



Norges miljø- og  
biovitenskapelige  
universitet

Masteroppgave 2016 30 stp

Institutt for matematiske realfag og teknologi

## **En evaluering av passasjerutviklingen forårsaket av NSBs ruteplanendring i 2012/2015**

An evaluation of the development in passenger demand caused by NSB's timetabling change in 2012/2015.

Torbjørn Aasen Stigen  
Industriell Økonomi



## Forord

Følgende oppgave markerer avslutningen av min studietid ved Norges miljø- og biovitenskapelige universitet innen industriell økonomi. I løpet av siste studiesemester utarbeides en masteroppgave med et arbeidsomfang på 30 studiepoeng hvor man spesialiseres i et ønsket tema. Oppgaven er utarbeidet våren 2016 over en tidsperiode på 18 uker.

Valget av oppgavens tema er gjort med bakgrunn i min økende interesse for transport og infrastruktur gjennom tidligere sommerpraksis hos NSB. Oppgaven er initiert av NSB som har bistått med god veiledning og støtte. Jeg ser på oppgavens problemstilling som interessant og samfunnsnyttig. Arbeidet har til tider vært utfordrende, men samtidig lærerikt. Kunnskapen jeg har tilegnet meg gjennom oppgaven vil jeg dra nytte av i fremtidig yrkeskarriere.

Jeg ønsker å rette en spesiell takk til Halvor Schrøder Hansen fra NSB trafikk og plan. Oppgaven ville ikke fremstått slik den er uten hans veiledning og hjelp. Flere fagfolk fra NSB trafikk og plan må også takkes. Min veileder ved NMBU, Tor Krisitan Stevik takkes for gode råd og kritiske tilbakemeldinger i løpet av arbeidsperioden. Nils Olsson fra NTNU må takkes for gode innspill vedrørende oppgaven. Tilslutt rettes en takk til familie, venner og kjæreste for støtte og oppmuntring gjennom arbeidet. God lesning!

Ås, 10. mai 2015

---

Torbjørn Aasen Stigen

## Sammendrag

Offentlige infrastrukturprosjekter krever enorme ressursinnsatser og har som formål å forbedre landets trafikktilbud. Tross prosjektenes ressurstyngde presiserer tidligere studier at evalueringer av prosjektenes faktiske nytteeffekter er mangelvare (Flyvbjerg et al. 2005; Nicolaisen & Driscoll 2014; Pickrell 1990). Studiene oppfordrer til videre utførelser av effektevalueringer for å etablere et solid empirisk fundament rundt temaet. Trafikkprognoser er en hjørnestein i ex ante evalueringer av prosjekters nytteverdi og har stor påvirkning for hvilke prosjekter som realiseres. En rød tråd fra tidligere forskning er at jernbanerelaterte trafikkkprognoser virker å være overestimerte, noe som gir et feilaktig bilde av nytte-kostnad forholdet ved hvert enkelt prosjekt (Flyvbjerg et al. 2005; Nicolaisen & Driscoll 2014; Pickrell 1990).

Den faktiske nytten av jernbaneprosjekter synliggjøres gjennom en mer attraktiv ruteplan med eksempelvis hyppigere avganger og kortere reisetider. Denne oppgavens formål har vært å evaluere effekten en større ruteendring har hatt for Østlandets passasjerutvikling. Endringen ble gjennomførbar i forbindelse med ferdigstilling av nytt dobbeltspor mellom Asker og Lysaker. Ruteendringen er blitt gradvis implementert av NSB i 2012 og 2015 (NSB 2011).

For å utføre nødvendig analyser er det benyttet multipl lineær regresjon med hensikt om å estimere sammenhengen mellom forklaringsvariabler tilknyttet ruteendringen og passasjerutviklingen. Det er utført analyser med ulike antakelser om til hvilke tidspunkter nytten ved forbedret ruteplan realiseres.

Oppgavens resultater viser signifikante sammenhenger på 90 % nivå mellom passasjerutvikling og avgangsfrekvens, reisetid og befolkningsvekst. Effekten av ny ruteplan viser seg å være generelt høyere enn trafikkkprognosenes beregninger. Resultatene varierer mye under hvilke antakelser analysen er basert på. Oppgavens resultater viser at ruteplanens effekt er høyere under antakelse om en gradvis effektrealisering enn ved umiddelbar effektrealisering. Dette indikerer at det finnes en treghet i markedets respons på tilbudsendringer.

## **Abstract**

Public infrastructure projects aim to improve the public transport system, but require enormous sums of money from taxpayers. Despite the projects' huge resource allocation, previous studies mention limited knowledge of the true effects of infrastructure projects (Flyvbjerg et al. 2005; Nicolaisen & Driscoll 2014; Pickrell 1990). Therefore, these studies recommend future ex-post evaluations of infrastructure projects to establish more empirical knowledge within the field. Demand forecasts are a cornerstone for the cost - benefit analysis of infrastructure projects and play an important role in the final choice of a project alternative. A general perception is that railway-related demand forecasts tend to be overestimated which gives an optimistic picture of the cost -benefit relationship (Flyvbjerg et al. 2005; Nicolaisen & Driscoll 2014; Pickrell 1990).

For railway projects, the effect of infrastructure improvements gets realized through improved timetables (higher train frequencies, reduced travel time etc.). The purpose of this thesis is to evaluate the effects on the passenger demand caused by a major timetable change. A new double-track section between Asker and Lysaker made the change possible, and it was gradually implemented in 2012 and 2015 (NSB 2011).

Multiple linear regression is used to estimate the relationship between passenger demand development and the explanatory variables linked to the timetable change. The thesis is based on several analyses with different assumptions of when the true effects of the change are realised.

The results show, with 90 % significance, a relationship between passenger demand and train frequency, travel time and population growth. The effect of the timetable change proves to be generally higher than assumed in demand forecasts, but with varying results for the different analyses. Generally, the results indicate higher effects of the timetable change under the assumption of a delayed realization of the effects compared to immediate realization. This indicates that the market responds gradually to changes.

# Innholdsfortegnelse

<b>1. INTRODUKSJON</b> .....	<b>1</b>
1.1 KORT OM JERNBANETRAFIKK.....	1
1.2 KORT OM RUTEPLANENDRINGEN 2012.....	2
1.3 OPPGAVENS FORMÅL OG PROBLEMSTILLING.....	6
1.4 OPPGAVENS OPPBYGGING.....	7
<b>2. TEORETISK RAMMEVERK</b> .....	<b>8</b>
2.1 EVALUERING AV PROSJEKTER.....	8
2.2 TRAFIKKBREGNINGER FOR INFRASTRUKTURPROSJEKTER.....	12
2.2.1 Trafikkbelastninger fra tidligere litteratur.....	12
2.3 TRAFIKKBREGNINGER I FORKANT AV RUTEENDRING 2013 OG 2015.....	17
2.4 TIDLIGERE EX POST EVALUERINGER AV STØRRE INFRASTRUKTURPROSJEKTER.....	18
<b>3. METODIKK</b> .....	<b>23</b>
3.1. KORT OM KVANTITATIV OG KVALITATIV FORSKNINGSMETODE.....	23
3.2 MULTIVARIAT ANALYSE.....	24
3.2.1 <i>Multipel lineær regresjon</i> .....	24
3.3. OPPGAVENS STRATEGI.....	31
3.4 ANALYSEVERKTØY – PYTHON PROGRAMMERING.....	31
<b>4. DATAANALYSE</b> .....	<b>33</b>
4.1 VALG AV STATISTISK MODELL.....	33
4.1.1 <i>Presentasjon av variablene i modellen</i> .....	34
4.1.2 <i>Utelatte variabler</i> .....	35
4.2 OMRÅDEINDELING.....	39
4.3 DATAINNHEMTING OG DATASYSTEMATISERING.....	41
4.3.1 <i>Passasjertellinger og avgangsfrekvens</i> .....	42
4.3.2 <i>Pris/Nytt sonesystem i Oslo og Akershus</i> .....	43
4.3.3 <i>Kjøretidsendringer</i> .....	44
4.3.4 <i>Befolkningsvekst</i> .....	45
4.4 GJENNOMFØRING AV ANALYSEN.....	47
4.5 USIKKERHETSHÅNDTERING.....	51
<b>5. RESULTATER</b> .....	<b>55</b>
5.1 RESULTATER MED VEKTEDE OBSERVASJONER.....	55
5.1.1 <i>Umiddelbar effektrealisering med vektete observasjoner</i> .....	55
5.1.2 <i>Forsinkelsestilfelle 1 med vektete observasjoner</i> .....	58
5.1.3 <i>Forsinkelsestilfelle 2 med vektete observasjoner</i> .....	59
5.2 RESULTATER MED IKKE - VEKTEDE OBSERVASJONER.....	61
5.2.1 <i>Umiddelbar effektrealisering med ikke - vektete observasjoner</i> .....	61
5.2.2 <i>Forsinkelsestilfelle 1 med ikke - vektete observasjoner</i> .....	62
5.2.3 <i>Forsinkelsestilfelle 2 med ikke - vektete observasjoner</i> .....	64
5.2.4 <i>Estimerte koeffisienter ved regresjonsanalyser uten pris</i> .....	66
<b>6. DISKUSJON</b> .....	<b>67</b>
6.1 DISKUSJON RUNDT REGRESJONSMODELLENE.....	67
6.1.1 <i>Avgangsfrekvens</i> .....	67
6.1.2 <i>Reisetid</i> .....	68
6.1.3 <i>Pris</i> .....	69
6.1.4 <i>Befolkningsvekst</i> .....	70
6.1.5 <i>Modellene som helhet</i> .....	71
6.2 SAMMENLIKNING MOT PROGNOSE.....	72
6.3 ET KRITISK BLIKK PÅ OPPGAVEN.....	77
6.4 GENERELL DISKUSJON.....	79

6.5 VIDERE ARBEID.....	80
<b>7. KONKLUSJON.....</b>	<b>82</b>
<b>8. LITTERATURLISTE.....</b>	<b>84</b>
<b>9. LISTE OVER VEDLEGG.....</b>	<b>88</b>

## Liste over figurer

FIGUR 1: NORGES JERNBANENETT IDAG (SVINGHEIM 2011).....	1
FIGUR 2: PRINSIPPER FOR UTVIKLING AV RUTEPLANEN (NSB 2011, P. 2) .....	3
FIGUR 3: GRUNNRUTETILBUD FRA DESEMBER 2012 (NSB 2011, P. 11).....	4
FIGUR 4: GRUNNRUTETILBUD VED FULLSTENDIG MODELL (NSB 2011, P. 6) .....	5
FIGUR 5: EVALUERING I ULIKE FASER AV ET PROSJEKT (OLSSON 2005, P.16 ).....	9
FIGUR 6: REFERANSEGRUNNLAG OG REFERANSEPUNKTER I EX POSTEVALUERINGER (OLSSON ET AL. 2015, P.77) .....	11
FIGUR 7: SAMMENHENGEN MELLOM GK OG ETTERSSPØRSEL (OLSSON ET AL. 2015, P. 16) .....	13
FIGUR 8: BEREGNEDE VIRKNINGER AV NY RUTEPLAN 2013 (DYBDAHL ET AL. 2011, P. 9) .....	17
FIGUR 9: BEREGNEDE VIRKNINGER AV NY RUTEPLAN 2015 (DYBDAHL ET AL. 2011, P. 18) .....	18
FIGUR 10: PROGNOSEFEIL I INFRASTRUKTURPROSJEKTER (FLYVBJERG ET AL. 2005, P. 134) .....	19
FIGUR 11: ILLUSTRASJON AV MINSTE KVADRATERS METODE (LØVÅS 2004, P. 273) .....	26
FIGUR 12: OBSERVASJONER I REGRESJONSANALYSEN PÅ MATRISEFORM (RENCHER 2002, P. 324) .....	26
FIGUR 13: EKSEMPEL PÅ OUTPUT (LØVÅS 2004, P.289) .....	27
FIGUR 14: EKSEMPEL PÅ RESIDUALPLOT (LØVÅS 2004, P.282).....	29
FIGUR 15: LINJEKART PÅ ØSTLANDET (NSB 2016) .....	41
FIGUR 16: DATAFLYT I ARBEIDET .....	41
FIGUR 17: SONEKART (RUTER 2016).....	43
FIGUR 18: AVGANGER I 2015 FOR LILLESTRØM (KNUTEPUNKT) .....	44
FIGUR 19: SKJERMBILDE FRA GIS .....	46
FIGUR 20: GRUNNKRETSE TILDELT STASJONENE ALNA OG ASKER.....	46
FIGUR 21: SAMMENSTILLING AV INPUT DATA .....	47
FIGUR 22: UTVIKLING MELLOM 2010-2015 VED UMIDDELBAR EFFEKTREALISERING (VEKTET).....	57
FIGUR 23: UTVIKLING MELLOM 2010-2015 VED FORSINKELSESTILFELLE 1 .....	59
FIGUR 24: UTVIKLING MELLOM 2010-2015 VED FORSINKELSESTILFELLE 2 (VEKTET) .....	60
FIGUR 25:UTVIKLING MELLOM 2010-2015 VED UMIDDELBAR EFFEKTREALISERING (IKKE -VEKTET)..	62
FIGUR 26: UTVIKLING MELLOM 2010 - 2015 VED FORSINKELSESTILFELLE 1 (IKKE -VEKTET).....	64
FIGUR 27: UTVIKLING MELLOM 2010 - 2015 VED FORSINKELSESTILFELLE 2 (IKKE -VEKTET) .....	65
FIGUR 28: SAMMENSTILLING AV UTVIKLINGS PLOT (IKKE -VEKTET).....	65
FIGUR 29: SAMMENSTILLING AV UTVIKLINGS PLOT (VEKTET) .....	66
FIGUR 30: PRIS-PASSASJERVEKST FOR RUTEROMRÅDENE 2011-2012 .....	69



## Liste over tabeller

TABELL 1: OPPGAVENS OPPBYGNING .....	7
TABELL 2: KJENNETEGN VED KVALITATIV OG KVANTITATIV FORSKNING (SILVERMAN 2011, P.5) .....	23
TABELL 3: MODELLENS VARIABLER .....	34
TABELL 4: PUNKLIGHETSOVERSIKT OVER ØSTLANDET .....	37
TABELL 5: OMRÅDEOVERSIKT .....	39
TABELL 6: PRISREDUKSJON FOR RUTEROMRÅDER .....	44
TABELL 7: OUTPUT VED UMIDDELBAR EFFEKTREALISERING (VEKTET) .....	56
TABELL 8: BEREGNEDE EFFEKTER VED UMIDDELBAR EFFEKTREALISERING (VEKTET) .....	56
TABELL 9: FORSINKELSESKOEFFISIENTER 1 .....	58
TABELL 10: OUTPUT VED FORSINKELSESTILFELLE 1 (VEKTET) .....	58
TABELL 11: BEREGNEDE EFFEKTER VED FORSINKELSESTILFELLE 1 (VEKTET) .....	58
TABELL 12: FORSINKELSESKOEFFISIENTER 2 .....	59
TABELL 13: OUTPUT VED FORSINKELSESTILFELLE 2 (VEKTET) .....	60
TABELL 14: BEREGNEDE EFFEKTER VED FORSINKELSESTILFELLE 2 (VEKTET) .....	60
TABELL 15: OUTPUT VED UMIDDELBAR EFFEKTREALISERING (IKKE - VEKTET) .....	61
TABELL 16: BEREGNEDE EFFEKTER VED UMIDDELBAR EFFEKTREALISERING (IKKE - VEKTET) .....	62
TABELL 17: OUTPUT VED FORSINKELSESTILFELLE 1 (IKKE - VEKTET) .....	63
TABELL 18: BEREGNEDE EFFEKTER VED FORSINKELSESTILFELLE 1 (IKKE - VEKTET) .....	63
TABELL 19: OUTPUT VED FORSINKELSESTILFELLE 2 (IKKE -VEKTET) .....	64
TABELL 20: BEREGNEDE EFFEKTER VED FORSINKELSESTILFELLE 2 (IKKE - VEKTET) .....	65
TABELL 21: SAMMENLIKNING AV RESULTATER OG PROGNOSE (VEKTET) .....	73
TABELL 22: SAMMENLIKNING AV RESULTATER OG PROGNOSE (IKKE - VEKTET) .....	74
TABELL 23: OBSERVERT TOTALVEKST MOT ANTATT TOTALVEKST FRA PROGNOSE .....	75

# 1. Introduksjon

## 1.1 Kort om jernbanetrafikk

Norges første jernbanestrekning ble åpnet i 1854. Jernbanestrekket på 68 km mellom Eidsvoll og Kristiania fikk navnet Hovedbanen. Økt handel og samferdsel i 1840-årene skapte grobunn og interesse for å utvikle jernbane i Norge. Første året hovedbanen var i drift ble 128 000 passasjerer og 83 000 tonn trelast fraktet. I dag er jernbanen vesentlig mer utbygd og en viktig brikke i Norges transportnettverk (Svingheim 2008).



Figur 1: Norges jernbanenett idag (Svingheim 2011)

Jernbanen er avgjørende i infrastrukturen for de fleste land. Jernbanetrafikk anses som en miljøvennlig og effektiv transportmetode og kan inndeles i person- og godstrafikk. Olsson & Veiseth (2011) forklarer at sammenliknet med alternative transportmetoder kreves langt mindre energi for å flytte en gitt masse på skinner sammenliknet med veg.

Jernbanen drar derfor nytte av stordriftsfordelene den har. Kostnaden knyttet til frakt av en passasjer eller et tonn gods faller proporsjonalt med økende antall passasjerer eller økt godsmengde. Jernbanens tekniske innretninger deles hovedsakelig inn i infrastrukturen (signalanlegg, skinner, stasjoner etc.) og det rullende materiell (persontog, lokomotiver, godstog etc.). Jernbanetrafikk har lav fleksibilitet sammenliknet med transportalternativer i luft og på vei. God planlegging og tett samarbeid mellom alle aktørene i jernbanesektoren er derfor nødvendig. Norges jernbanesektor er i dag delt opp i tre hovedorganer: jernbaneverket, jernbanetilsynet og trafikkutøverne. Jernbaneverkets hovedoppgave er å bygge ut og vedlikeholde infrastrukturen, men har også ansvarsområder innen strømforsyning, signalanlegg og operativ togledelse. Trafikkutøverne frakter personer eller gods og inngår avtaler med jernbaneverket om sporbruk. De største trafikkutøverne i Norge er NSB AS og flytoget innen persontrafikk. CargoNet AS er størst innen godstrafikk. Jernbanetilsynets oppgave er oppfølging av sikkerheten på jernbanen og at alle aktørene følger jernbanens lovverk. Et tett samarbeid mellom aktørene gjør dagens jernbane til det den er (Olsson & Veiseth 2011).

## **1.2 Kort om ruteplanendringen 2012**

En ruteplan fra NSB setter rammene for tilbudet på Norges jernbane. Derfor utvikles det jevnlig nye ruteplaner. En gjennomtenkt og tilpasset ruteplan er vesentlig for å utnytte eksisterende infrastruktur og for å tilfredsstille markedets etterspørsel best.

Ruteplanlegging er viktig for NSB og har stor betydning for selskapets resultater. Et forbedret togtilbud synliggjøres gjennom en ny ruteplan og er gjennomførbart ved økt kapasitet på jernbanen. Dette forekommer ved ferdigstilling av ny infrastruktur eller nytt togmateriell. En ruteplan kan beskrives som en optimalisert plan for å tilby et best mulig togtilbud utfra tilgjengelig infrastruktur og ressurser. Siden 2006 arbeidet NSB med å utvikle en ny ruteplan for Østlandet som utgjør over 80 prosent av persontrafikken i Norge (NSB 2011). Tilbudsløftet ble mulig som følge av nytt ferdigstilt dobbeltspor mellom Asker og Lysaker. Nye tog fra Stadler AG i Sveits (FLIRT togene) var en forutsetning for gjennomføring av ny ruteplan. Ruteplanen ble utviklet med formål om å være et markedsorientert tilbud med vektning på enkelhet for kunden, jevn frekvens og økt tilbuds nivå (NSB 2007). De styrende prinsippene vist i figur 2 var retningsledende ved utviklingen.

Styrende prinsipper for utvikling av rutetilbudet
– Linjenettet skal være enkelt og ha få varianter – dette letter markedsføringen og gjør terskelen lavere for å bruke toget (viktig for å nå nye kunder)
– Jevn fordeling av avganger over timen prioriteres på viktige knutepunkt
– Ujevn fordeling over timen kan aksepteres for stasjoner utenfor et hovedmarked for å oppnå jevn fordeling på timen i hovedmarkedet
– Jevnt fordelt frekvens på dagtid prioriteres foran jevn frekvens i rushtiden når frekvensen på dagtid er fire tog i timen eller høyere
– I bynære områder prioriteres høy frekvens foran kortere reisetid
– Det aksepteres ikke at man "hopper over" en avgang i tilbudet
– Det aksepteres at ekspresstog bryter med prinsippet om fast stoppmønster og jevn frekvens for knutepunktstoppende tog
– Det aksepteres nedleggelse av mindre holdeplasser dersom det medfører kortere reisetid for andre reisende og at den totale etterspørselen på linjen opprettholdes eller økes på sikt
– I utgangspunktet utvides rushtidens varighet i det bynære området til tre timer og til to timer utenfor det bynære
– Det skal tilstrebes gode overgangsmuligheter i knutepunkter mellom ulike delmarkeder
– Linjer fra nordøst til vest etableres i rutemodellen slik at man kan videreføre mest mulig av løsningen også etter at nytt dobbeltspor Oslo-Ski er bygget.

Figur 2: Prinsipper for utvikling av ruteplanen (NSB 2011, p. 2)

"Et hovedprinsipp ved konseptet er at ulike linjer samordnes for å oppnå synergieffekt av den totale innsatsen på felles strekninger. Knutepunktstoppende tog settes i system slik at de ulike linjene til sammen gir ti minutters intervall mellom avgangene på de store stasjonene mellom Asker og Lillestrøm" (NSB 2011, p. 4).

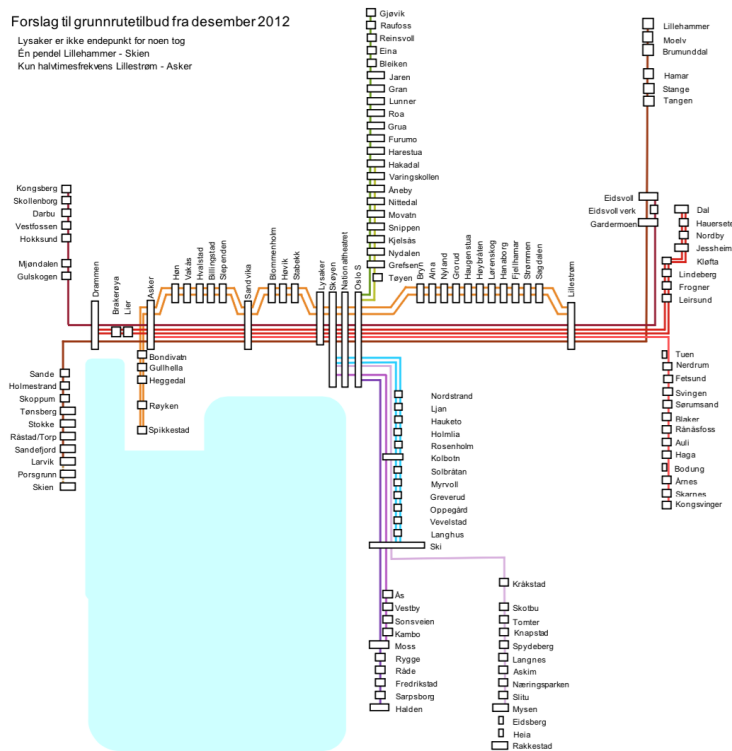
Ruteplanen ga tilnærmet et dobbelt så godt tilbud sammenliknet med tidligere mellom Asker og Lillestrøm. Kapasitetsanalyser identifiserte også hvilke mindre infrastrukturforetak som var nødvendig for å oppnå full utnyttelse av nye tog og ruteplanen. Et tilbudsløft på Østfoldbanen ble avklart å ikke være mulig før Follobanens ferdigstillelse (NSB 2011).

