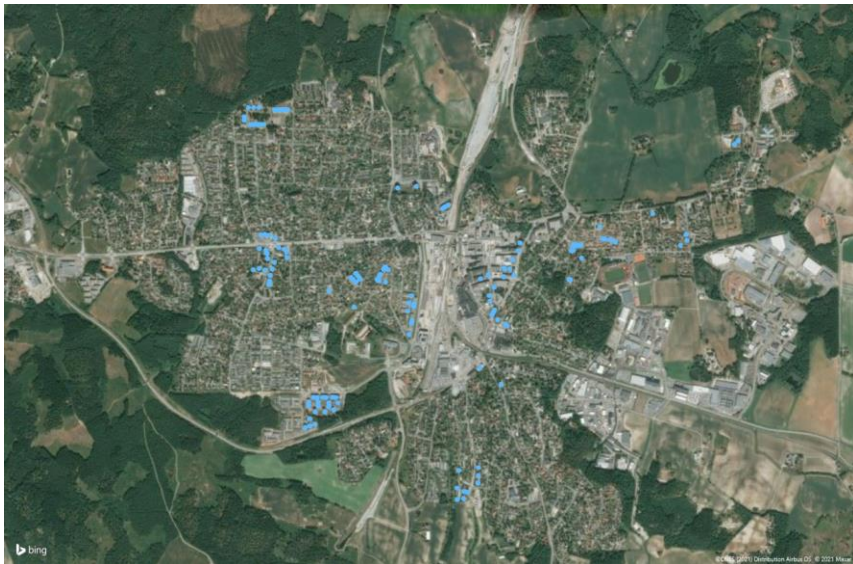
Som poengtert i oppgaven er det en forholdsvis liten andel små leiligheter i kort avstand til Sandvika stasjon, og sentrumsområdet domineres av mellomstore og store leiligheter.

Vedlegg 3

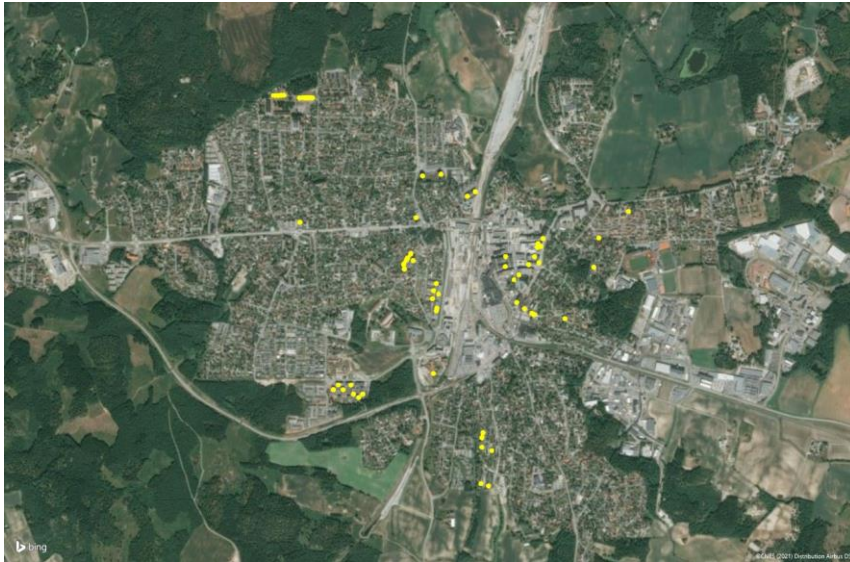
Observasjoner for Ski fordelt etter kategorier. Rød er liten leilighet, blå er mellomstor leilighet og gul er stor leilighet.