De største forbedringene ved ny ruteplan oppsummeres av NSB (2011) i punktene under:

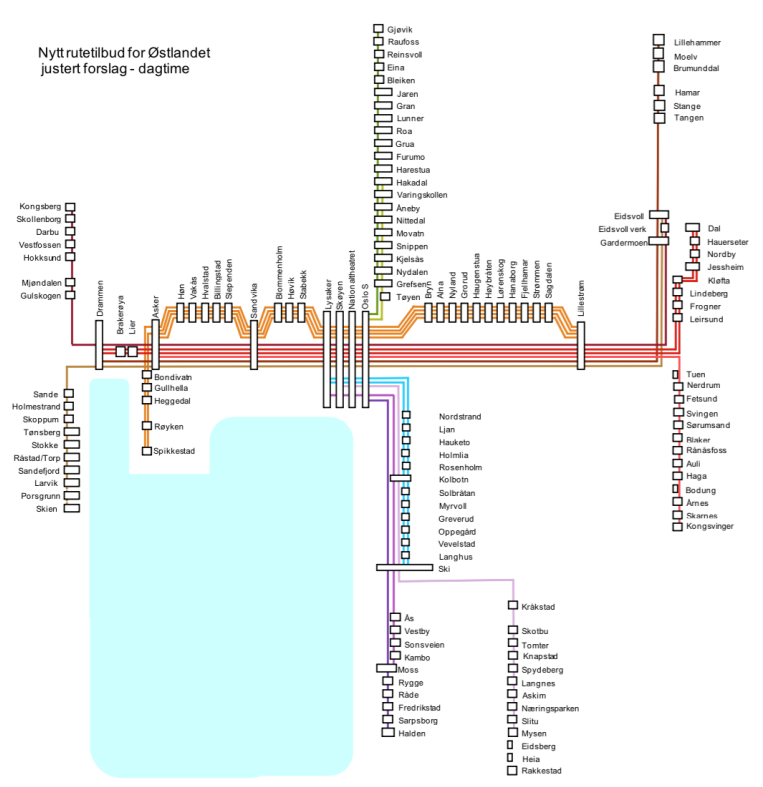
- Jevn frekvens hvert tiende minutt med knutepunkt stoppende tog mellom Asker – Lillestrøm.
- Økt frekvens fra tre til fem tog pr time i grunnrute mellom Drammen – Oslo.

- Dobbel frekvens i grunnruten for fullt stoppende tog Asker – Lillestrøm.
- Dobbel frekvens i grunnruten på alle stasjoner Leirsund – Jessheim – Dal.
- Dobbel frekvens i grunnruten på Eidsvoll verk.
- Tidsbesparelser pga. ny infrastruktur mellom Sandvika – Lysaker og Tønsberg – Barkåker.
- Flere tog gir bedre setekapasitet i rushtidene (p. 9)

Nye ruteplaner stiller mange krav og forutsetninger til infrastrukturen. Derfor innføres ofte ruteplanendringer trinnvis. For dette tilfellet gjorde bygging av nytt vendeanlegg på Høvik stasjon det vanskelig å innføre den fullstendige ruteplanen i 2012. Utsatt åpning av stasjonen førte til at vendeanlegget først ble brukt i 2015 (Nordli 2015). I desember 2012 ble mest mulig av ruteplanen innført, men antall tog mellom Oslo S og Lysaker forble på nivå som tidligere. I 2015 ble endelig rutemodell implementert (NSB 2011). Figur 3 og 4 viser ruteplanene for 2012 og 2015. En linje i figur 3 og 4 representerer et tog som går hver time. Grunnet flere tog i 2015 er det dermed flere linjer i figur 4.



Figur 3: Grunnrutetilbud fra desember 2012 (NSB 2011, p. 11)



Figur 4: Grunnrutetilbud ved fullstendig modell (NSB 2011, p. 6)

### **1.3 Oppgavens formål og problemstilling**

Denne oppgaven har som formål å gi en bedre forståelse av betydningen innføring av nye ruteplaner har for NSBs passasjervekst. Med tanke på de enorme ressursinnsatsene som nedlegges i store infrastrukturprosjekter og til utvikling av ruteplaner er det nyttig å kartlegge effektene dette gir.

Trafikkprognoser for infrastrukturprosjekter påvirker sterkt hvilke prosjekter som blir prioritert da de danner fundamentet for en nytte-kostnadsberegning. Med bakgrunn i dette understreker tidligere litteratur nødvendigheten av å gjennomføre effektevalueringer av prosjekter i ettertid for å tilegne økt empirisk kunnskap av prognosenes kvalitet (Flyvbjerg et al. 2005; Nicolaisen & Driscoll 2014; Pickrell 1990). Oppgaven ønsker å gi bedre kunnskap om usikkerheten ved trafikkprognosene som er gjort i forbindelse med ruteendringen. Funnene i oppgaven er ment å gi NSB et bredere grunnlag i videre arbeid med utvikling av ruteplaner og trafikkprognoser. Gjennom kvantitative analyser og en diskusjon rundt resultatene i oppgaven forsøker jeg å belyse følgende problemstilling.

#### **Hvilken effekt har større tilbudsendringer ved togtilbudet i 2012 og 2015 hatt for passasjerutviklingen på Østlandet og hvordan samsvarer dette med trafikkprognosene?**

For å besvare problemstillingen kreves en analyse av ruteplanen og trafikkutviklingen som er observert. Virkningen som endringene i ruteplanen har hatt for passasjerutviklingen må estimeres og isoleres fra andre ytre virkninger. Følgende forskningsspørsmål er videre definert for å besvare problemstillingen.

#### **Forskningsspørsmål 1:**

Hvilken effekt har tilbudsendringene i 2012 og 2015 hatt for passasjerutviklingen på Østlandet?

#### **Forskningsspørsmål 2:**

Hvilke faktorer har i størst grad drevet passasjerveksten?

#### **Forskningsspørsmål 3:**

Til hvilken grad stemmer effektberegningene i oppgaven med trafikkprognosene?

## 1.4 Oppgavens oppbygging

Tabell 1 gir en oversikt over hovedelementene som inngår i oppgaven og elementenes innhold og hensikt.

Tabell 1: Oppgavens oppbygning

<b>Del</b>	<b>Innhold</b>
<b>Teori</b>	Teorikapittelet belyser relevant teori fra tidligere forskning rundt oppgavens tematikk. Kapittelet har som formål å presentere litteratur som danner et diskusjonsgrunnlag for oppgavens resultater.
<b>Metode</b>	Metodekapittelets hensikt er å beskrive hvilke grunnleggende metoder og verktøy som er valgt for å angripe problemstillingen. Leseren skal få et bedre grunnlag for å forstå oppgavens analyse og resultatene som blir fremlagt.
<b>Dataanalyse</b>	Kapittelet som omfatter dataanalyse ønsker å gi leseren en oversikt over den trinnvise fremgangen i analysearbeidet. Dette inkluderer hvordan data er innhentet og en forklaring av hvordan beregningene er utført. Kapittelet forklarer hvordan den analytiske modellen er oppbygd.
<b>Resultater</b>	Kapittelet presenterer oppgavens resultater. Resultatene blir beskrevet og presentert objektivt.
<b>Diskusjon</b>	Diskusjonskapittelet sammenlikner resultatene presentert mot tidligere litteratur og forskning beskrevet i teorikapittelet. Egne refleksjoner vil også fremkomme i kapittelet.



## 2. Teoretisk rammeverk

Kapittelets formål er å presentere relevant teori for oppgavens problemstilling. Teksten presentert i kapittelet skal brukes som diskusjonsgrunnlag for oppgavens resultater. Innledningsvis tar kapittelet for seg noe grunnleggende vedrørende evaluering av prosjekter. Deretter blir tidligere litteratur rundt trafikkberegninger og trafikkprognoser presentert. Trafikkprognosene for oppgavens aktuelle prosjekt blir fremlagt i kapittelet.

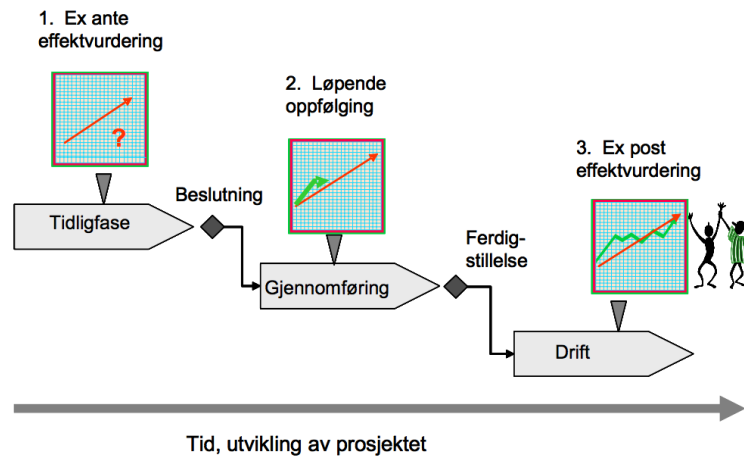
### 2.1 Evaluering av prosjekter

”Prosjekter evalueres på ulike tidspunkt under eller etter gjennomføringen. I noen tilfeller gjøres dette når et alvorlig problem har oppstått, for å finne ut hva som bør gjøres. I andre tilfeller ønsker en å trekke lærdom av prosjektet på godt og ondt.” (Samset 2008, p. 307)

Evaluering av prosjekter er nødvendig ved flere tidspunkt av en prosjektgjennomføring. Å anslå et prosjekts vellykkethet er først mulig i etterkant når resultatene er kommet. Evalueringer før og under prosjekter er likevel nyttig for å anslå om prosjektene bør iverksettes og om korrigerende tiltak bør gjøres underveis. Hva som er et vellykket prosjekt er diskutabelt. Hvilke kriterier må oppfylles for at prosjektet blir definert som vellykket? Dette vil variere for hvert enkelt prosjekt. En prosjektevaluering bør inneholde en vurdering av lønnsomhet og nytte av prosjektet på lengre sikt (Samset 2008).

#### **Ex ante og ex post evaluering**

Prosjektevalueringer gjøres før, under og etter gjennomføring. En evaluering i forkant av et prosjekt kalles en ex ante evaluering og beskriver forventningene som finnes til prosjektet. En underveisevaluering vil vurdere hvordan det ser ut til at prosjektet blir. Evalueringer i etterkant av gjennomføringen kalles ex post evalueringer og ser på hvordan prosjektet faktisk ble (Olsson 2005). Figur 5 gir oversikt over til hvilke tidspunkter de ulike evalueringstypene finner sted i prosjektskyklusen.



Figur 5: Evaluering i ulike faser av et prosjekt (Olsson 2005, p.16 )

Det finnes vesentlige sammenhenger mellom ex ante og ex post evalueringer. Den metodiske fremgangsmåten for evalueringene er i stor grad lik da valg av vurderingskriterier ofte er tilsvarende. Ex ante vurderinger er avhengig av at det finnes tidligere ex post evalueringer fra lignende prosjekter som kan benyttes til å estimere virkningene av nye liknende prosjekter. Ex post evalueringer er med andre ord med på å forbedre kvaliteten til fremtidige ex - ante evalueringer. Likevel kan generell teori og empirisk kunnskap kan også benyttes i gjennomføring av ex ante evalueringer (Olsson 2005).

Det finnes flere tilnæringsmetoder for å gjennomføre ex post evalueringer. Disse kan med bakgrunn i Olsson et al. (2010) kategoriseres som:

- Samfunnsøkonomiske evalueringer
- Bedriftsøkonomiske evalueringer
- Holistisk evaluering: helhetsevaluering, flere tilnæringsmåter,
- Prestasjonsmålt evaluering: Evaluerer nøkkelparametere som er essensielle for prosjektets vellykkethet

Valget av tilnæringsmetode avhenger av perspektivet på evalueringen og hva som er aktuelt å studere. Denne oppgaven har en tilnærming som siste punkt hvor en sentral nøkkelparameter for prosjektets grad av vellykkethet evalueres - passasjerutviklingen.

## **Referansepunkter og referansegrunnlag i evalueringsprosesser**

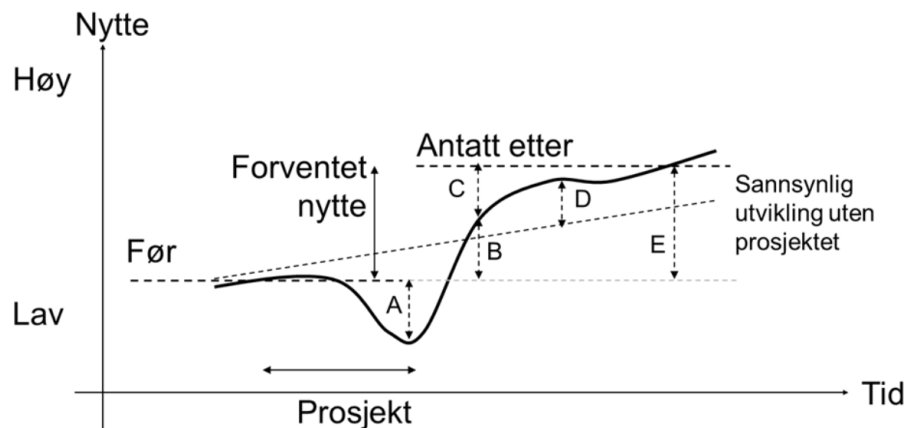
En ex post evaluering vurderer utfallet av et prosjekt etter ferdigstillelse. For å gjøre dette er det essensielt å måle utfallet opp mot noe - et referansegrunnlag. Flere tilnæringsmåter er interessante å bruke som referansegrunnlag. Med bakgrunn i Olsson et al. (2010) kan eksempler på referansegrunnlag nevnes å være:

- Situasjon før implementering mot situasjon etter implementering.
- Antatt utfall ved implementering mot faktisk utfall ved implementering
- Antatt utfall uten implementering mot faktisk utfall ved implementering
- Utfall av liknende prosjekter mot utfall av aktuelt prosjekt

Prosjektevalueringer med definerte referansegrunnlag krever videre at det finnes et referansepunkt. Når skal resultatene sammenliknes mot referansegrunnlaget etter prosjektgjennomføring? Olsson et al. (2010) ser nærmere på hvordan valg av referansepunkter påvirker evalueringresultatet og går i dybden på prestasjonsmålte ex post evalueringer av fire norske jernbaneprosjekter. Studiet konkluderer med at valget av referansepunkter har vesentlig betydning for utfallet av en ex post evaluering. Ulike valg kan være kritisk for om prosjektet blir definert som vellykket. Derfor kan valget av referansetidspunkt påvirke evalueringens utfall i en ønsket retning.

Olsson et al. (2010) konkluderer med at måling av prestasjonsgjennomsnittet over noen år før implementering av prosjektet for sammenlikning av prestasjonsgjennomsnittet over noen år etter implementering er gunstig for å gi en robust evaluering.

Det bemerkes også av Olsson et al. (2015) viktigheten av referansegrunnlag i ex post evalueringer. Sammenlikningsgrunnlaget bør for alle år være basert på samme periode for å unngå sesongvariasjoner i prestasjonen. Figur 6 viser forskjellige referansegrunnlag for en ex post evaluering med en prinsippsskisse over prestasjonsutviklingen før og etter implementering av et prosjekt.



Figur 6: Referansegrunnlag og referansepunkter i ex post evalueringer (Olsson et al. 2015, p.77)

A, B, C, D og E viser ulike referansesituasjoner for en ex post evaluering. Situasjon A tar utgangspunkt i en evaluering rett etter implementering av prosjektet. Prestasjonene kan ofte forverres ved dette tidspunktet grunnet "barnesykdommer". Situasjon B viser sammenlikning av forbedringen en stund etter prosjektets fullføring med situasjonen før, mens i situasjon C sammenliknes forventede effekter mot faktiske effekter ved samme tidspunkt som i B. I situasjon C har ikke all forventet forbedring blitt realisert og resultatet kan virke skuffende. Situasjon D sammenlikner hvordan det antas at det hadde vært uten prosjektet. Situasjon E gjør en sammenlikning lang tid etter gjennomføring. Figuren viser hvordan nytteverdien fra prosjekter ofte kommer frem over tid. Dette medfører at man opplever forskjellige resultater i en evalueringsprosess ved ulike referansetidspunkter. Ved et for tidlig referansetidspunkt vil man risikere å evaluere prosjektet før nytten er realisert, mens ved et sent referansetidspunkt (4-5 år etter avslutning) kan man risikere å måle effekter som ikke er forårsaket av prosjektet, men av andre årsaker (Olsson et al. 2015).

## 2.2 Trafikkberegninger for infrastrukturprosjekter

“Traffic engineers often compare traffic to a fluid, assuming that a certain volume must flow through the road system. But urban traffic may be more comparable to a gas that expands to fill available space.” (Litman 2015, p. 2)

Sitatet bemerker at når trafikktilbudet forbedres tenderer den totale trafikkmengden i urbane områder til å øke. Folk reiser generelt mer som følge av et mer attraktivt tilbud. En økt total trafikkmengde som følge av tilbudsendringer kalles generert eller induisert trafikk (Litman 2015).

Ved trafikkberegninger grunnet endringer i togtilbudet er det nødvendig å etablere en kvantitativ sammenheng mellom trafikketterspørsel og togtilbudet. For å kunne beregne nytteverdien av forbedringer i jernbanens infrastruktur må det utføres trafikkberegninger som estimerer nye trafikkstrømmer. Ulike trafikkberegningsmodeller benyttes. Jernbaneverket (2015) nevner at en god trafikkberegningsmodell bør inneholde alle relevante variabler som er med på å forklare trafikantenes valg. Trafikkmodeller bør ideelt sett kunne kvantifisere alle variabler som påvirker reisendes verdsetting av endringer i trafikktilbudet. Nasjonal transportplan har utviklet flere gode trafikkmodeller. Ved valg av modell bør det undersøkes om alle relevante variabler er inkludert i modellen. Dette kan være vanskelig å få til i praksis. Derfor er ofte variabler som har betydning dessverre utelatt fra trafikkberegningen (Jernbaneverket 2015).

Neste avsnitt går inn på ulike elastisitetstall mellom togtilbud og etterspørsel som brukes i trafikkberegninger og elastisitetstall som er funnet tidligere. I økonomien bruker man ofte elastisiteter mellom pris og etterspørsel. For eksempel betyr en priselastisitet på -2,5 at en prisreduksjon på 1,0 % gir økt etterspørsel på 2,5 %. Følgende avsnitt beskriver hvordan dette anvendes i trafikkberegninger.

### 2.2.1 Trafikkelastisiteter fra tidligere litteratur

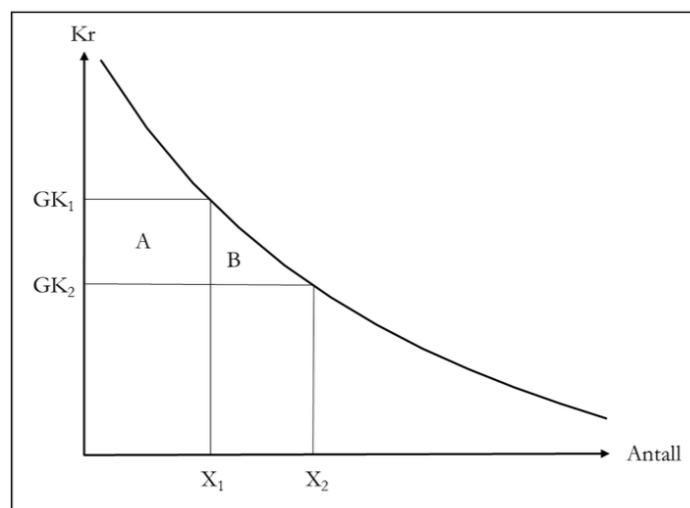
Jernbaneverket (2015) forklarer forenklet modell for beregning av trafikale konsekvenser. Modellen tar standpunkt i at økt antall trafikanter på jernbanen er forårsaket av et mer attraktivt togtilbud. En endring i togtilbudet kan forekomme gjennom flere avganger, kortere reisetid og mer pålitelig levering (punktlighet). Ved bruk av metoder for å verdsette disse elementene kan man kvantifisere togtilbudet etter

begrepet generaliserte kostnader. En reduksjon i generaliserte kostnader har sammenheng med et forbedret togtilbud og gir dermed positive utslag i trafikketerspørselen. Jernbaneverket (2015) nevner at følgende elementer bør inngå i begrepet generaliserte kostnader, men elementene som inngår varierer likevel i hvert prosjekt.

- Reisetid
- Gangtid til stasjon
- Ventetid
- Antall stopp
- Andre kostnader

Metodehåndboken tar utgangspunkt i standardverdier for trafikkelasiteter.

Trafikkelasitetene beskriver sammenhengen mellom generaliserte kostnader og etterspørselen. Standardverdiene er -1,2 og -1,5 for henholdsvis reiser under 50 km og reiser over 50 km (Jernbaneverket 2015). Dette betyr at en reduksjon i generaliserte kostnader på 1,0 % gir økt etterspørsel på 1,2 og 1,5 %. Olsson et al. (2015) forklarer også prinsippet rundt generaliserte kostnader. Figuren under illustrerer sammenhengen mellom generaliserte kostnader og passasjeretterspørsel.



Figur 7: Sammenhengen mellom GK og etterspørsel (Olsson et al. 2015, p. 16)

Trafikkelasiteter er i stor grad avhengig av hvilke komponenter som er lagt inn i begrepet generaliserte kostnader og ikke minst hvordan komponentene vektet. Det er store variasjoner i hvilke elastisiteter som faktisk er rette ved hvert enkelt tilfelle.

Trafikkberegninger er derfor svært sårbare for de forutsetninger som er lagt inn i beregningen (Aarhaug et al. 2013).

En forskningsrapport gjort av Fearnley et al. (2012) undersøker de store etterspørselsdriverne i jernbanemarkedet og hvordan etterspørselen responderer til endringer i disse. Med andre ord undersøkes elastisitetstallene til spesifikke faktorer som preger jernbanemarkedet. Dette ligger dermed på et mer detaljert nivå en generaliserte kostnader som omfatter flere faktorer under ett. Rapporten baseres på eksisterende litteratur fra utlandet og Norge. Det nevnes at passasjeretterspørselen preges av både eksterne og interne etterspørselsdrivere. Med eksterne drivere menes faktorer som ikke kan påvirkes direkte av togtilbudet, mens interne drivere er faktorer med direkte tilknytning til togtilbudet.

Av eksterne drivere nevner Fearnley et al. (2012) bruttonasjonalprodukt og sysselsetting som faktor med størst innflytelse med elastisitetstall mellom 1-1,3. FitzRoy & Smith (1995) nevner i sitt studie av jernbanetraffikk i fjorten europeiske land en elastisitet for bruttonasjonalprodukt på 0,83. Fearnley et al. (2012) nevner videre drivstoffkostnader og reisetid for bil med elastisitetstall på henholdsvis 0,2 og 0,3. Reisetider og pris med buss er også nevnt med elastisiteter på 0,25 og 0,2. Disse tallene er basert på togmarkedet i Storbritannia. Dette blir også tatt opp av FitzRoy & Smith (1995) som beregner en elastisitet for bensinpris på 0,57. Inntekt kan også anses å være en ekstern etterspørselsdriver. Hamre (2002), som ser på Norges jernbanemarked, beregner en elastisitet for personinntekt til å ligge på 0,162. Befolkningsvekst vil også tolkes som en ekstern etterspørselsdriver. Wijeweera & Charles (2013) sitt studie rundt hva som forårsaket en stor passasjerøkning i Melbournes togtraffikk fant blant annet en signifikant sammenheng mellom befolkningsvekst og passasjerutvikling. Elastisitetsverdien ble estimert til å ligger rundt 0,5. Studiet fant også en signifikant sammenheng mellom bensinpris og passasjerutvikling med en verdi på 0,06.

Av interne drivere forklarer Fearnley et al. (2012) at harde kvalitetsfaktorer som avgangsfrekvens, pris og reisetid har størst betydning for passasjerutviklingen. En tendens i elastisitetenes oppførsel er at langtidselastisiteter har høyere verdier enn korttidselastisiteter. Dette indikerer at driverne får større påvirkningskraft over tid.

Dette kan forklares av at markedet bruker tid på å reagere på tilbudsendringer (Fearnley et al. 2012). Denne observasjonen blir også støttet i Paulley et al. (2006) sitt studie av utviklingen av jernbanetransport i Storbritannias urbane områder, samt av Wijeweera & Charles (2013).

Hamre (2002) beregner reisetidselastisiteter på -0,651 for tjenestereiser og -0,366 for private reiser over 100 km, med et totalgjennomsnitt på -0,403. Disse tallene er beregninger med utgangspunkt i det norske togmarkedet og sier at eksempelvis en reisetidsreduksjon på 10 % gir 4,03 % flere reisende.

Elastisiteter vedrørende avgangsfrekvens er av Hamre (2002) beregnet å ligge på 0,394 for lengre tjenestereiser og 0,294 for lengre private reiser, med en total på 0,307.

FitzRoy & Smith (1995) beregner en frekvenselastisitet på 0,44 for europeisk jernbanetransport. En dobling i avganger vil med andre ord gi 44 % flere reisende.

Paulley et al. (2006) nevner en elastisitet for endringer i materiellkilometer på 0,75 (for kort tidshorison). Tallet hadde grunnlag i få målinger og bør tolkes med sunn skepsis.

Materiellkilometer kan sammenliknes med avgangsfrekvens da en dobling i avgangsfrekvens på en strekning vil medføre en dobling i materiellkilometer dersom avgangene går like langt som tidligere med samme togstørrelse.

Vedrørende prisendringer beregnes det av Hamre (2002) en priselastisitet på -0,301. En dobling i pris er med andre ord beregnet å redusere antall passasjerer med rundt 30 %.

Paulley et al. (2006) finner også priselastisiteter i størrelsesorden -0,09 til -1,02 med et gjennomsnitt på rundt -0,50. Det vises til at priselastisitetene virker å være dynamiske med en tendens til å øke over tid. Elastisitetene er dermed svært variable. Det bemerkes også at arbeidsrelaterte reiser ofte er mindre følsomme til prisendringer enn fritidsreiser.

Reisende med større valgfrihet av transportmetode, for eksempel bileiere, er også mer prissensitive enn de uten transportalternativer (Paulley et al. 2006).

Også Wijeweera & Charles (2013) estimerer en signifikant priselastisitet for togmarkedet i Melbourne. Elastisiteten for pris var likevel funnet å være svært lav med en verdi på 0,032 på kort sikt og 0,07 på langsikt. Dette avdekket at markedet var tilsynelatende lite prissensitivt, hvor en prisreduksjon på 10 % gav 0,32 % og 0,7 % flere reisende. Fouquet (2012) studerte hvordan pris og inntekt har påvirket transporttetterspørselen over et større tidsrom (1850-2010). Funnene til Fouquet indikerte at pris og inntekt har fått mindre påvirkningskraft på den totale trafikketterspørselen over tid og at elastisitetene synker i verdi når utviklingsland blir



industrialiserte. Fouquet (2012) nevner inntekts og priselastisiteter på henholdsvis 3,1 og -1,5 i midten av 1900 tallet, mens i 2010 var elastisitetenes verdier sunket til 0,8 og -0,6. Punktlighet kan også tolkes som en intern etterspørselsdriver i togmarkedet da det definerer togets leveringsevne. Fearnley et al. (2012) trekker frem punktlighetens påvirkning på antall reisende. Overordnet bemerkes det at dårlig punktlighet oppfattes som en belastning og derfor er viktig. Likevel nevnes det svært varierende resultater vedrørende punktlighetens alvorlighetsgrad for passasjerutviklingen og gir signaler på at det er å forvente at punktlighetsendringer ikke gir store utslag i etterspørselen. Logikken bak disse funnene kan kobles opp mot at mange togreisende, tross stor irritasjon over svak punktlighet, ikke har andre transportalternativer å spille på når togets leveringsevne først svikter (Fearnley et al. 2012).

Fearnley & Bekken (2005) går i dybden på dynamikken i etterspørselastisitetene og hvordan disse utvikler seg over tid. Det blir bemerket at det fra tidligere er liten kjennskap til i hvilken grad elastisitetsverdiene utvikler på sikt. De samlet inn elastisitetsverdier for kortsikt og langsikt fra ulike studier og sammenliknet forholdstallene mellom disse. Kort sagt finner Fearnley & Bekken(2005) et gjennomsnittlig forhold mellom elastisitetsverdiene på 1,84. Dette indikerer at langtidselastisitetene er rundt 84 % høyere enn de kortsiktige elastisitetene. Med andre ord er kun litt over halvparten av effekten ved en endring realisert på kort tid. Kort sikt er antatt å være en effektrealisering innen ett år, mens langsikt kan være oppimot 10-20 år. Likevel nevnes det vanskeligheter med å se ytterligere forbedringer utover en tidsperiode på 5-7 år (Fearnley & Bekken 2005).

Som en generell oppsummering kan det nevnes at elastisitetstall er vanskelig å anslå presist. Ulik litteratur viser til svært forskjellige elastisiteter. Det er med andre ord store forskjeller for hva som er funnet i hvert enkelt marked og for hvert prosjekt. Det finnes usikkerhetsfaktorer knyttet til elastisitetstallene. Verdiene på elastisitetene er også dynamiske og har som tendens å øke over tid. Trafikkberegningsmodeller vil utelate enkelte variabler som påvirker de store trafikkstrømmene. Tidligere studier har funnet signifikante sammenhenger mot passasjeretterspørsel for mange ulike variabler, men det er store forskjeller i hva som gjelder for hvert marked og prosjekt. Metoden ved bruk av generaliserte kostnader er ofte brukt til å estimere effekter av togtilbudsendringer, men vekting av de ulike faktorene i metoden og hvilke komponenter som er med vil variere, og dermed påvirke resultatene. Konkurransforholdet mellom tog og alternative

transportalternativer vil også ha innvirkning på effektutslaget ved nye tiltak. Stort sett er det mange forhold som varierer (Jernbaneverket 2015).

På et generelt nivå kan det nevnes at harde faktorer som reisetid, frekvens, punktlighet, pris med mer spiller en viktig rolle for hvordan markedet verdsetter togtilbudet.

Potensialet i Norges togmarked kan i hovedsak realiseres gjennom reduserte reisetider og hyppigere togavganger, med fokus på de største byene. I tillegg eksisterer det mange eksterne faktorer som påvirker etterspørselen som er utenfor hva som kan kontrolleres av jernbanen (Fearnley et al. 2012).

### 2.3 Trafikkberegninger i forkant av ruteendring 2013 og 2015

Før innføring av ny ruteplan i desember 2012 ble det gjort trafikkberegninger for påfølgende år av selskapet Vista Analyse AS. Som grunnlag for trafikkberegningene ble kvaliteten på togtilbudet definert utfra generaliserte kostnader med standard elastisitetstall. Trafikkberegningene har tatt høyde for innføringen av nytt sonesystem med ruter, noe som ga utslag i billettpriser for ulike områder, og er derfor estimerer på de isolerte virkningene av ny ruteplan. Beregningene har tatt utgangspunkt i at all effekt av endringer i togtilbudet realiseres umiddelbart (Dybdahl et al. 2011).

Figur 8 viser trafikkprognosene ved innføring av nytt ruteopplegg i 2013.

Pr. år	Referanse 2008	Referanse 2013	Økning 2008-2013	Ny ruteplan 2013	Økning Ny ruteplan
<b>InterCitymodellen for Østlandet</b>					
Antall reiser	9 622 689	10 306 641	7,1 %	10 615 460	3,0 %
- herav regiontog	6 404 613	6 960 572	8,7 %	7 156 966	2,8 %
- herav lokaltog	3 218 076	3 346 069	4,0 %	3 458 494	3,4 %
Mill. personkm	736	793	7,7 %	815	2,8 %
<b>Emma Fredrik</b>					
Antall reiser	28 925 335	30 585 627	5,7 %	31 519 943	3,1 %
Mill. personkm	464	491	5,7 %	512	4,4 %
<b>Tilbringertrafikk, Gardermoen</b>					
Antall reiser	458 358	509 677	11,2 %	1 414 162	177,5 %
Mill. personkm	27	30	11,2 %	80	171,4 %
<b>SUM NSB</b>					
Antall reiser	39 006 382	41 401 945	6,1 %	43 549 564	5,2 %
Mill. personkm	1 227	1 313	7,0 %	1 407	7,2 %

Figur 8: Beregnede virkninger av ny ruteplan 2013 (Dybdahl et al. 2011, p. 9)

Figuren viser trafikkprognoser for ulike trafikktypene på Østlandet hver for seg og en samlet vekst for NSB. Referanse 2013 er estimert antall reiser og personkm uten innføring av nytt ruteopplegg. Ny ruteplan 2013 er estimert trafikkmengde i 2013 ved innføring av ny ruteplan. Den prosentvise forskjellen mellom ny ruteplan 2013 og referanse 2013 gir estimert virkning av ny ruteplan. Prognosene estimerer at innføring

av ny ruteplan gir en vekst på 5,2 % i antall reiser og 7,2 % i personkm for total trafikk på Østlandet. Mest merkbart i trafikkberegningene er den betydelige økningen i tilbringertrafikk til og fra Gardermoen på hele 177,5 % i antall reiser. Dette er forklart som årsak av redusert reisetid og høyere opplevd avgangsfrekvens som følge av bedre fordelte avganger over timen. Tabell 9 viser estimerte virkninger av fullstendig implementert ruteplan i 2015 og følger samme oppsett som figur 8. Prognosene gir en totalvekst på 20,9 % i antall reiser og 17,6 % i personkm (Dybdahl et al. 2011).

Pr. år	Referanse 2008	Referanse 2015	Økning 2008-2015	Alternativ E 2015	Økning R2015
<b>InterCitymodellen for Østlandet</b>					
Antall reiser	9 622 689	11 011 923	14,4 %	12 388 581	12,5 %
- herav regiontog	6 404 613	7 515 249	17,3 %	8 469 622	12,7 %
- herav lokaltog	3 218 076	3 496 674	8,7 %	3 918 959	12,1 %
Mill. personkm	736	859	16,8 %	942	9,7 %
<b>Emma Fredrik</b>					
Antall reiser	28 925 335	31 276 128	8,1 %	37 074 608	18,5 %
Mill. personkm	502	543	8,1 %	611	12,7 %
<b>Tilbringertrafikk, Gardermoen</b>					
Antall reiser	458 358	515 838	12,5 %	2 287 549	343,5 %
Mill. personkm	27	28	6,8 %	129	354,4 %
<b>SUM NSB</b>					
Antall reiser	39 006 382	42 803 890	9,7 %	51 750 738	20,9 %
Mill. personkm	1 264	1 430	13,1 %	1 683	17,6 %

Figur 9: Beregnede virkninger av ny ruteplan 2015 (Dybdahl et al. 2011, p. 18)

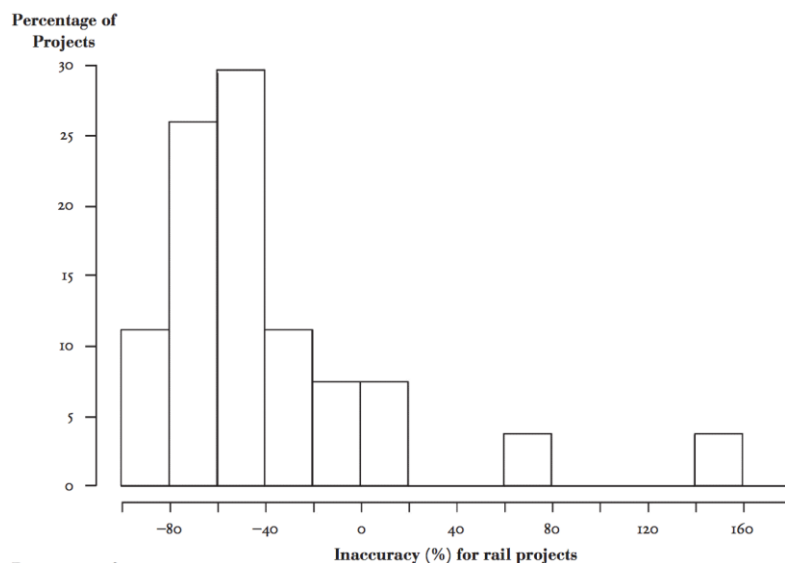
## 2.4 Tidligere ex post evalueringer av større infrastrukturprosjekter

Følgende delkapittel omhandler tidligere studier som tar for seg ex post evalueringer av større jernbaneprosjekter. Studiene stiller prosjektenes faktiske resultater opp mot deres trafikkprognoser. Som nevnt er trafikkprognoser avgjørende i estimeringen av nytte-kostnadsforholdet i infrastrukturprosjekter og har dermed påvirkning ved valg av prosjektalternativ. Kunnskap rundt prognosenes pålitelighet er viktig for å kunne gi realistiske estimater av prosjektenes faktiske nytte. Flyvbjerg et al. (2005) skriver følgende om trafikkberegninger knyttet opp mot større infrastrukturprosjekter.

“Despite the enormous sums of money spent on transportation infrastructure, surprisingly little systematic knowledge exists about the costs, benefits, and risks involved. The literature lacks statistically valid answers to the central and self-evident question of whether transportation infrastructure projects perform as forecasted.” (Flyvbjerg et al. 2005, p. 131)

Sitatet understreker at det eksisterer lite empirisk kunnskap rundt hvordan store prosjekter presterer i forhold til forventet. Tematikken vedrørende prognosers usikkerhet har likevel fått noe større oppmerksomhet de siste årene (Nicolaisen & Driscoll 2014). Flyvbjerg et al. (2005) er det første studiet med statistisk signifikans som har undersøkt treffsikkerheten til større infrastrukturprosjekters trafikkprognoser. Studiet er det mest omfattende innen temaet. 210 infrastrukturprosjekter fordelt over fjorten nasjoner er inkludert. Omfanget av investerte ressurser for de enkelte prosjekter er i stor grad avhengig av trafikkprognosene. Som eksempel nevnes prosjektet Skytrain fra Bangkok med en kostnadsramme på to milliarder dollar. I prosjektets etterkant ble det klart at trafikkprognosene var 250 % høyere enn faktisk trafikk. Skytrain ble derav svært overdimensjonert og medførte store finansielle tap for de involverte partene (Flyvbjerg et al. 2005).

Figur 10 gir en oversikt over prognosefeil for prosjektene brukt i studiet. Den store tyngden av prosjektene har en negativ unøyaktighet, hvilket betyr at trafikkprognosene var overestimerte.



Figur 10: Prognosefeil i infrastrukturprosjekter (Flyvbjerg et al. 2005, p. 134)

Studiet konkluderer med at trafikkprognoser for jernbaneprosjekter er misledende. Ni av ti jernbanerelaterte trafikkprognoser overestimeres med 106 %. Faktisk trafikk ble dermed estimert å gjennomsnittlig være 51,4 % lavere enn antatt i prognosene, noe som også synes i figur 10. Dette gir overestimerte nytte – kost forhold i prosjektene hvor økonomiske tap medfølger. Trafikkprognoser er heller ikke funnet å bli vesentlig mer

presise over perioden prosjektene i studiet har blitt utført, et tidsintervall på 30 år (Flyvbjerg et al. 2005).

Pickrell (1990) undersøkte også, på et tidligere tidspunkt, hvilke resultater større jernbaneprosjekter ga sammenliknet med deres ex ante evalueringer. Pickrell brukte ti prosjekter fra Amerika og sammenliknet trafikk og kostnadsprognosene opp mot faktiske resultater. Pickrell benyttet seg av prognoser gjort i forkant av endelig valgt prosjektalternativ. Dermed hadde prognosene stor påvirkningskraft for valg av prosjekt. Tidsintervallet for prosjektene strak seg over perioden 1971 – 1987. Pickrell (1990) sine funn viste at trafikkprognosene var overestimerte for alle prosjektene inkludert i analysen, med en gjennomsnittlig overestimering på 65 %. Prosjektet med best prognose var overestimert med 28 %. Det nevnes også problematikken for å vite om andre prosjekter ville blitt valgt til fordel for prosjektene som ble realisert, da det er umulig å finne prognosefeilene ved de urealiserte prosjektene (Pickrell 1990).

Studiet medførte opphetede diskusjoner vedrørende trafikkprognosers troverdighet. Button (2009) gjorde derfor et nytt studium som fulgte opp funnene til Pickrell. Med bakgrunn i 47 infrastrukturprosjekter i perioden 1972 – 2005 ble det undersøkt om den såkalte Pickrell effekten var like sterkt fremtredende utover 1990 og 2000 tallet. Button sine funn ga indikasjoner på tydelige forbedringer i prognosene vedrørende trafikketerspørsel, men kostnadsestimatene hadde ikke fremtredende tegn til forbedringer. Pickrell sitt studie satte på flere måter trafikkprognosers usikkerhet på kartet, noe som også kan ha initiert forbedringene observert i ettertid (Button 2009).

Givoni & Dobruszkes (2013) og Beria & Grimaldi (2011) gikk i dybden på ex post evalueringer av høyhastighetsjernbane, et tema som det tidligere har vært lite kjennskap til, tross stor satsing på høyhastighetstog verden rundt. Givoni & Dobruszkes (2013) bemerker viktigheten av ex post evalueringer av jernbaneprosjekter med bakgrunn i Flyvbjerg et al. (2005) sitt studie og trekker frem høyhastighetsprosjekter i Taiwan, Frankrike og England som eksempler med overestimerte trafikkprognoser. Som eksempel nevnes *Eurostar* sitt høyhastighetstilbud over den engelske kanal med en prognose for faktiske passasjerer på 25 millioner i 2006. Faktisk trafikk i 2011 lå rundt 9,7 millioner (Givoni & Dobruszkes 2013). Høyhastighetsjernbane danner et helt nytt infrastrukturtilbud hvor all trafikk er et resultat av enten overført trafikk fra andre eksisterende transporttilbud (som fly, veg, tidligere toglinjer) eller ny induisert trafikk.

Givoni & Dobruszkes (2013) forsøkte i sitt studie å skille hva som kan forventes av overført trafikk og induisert trafikk. Funnene ga grunner til å anta at rundt 20 % av etterspørselen for høyhastighetsjernbane to til fire år etter ferdigstillelse kan tolkes som induisert etterspørsel. Den resterende andelen på rundt 80 % tolkes dermed som overført trafikk. Det presiseres at det her er store usikkerheter i beregningene. Hva som er induisert og overført trafikk hevdes å være vesentlig for å kunne forsvare utbyggingsprosjekter i den størrelsen, da dette viser til om prosjektene faktisk gir økt trafikkmengde eller kun tolkes som et bedre alternativ til tidligere brukte transportmetoder (Givoni & Dobruszkes 2013).

Beria & Grimaldi (2011) utførte en tidlig ex post evaluering av høyhastighetsjernbane i Italia. Italiensk høyhastighetsjernbane ble besluttet å bygges i 1990 som et nytt jernbanenett isolert fra det eksisterende. Dette skulle sikre raske forbindelser mellom Italias større byer. Første del av høyhastighetsnettet ble fullstendig ferdigstilt i desember 2009 og hadde tilknytning mellom byene Turin og Salerno. Beria & Grimaldi (2011) studerte trafikkprognosene opp mot resultatene i årene etter åpning. Artikkelen nevner at fullstendig etterspørselseffekt ved endringer i togtilbudet først vil være realisert tidligst fem år i etterkant av implementering. Av den grunn inneholder ikke evalueringens datagrunnlag tilstrekkelig med etterspørselsdata. Studiet anbefaler en mer robust gjennomgang når forutsetningene er tilgjengelig. Likevel konkluderes det med at trafikkprognosene tilsynelatende var svært optimistiske med forventning om en sterk passasjeretterspørsel. Til tross for stor samfunnsnytte ved innføring av høyhastighetslinjene viser funnene at nytte- kostnadsforholdet for prosjektet var sterkt overvurdert i forkant (Beria & Grimaldi 2011).

I de siste tiårene har problematikken rundt prognoseusikkerhet fått langt mer oppmerksomhet enn tidligere. Nicolaisen & Driscoll (2014) gjorde derfor et litterært dypdykk i temaet og undersøkte de mest omfattende studiene rundt trafikkprognosers usikkerhet. Nicolaisen & Driscoll (2014) inkluderer de studier som ser på et mer omfattende omfang av jernbane og vegprosjekter, blant annet Flyvbjerg et al. (2005), Pickrell (1990) og Button (2009). Derfor har artikkelen utelukket mindre ex post analyser som undersøker kun enkelte prosjekter, eksempelvis Beria & Grimaldi (2011). Gjennom sin litteraturgjennomgang finner Nicolaisen & Driscoll (2014) klare signaler på at trafikkprognoser opp gjennom har vært problematiske for både vei- og jernbanetrafikk. I motsetning til jernbaneprosjekter sin tendens til overestimering av

fremtidig trafikk viser veiprosjekter noe av det motsatte. Det er nevneverdig at studiet viser indikasjoner på at nyere prognoser virker å ha lavere usikkerhet enn tidligere. Nyere prosjekter rapporterer til lavere standardavvik ved prognosefeilene. Dette er i samsvar med funnene til Button (2009). Selv med positive trender nevnes det at empiriske studier angående prognoseusikkerhet fortsatt er mangelvare. Derfor oppfordres det sterkt til tilsvarende forskning i fremover. Den positive trenden i prognoseusikkerhet kan til en viss grad begrunnes i det økende fokuset rundt temaet og et sterkt økende fokus på forbedring av kollektivtrafikk i mange land. Dette gir planleggere verdifull erfaring innen temaet og er sannsynlige årsaker til mer robuste prognoser (Nicolaisen & Driscoll 2014).