Figur 73 Oversikt over små leiligheter i Ski



Figur 74 Oversikt over mellomstore leiligheter i Ski



Figur 75 Oversikt over store leiligheter i Ski

Som man kan se, er det svært få leiligheter i kategorien liten, og disse er i tillegg konsentrert i klynger lengre unna togstasjonen. Man kan også se at mellomstore og store leiligheter ligger i klynger, men da det er flere observasjoner i disse kategoriene

Vedlegg 4

Regresjon med interaksjon mellom avstand til togstasjon (Meter), leilighetsstørrelse og by.

```
Call:
lm(formula = log(KvmF) ~ Meter * Storrelse * Kommune + Aar +
    Byggeaar + Etasje, data = dta4)

Residuals:
    Min       1Q   Median       3Q      Max
-0.62032 -0.10455 -0.00221  0.10359  2.30521

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    -1.133e+02  9.003e-01 -125.904 < 2e-16 ***
Meter           -2.538e-04  1.815e-05  -13.981 < 2e-16 ***
StorrelseMellomstor -3.306e-01  2.815e-02  -11.743 < 2e-16 ***
StorrelseStor   -2.926e-01  2.815e-02  -10.394 < 2e-16 ***
KommuneLILLESTROM -2.398e-01  2.805e-02   -8.549 < 2e-16 ***
KommuneNORDRE FOLLO -3.474e-01  4.796e-02   -7.245 4.70e-13 ***
Aar              6.078e-02  4.473e-04  135.882 < 2e-16 ***
Byggeaar         1.072e-03  4.829e-05   22.193 < 2e-16 ***
Etasje           9.481e-03  1.287e-03    7.364 1.95e-13 ***
Meter:StorrelseMellomstor 1.256e-04  2.041e-05    6.155 7.87e-10 ***
Meter:StorrelseStor     1.337e-05  2.003e-05    0.667  0.505
Meter:KommuneLILLESTROM 1.438e-04  2.020e-05    7.119 1.18e-12 ***
Meter:StorrelseMellomstor:KommuneLILLESTROM 2.812e-04  5.306e-05    5.299 1.19e-07 ***
StorrelseMellomstor:KommuneLILLESTROM 2.207e-01  3.077e-02    7.171 8.09e-13 ***
StorrelseStor:KommuneLILLESTROM 2.026e-01  3.210e-02    6.312 2.90e-10 ***
StorrelseMellomstor:KommuneNORDRE FOLLO 2.331e-01  5.095e-02    4.576 4.82e-06 ***
StorrelseStor:KommuneNORDRE FOLLO 2.349e-01  5.409e-02    4.342 1.43e-05 ***
Meter:StorrelseMellomstor:KommuneLILLESTROM -1.987e-04  2.297e-05   -8.654 < 2e-16 ***
Meter:StorrelseStor:KommuneLILLESTROM -1.012e-04  2.431e-05   -4.164 3.16e-05 ***
Meter:StorrelseMellomstor:KommuneNORDRE FOLLO -2.710e-04  5.594e-05   -4.845 1.29e-06 ***
Meter:StorrelseStor:KommuneNORDRE FOLLO -3.529e-04  6.042e-05   -5.841 5.40e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.164 on 8464 degrees of freedom
Multiple R-squared:  0.745,    Adjusted R-squared:  0.7444
F-statistic: 1236 on 20 and 8464 DF,  p-value: < 2.2e-16
```

Figur 76: Log-Level regresjonsmodell med tre-veis interaksjon mellom Meter: Storrelse og By

Som nevnt vil underproblemstilling 3 kunne besvares og testes gjennom denne modellen. Det som har vært utfordrende er å tolke estimatene på bakgrunn av tre-veis interaksjon. Vi kan se at alle variablene er signifikante på 99% nivå, utenom «Meter:StorrelseStor». Ut ifra hypotesetesting vil man kunne forkaste 0-hypotesene. Utfordringen ligger som nevnt i tolkningen av estimatene, da vi med vårt kunnskapsnivå ikke kan si noe om forholdene mellom de ulike leilighetsstørrelsene og kommunene. Modellen i sin helhet presenteres likevel for å skape åpenhet, og som en mulighet for de som kan tolke den. Da denne regresjonsmodellen har med alle variabler inkludert i denne masteroppgaven, vil en eventuell tolkning av estimatene trolig gi noe ulike resultater, men vi ser ikke for oss at det vil endre konklusjonen nevneverdig.



Norges miljø- og biovitenskapelige universitet
Noregs miljø- og biovitenskapelige universitet
Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås
Norway