Årsakene til prognosefeilene er vanskelige å anslå presist, men Sanko et al. (2013) forsøkte å utarbeide mer kunnskap vedrørende dette under et case studie rettet mot *Tokadai Line*, en jernbane linje på 7,4 km i utkanten av Nagoya i Japan. Årsakene til prognosefeil varierer, men kan i hovedsak forklares av feil og usikkerheter i input dataen til trafikkberegningsmodellen, eller av modellen i seg selv. Sanko et al. (2013) gjorde funn som antyder at beregningsmodellene gir større utslag på prognosefeilene enn input dataen i til modellen. Selvom usikkerheten til input dataen skulle inneha en generelt større usikkerhet enn modellen er ikke det samsvarende med at denne usikkerheten gir større usikkerhet for selve trafikkberegningen og prognoseusikkerheten (Sanko et al. 2013). Med andre ord kan dette tolkes som at forutsetningene til beregningsmodellen er vel så viktig for utslaget i prognoseusikkerheten som dataen man putter inn i modellen for gjennomføring av beregninger.

Litteraturen presentert i delkapittelet viser en generell trend for at trafikkprognoser til større jernbaneprosjekter ofte er overestimerte og preget av stor usikkerhet. Overvurderinger av nytteforholdet for de enkelte prosjektene er følger av dette. Likevel viser enkelte studier til tegn på forbedringer i trafikkprognoser i nyere tid (Button 2009; Nicolaisen & Driscoll 2014). Med tanke på variasjonen i elastisitetstallene som er blitt funnet fra tidligere litteratur er det ikke unaturlig at det er stor variasjon i hvilke effekter som blir beregnet for ulike prosjekter.

### 3. Metodikk

Følgende kapittel ønsker å gi et overordnet innblikk over hva slags metodikk og overordnet strategi som er valgt for å angripe oppgavens problemstilling. Kapittelet forklarer hvorfor enkelte metoder er valgt og grunnleggende om den statistiske metoden som er anvendt for oppgavens analyser. Det blir kort forklart rundt valget av analyseverktøyet som er brukt – Python programmering.

#### 3.1. Kort om kvantitativ og kvalitativ forskningsmetode

Generelt deles all forskning inn i to hovedtyper, kvantitativ og kvalitativ forskningsmetode. Disse metodene avviker endel fra hverandre ved ulike kjennetegn. Et kjennetegn ved kvantitativ forskning er at det i stor grad benytter seg av datagrunnlag basert på tall (kvantitativt) mens kvalitativ forskning benytter datagrunnlag basert på ord, for eksempel intervjuer. Kvantitativ forskning tar oftere utgangspunkt i en hypotese som testes, mens kvalitativ forskning bruker å indusere hypoteser underveis i studiet (Silverman 2011). Tabell 2 nevner egenskaper som er hevdet å kjennetegne forskningsmetodene. Tabellen er oversatt og hentet fra Silverman (2011).

Tabell 2: Kjennetegn ved kvalitativ og kvantitativ forskning (Silverman 2011, p.5)

<b>Kvalitativ</b>	<b>Kvantitativ</b>
Myk	Hard
Fleksibel	Låst
Subjektiv	Objektiv
Case studier	Undersøkelser
Spekulativ	Hypotese testing

Valg av forskningsmetode avhenger av problemstillingen. For denne oppgaven er det nødvendig å bearbeide mengder av kvantitativ data. Tidligere forskning innen oppgavens tema er gjort gjennom kvantitative analyser. For oppgaven er det derfor naturlig å basere oppgavens analyser på kvantitativ metode. Silverman (2011) nevner fem hovedmetoder for å bruke kvantitativ forskningsmetode. En av disse er beskrevet som dataanalyse av større mengder tidligere innhentet data, noe som kjennetegner oppgaven.



## 3.2 Multivariat analyse

Rencher (2002) forklarer at multivariat analyse eller multivariabel analyse gir muligheten til å studere komplekse problemer hvor flere variabler som forårsaker endringer i en annen variabel. Vi skiller mellom variablene som forklarer en variabel og variabelen som blir forklart da de inndeles i to kategorier. Disse kategoriene er henholdsvis **forklaringsvariabler** og **responsvariabler**. Endringer i forklaringsvariablene vil forårsake endringer i responsvariabelen. Med andre ord beskriver forklaringsvariablene adferden til responsvariabelen. Multivariate analyser studerer den samlede effekten forklaringsvariablene gir en responsvariabel, og effekten til hver enkelt variabel på responsvariabelen når den er i samspill med andre forklaringsvariabler. Multivariat analyse er en kategori av flere statistiske metoder med formål om å løse problemstillinger med responsvariabler påvirket av forklaringsvariabler. Opprinnelig ble multivariat analyse mest anvendt innen biologisk forskning, men i senere tid er metodene anvendt innen flere vitenskapelige grener. Moderne datamaskiner har gjort multivariat analyse brukervennlig og gjennomførbart uten betydelige kostnader (Rencher 2002). Følgende delkapittel beskriver den multivariate metoden brukt i oppgaven.

### 3.2.1 Multippel lineær regresjon

“Multippel lineær regresjon hjelper oss å finne den beste lineære sammenhengen mellom responsen og de ulike forklaringsvariablene. Multippel regresjon er et av statistikkfagets mest nyttige – og mest brukervennlige verktøy“  
(Løvås 2004, p. 288)

Som Løvås (2004) beskriver er regresjonsanalyse et hendig verktøy for å angripe problemstillinger bestående av multivariat data. I henhold til Rencher (2002) kan regresjonsanalysen inndeles i tre hovedtyper utfra antall variabler som inngår i analysen:

1. **Enkel lineær regresjon:** En forklaringsvariabel og en responsvariabel.
2. **Multippel lineær regresjon:** Flere forklaringsvariabler med en responsvariabel, også kalt univariat multippel regresjon.
3. **Multivariat multippel lineær regresjon:** Flere forklaringsvariabler og flere responsvariabler, også kalt multivariat multippel regresjon.

Dette delkapittelet går inn på type nummer to, multippel lineær regresjon som antar at den virkelige sammenhengen mellom forklaringsvariablene og responsvariabelen kan beskrives som en lineær matematisk modell. Modellen kan uttrykkes som i likning 3.1 og avhenger av hvor mange forklaringsvariabler som inngår i regresjonen (Løvås 2004).

$$Y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n \quad 3.1$$

Y representerer responsvariabelen som er avhengig av flere forklaringsvariabler, x-variablene. Forklaringsvariablenes tilhørende koeffisienter,  $\beta$  verdiene, forklarer sammenhengen mellom hver enkelt forklaringsvariabel og responsvariabelen.  $\beta$ -verdiene er ukjente størrelser som blir estimert i regresjonsanalysen. Konstantleddet  $\alpha$  beskriver responsvariabelens verdi ved ingen endring i forklaringsvariablene.

Regresjonsanalysens formål er å estimere de ukjente koeffisientenes verdier. For å få til dette kreves innsamling av observasjoner som kan knyttes opp mot en antatt modell. Observasjonene som samles inn kan beskrives matematisk som i likning 3.2. (Løvås 2004)

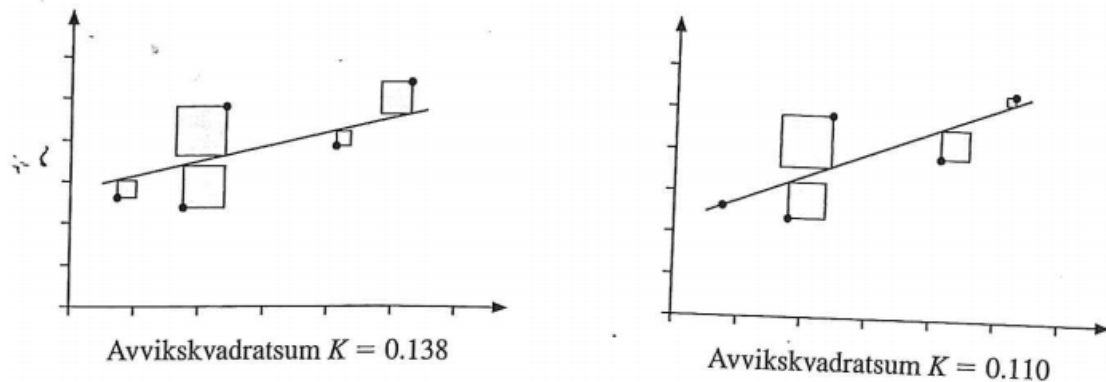
$$Y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \dots + \beta_n x_n + e_i \quad 3.2$$

Likning 3.2 beskriver hvordan en vilkårlig observasjon antas å være avhengig av ulike verdier til forklaringsvariablene. Det bakerste leddet  $e_i$  representerer et feilledd som forklarer avviket mellom den estimerte verdien av observasjonen og den faktiske observasjonen. Dette feilleddet kalles residualer. Etter innsamling av observasjoner er det mulig å anvende dataprogrammer for å gjennomføre regresjonsanalysen. Som resultat estimeres verdier for  $\beta$ -koeffisientene og konstantleddet  $\alpha$ , samt flere tall med indikasjoner på hvor god regresjonsmodellen er (Løvås 2004).

### **Minste kvadraters metode**

Multippel lineær regresjon estimerer en lineær modell som passer best til datasettet som leveres. Den beste modellen defineres utfra *minste kvadraters metode* (*Ordinary least squares*). Figur 11 viser plot av to regresjonslinjer som har bakgrunn i samme datasett, men linjene er ulike. Avstanden fra observasjonspunktet og ned til regresjonslinjen er observasjonens residual. Kvadratene ved hvert målepunkt viser

arealet av residualene kvadrert. Minstekvadraters metode estimerer koeffisientverdier som gir den totalt minste kvadratsummen. Kvadratsummen er arealsummen av kvadratene for alle observasjonene. (Løvås 2004)



Figur 11: Illustrasjon av minste kvadraters metode (Løvås 2004, p. 273)

I følge minste kvadraters metode er regresjonslinjen til høyre i figur 11 en bedre regresjonsmodell da avvikskvadratsummen er lavere enn i modellen til venstre. Dataobservasjonene som leveres inn i multipl linear regresjons kan fremstilles som ulike matriser. Datasettet som brukes i regresjonsanalysen kan derfor være lignende som oppsettet i figur 12.

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1q} \\ 1 & x_{21} & x_{22} & \cdots & x_{2q} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nq} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_q \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}.$$

Figur 12: Observasjoner i regresjonsanalysen på matriseform (Rencher 2002, p. 324)

Til venstre er observasjonene til responsvariabelen i en matrise. På andre side følger matrisen med observasjoner for hver forklaringsvariabel som er tilknyttet observasjonen i responsvariabelen. Denne matrisen multipliseres med en koeffisientmatrise bestående av de ukjente koeffisientene som blir estimert i analysen. Residualene til observasjonene er lagt ved i en egen matrise ytterst til høyre.

### Hvor god er den estimerte modellen?

Etter gjennomført multipl lineær regresjonsanalyse er det nødvendig å vurdere hvor godt modellen faktisk gjenspeiler virkeligheten. I tillegg til estimerer på forklaringsvariablenes koeffisienter gir regresjonens resultater ut tall som indikerer hvor god modellen er. Under er det vist et output av en multipl lineær regresjonsmodell gjort i Excel.

<i>Regresjonsstatistikk</i>					
Multipl R	0,9682				
R-kvadrat	0,9373				
Justert R-kvadrat	0,7493				
Standardavvik s	0,1335				
Observasjoner	5				

<i>Variansanalyse</i>					
	<i>fg</i>	<i>SK</i>	<i>GK</i>	<i>F</i>	<i>Signifikans-F</i>
Regresjon	3	0,2664	0,0888	4,9853	0,3154
Residualer	1	0,0178	0,0178		
Totalt	4	0,2842			

	<i>Koeffisienter</i>	<i>Standardfeil</i>	<i>t-Stat</i>	<i>P-verdi</i>
Skjæringspunkt	-2,0010	1,3808	-1,4491	0,3845
Motor	0,0093	0,0025	3,6594	0,1698
Vekt	0,0019	0,0011	1,6924	0,3398
Alder	-0,0061	0,0067	-0,9011	0,5331

Figur 13: Eksempel på output (Løvås 2004, p.289)

Fra toppen av figur 13 er første tall verdt å merke seg *R – kvadrat*.

R- kvadrat forklarer hvor mye av den observerte variasjonen til responsvariabelen som forklares av den estimerte modellen. En lav verdi for R- kvadrat indikerer at regresjonsmodellen beskriver lite av de endringene som er observert i responsvariabelen. Dette er ofte grunnet at ikke alle forklaringsvariabler som påvirker responsvariabelen er inkludert. Fjerning av forklaringsvariabler med liten betydning for responsvariabelen vil derfor ikke gi store utslag på R-kvadrat (Løvås 2004). I dette eksempelet er R – kvadrat 0,9373. Det betyr nesten 94 % av endringer observert i responsvariabelen forklares av regresjonsmodellen. Nederste delen av figuren viser estimerer på forklaringsvariablenes koeffisienter, deres tilhørende standardavvik, t-verdier og p- verdier. T- og p- verdiene henger tett sammen og viser sannsynligheten for at endringer i forklaringsvariabelen gir utslag på responsvariabelen. Kort forklart er det to typer hypoteser vedrørende forklaringsvariablene. Null hypotesen ( $H_0$ ) er hypotesen regresjonsmodellen tar som utgangspunkt og antar at det ikke eksisterer en

sammenheng mellom forklarings- og responsvariabelen. Den andre hypotesen ( $H_1$ ) antar at det finnes en sammenheng. Dersom  $H_0$  gjelder vil koeffisienten mellom den aktuelle forklarings- og responsvariabelen være lik null, mens dersom  $H_1$  gjelder vil koeffisienten ikke være lik null. Hypotesene kan beskrives som i likning 3.3 og 3.4 (Løvås 2004).

$$H_0: \beta = 0, \text{ingen sammenheng} \quad 3.3$$

$$H_1: \beta \neq 0, \text{det finnes en sammenheng} \quad 3.4$$

T verdien tilknyttet hver enkelt koeffisient regnes ut ved å dividere estimert koeffisient på tilhørende standardavviket.

$$T = \frac{\hat{\beta}}{SE(\hat{\beta})} \quad 3.5$$

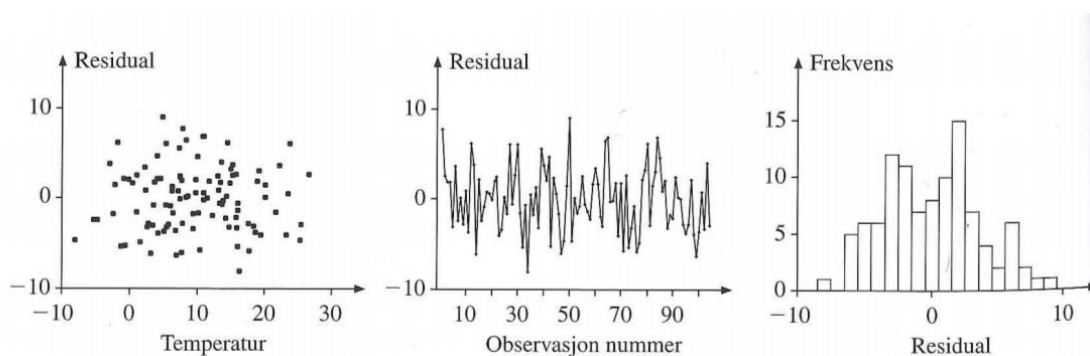
P verdien er basert på T verdien og forklarer sannsynligheten for å observere gitt T-verdi under forutsetning at  $H_0$  er sann. Høye T- verdier genererer lave p-verdier (Løvås 2004). Dersom vi tar utgangspunkt i eksempelet i figur 13 og tallene tilhørende forklaringsvariabelen *motor* ser vi at T - verdien er tilnærmet 3,66 med en tilhørende p - verdi på  $\approx 0,17$ . Det er dermed 17 prosent sannsynlig å observere denne verdien dersom  $H_0$  gjelder. Med andre ord er det 83 % sannsynlig at  $H_1$  gjelder, som likning 3.6 viser.

$$\text{Sannsynlighet for } H_1 = 1 - P_{\text{verdi}} = 1 - 0,17 = 83 \% \quad 3.6$$

En forklaringsvariabel kalles signifikant dersom det er estimert en tilstrekkelig lav P-verdi for variabelen. Det er vanlig å bruke signifikansnivåer på 5 % eller 10 %, noe som betyr at p verdien må ligge under 0,05 eller 0,10 for at man skal konkludere med at det er sammenheng mellom forklaringsvariabelen og responsvariabelen. P - verdiene for koeffisientene og R-kvadrat er gode mål på hvor godt den lineære regresjonsmodellen beskriver den faktiske situasjonen (Løvås 2004).

## Undersøkelse av modellens forutsetninger

Ved en regresjonsanalyse antas det at modellens residualer har tre ulike egenskaper. Residualenes varians må være konstant og uavhengig av forklaringsvariablenes verdi. Med andre ord er ikke residualene *heteroskedastiske*. Videre antas det at residualene er uavhengige av hverandre - de ikke er *autokorrelerte*. Tilslutt antas det at residualene er tilnærmet normalfordelte med en forventet verdi på null (Løvås 2004; Brooks 2008). For å kontrollere at forutsetningene er tilstede er det hensiktsmessig å produsere residualplot. Dersom forutsetningene til residualene ikke stemmer kan koeffisientenes standardavvik og P verdier bli feilaktige, men de estimerte koeffisientverdiene vil forbli like. Dette kan igjen påvirke hypotesetestingen vedrørende forklaringsvariablenes signifikans. Forutsetningen som er minst alvorlig er at residualene er normalfordelte. Det er tilstrekkelig at residualplotet er tilnærmet normalfordelt. Dersom residualenes størrelse er avhengig av størrelsen på forventet verdi i responsvariabelen vil dette kunne sees ved hjelp av et spredningsplot. For å få et grafisk bilde på om residualene er uavhengige av hverandre lages det plot som viser residualenes verdi i rekkefølgen de er registrert. Tilfeldige hopp i residualenes verdier indikerer at det ikke er noen sammenheng. Residualenes uavhengighet av hverandre er den viktigste forutsetningen for regresjonsmodellen (Løvås 2004). Figur 14 under viser residualplot som kan brukes til undersøkelse.



Figur 14: Eksempel på residualplot (Løvås 2004, p.282)

Plottet til venstre viser et spredningsplot som undersøker om residualene er uavhengige av forklaringsvariablene. Neste plot sjekker residualenes uavhengighet av hverandre - residualene synes å hoppe vilkårlig. Siste plot viser en tilnærmet normalfordeling av residualenes verdier. Forutsetningene til residualene ser ut til å stemme i dette tilfellet (Løvås 2004). En annen test for å undersøke residualenes uavhengighet er en Durbin

Watson (DW) – test (Brooks 2008). Ved gjennomføring av en slik test estimeres et DW tall. Sammenlikning av DW - tallet med en øvre og nedre kritisk verdi brukes for å kartlegge residualenes uavhengighet. De kritiske verdiene hentes fra en DW – tabell og avhenger av analysens observasjoner og antall forklaringsvariabler. Dersom DW- tallet er lavere enn nedre kritisk verdi kan man anta at residualene er autokorrelerte (avhengige). Om DW – tallet er høyere enn øvre kritisk verdi kan man se bort fra autokorrelasjon. En verdi mellom øvre og nedre grense gjør testen usikker (Brooks 2008). Denne analysen inkluderer 90 observasjoner og fire forklaringsvariabler (se kapittel 4.1.1). De nedre og øvre grensene for DW-tallet er for denne analysen er henholdsvis 1,43 og 1,61 ved et signifikansnivå på 99 % (Brooks 2008). Dersom residualene er autokorrelerte kan man med andre ord ikke stole på koeffisientens standardavvik og p – verdier (signifikanstester). Derfor trengs nye modifiserte standardavvik som tar høyde for autokorrelasjon. Dette kan gjøres gjennom en Newey – West test som estimerer HAC (Heteroschedasticity and autocorrelation consistent) standardavvik. De modifiserte standardavvikene tar høyde for at residualene er autokorrelerte og i tillegg eventuelle problemer dersom de er heteroskedastiske (Brooks 2008).

### **Multikorrelasjon**

Multikorrelasjon kan forekomme i regresjonsanalyser. Det er korrelasjon mellom de ulike forklaringsvariablene. En perfekt lineær sammenheng mellom to forklaringsvariabler vil gi en korrelasjon lik 1. Multikorrelasjon gir ingen brudd på regresjonsanalysens forutsetninger, men kan gi vanskeligheter for å skille effekten hver enkelt forklaringsvariabel har på responsvariabelen. Ved sterk korrelasjon mellom to forklaringsvariabler kan det være vanskelig å skille hvilken variabel som har en sterkest effekt på responsvariabelen. Typiske trekk ved multikorrelasjon er at forklaringsvariablenes T verdier er lave og ikke signifikante samtidig som modellens R – kvadrat er høy. Ulogiske fortegn på forklaringsvariablene er også vanlig. Dersom modellen har multikorrelasjon ved seg kan en løsning være å overse problemet dersom forklaringsvariablene er signifikante. En annen mulighet er å fjerne en av de korrelerte variablene dersom dette ikke er kritisk for hva man er ute etter i analysen (Brooks 2008).

### 3.3. Oppgavens strategi

Passasjerutviklingen de foregående årene kan knyttes opp mot flere variabler.

Forbedret rutetilbud for store deler av Østlandet er kun en av flere årsaker til den totale passasjerveksten. Andre kjente variabler som trolig har hatt en positiv innvirkning på passasjerveksten er:

- Nytt sonesystem med Ruter (medført prisendringer, se vedlegg 2)
- Befolkningsvekst

Også flere variabler enn disse påvirker reisealternativet folk velger. En oversikt over variabler som er utelatt fra analysen og hvorfor er beskrevet i delkapittel 4.1.2.

Variablene nevnt over er de som best kan knyttes opp og kvantifiseres til ulike lokale områder, sammen med variablene knyttet til forbedret rutetilbud. Oppgavens formål er å estimere den totale effekten ruteendringen har hatt for passasjerutviklingen. For å få til det er det nødvendig å isolere effekten av ruteendringen fra andre variabler som har påvirket passasjerveksten. Derfor har det vært hensiktsmessig å gjøre en geografisk avgrensning over områder som har opplevd større endringer i togtilbudet. Dersom en kvantitativ analyse skal gjennomføres er det med andre ord nødvendig å kunne kvantifisere togtilbudet og andre variabler som antas å påvirke passasjerutviklingen opp mot ulike områder. Med bakgrunn i dette er Østlandet systematisk inndelt i 18 områder. En mer detaljert beskrivelse av områdene fremkommer i delkapittel 4.2. Et område er en kombinasjon av enkelte stasjoner og toglinjer. Hvert område er påvirket av ulike variabler, men stasjonene i et område har opplevd en lik endring i disse. Veksten av reisende i ett område antas derfor å være en effekt av endringer i variablene som inntreffer. Datagrunnlaget for analysen er basert på årene 2010 – 2015. Endringer i de ulike variablene og passasjertallene er derfor kvantifisert og sortert ut til hvert område for hvert år. Med bakgrunn i dette er det gjort regresjonsanalyser som har lagt grunnlaget for å kunne estimere ruteendringens effekt. Mer detaljerte beskrivelser av analysearbeidet er beskrevet i kapittel 4.

### 3.4 Analyseverktøy – Python programmering

Denne oppgaven har krevd mye databearbeidelse. En stor tidsperiode av arbeidsperioden har blitt brukt til dette. For å automatisere ulike prosesser har det derfor vært formålstjenlig og tidsbesparende å utvikle datascript for å sortere og analysere data. Valgt programmeringsspråk er Python. Valget er gjort med bakgrunn



noe personlig erfaring med Python fra tidligere databehandlingskurs på NMBU, samt gode muligheter for veiledning rundt programmeringen i NSB. Python er opplevd å ha et brukervennlig grensesnitt og er anvendbart til flere formål. Det har vært praktisk å kunne benytte samme program til både databehandling, gjennomføring av analyser, beregninger og plotting av grafer. For utførelse av regresjonsanalysen i Python er det nedlastet og anvendt en statistisk applikasjon kalt statmodels. Mer detaljert rundt funksjonene til datascriptene og beregningene blir beskrevet i kapittel 4. Datascriptene er også vedlagt som vedlegg 4 og 5.

## 4. Dataanalyse

Følgende kapittel ønsker å gi leseren en systematisk og ryddig oversikt over progresjonen i det analytiske arbeidet. Begrunnelser for valg som er gjort og antakelser i analysen blir forklart. Først begrunnes hvorfor det er valgt å anvende multippel lineær regresjon og hvilke variabler som er inkludert i modellen. Utelatte variabler blir også beskrevet. Videre forklares områdeinndelingen som er gjort. Kapittelet går så trinnvis gjennom alt fra datainnhenting til beregning av endelige resultater. Kapittelets siste del forklarer hvordan usikkerheten ved datagrunnlaget er forsøkt håndtert.

### 4.1 Valg av statistisk modell

Nødvendig datagrunnlag for å gjennomføre analyser av ruteplanendringen kan sees på som et multivariat datasett bestående av flere forklaringsvariabler og en responsvariabel. Innhentet passasjerdata legger grunnlaget for observasjoner knyttet til responsvariabelen - passasjerutviklingen. En kvantifisering av utviklingen i togtilbudet og andre eksterne faktorer definerer forklaringsvariablene og deres verdier.

Jernbaneverket (2015) har definert sammenhengen mellom passasjervekst og togtilbud ved bruk av generaliserte kostnader og standard elastisitetsverdier. Matematisk kan dette presenteres som følgende i likning 4.1

$$\Delta E = \varepsilon x \Delta GK \quad 4.1$$

Her er  $\Delta E$  endring i passasjerretterspørsel,  $\varepsilon$  en standard elastisitet og  $\Delta GK$  endring i generaliserte kostnader. Trafikkprognosene til Dybdahl et al. (2011) i forkant gjort av ruteendringen er også basert på jernbaneverkets metodehåndbok ved bruk av generaliserte kostnader. Likning 4.1 viser et lineært forhold mellom etterspørsel og tilbud da elastisiteten er antatt å være konstant. I tillegg er det verdt å nevne at tidligere studier som FitzRoy & Smith (1995) har utført sine analyser ved bruk av regresjonsanalyse og minstekvadraters metode (OLS) for å kartlegge elastisitetsforhold i jernbanemarkedet.

Med bakgrunn i et multivariat datasett og et antatt lineært forhold mellom tilbud og etterspørsel er det valgt å benytte multippel lineær regresjonsanalyse også for denne oppgaven.

#### 4.1.1 Presentasjon av variablene i modellen

Tabell 3 gir en oversikt over de ulike variablene som er inkludert i analysen og på hvilket format deres verdier er presentert. Mer detaljert rundt hvordan datagrunnlaget er innhentet forklares senere i kapittelet.

Tabell 3: Modellens variabler

Navn (forkortelse)	Variabeltype	Format
Passasjervekst (PV)	Responsvariabel	Relativ passasjervekst siden 2010.
Avgangsfrekvens (AF)	Forklaringsvariabel	Relativ økning i togavganger fra 2010.
Reisetid (RT)	Forklaringsvariabel	Absolutt endring i kjøretid siden 2010 (minutter).
Pris (PR)	Forklaringsvariabel	Absolutt endring i pris (Grunnet nytt sonesystem)
Befolkningsvekst (BV)	Forklaringsvariabel	Relativ befolkningsvekst siden 2010.

En forbedring av togtilbudet skjer i hovedsak gjennom økt avgangsfrekvens og redusert kjøretid. Variablene passasjervekst, avgangsfrekvens og befolkningsvekst er representert som relativ utvikling siden 2010. For reisetidsendringer og pris er det valgt å benytte absolutte tall. Dette er gjort for å unngå at for eksempel små tidsbesparelser på 1 min for reiser som er korte i utgangspunktet skal få samme betydning som større tidsbesparelser på opp mot 10 min på lengre reiser. Samme prinsipp gjelder også for hvorfor pris er valgt å presenteres absolutt. Strekninger som har opplevd store reisetidsbesparelser (eller prisenbesparelser) er tenkt å kunne forklare passasjerutviklingen grunnet reisetidsreduksjoner bedre enn svært små endringer på korte reiser hvor dette knapt er merkbart. Ved bruk av relative tall ville disse to

ytterpunktene få lik påvirkning for betydningen av reisetid. Derfor er dette unngått, og absolutte tall er benyttet.

For å kunne gjennomføre en regresjonsanalyse er det nødvendig med observasjoner som kan tilknyttes den antatte modellen. For denne analysen vil en observasjon være observert passasjerutvikling i et område et år. Denne utviklingen antas å være forårsaket av endringer i forklaringsvariablene. Observasjonene forsøkes å beskrives av regresjonsanalysen som følgende (forkortelser på variablene brukes):

$$PV_{i,y} = \beta_1 AF_{i,y} + \beta_2 RT_{i,y} + \beta_3 PR_{i,y} + \beta_4 BV_{i,y} + \varepsilon_{i,y} \quad 4.2$$

Her er  $PV_{i,y}$  passasjerutviklingen for området  $i$ , året  $y$  siden 2010. Videre viser  $\beta$ -koeffisientene estimert sammenheng mellom de ulike forklaringsvariablene og responsvariabelen.  $AF_{i,y}$ ,  $RT_{i,y}$ ,  $PR_{i,y}$ ,  $BV_{i,y}$  er verdiene for forklaringsvariablenes endringer siden 2010 for området  $i$  og året  $y$ , likt som for passasjerutviklingen.  $\varepsilon_{i,y}$  er avviket mellom estimert og faktisk verdi, altså observasjonens residual.

Totalt inngår det 90 observasjoner i analysen. Østlandet er delt opp i 18 områder og det er 5 år med observasjoner etter referanseåret 2010. Observasjonene i 2010 vil kun ha verdier lik null og benyttes ikke da disse observasjonene ikke påvirker resultatene i noen grad.

#### 4.1.2 Utelatte variabler

Som i de fleste regresjonsanalyser utelukkes enkelte variabler fra analysen til tross for at de antakeligvis forårsaker endringer i responsvariabelen. Hvorfor dette gjøres vil avhenge. Generelt sett er det avgjørende at forklaringsvariablene kan knyttes opp mot hver enkelt observasjon. Hvis ikke kan variablene medføre støy i modellen i stedet for å forklare den. Det er mest ryddig å bruke få variabler som er klart knyttet opp mot hver enkelt observasjon for å unngå en unødvendig komplisert modell. Dette er spesielt nødvendig for dette tilfellet hvor usikkerheten i tallene knyttet til responsvariabelen også inneholder usikkerhet (se 4.3 og 4.5). En konsekvens av å utelate enkelte variabler er at effektene (koeffisientene) til variablene som inngår også vil inkludere effekten av de utelatte variabler som har en tilnærmet lik utvikling (er korrelerte) som de inkluderte variablene. Dette delkapittelet nevner variablene som antas å ha påvirkning på passasjerutviklingen, men likevel er utelatt fra analysen.

## **Nye tog**

Nytt moderne togmaterieell fra Stadler AS har trolig gitt reisende en bedre togopplevelse og bidratt til muligheten for økt setekapasitet for de enkelte avgangene som nevnt i kapittel 1.2. Men innføringen av togmateriellet har skjedd i takt med ruteendringen og er sterkt korrelert med de toglinjer som har fått et økt togtilbud, da togene har vært en forutsetning for tilbudsøkningen (NSB 2011). Dette gjør det vanskelig for analysen å skille mellom effekten av eksempelvis redusert kjøretid og nye tog. Av den grunn er det valgt å utelate innføring av nye tog som en egen forklaringsvariabel. Effekten nye tog har gitt er det dermed rimelig å anta at blir plukket opp gjennom de andre forklaringsvariablene knyttet til togtilbudet da dette henger såpass tett sammen.

## **Bruttonasjonalprodukt, sysselsetting og inntekt**

BNP, sysselsetting og inntekt er faktorer som blant annet Fearnley et al. (2012), FitzRoy & Smith (1995) og Hamre (2002) nevner å ha påvirkning for passasjerutviklingen. Likevel er dette utelatt fra analysen. Årsaken er vanskeligheten for å skalere disse variablene ned mot ønsket detaljnivå til hvert område. Det er dermed vanskelig å knytte variablene opp mot observasjonene som inngår i analysen. Likevel er det naturlig å anta at en sterk vekst i disse variablene er med på å trekke den totale passasjerveksten opp.

## **Punktlighet**

Dårlig punktighet er en belastning for reisende, til tross for at Fearnley et al. (2012) nevner at man kan forvente små endringer i etterspørselen ved punktlighetsendringer. Tabell 4 viser årlig total endestasjonspunktighet for alle linjer på Østlandet mellom 2010 til 2015. Linjene som er markert med kursiv er linjene som ble erstatt av en ny linje fra 2013. Den historiske punktighetsdataen er innhentet fra NSBs database. Tabellen ønsker å gi et bilde av den generelle utviklingen i punktighet på Østlandets linjer. Den totale punktigheten for Østlandet har hatt en svak positiv utvikling de siste seks årene. For hver enkelt linje finnes det mer variasjon i den årlige punktigheten, men tabellen viser til generelt jevne punktighetsnivåer over 80 %. Det største unntaket er linjen Drammen – Lillehammer (R10). Punktighet er valgt å utelates fra analysen da forsinkelser slår inn såpass sporadisk og tilfeldig for aktuelle linjer. Det er derfor vanskelig å knytte punktighet opp mot enkelte områder, spesielt områdene bestående av flere linjer (se delkapittel 4.2). Grunnet relativt stabile punktighetsnivåer for de

fleste linjer antas det i analysen at punktlighetsutviklingen ikke påvirker passasjertallene i de enkelte områdene nevneverdig.

Tabell 4: Punktlighetsoversikt over Østlandet

<b>Linje</b>	<b>2010</b>	<b>2011</b>	<b>2012</b>	<b>2013</b>	<b>2014</b>	<b>2015</b>
Asker – Kongsvinger <i>Skøyen – Kongsvinger</i>	83,8%	88,9 %	90,8 %	81,6%	85,7%	86,6%
Drammen – Dal <i>Skøyen – Dal</i>	80,7 %	86,8 %	89,4 %	84,5 %	84,7%	89%
Drammen – Lillehammer <i>Oslo - Lillehammer</i>	64,9 %	65,6 %	81,0 %	74,2%	84,1%	84,4%
Eidsvoll – Skien <i>Oslo – Skien</i>	74,5%	82,6%	79,4 %	86,3%	81,7%	84%
Oslo – Halden	86,3%	86,2%	90,1%	88,0%	83,0%	86,7%
Kongsberg – Eidsvoll	75,9 %	82,0 %	84,1 %	81,0%	88,9%	87,4%
Skøyen – Mysen	87,1%	83,9%	91,3 %	91,9%	88,5%	91,7%
Spikkestad – Lillestrøm <i>Drammen – Lillestrøm</i>	77,0 %	84,8 %	85,4 %	93,9%	94,5%	93,7%
Skøyen – Moss <i>Spikkestad – Moss</i>	79,5 %	83,6 %	89,0%	88%	86,8%	85,3%
Skøyen – Ski	88,5%	87,9%	90,4 %	90,7%	89,1%	88,5%
<b>Total punktlighet NSB</b>	<b>80,6 %</b>	<b>84,1 %</b>	<b>87,4 %</b>	<b>87,2 %</b>	<b>87,6%</b>	<b>88,7 %</b>

### **Konkurransen forhold mot andre transportalternativer**

Hvilket transportalternativ folk velger avhenger av de ulike transporttilbudenes attraktivitet relativt til hverandre. Hvor mange som velger å reise med tog påvirkes av hvor attraktivt togtilbudet er sammenliknet med andre alternativer. Bil og buss er trolig togets største konkurrenter da begge alternativer egner seg for arbeids og fritidsreiser på Østlandet. Flytrafikk kan også anses som en konkurrent for togtrafikken ved spesielt lengre reiser, men i denne oppgaven som undersøker trafikken på Østlandet er det av mindre interesse. Fearnley et al. (2012) og Wijeweera & Charles (2013) nevner elastisiteter for drivstoffkostnader, pris og reisetid for bil og buss. Dette indikerer en

sammenheng ved at endringer i buss og biltilbudet påvirker tog etterspørselen. Togets relative attraktivitetsutvikling sammenliknet med andre transportalternativer er utelatt fra analysen. Bensinpriser vil være en variabel som har utviklet seg mer eller mindre likt over hele Østlandet og kan derfor ikke detaljeres ned mot hvert område. Bomstasjoner og reisetidsendringer kunne i større grad blitt knyttet opp mot spesifikke områder. Grunnet oppgavens tidsbegrensing er dette likevel utelatt. I tillegg er det verdt å nevne at til hvilken grad ulike linjer er belastet med buss for tog som følge av arbeid på jernbanen heller ikke er inkludert.

## 4.2 Områdeinndeling

Følgende delkapittel viser Østlandets områder. Inndelingen er gjort med hensikt om å dele ulike stasjoner i områder utfra hvilke faktorer som er antatt å drive passasjerutviklingen. Inndelingen er gjort selv, men med god veiledning fra NSB om hva som er en hensiktsmessig inndeling med tanke på stasjoner og tilbudsendringer. Tabell 5 viser en oversikt over områdene. Stasjonene, toglinjene og forklaringsvariablene som har endret seg i større grad inngår i tabellen. Hvert område er tildelt en hovedstasjon (uthevet). Hovedstasjonen fungerer som et referansepunkt for hele området for å beregne reisetidsendringer mellom området (hovedstasjonen) og Oslo S, samt prisendringer ved nytt sonesystem av Ruter. Reisetidsendringene og prisendringene er så å si tilsvarende for de andre stasjonene i området.

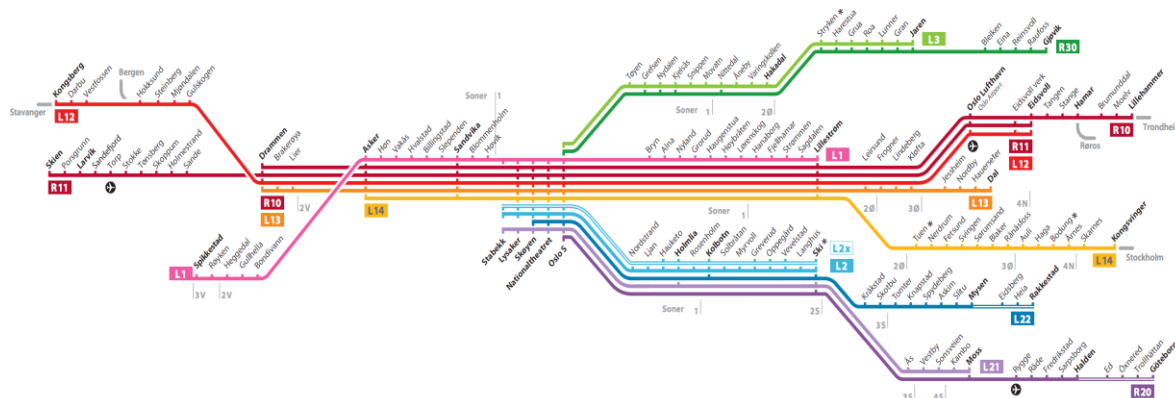
Tabell 5: Områdeoversikt

Nr	Navn	Variabler	Stasjoner	Linjer
1	Asker – Oslo (lokaltog)	PR, AF, RT, BV	Asker, Høn, Vakås, Hvalstad, <b>Billingsstad</b> , Slependsen	L1
2	Asker – Oslo (knutepunktstog)	PR, BV, AF, RT	<b>Asker</b> , Sandvika	R10, R11, L12, L13, L14
3	Lillestrøm –Oslo (lokaltog)	PR, BV, AF	Bryn, Alna, Nyland, Grorud, <b>Haugenstua</b> , Høybråten, Lørenskog, Hanaborg, Fjellhamar, Strømmen, Sagdalen, Lillestrøm	L1
4	Lillestrøm – Oslo (knutepunktstog)	PR, BV, F	<b>Lillestrøm</b>	R10/11,L1 2, L13, L14
5	Oslo S – Ski (lokaltog)	PR, BV	Nordstrand, Ljan, Hauketo, <b>Holmlia</b> , Rosenholm, Kolbotn, Solbråtan, Myrvoll, Greverud, Oppegård, Vevelstad, Langhus	L2
6	Ski – Sonsveien	PR, BV	<b>Ås</b> , Vestby, Kambo, Sonsveien	L21
7	Drammen -	RT, BV	<b>Kongsberg</b> , Darbu,	L12



	Kongsberg		Vestfossen, Mjøndalen, Hokksund, Gulskogen	
8	Skarnes - Kongsvinger	BV, AF	<b>Kongsvinger</b> , Skarnes	L14
9	Asker - Spikkestad	BV, PR, AF	<b>Spikkestad</b> , Røyken, Hallenskog, Heggedal, Gullhella, Bondivann	L1
10	Lillestrøm - Eidsvoll	BV, PR, AF	<b>Eidsvoll</b> , Eidsvoll verk, Gardermoen	R10, R11, L12
11	Lillestrøm - Årnes	BV, PR, RT	<b>Årnes</b> , Bodung, Haga, Auli, Rånåsfoss, Blaker, Sørumsand , Svingen, Fetsund, Nerdrum, Tuen	L14
12	Moss - Halden	BV	Moss, Rygge, Råde, <b>Fredrikstad</b> , Sarpsborg, Halden	R20
13	Drammen-Asker	BV, RT, AF	<b>Drammen</b> , Lier, Brakerøya	R10, R11, L13, L12
14	Lillestrøm - Dal	BV, PR, RT, AF	Leirsund, Lindeberg, Jessheim, Nordby, Hauer seter, <b>Dal</b>	L13
15	Ski (knutepunktstog)	BV, PR	<b>Ski</b>	L21, L22, R20
16	Drammen - Skien	BV, RT	Skien, Porsgrunn, Larvik, Sandefjord, Torp, Stokke, <b>Tønsberg</b> , Skoppum, Holmestrand, Sande	R10
17	Ski - Mysen	BV	Kråkstad, Skotbu, Tomter, Knapstad, Spydeberg, <b>Askim</b> , Slitu, Mysen	L22
18	Tangen - Lillehammer	BV, RT	Lillehammer, Moelv, Brumunddal, <b>Hamar</b> , Stange, Tangen	R10

Områdene mellom Asker og Lillestrøm (område 1-4) har i størst grad opplevd en tilbudsøkning, men også andre områder har opplevd tilbudsendringer i form av kjøretidsbesparelser og økt avgangsfrekvens. Det antas at hovedandelen av trafikken går til og fra områdene utenfor Oslo mot de store stasjonene i Oslo (Lysaker, Skøyen, Nationaltheatret, Oslo S). Disse stasjonene er ikke tatt med i områdene for å unngå dobbelttelling av reisende (forklares nærmere i delkapittel 4.3). Noen områder er splittet opp i lokal og knutepunktstoppende trafikk for å skille mellom tilbudsendringer som har skjedd på forskjellige linjer. Figur 15 viser linjekartet for Østlandet.



Figur 15: Linjekart på Østlandet (NSB 2016)

### 4.3 Datainnhenting og datasystematisering

Følgende delkapittelet beskriver hvilken data som er brukt i regresjonsanalysen. Delkapittelet beskriver nærmere hvordan verdiene til forklaringsvariablene i de ulike områdene er bestemt og hvordan dataen er bearbeidet. All data er generelt blitt hentet inn internt i NSB.



Figur 16: Dataflyt i arbeidet

#### 4.3.1 Passasjertellinger og avgangsfrekvens

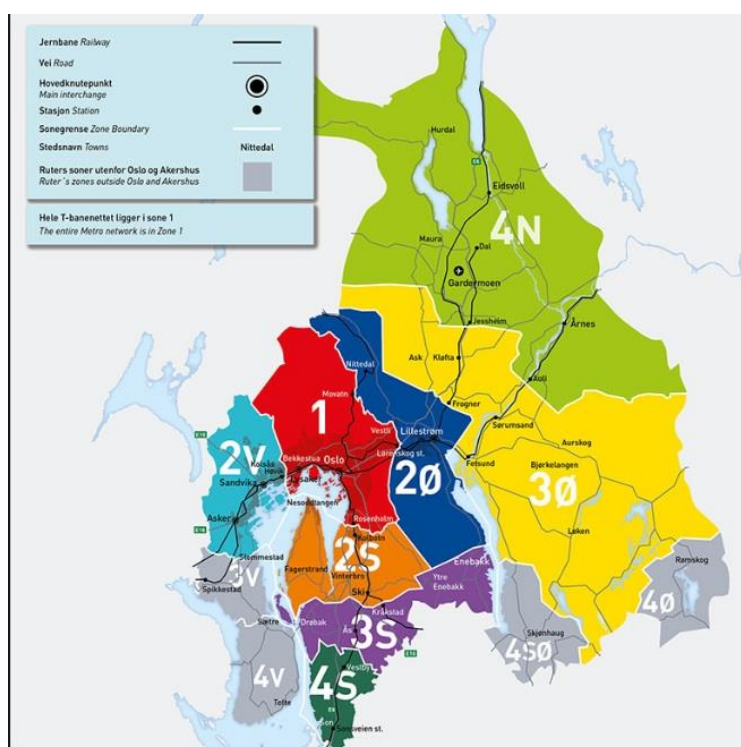
NSB gjennomfører en omfattende passasjertelling en hverdag (tirsdag) hver høst. Passasjertellingene måler avstigende og påstigende på alle stasjoner for alle tog på Østlandet. Med andre ord kartlegges all trafikk på Østlandet en hel dag. Rådataen til passasjertellingene viser avstigende og påstigende på alle stasjoner som gjelder for et tognummer. Et tognummer representerer et faktisk kjørt tog som går fra stasjon A til stasjon B. Tognumrene er kategorisert inn i tognummer serier. Dette betyr at for eksempel tognummerserien 7100 – 7300 består av tog som går på linje L1 mellom Spikkestad og Lillestrøm. Et tognummer gjelder for hele linjen og vil derfor kunne bevege seg gjennom flere av områdene som er definert i denne oppgaven. Tognumrene er mer eller mindre like for hvert år med unntak av små avvik eller innføring av nye tognummerserier. Antall tog som har kjørt inn og ut av et område vil tilsvare antallet tognumre som faktisk er telt i det representative området. En økning i antall telte tognummer i et område er dermed et godt mål på utviklingen i avgangsfrekvensen. Passasjertellingsdataen er derfor benyttet til å beskrive både passasjer- og frekvensutviklingen i de ulike områdene. Som nevnt er det unngått å bruke tellinger for de største stasjonene i Oslo. Dette er gjort for å unngå dobbelttelling av passasjerer. Det antas, etter anbefaling fra fagpersoner i NSB, at majoriteten av trafikkstrømmene foregår til/fra områdene rundt Oslo og inn mot Oslo. Ved å telle en påstigende eller avstigende som en reisende unngås det at samme passasjer telles på nytt som avstigende eller påstigende i Oslo. Passasjertellingene er gjennomført manuelt av konduktører. Siden flere går av eller på ved knutepunktstasjonene i Oslo er det også vanskeligst for konduktør(ene) å telle her. Derav er det høyere telleusikkerhet ved disse stasjonene. Ved å unngå å benytte tellingene fra disse stasjonene vil man også unngå de mest usikre passasjertellingene. Det benyttes heller tellinger for stasjonene med færre avstigende og påstigende per stasjon utenfor Oslo. Det blir forklart litt mer rundt hvordan usikkerheter ved oppgaven håndteres i delkapittel 4.5.

Formatene passasjertellingene er levert på har variert og består av store mengder data som trenger god systematisering. Det er utviklet et script i Python for å utføre dette (vedlegg 4). Kort forklart leser scriptet inn all rådata og sorterer antall reisende til hvert område for hvert år. Hvert område har blitt tildelt sine stasjoner og nødvendige tognummerintervaller. Som resultat leverer scriptet ut to filer. En fil består av absolutte tall på antall reisende og antall telte tog i alle områdene over de ulike årene. Totale

reiser på Østlandet for hvert år er også summert opp. Den andre filen gir samme informasjon, men med relativ utvikling siden 2010 for områdene.

#### 4.3.2 Pris/Nytt sonesystem i Oslo og Akershus

Ny struktur i Ruters sonesystem medførte prisendringer for flere reisende. Denne endringen ble innført oktober 2011 (vedlegg 2). Siden passasjertellingene også gjennomføres på høsten antas det at første år nytt sonesystem påvirker passasjerantallet er fra høsten 2012, ca. ett år etter faktisk innføring. Figur 17 viser nytt sonekart etter 2012. Generelt inneholdt det eldre sonesystemet flere soner. Ved nytt sonesystem ble soneinndelingen forenklet ved bruk av færre soner.



Figur 17: Sonekart (Ruter 2016)

Prisreduksjonen har vært ulik for områdene som ligger i rutersonene. For å estimere prissjokket i de relevante områdene er det valgt å se på endring i absolutt pris på månedskort mellom områdenes hovedstasjon og Oslo S mellom 2011 og 2012. Det utelates å se på små årlige prisjusteringer da vi hovedsakelig er ute etter å kartlegge effekten av et prissjokk. Før nytt sonesystem ble innført var det mulig å kjøpe enten personlige eller upersonlige månedskort til ulik pris. Personlige månedskort ble benyttet av kun en person, mens upersonlige kunne lånes bort. Siden månedskortene etter 2012 har vært upersonlige er det valgt å sammenlikne priser i 2012 med priser på upersonlige månedskort i 2011. Prisene på månedskort i ruter områdene med nytt og

gammelt sonesystem finnes i vedlegg 2. Tabell 6 gir en oversikt over prisreduksjonene i områdene som inngår i ruters sonesystem fra 1. februar 2011 til 1. februar 2012.

Tabell 6: Prisreduksjon for ruterområder

Område	Gammel sone	Ny sone	Prisreduksjon (kr)
Oslo – Ski (lokal)	2	1	360
Lillestrøm – Årnes	10	3	480
Asker - Spikkestad	7	3	160
Lillestrøm (knutepunkt)	4	2	260
Asker – Oslo (lokal)	4	2	260
Asker – Oslo (knutepunkt)	4	2	260
Lillestrøm - Dal	10	4	480
Ski (knutepunkt)	4	2	260
Lillestrøm - Eidsvoll	12	4	480
Ski – Sonsveien	6	3	40
Lillestrøm – Oslo (lokal)	2	1	360

#### 4.3.3 Kjøretidsendringer

For å kartlegge kjøretidsendringer mellom områdene og Oslo S er det brukt relevante rutetabeller<sup>1</sup> mellom 2010 og 2015. For hvert område er det laget et Excel ark som viser avganger for hvert år i rushtid (definert mellom 1530 – 1630) og grunnrutetid (definert som mellom 1200-1300). Figuren under viser hvordan Excel arket for Lillestrøm knutepunkt ser ut.

Lillestrøm (knutepunkt)		2015						
Morgenrush	Skal med?(0/1)	1	1	1	1	1	1	1
	Lillestrøm	07:26	07:36	07:46	07:56	08:06	08:16	08:20
	Oslo	07:36	07:46	07:56	08:06	08:16	08:26	08:30
	Reisetid	00:10	00:10	00:10	00:10	00:10	00:10	00:10
	mellomregning	0,00694444	0,00694444	0,00694444	0,00694444	0,00694444	0,00694444	0,00694444
Ettermiddagsrush	Skal med?(0/1)	1	1	1	1	1	1	1
	Oslo	15:32	15:34	15:44	15:54	16:04	16:14	16:24
	Lillestrøm	15:44	15:45	15:55	16:05	16:15	16:25	16:35
	Reisetid	00:12	00:11	00:11	00:11	00:11	00:11	00:11
	mellomregning	0,00833333	0,00763889	0,00763889	0,00763889	0,00763889	0,00763889	0,00763889
Grunnrute	Skal med?(0/1)	1	1	1	1	1	1	1
	Lillestrøm	12:16	12:26	12:56	12:36	12:06	12:46	12:46
	Oslo	12:26	12:36	13:06	12:46	12:16	12:56	12:56
	Reisetid	00:10	00:10	00:10	00:10	00:10	00:10	00:10
	mellomregning	0,00694444	0,00694444	0,00694444	0,00694444	0,00694444	0,00694444	0,00694444
Grunnrute	Skal med?(0/1)	1	1	1	1	1	1	1
	Oslo	12:34	12:24	12:44	12:54	12:04	12:14	12:14
	Lillestrøm	12:45	12:35	12:55	13:05	12:15	12:25	12:25
	Reisetid	00:11	00:11	00:11	00:11	00:11	00:11	00:11
	mellomregning	0,00763889	0,00763889	0,00763889	0,00763889	0,00763889	0,00763889	0,00763889
Ant grunnrute	6							
Ant rush	7,0							
Gjennomsnitt reisetid rush	11							
Gjennomsnitt reisetid grunnrute	10,5							

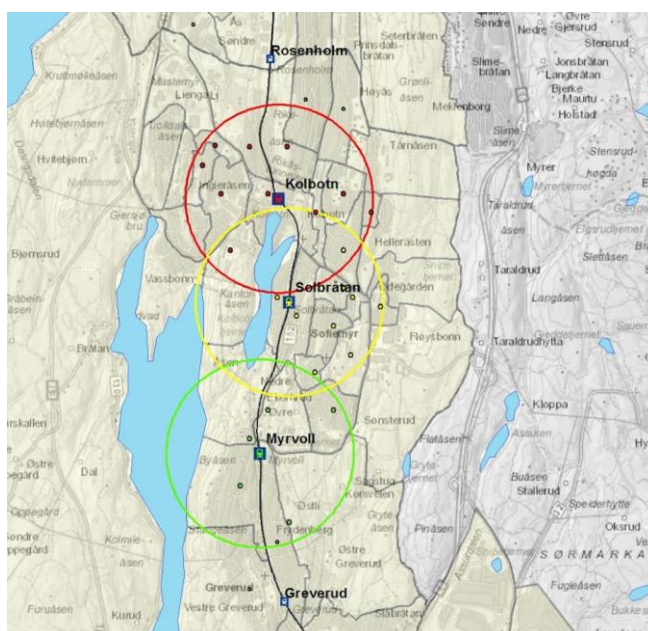
Figur 18: Avganger i 2015 for Lillestrøm (knutepunkt)

<sup>1</sup> En oversikt over hvilke rutetabeller som er brukt finnes i vedlegg 3.

Dokumentet gir i tillegg en oversikt over avgangsfrekvenser for områdene. Dette brukes til å sammenlikne med frekvensutviklingen som er kartlagt av Python scriptet (vedlegg 4) som en kontroll på at utviklingen ser forholdsvis lik ut. Kjøretidene mellom områdene og Oslo S defineres som gjennomsnittlig reisetid for både grunnrute og rushrute og er beregnet for hvert område hvert år. Endringene i reisetid finnes deretter utfra disse beregnede kjøretidene.

#### **4.3.4 Befolkningsvekst**

Det er beregnet en relativ befolkningsvekst for hvert område siden 2010 frem til 2015. Dette underkapittelet forklarer hvordan befolkningsutviklingen for hvert område er beregnet. Befolkningsdataen brukt er hentet inn fra Statistisk sentralbyrå og inneholder befolkningstall i landets grunnkretser (SSB 2016). For å kartlegge befolkningen i et område er det valgt å se på befolkningen i grunnkretsene som ligger rundt områdenes stasjoner. Ved hjelp av programvaren GIS (geografisk informasjons system) kartlegges hver grunnkrets sitt befolkningstyngdepunkt. Generelt antas det at majoriteten av reisende bor innen en kilometers omkrets fra togstasjonen. Denne antakelsen er også en generell oppfatning av fagpersoner i NSB. Derfor benyttes kun befolkningsdata fra de grunnkretser som har et befolkningstyngdepunkt liggende innenfor en kilometer fra områdenes stasjoner. Noen stasjoner ligger nære hverandre. Dette gjør at enkelte grunnkretser sitt befolkningstyngdepunkt kan ha flere stasjoner innen en kilometers avstand. Ved et slikt tilfelle tildeles befolkningen i grunnkretsen den stasjonen som ligger nærmest grunnkretsens befolkningstyngdepunkt. Figur 19 viser hvordan ulike befolkningstyngdepunkt kan ligge innenfor en kilometers radius fra flere stasjoner. Prikkene på skjermen viser befolkningstyngdepunkt for ulike grunnkretser. Fargen på punktene indikerer hvilken stasjon de er tildelt.



Figur 19: Skjerm bilde fra GIS

Figur 20 viser et skjerm bilde fra Excel over grunnkretser som er tildelt ulike stasjoner. Det er også informasjon om distanse fra grunnkretsens befolkningstetthetspunkt til stasjonen.

GKRETS_ID	Navn	stasjonsnr	Fylke	kommune	Distance	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
3013701	Alna	200100	Oslo	Oslo	916,9	4	9	17	23	24	18	25	18	18	22	31	30
3013712	Alna	200100	Oslo	Oslo	942,2	64	69	127	188	206	195	213	193	94	108	81	61
3014108	Alna	200100	Oslo	Oslo	820,1	1414	1500	1535	1567	1579	1616	1535	1568	1661	1629	1645	1647
3014109	Alna	200100	Oslo	Oslo	996,0	1911	1882	1855	1907	1920	2001	2005	2108	2116	2161	2143	2163
3014110	Alna	200100	Oslo	Oslo	867,8	1642	1631	1671	1641	1633	1644	1715	1787	1784	1790	1819	1818
3014207	Alna	200100	Oslo	Oslo	841,3	308	317	286	346	350	378	383	417	402	385	389	381
3014209	Alna	200100	Oslo	Oslo	839,5	56	52	47	46	35	26	34	32	18	23	24	16
2200804	Asker	1413000	Akershus	Asker	620,8	292	302	298	275	288	301	297	294	331	335	346	343
2200806	Asker	1413000	Akershus	Asker	909,0	635	632	622	603	614	619	614	625	613	635	640	653
2200807	Asker	1413000	Akershus	Asker	518,3	287	290	288	308	367	448	456	449	463	446	466	457
2200808	Asker	1413000	Akershus	Asker	782,7	456	447	457	472	475	481	482	462	479	558	644	663
2200901	Asker	1413000	Akershus	Asker	616,7	354	354	352	361	363	356	366	354	359	358	359	371
2200902	Asker	1413000	Akershus	Asker	599,9	395	402	408	402	397	401	403	415	428	445	446	454
2200903	Asker	1413000	Akershus	Asker	942,3	1081	1069	1106	1104	1096	1116	1149	1185	1242	1272	1297	1286
2200904	Asker	1413000	Akershus	Asker	139,6	150	140	138	153	282	293	291	280	272	279	291	282
2200906	Asker	1413000	Akershus	Asker	919,7	8	7	17	17	9	6	8	8	7	8	8	8
2200907	Asker	1413000	Akershus	Asker	233,3	375	371	375	399	401	408	420	415	426	421	401	391
2201205	Asker	1413000	Akershus	Asker	602,6	632	635	652	673	668	716	725	758	730	792	769	757
2201208	Asker	1413000	Akershus	Asker	800,6	304	306	311	316	334	379	378	411	409	412	404	398

Figur 20: Grunnkretser tildelt stasjonene Alna og Asker

Hver stasjon blir tildelt et totalt befolkningstall ved å summere opp befolkning i alle grunnkretsene som er tildelt stasjonen. Utfra dette beregnes et befolkningstall per område ved å summere opp befolkningen i områdets stasjoner. Tilslutt beregnes den relative befolkningsutviklingen fra 2010-2015 for hvert område.

All data vedrørende befolkningsutvikling, prisutvikling og reisetid sammenstilles tilslutt i et Excel ark som sammen med filene rundt passasjer og frekvensutvikling brukes til gjennomføring av regresjonsanalysen. Figur 21 viser sammenstillingen av input data for

tre av områdene. Positive tall betyr reduksjon i reisetid, reduserte priser og økt relativ utvikling i befolkning for områdene.

Område		2010	2011	2012	2013	2014	2015
<b>Oslo S - Ski (Lokal)</b>							
	Tid	0	0	0	-1	-1	-1
	Ruter	0	0	360	360	360	360
	Befolkningsvekst	0	1,4648941	3,5063954	4,9724883	6,4361836	7,9514259
<b>Asker - Spikkestad</b>							
	Tid	0	0	1	0	-1	-1
	Ruter	0	0	160	160	160	160
	Befolkningsvekst	0	2,1766429	3,4323985	4,7778509	8,1444717	10,661963
<b>Drammen - Kongsberg</b>							
	Tid	0	1	1	6	6	6
	Ruter	0	0	0	0	0	0
	Befolkningsvekst	0	2,5002991	4,6117957	6,9685369	8,9065678	10,898433

Figur 21: Sammenstilling av input data

#### 4.4 Gjennomføring av analysen

Regresjonsanalysen er gjort i Python gjennom et annet datascript (vedlegg 5). Dette scriptet benytter seg av en statistisk applikasjon som utfører selve analysen. Scriptet leser inn all input data og formaterer den til ulike observasjoner med form som i likning 4.2, men uten koeffisienter da analysen estimerer disse.

Noen områder har hatt flere reisende enn andre. I denne oppgaven er det utført analyser som estimerer koeffisienter både med og uten å ta hensyn til ulik fordeling av antall reisende i områdene. Det er tenkt ut fordeler og ulemper ved begge tilfellene. Ved å ta hensyn til dette er alle områdene tildelt et vektaltall avhengig av hvor stor trafikkandel området utgjorde av totaltrafikken for det gjeldende året. Vektaltallet sier hvor stor påvirkning områdes observasjon har for analysens resultater. Dette betyr at et område med dobbelt så mye trafikk som et annet også har dobbel påvirkning på analysens koeffisient resultater. Denne oppgaven studerer den totale passasjerutviklingen på Østlandet grunnet tilbudsendingen. Metoden ved tildeling av vektning til ulike områder virker logisk ved å se på Østlandets totaltrafikk som en stor felles masse. En observasjon som inneholder 10 % av totaltrafikken burde ha en større betydning for de estimerte koeffisientresultatene enn en observasjon bestående av 1 %



av totaltrafikken. Dette høres ut som en rettferdig fordeling. Likevel er en ulempe ved bruk av denne fremgangsmåten at det i praksis kun er få observasjoner med stor trafikkandel som får en vesentlig betydning for koeffisientresultatene. Et litt dramatisk eksempel er å si at områdene som har opplevd tilbudsendringer i form av økt frekvens og redusert reisetid er de med lavest trafikkandel og vekt, mens områdene med stor vekt kun har opplevd befolkningsvekst. Ved en slik situasjon vil ikke analysen legge spesielt godt til rette for å estimere elastisiteter knyttet til tilbudsendringene. Dersom det i tillegg er mye støy i observasjonene med høy vekt kan dette gi uheldige utslag for resten av modellen. Ved å studere totaltrafikken utvikling er det gjort en antakelse om at alle områder responderer likt til endringer i modellens forklaringsvariabler. Ved å anta dette må det videre antas at områdene, uavhengig av trafikkandel, har lik relativ utvikling ved tilsvarende endringer. Det antas med andre ord å være like koeffisienter for alle områder. Dette gjør at hver enkelt observasjon er like "interessant" for analysen, som igjen gjør det mer naturlig å gi lik vekt til alle observasjoner uavhengig av trafikkandel. En fordel ved å unngå vektlegging er at flere observasjoner i praksis har god påvirkningskraft, med andre ord er det et bredere observasjonsgrunnlag som ligger til grunn for koeffisient resultatene. Med bakgrunn i usikkerheten rundt hvilken av metodene som er best egnet er det valgt å utføre analyser både med og uten vektning av observasjonene.

Videre gjennomføres regresjonsanalysen og gir ut estimerte koeffisienter for forklaringsvariablene befolkningsvekst, pris, tid og avgangsfrekvens. Det er valgt å gjennomføre analysen uten bruk av et konstant ledd  $\alpha$ . Det antas at passasjerutviklingen står på stedet hvil i et område dersom ingen av forklaringsvariablene endrer seg der. Derfor skal verdien av konstantleddet være null. Med estimerte verdier for sammenhengen mellom endringer i forklaringsvariablene og passasjerutviklingen er det mulig å kartlegge effekten av ruteendringen.

For å gjøre dette beregner scriptet først ut en relativ passasjerutvikling for alle områdene kun grunnet ruteendringen og grunnet andre faktorer. Likningene 4.3 og 4.4 beskriver hvordan den relative utviklingen for hvert område beregnes for et område, et år.

$$\text{Effekt ruteplan for år } y \text{ og område } i: \Delta PV_{y,i} = \beta_1 \Delta AF_{y,i} + \beta_2 \Delta RT_{y,i} \quad 4.3$$

Her representerer  $\beta_1$  koeffisienten mellom avgangsfrekvens og passasjerutvikling mens  $\beta_2$  koeffisienten mellom passasjerutvikling og kjøretidsendringer.  $\Delta AF_{y,i}$  og  $\Delta RT_{y,i}$  er endringer i avgangsfrekvens og reisetid for året  $y$  siden 2010.  $\Delta PV_{y,i}$  blir dermed beregnet relativ passasjervekst for område  $i$  ved året  $y$  siden 2010 grunnet tilbudsendringer for likning 4.3, og grunnet andre faktorer i likning 4.4.

$$\text{Effekt andre faktorer i året } y \text{ for område } i: \Delta PV_{y,i} = \beta_3 \Delta PR_{y,i} + \beta_4 \Delta BV_{y,i} \quad 4.4$$

$\beta_3$  er koeffisient mellom prisutvikling og passasjervekst mens  $\beta_4$  er koeffisient mellom befolkningsvekst og passasjervekst.  $\Delta PR_{y,i}$  og  $\Delta BV_{y,i}$  er endringer i pris og befolkningsvekst. Totalutviklingen i et område kan beskrives som en sammenslåing av likning 4.3 og 4.4 hvor effekter av ytre faktorer og effekter av ruteplanen virker inn. For alle områder brukes likning 4.3 for å estimere relativ passasjerutvikling siden 2010 grunnet tilbudsendringer. Ved å multiplisere den relative veksten med faktisk antall reisende i området for 2010 beregnes et absolutt tall på økt antall reisende grunnet endret tilbud for det området ved et ønsket år. Ved en summasjon av absolutt økning i antall reisende grunnet tilbudsendringer for alle områdene beregnes det en absolutt totalvekst i reisende. Total økning i reisende for år  $y$  grunnet tilbudsending beskrives i likning 4.5.

$$\text{Total passasjerøkning år } y: PV_y = \sum_{i=1}^{18} \Delta PV_{y,i} \times N_i \quad 4.5$$

For likning 4.5 er  $\Delta PV_{y,i}$  relativ passasjervekst siden 2010 ved året  $y$  og for området  $i$  kun grunnet tilbudsending ved bruk av likning 4.3, mens  $N_i$  er faktisk reisende i område  $i$  for året 2010. Etter å ha beregnet total passasjerøkning for årene estimeres det videre hva som er effekten av ny ruteplan for alle år etter innføring, altså 2013, 2014 og 2015. Effekten blir beregnet ved å dividere den totale passasjerøkningen for året (fra likning 4.5) på antatt passasjerantall uten innføring av ruteplan. Det har skjedd små tilbudsendringer mellom 2010 til 2012 før innføringen av ny ruteplan som gir en liten estimert effekt. Denne effekten er ikke del av ny ruteplan i desember 2012 og blir kalt  $PV_{2012} \cdot PV_{2012}$  beregnes også ved bruk av likning 4.3 og 4.5. Antatt passasjerantall uten innføring av ny ruteplan beregnes å være faktisk totalt telte passasjerer for gjeldene år fratrukket estimert total passasjerøkning for året grunnet tilbudsendringer.

Beregningen vises i likning 4.6.  $N_y$  er faktisk telte passasjerer for året  $y$ .  $PV_y$  er absolutt vekst i antall reisende grunnet tilbudsendringer for året  $y$  og er basert på likning 4.5.

$$\text{Effekt år } y: E_y = \frac{PV_y - PV_{2012}}{N_y - PV_y} \times 100 \% \quad 4.6$$

Denne beregningen gir effektberegninger på samme oppsett som trafikkprognosene gjort av Dybdahl et al. (2011), nemlig en effekt beregnet utfra hvor mye høyere trafikken er ved innføring av ny ruteplan enn antatt uten endringer.

En generell oppfatning er at effekten ved endringer i et togtilbud gradvis realiseres (Fearnley & Bekken 2005; Fearnley et al. 2012; Olsson et al. 2015; Paulley et al. 2006; Wijeweera & Charles 2013; Beria & Grimaldi 2011). Det finnes altså en treghet i markedets respons til endringer. Dette er forsøkt å ta hensyn til ved å gjennomføre analysen ved bruk av forsinkelseskoeffisienter. Forsinkelseskoeffisientene forteller noe om hvor mye av effekten som er antatt realisert ved ulike år etter en endring. Da det er fem år som inngår i analysen er det definert fem ulike forsinkelseskoeffisienter;  $C_1, C_2, C_3, C_4, C_5$ .  $C_1$  definerer hvor stor andel av nytten som antas realisert samme året som endringen skjedde.  $C_2$  definerer andel nytte realisert to år etter endringen osv. Passasjertellingene er gjort på høsten og de større endringene har skjedd i desember for de ulike årene. Dette betyr at endringen har vært realisert i ca. trekvarter år før første telling er blitt gjort. Dermed vil en verdi på  $C_1$  lik 1 bety at all effekt er antatt realisert etter trekvarter år. For å forsøke å implementere en forsinkelseeffekt i analysen er det valgt å gi en treghet i realiseringen av endringer i forklaringsvariablene som antas å ha en forsinkelse ved seg: frekvens, reisetid og pris. Med en treghet i realiseringen av forklaringsvariabelen menes at dersom for eksempel avgangsfrekvensen økes med 50 % for første året og verdien på  $C_1$  er lik 0,8 vil verdien for avgangsfrekvensen være økt til 40 % for første året etter endringen. Dersom det ikke skjer noen ytterligere økning i frekvensen og verdien for  $C_2$  er lik 0,9 vil videre verdien være 45 % høyere enn opprinnelig for neste år. Forklaringsvariablenes verdier ved bruk av forsinkelseskoeffisienter er beregnet ved bruk av likning 4.7.

$$V_y = \sum_{i=0}^{y-1} C_{i+1} (V_{y-i} - V_{y-i-1}) \quad 4.7$$

For likning 4.7 er  $V_y$  forklaringsvariabelens verdi justert for forsinkelser ved året  $y$ .  $C_{i+1}$  er forsinkelseskoeffisienten som varierer ettersom hvor mange år det er siden endringen  $V_{y-i} - V_{y-i-1}$  ble implementert.  $V_{y-i} - V_{y-i-1}$  er endringen i forklaringsvariabelens verdier mellom de ulike årene i forkant av det aktuelle året som blir sett på. For at likning 4.7 skal fungere på rett måte må verdiene for  $y$  ved året 2010 være lik 0, 2011 lik 1, 2012 lik 2 osv. Likning 4.7 håndterer at flere endringer skjer underveis i årene 2010-2015, og det er kontrollert at likningen gir rette verdier ved ulike beregninger.

Analysen gjennomføres ved antakelse av at all effekt realiseres samme år som en endringen og under to ulike forsinkelsestilfeller både med vektete og ikke - vektete observasjoner. Totalt seks analyser. Forsinkelseskoeffisientenes verdier i de ulike forsinkelsestilfellene er kun antakelser på hva som kan være realistiske verdier og har først og fremst hensikt å studere hvordan dette påvirker analysens resultater i forhold til en antakelse om umiddelbar effektrealisering. I tillegg til effektberegninger plotter scriptet hvordan total passasjerutvikling for hele Østlandet har vært, hvordan estimert totalutvikling er for årene (basert på analysens koeffisienter), beregnet utviklingen kun av tilbudsending og beregnet utvikling kun grunnet befolkningsvekst. For analysene med forsinkelsestilfeller blir estimert totalutvikling og utvikling kun av tilbudsending ekstrapolert frem til 2019 for å illustrere modellens antatte utvikling inntil all effekt av siste endring i 2015 er realisert.

Tre residualplot fremkommer også for å undersøke modellens forutsetninger. Disse finnes, sammen med korrelasjonsmatrisene mellom forklaringsvariablene i vedlegg 1.

#### **4.5 Usikkerhetshåndtering**

I kapittel 2 ble det nevnt usikkerhet i forhold til trafikkprognoser i forkant av prosjektgjennomføringer. Usikkerhet er ofte tydelig i tidligfasen av prosjekter, hvor viktige og avgjørende beslutningssituasjoner inntreffer (Austeng et al. 2005). Ofte blir usikkerhet knyttet opp til selve gjennomføringen av et prosjekt. Likevel har jeg gjennom oppgaven erfart at usikkerhet også i stor grad kan relateres til evalueringer av prosjekter. Usikkerhet vil finnes i alt forskningsarbeid med varierende alvorlighetsgrad. Dette gjelder også for denne oppgaven hvor passasjertellingene innehar mer usikkerhet enn ønskelig. Usikkerhet bør forsøkes å håndteres. For å kunne gjøre dette er det avgjørende å kjenne kildene til usikkerheten. I henhold til Austeng et al. (2005) kan

usikkerhetens kilder grovt deles inn i kategoriene konseptuell usikkerhet, operasjonell usikkerhet, kontekstuell usikkerhet og scenariell usikkerhet. Med konseptuell usikkerhet menes usikkerhet som kan tilknyttes analysen og den omkringliggende modellen. Operasjonell usikkerhet betegner den usikkerhet som mer kan relateres til selve gjennomføringen av arbeidet – en indre usikkerhet. Kontekstuell usikkerhet kan relateres til den usikkerhet som finnes i arbeidet/prosjektets omgivelser. Scenariell usikkerhet kan tolkes som usikkerhet relatert til de scenariene som viktige beslutninger bygges på, deres pålitelighet og relevans (Austeng et al. 2005).

Noen av de større usikkerhetskildene for denne oppgaven kan i større grad kobles opp mot usikkerhetskategoriene nevnt. Hovedsakelig kan oppgavens usikkerhet grovt deles inn i følgende:

- Usikkerhet ved rådataen.
- Usikkerhet ved analysegjennomføring.
- Usikkerhet ved antakelsene for oppgaven.

Delkapittel 4.5 har som formål å beskrive hvordan usikkerheten under disse kategoriene er forsøkt å håndteres, samt nevne eventuelle anbefalinger til hvordan å etterprøve kvaliteten av datagrunnlaget i et videre arbeid.

### **Usikkerhet ved rådataen**

Passasjertellingene som benyttes i oppgaven inneholder usikkerhet. Ved mottakelse av passasjertellingene ble det presisert at datagrunnlaget er sårbart for variasjoner i trafikken over tid og må brukes med forsiktighet. Av den grunn må resultatene som fremkommer benyttes med en sunn skepsis.

Ideelt sett burde dataskaleringen vært mer detaljert med hyppigere tellinger. Det ville gitt et mer robust datagrunnlag med et mer omfattende innblikk i trafikken. Det er dessverre lite å gjøre for å redusere selve usikkerheten til de faktiske tellingene. Likevel er det gjort tiltak for å unngå å inkludere mer usikkerhet enn nødvendig. Først bør det igjen nevnes at passasjertellingene ved de største stasjonene i Oslo (Oslo S, Nationaltheatret, Skøyen, Lysaker) er utelatt fra analysen som følge av antakelsene om reisemønstrene beskrevet i 4.3.1. Disse tellingene er trolig mest usikre da flest passasjerer stiger av og på her. Utelatelsen av disse tellingene er derfor til fordel for oppgavens usikkerhet.

Det er også utført en test for å avdekke åpenbare feiltellinger i datagrunnlaget. Scriptet som sorterer passasjertellingene (vedlegg 4) leverer et oversiktsdokument for alle områdene med informasjon over hvilke tognummer som er telt for alle år og hvor mange avstigende og påstigende som er telt i området for hvert tognummer. Disse dokumentene er gjennomgått for å undersøke at alle tognummerserier som skal være med i et område er telt, samt avdekke åpenbare feiltellinger. Åpenbare feiltellinger kan for eksempel sees dersom et tognummer har 300 avstigende og 300 påstigende i et område et år, mens påfølgende år er det ingen avstigende eller påstigende for samme tognummer. Et slikt funn gir grunn til å undre seg over tellingens kvalitet.

NSB har tildels begynt med automatiske passasjertellinger for enkelte togsett som registrerer avstigende og påstigende likt de manuelle tellingene. Jeg har blitt forklart at automatisk tellesystem så langt kun er brukt på et fåtall av togene og kvaliteten på tellingene er ikke "friskmeldt". Derfor er det rådet fra NSB å ikke benytte disse tellingene i analysen. En anbefaling for få bedre kjennskap til den faktiske usikkerheten ved de manuelle passasjertellingene som er brukt, er å sammenlikne fremtidige togs automatiske tellinger mot deres manuelle tellinger. Dette vil ikke gi en oversikt over usikkerheten til tellingene inkludert i oppgaven, men kan forhåpentligvis gi en pekepinn over hvor usikre de manuelle tellingene generelt er og har vært. Dette kan igjen brukes til å vurdere usikkerheten ved oppgavens resultater.

### **Usikkerhet ved analysegjennomføring**

Feil under bearbeidelse av datagrunnlaget og ved gjennomføring av analysen kan forekomme og medfører unødvendig usikkerhet i resultatene. For denne oppgaven er det forsøkt å kontrollere at feil underveis i arbeidet er blitt rettet opp i og generelt unngås. Dette er i hovedsak gjort ved å kontinuerlig gjøre kontrollsjekker av om scriptene som er utviklet arbeider som tiltenkt. Dette gjelder sortering av rådataen og beregninger som gjøres i scriptet. Stikkprøver er utført hvor beregningene gjøres manuelt og sammenliknes med beregningene gjort av scriptet. For å kontrollere at regresjonsanalysen i Python utføres på rett måte er analysen kontrollsjekket ved å bruke regresjonsverktøyet i Excel. Resultatene har vært identiske og bekrefter at den statiske applikasjonen som er brukt jobber likt som andre regresjonsverktøy. Alt arbeid som er utført i forbindelse med analysene er blitt overlevert til dyktige fagfolk i NSB som

kan utføre videre og mer detaljerte analyser fremover dersom det skulle bli ønskelig og nødvendig.

### **Usikkerhet ved antakelsene for oppgaven**

For at oppgaven skulle være gjennomførbar i tidsperioden som er blitt tildelt har det vært nødvendig å gjøre antakelser. Antakelser er ofte forenklinger av virkeligheten og setter begrensninger for resultatenes gyldighetsområde. Mange antakelser øker også usikkerheten ved hvor godt resultatene gjenspeiler virkeligheten. For denne oppgaven er det forsøkt å beskrive så godt som mulig hvilke antakelser og fremgangsmetoder som ligger til grunn for resultatene som blir presentert. Dette er viktig for å kunne klarlegge forutsetningene for analysen ved videre arbeid. Antakelsene er grundigere beskrevet i kapittel 6.3. Totalt er det gjort seks ulike analyser, alle med ulike forutsetninger og antakelser. Dette er gjort for å forsøke å oppnå et bredt evalueringsgrunnlag. Ved gjennomføring av kun en analyse ville oppgavens resultater være svært avhengig av de forutsetninger som ligger til grunne for kun denne. Flere analyser gir et mer helhetlig bilde og avdekker generelle trender og variasjoner i resultatene.

Utførelsen av testene og kontrollene som er beskrevet, samt en grundig forklaring av under hvilke forutsetninger resultatene for oppgaven er basert på, har som formål å redusere unødvendig usikkerhet ved analysens gjennomføring og redegjøre for hvordan ting er gjort. Dette er ment å gi en god oversikt over grunnlaget til resultatene og hvordan disse bør brukes, samt en oversikt som danner et utgangspunkt for eventuelt videre arbeid.

## 5. Resultater

Dette kapittelet presenterer og beskriver resultatene som er fremkommet. Det er gjort totalt seks ulike analyser. Tre analyser med forskjellige antakelser om når effekten av endringer realiseres er gjort under bruk av både vektete og ikke vektete observasjoner. Et tilfelle antar umiddelbar effektrealisering det første året. To andre tilfeller er basert på ulike forsinkelseskoeffisienter. Resultatene blir presentert i form av grafer og tabeller for deretter å bli kort beskrevet. I analysene er det fremkommet tegn til noe autokorrelerte residualer da det er estimert et DW tall litt lavere enn nedre kritisk verdi (1,43) for analysene. DW tallene har variert i intervallet 1,1-1,4. Derfor er det gjort en Newey West test som har estimert modifiserte standardfeil og P – verdier. De modifiserte standardfeilene og P- verdiene er markert med en stjerne(\*) i tabellene ved siden av de originale verdiene. Generelt er P – verdiene noe høyere etter å ha tatt høyde for autokorrelasjon og eventuell heteroskedastisitet, men problemet har ikke medført vesentlige endringer i analysens resultater og signifikanstester. Variablene blir betraktet som signifikante utfra et 90 % nivå, men variablenes P – verdier gir ofte signifikans også på 95 % og 99 % nivå. En grundigere diskusjon av resultatene gjøres i kapittel 6. Med tanke på at togtrafikken i Norge er i en omorganiseringsperiode hvor NSB blir konkurranseutsatt, har NSB ytret et ønske om å holde de eksakte koeffisientestimatene skjult utover en standard konfidensialitetsperiode på 5 år. Derfor er disse skjult for første tilfelle ved å erstatte tallene med bokstaver. Utviklingen i koeffisientverdiene for de ulike analysene blir likevel presentert i forhold til hverandre. Dette muliggjør for en generell diskusjon rundt koeffisientenes utvikling i kapittel 6.

### 5.1 Resultater med vektete observasjoner

Følgende delkapittel presenterer de resultater som er fremkommet ved å tildele alle observasjonene ulik vekt. Resultatene for de forskjellige tilfellene blir presentert med samme oppsett.

#### 5.1.1 Umiddelbar effektrealisering med vektete observasjoner

Resultatene som er fremkommet under antakelsen av at all effekt av ruteplanens endringer er realisert i løpet av første året presenteres her. Tabell 7 og 8 viser regresjonsanalysens resultater og estimerte effekter av ruteplanen for de ulike årene i etterkant.



Tabell 7: Output ved umiddelbar effektrealisering (vektet)<sup>2</sup>

Forklaringsvariabler	Koeffisient	Standardavvik	P-verdi	R <sup>2</sup>
Avgangsfrekvens	A	0,112/0,232*	0,000/ 0,015*	0,807
Reisetid	B	0,852/1,625*	0,000/ 0,061*	
Pris	C	0,014/0,017*	0,501/ 0,592*	
Befolkningsvekst	D	0,598/0,769*	0,000/ 0,000*	

Tabell 7 viser estimerte verdier for forklaringsvariablenes koeffisienter med tilhørende standardavvik og p - verdier. Modellens R kvadrat forklarer at 80,7 % av observert variasjonen i passasjerveksten forklares av modellens variabler. P - verdiene til forklaringsvariablene viser sannsynligheten for at det finnes en sammenheng mellom dem og passasjerveksten basert på datagrunnlaget. Regresjonsanalysen viser signifikante sammenhenger mellom passasjerutviklingen og forklaringsvariablene avgangsfrekvens, reisetid og befolkningsvekst på et 90 % nivå. Avgangsfrekvens og befolkningsvekst er også signifikant på 95 % nivå, men ikke reisetid. Det estimeres at en prosent økning i avgangsfrekvens gir A % passasjervekst, ett minutt reisetidsbesparelse gir B % passasjervekst og at en befolkningsvekst på en prosent gir D % passasjervekst. P- verdien for pris viser at datagrunnlaget ikke finner en signifikant sammenheng mellom et prissjokk og passasjerutviklingen.

Tabell 8 viser effektberegningene for årene etter ruteendring. Tallene er basert på likning 4.6.

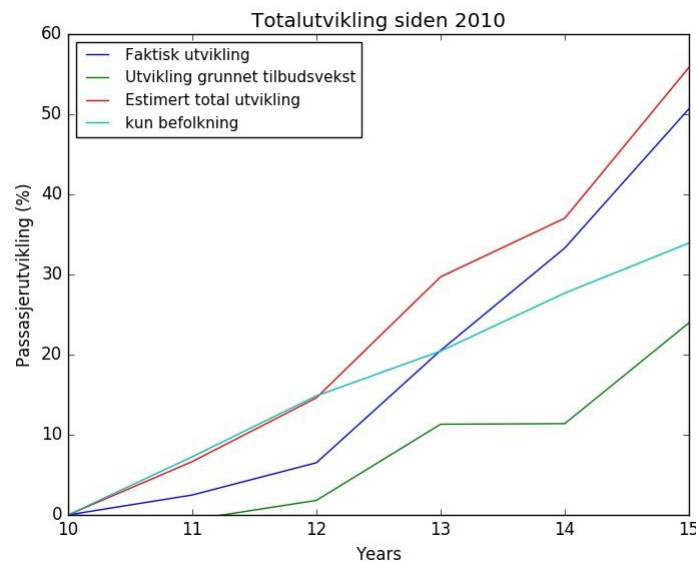
Tabell 8: Beregnede effekter ved umiddelbar effektrealisering (vektet)

År	2013	2014	2015
Beregnet effekt	8,70 %	7,86 %	17,48 %

<sup>2</sup> Koeffisientene presenteres med bokstaver for første analyse etter ønske fra NSB.

Beregningene viser at antall reisende var 8,70 % prosent høyere i 2013 grunnet innføring av ny ruteplan. For 2014 var passasjertallet 7,86 % høyere enn uten ny ruteplan. Årsaken til nedgangen fra 2013 til 2014 er at estimert passasjerøkning grunnet ruteendring er lik for begge årene da det ikke er noen videre tilbudsøkning, men tallet divideres på et høyere totalt passasjertall i 2014. For 2015 er beregnet effekt av ny ruteplan 17,48 %. Effekthoppet skyldes ytterligere tilbudsøkning ved full implementering.

Figur 22 viser et plot med grafer for faktisk passasjerutvikling, estimert passasjerutvikling, passasjerutvikling grunnet kun tilbudsvekst og kun grunnet befolkningsvekst. Grafene for faktisk og estimert utvikling viser en mer beskjeden vekst i forkant av ruteendringen som følge av kun befolkningsvekst enn i etterkant av hvor veksten er brattere. Disse følger hverandre mer eller mindre, men nivået på estimert totalutviklingen ligger hele veien noe høyere enn faktisk utvikling. Vekst grunnet befolkningsutvikling stiger nesten konstant og følger den estimerte totalutviklingen frem til ruteendringen. Estimert totalvekst har en brattere stigning før ruteendringen enn faktisk utvikling, noe som antyder en overestimering av befolkningsvekstens betydning for passasjerutviklingen.



Figur 22: Utvikling mellom 2010-2015 ved umiddelbar effektrealisering (vektet)

### 5.1.2 Forsinkelsestilfelle 1 med vektete observasjoner

Følgende analyse opererer med forsinkelseskoeffisienter som vist i tabell 9.

Tabell 9: Forsinkelseskoeffisienter 1

C <sub>1</sub>	C <sub>2</sub>	C <sub>3</sub>	C <sub>4</sub>	C <sub>5</sub>
0,66	0,83	1,0	1,0	1,0

Forsinkelseskoeffisientene antar at 66 % av effekten er realisert innen første året, 83 % etter år to og 100 % innen tre år. Resultatene fra analysen ved forsinkelsestilfelle 1 er vist i tabell 10.

Tabell 10: Output ved forsinkelsestilfelle 1 (vektet)

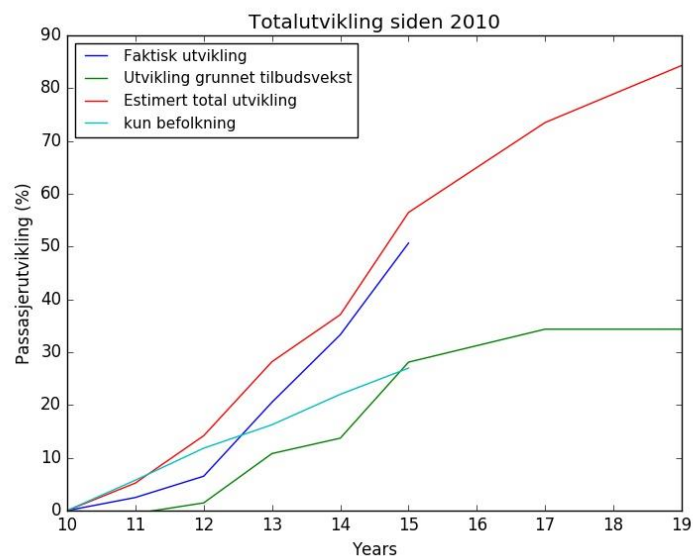
Forklaringsvariabler	Koeffisient	Standardavvik	P-verdi	R <sup>2</sup>
Avgangsfrekvens	A + 0,263	0,141/0,250*	0,000/ 0,001*	0,825
Reisetid	B + 1,086	0,941/2,025*	0,000/ 0,042*	
Pris	C + 0,015	0,016/0,022*	0,698/ 0,783*	
Befolkningsvekst	D - 0,657	0,596/0,795*	0,000/ 0,000*	

Også her finner analysen signifikante sammenhenger for avgangsfrekvens, reisetid og befolkningsvekst. P verdiene er i dette tilfellet så lave at disse variablene også er signifikante på 95 % nivå. Generelt er verdiene for forklaringsvariablene som antas å ha en forsinkelse i effektrealiseringen høyere enn ved umiddelbar realisering. Pris forblir ikke signifikant. Koeffisienten for befolkningsvekst er noe lavere. R kvadrat ligger mer eller mindre på samme nivå med en verdi på 0,825.

Tabell 11: Beregnede effekter ved forsinkelsestilfelle 1 (vektet)

År	2013	2014	2015
Beregnet effekt	8,51 %	10,25 %	21,75 %

Tabell 11 viser effekter av ruteendringen for 2013,2014 og 2015. I motsetning til ved forrige tilfelle stiger effekten av endringen her hele veien da analysen antar en gradvis realisering av effekten.



Figur 23: Utvikling mellom 2010-2015 ved forsinkelsestilfelle 1

Som følge av en lavere koeffisient for befolkningsvekst stiger grafen for estimert totalutvikling i figur 23 noe mindre i forkant av ruteendringen enn i figur 22. Ellers er utviklingen forholdsvis lik. Det er ekstrapolert en utvikling frem til 2019 for estimert totalutvikling og utvikling grunnet tilbudsvekst. Figur 23 viser en svak videre vekst frem til 2017 grunnet kun tilbudsvekst. Etter dette er all effekt realisert og veksten videre i estimert totalutvikling er forårsaket av befolkningsvekst.

### 5.1.3 Forsinkelsestilfelle 2 med vektete observasjoner

I dette tilfellet antas det at effekten ved endringer realiseres slik koeffisientene i tabell 12 viser. 50 % av effekten er realisert innen første året (trekvarter år), 75 % etter to år og 100 % etter tre år.

Tabell 12: Forsinkelseskoeffisienter 2

C <sub>1</sub>	C <sub>2</sub>	C <sub>3</sub>	C <sub>4</sub>	C <sub>5</sub>
0,50	0,75	1,0	1,0	1,0

Tabell 13: Output ved forsinkelsestilfelle 2 (vektet)

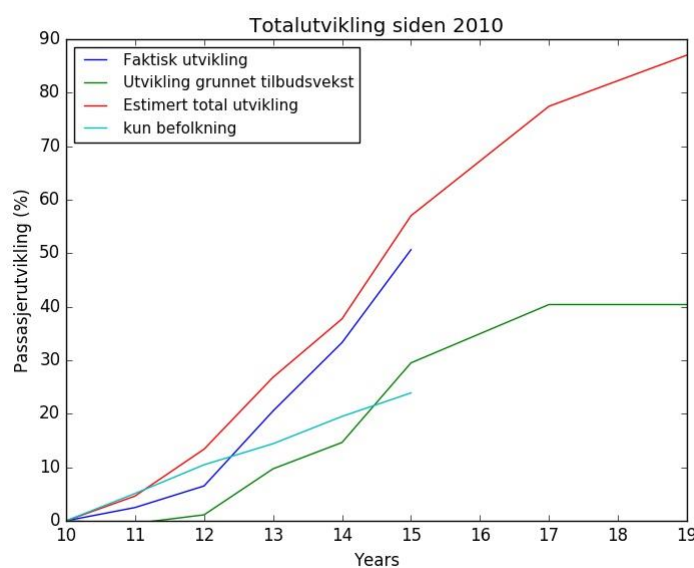
Forklaringsvariabler	Koeffisient	Standardavvik	P-verdi	R <sup>2</sup>
Avgangsfrekvens	A + 0,422	0,159/ 0,248*	0,000/ 0,000*	0,834
Reisetid	B + 1,609	0,968/ 2,222*	0,000/ 0,037*	
Pris	C + 0,025	0,017/0,024*	0,327/ 0,504	
Befolkningsvekst	D - 0,936	0,578/0,783*	0,000/ 0,005*	

Igjen estimeres signifikante koeffisientverdier for avgangsfrekvens, reisetid og befolkningsvekst ned mot 95 % nivå. Koeffisientenes verdier er ytterligere høyere enn i forsinkelsestilfelle 1, med unntak av koeffisienten til befolkningsvekst som er lavere. Variabelen pris har fortsatt ingen signifikant verdi.

Tabell 14: Beregnede effekter ved forsinkelsestilfelle 2 (vektet)

År	2013	2014	2015
Beregnet effekt	7,76 %	11,39 %	23,44 %

Som for forsinkelsestilfelle 1 viser tabell 14 at effekten blir gradvis innført.



Figur 24: Utvikling mellom 2010-2015 ved forsinkelsestilfelle 2 (vektet)

Estimert total utvikling i figur 24 ligger også her noe høyere enn faktisk utvikling, men er plassert tettere enn tidligere. Det er synlig at befolkningsutviklingen har en lavere estimert koeffisient i dette tilfellet da estimert totalvekst i forkant av ruteendringen er lavere enn i de tidligere tilfellene. Grafen for estimert total utviklingen stiger også i et mindre hakkete mønster enn de andre tilfellene som følge av en mer jevn fordeling av effektrealiseringen. Den ekstrapolerte utviklingen etter 2015 viser også at en større andel av effekten fra endringen i 2015 realiseres i de påfølgende årene. Dette bærer både grafen for estimert total utvikling og utvikling grunnet tilbudsvekst preg av.

## 5.2 Resultater med ikke - vektete observasjoner

Følgende delkapittel presenterer resultatene som er fremkommet under antakelsen av at alle observasjonene er likt vektet. Oppsettet er tilsvarende som i delkapittel 5.1 og forsinkelsestilfellene opererer med like forsinkelseskoeffisienter som tidligere.

### 5.2.1 Umiddelbar effektrealisering med ikke - vektete observasjoner

Følgende resultater er fremkommet under antakelse om umiddelbar effektrealisering med ikke-vektede observasjoner.

Tabell 15: Output ved umiddelbar effektrealisering (ikke - vektet)

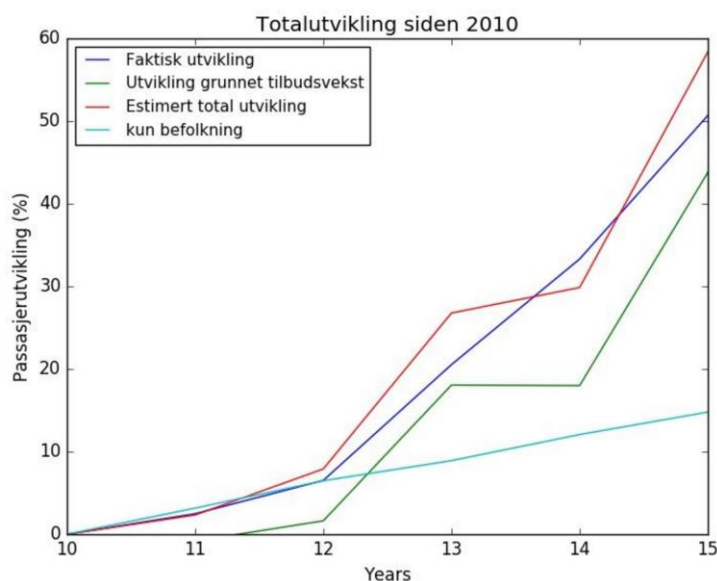
Forklaringsvariabler	Koeffisient	Standardavvik	P-verdi	R <sup>2</sup>
Avgangsfrekvens	A + 0,608	0,068/ 0,196*	0,000/ 0,000*	0,836
Reisetid	B + 0,057	1,257/ 1,188*	0,014/ 0,010*	
Pris	C + 0,008	0,020/ 0,016*	0,964/ 0,957*	
Befolkningsvekst	D - 1,788	0,751/ 0,628*	0,069/ 0,030*	

Som for analysene med vektete observasjoner er det også her estimert signifikante sammenhenger mellom avgangsfrekvens, reisetid og befolkningsvekst på 95 % nivå. Størst forskjell fra tidligere er det for verdiene til avgangsfrekvens og befolkningsvekst. Sammenliknet med analysene med vektete observasjoner er koeffisienten for avgangsfrekvens høyere enn tidligere og for befolkningsvekst lavere. For reisetid og pris er ikke verdiene vesentlig forskjellige. Neste tabell viser effektberegningene.

Tabell 16: Beregnede effekter ved umiddelbar effektrealisering (ikke - vektet)

År	2013	2014	2015
Beregnet effekt	16,03 %	14,19 %	39,44 %

En høyere estimert koeffisient for avgangsfrekvens gir tydelig påvirkning på estimert effekt av ruteendringen. Effekten for alle år er langt høyere enn situasjonene med vektete observasjoner. Neste figur viser utviklingsplottet for situasjonen.



Figur 25:Utvikling mellom 2010-2015 ved umiddelbar effektrealisering (ikke -vektet)

Lavere estimat av koeffisienten for befolkningsvekst gir et bedre treff mellom estimert utvikling og faktisk utvikling i forkant av ruteendringen. I etterkant av 2012 kan man observere en tydelig hakkete realisering av ruteendringens effekt som følge av antakelsen om umiddelbar realisering.

### 5.2.2 Forsinkelsestilfelle 1 med ikke - vektete observasjoner

Forsinkelsestilfelle 1 med ikke-vektede observasjoner benytter samme forsinkelseskoeffisienter som ved vektete observasjoner. Tabell 17 gir resultatene fra regresjonsanalysen.

Tabell 17: Output ved forsinkelsestilfelle 1 (ikke - vektet)

Forklaringsvariabler	Koeffisient	Standardavvik	P-verdi	R <sup>2</sup>
Avgangsfrekvens	A + 0,902	0,070/0,163*	0,000/ 0,000*	0,881
Reisetid	B + 1,276	1,265/1,637*	0,000/ 0,009*	
Pris	C + 0,020	0,020/0,020*	0,593/ 0,586*	
Befolkningsvekst	D - 2,202	0,667/0,678*	0,151/ 0,157*	

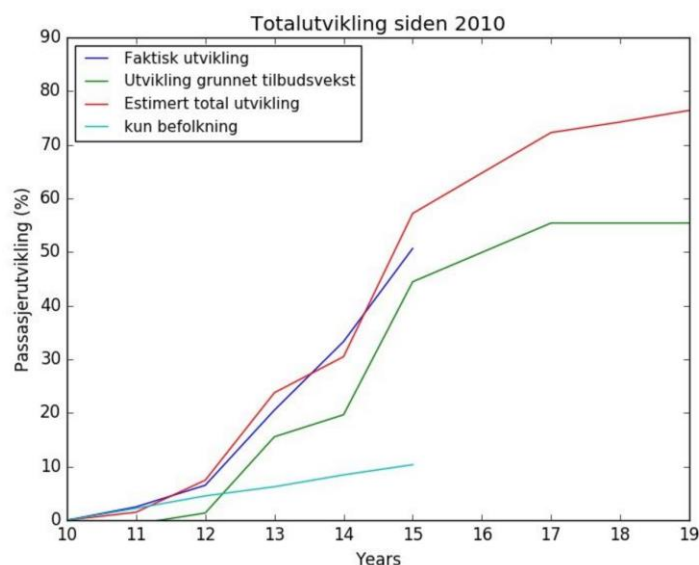
Som for analysene med vektete observasjoner estimeres det høyere verdier for koeffisientene ved antakelse av forsinket effektrealisering med unntak av befolkningsveksten som faller noe. R kvadrat øker også noe. I dette tilfellet er det funnet signifikante koeffisienter for kun avgangsfrekvens og reisetid. Befolkningsvekstens p verdi er ikke signifikant på 90 % nivå, men ville vært det på 80 % nivå.

Tabell 18: Beregnede effekter ved forsinkelsestilfelle 1 (ikke - vektet)

År	2013	2014	2015
Beregnet effekt	13,54 %	16,11 %	40,60 %

Effektene er også for dette tilfellet langt høyere enn ved bruk av vektete observasjoner som følge av en spesielt høy koeffisient for avgangsfrekvens.





Figur 26: Utvikling mellom 2010 - 2015 ved forsinkelsestilfelle 1 (ikke -vektet)

### 5.2.3 Forsinkelsestilfelle 2 med ikke - vektete observasjoner

Siste analyse bruker samme forsinkelseskoeffisienter som for vektete observasjoner.

Resultatene for analysen vises i tabell 19.

Tabell 19: Output ved forsinkelsestilfelle 2 (ikke -vektet)

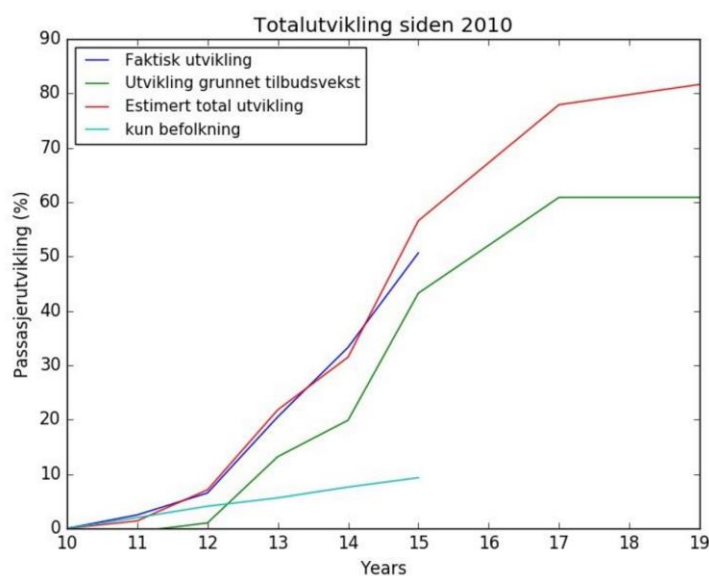
Forklaringsvariabler	Koeffisient	Standardavvik	P-verdi	R <sup>2</sup>
Avgangsfrekvens	A + 1,041	0,074/0,151*	0,000/ 0,000*	0,888
Reisetid	B + 1,839	1,313/1,910*	0,000/ 0,012*	
Pris	C + 0,027	0,021/0,021*	0,385/ 0,402*	
Befolkningsvekst	D - 2,298	0,650/0,698*	0,184/ 0,215*	

Igjen estimeres det kun signifikante verdier på et 95 % nivå mellom avgangsfrekvens og reisetid. Befolkningsvekst er ikke signifikant på 90 % nivå. Pris forblir også ikke signifikant. Koeffisientene er ytterligere høyere enn ved forsinkelsestilfelle 1, likt det som ble observert for de vektete analysene.

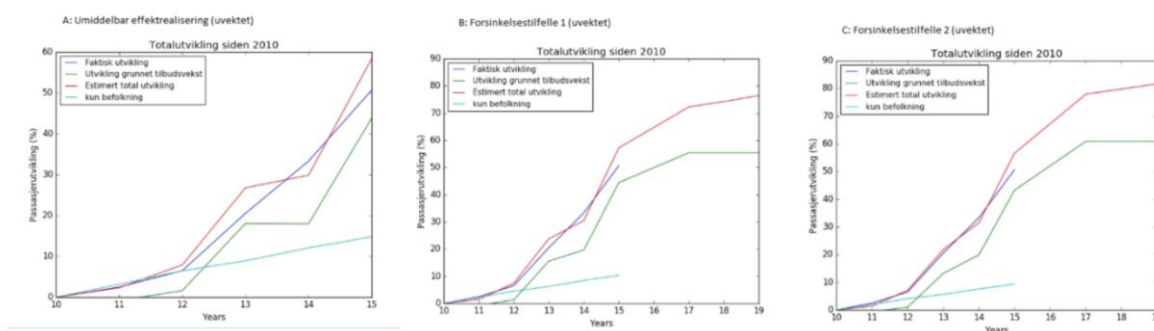
Tabell 20: Beregnede effekter ved forsinkelsestilfelle 2 (ikke - vektet)

År	2013	2014	2015
Beregnet effekt	11,35 %	16,65 %	39,34 %

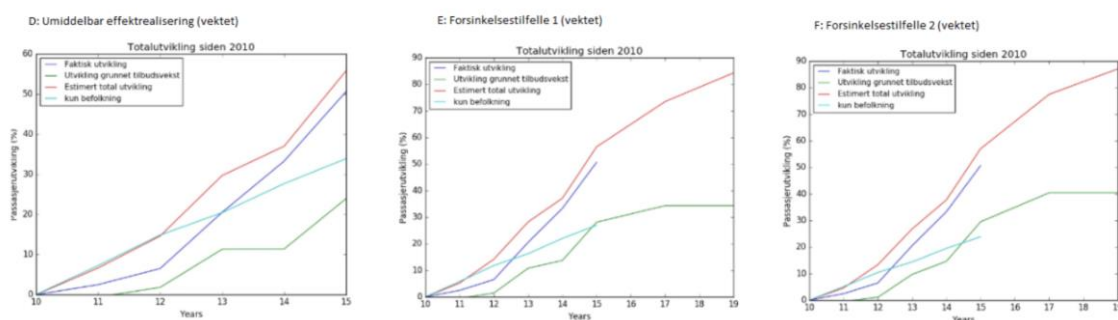
Figur 27 viser en svært sammenfallende estimert utvikling og faktisk utvikling. Kurvene nærmest overlapper hverandre hele veien frem til 2015. Figur 28 og 29 viser en sammenstilling av alle utviklingsplottene for både vektete og ikke vektete analyser.



Figur 27: Utvikling mellom 2010 - 2015 ved forsinkelsestilfelle 2 (ikke - vektet)



Figur 28: Sammenstilling av utviklingsplot (ikke - vektet)



Figur 29: Sammenstilling av utviklingsplot (vektet)

Ved å studere figur 28 og 29 kan det generelt observeres at plottene som har tatt utgangspunkt i ikke - vektete observasjoner overlapper bedre enn ved bruk av vektete observasjoner. Alle figurene viser også tydelige forskjeller på hvordan grafene blir mindre hakkete ved antakelse om en forsinket effekt realisering. Ved å sammenlikne vekst kun grunnet tilbudsendring er det tydelige forskjeller på hvor grafene ender ved bruk av vektete eller ikke vektete observasjoner. For ikke vektete observasjoner varierer veksten siden 2010 grunnet tilbudsendring mellom ca. 45 % - 60 %, mens for vektete observasjoner mellom ca. 25 % - 40 %. Med andre ord er det store variasjoner som også gir påfølgende variasjoner i effektberegningene.

#### 5.2.4 Estimerte koeffisienter ved regresjonsanalyser uten pris

Ingen av analysene har funnet signifikante sammenhenger mellom pris og passasjervekst. Derfor er det hensiktsmessig å kontrollere at analysene ikke beregner helt annerledes elastisiteter for de signifikante variablene dersom pris utelukkes. Alle analysene er gjort på nytt uten pris og har gitt nye beregnede elastisiteter. Resultatene fra regresjonsanalysene med ekskludering av variabelen pris har ikke gitt store utslag på de estimerte elastisitetene for avgangsfrekvens og reisetid. Utelukkelsen har dermed ikke stor påvirkning på hva som er beregnet effekt av endringer i ruteplanen. Befolkningsveksten har fått noe større utslag i sin estimerte koeffisient som følge av utelatelse av pris. Dette er ikke så unaturlig da korrelasjonsmatrisene i vedlegg 1 viser sterkest korrelasjon mellom pris og befolkningsvekst. Dette gjør at effekten pris har blitt tildelt i analysen "lettest" slår over på variabelen for befolkningsutvikling.

## 6. Diskusjon

Dette kapittelet går i dybden på resultatene. Hensikten er å stille informasjon presentert i kapittel 2 opp mot resultatene fra kapittel 5. En diskusjon rundt regresjonsresultatene og effektberegningene gir grunnlag for å besvare forskningsspørsmål 1 og 2. Deretter drøftes resultatene opp mot prognosene i forkant av ruteendringen, som gir mulighet til å besvare forskningsspørsmål 3. Kapittelet trekker inn kritiske elementer ved oppgaven, samt en helhetlig diskusjon tilslutt. Forslag til videre arbeid blir også nevnt.

### 6.1 Diskusjon rundt regresjonsmodellene

Kapittel 6.1 diskuterer regresjonsanalysen. Først blir resultatene for modellens forklaringsvariabler diskutert. Koeffisientestimatene til forklaringsvariablene kan tolkes som elastisiteter mellom forklaringsvariabelen og passasjeretterspørselen. Deretter diskuteres det litt helhetlig rundt regresjonsmodellen.

#### 6.1.1 Avgangsfrekvens

For alle analysene, uavhengig av forutsetninger, er det estimert signifikante sammenhenger mellom avgangsfrekvens og passasjerutvikling. Elastisitetsverdiene har variert mellom  $A$  til  $A + 0,422$  for forsinkelsestilfellene under antakelse om vektete observasjoner. For analysene med ikke - vektete observasjoner har elastisitetene fått verdier mellom  $A + 0,608$  til  $A + 1,041$ . Resultatene viser en trend på at elastisiteten øker under antakelse om en langsom effektrealisering. Med andre ord sier resultatene at økt avgangsfrekvens har større betydning for etterspørselen under antakelse om at effekten realiseres over en lengre tidsperiode. Denne indikasjonen er i samsvar med tidligere forskning som nevner at elastisitetsverdiene stiger over tid (Fearnley & Bekken 2005; Fearnley et al. 2012; Paulley et al. 2006; Wijeweera & Charles 2013). Generelt er elastisiteten estimert høyere under analysene med ikke - vektete observasjoner. Elastisitetstallene for avgangsfrekvensen er svært sterke og viser at markedet er sensitivt til endringer i variabelen. Hamre (2002) nevner en total frekvenselastisitet på 0,307 for lengre norske togreiser. Paulley et al. (2006) gir en frekvenselastisitet på 0,75, mens FitzRoy & Smith (1995) finner en elastisitet på 0,44. Disse tallene sier i seg selv at det ikke finnes noe fasitsvar for frekvenselastisiteten. Det er store forskjeller for hvert tilfelle. Elastisitetene for denne oppgaven er i de fleste tilfeller høyere enn alle tallene nevnt i disse studiene, og viser til at markedet har vært sensitivt til endringer i

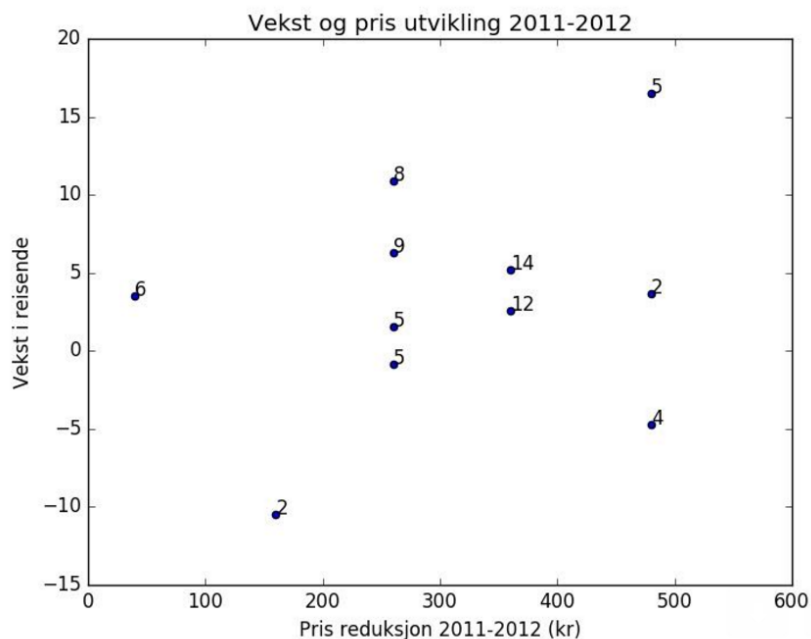
frekvensen. Elastisitetene observert ved bruk av ikke vektete observasjoner er spesielt høye og sier at en dobling av avgangsfrekvensen mer enn dobler antall reisende. Dette betyr at innføring av en ny togavgang ikke bare vil fylles opp i samme grad som tidligere avganger, men at alle togene også generelt får en høyere fyllingsgrad! Dette høres optimistisk ut. Tallene kan forsvares med at analysen ikke er tilstrekkelig detaljert. Det betyr at effekten av nye tog sannsynligvis er innbakt i elastisiteten til avgangsfrekvens. Antall ekstra seter som er tilbudt som følge av flere tog, sammenliknet med tidligere er heller ikke inkludert og har trolig en forholdsvis stor påvirkning på elastisiteten. Dersom denne variabelen skulle blitt mer detaljert for denne analysen kunne man eksempelvis forsøkt å skalere det ned mot økning i antall seter per område.

### **6.1.2 Reisetid**

For alle analysetilfeller er det estimert en signifikant sammenheng mellom reisetid og passasjerutvikling. Estimaten gir at en kjøretidsreduksjon på 1 minutt gir passasjervekst på rundt B til B + 2 % avhengig av effektrealiseringstilfellet. For reisetid er ikke variasjonen i elastisiteten så ulik ved bruk av vektete observasjoner eller ikke. Reisetidselastisiteten får også en høyere verdi under antakelsen om forsinket effektrealisering. Det er vanskelig å sammenlikne reisetidselastisitetene direkte opp mot tidligere litteratur da denne oppgaven bruker en elastisitet som fungerer på absolutte kjøretidsendringer. Det er heller gjort en indirekte sammenlikning opp mot tidligere litteratur hvor det er tatt utgangspunkt i tre ulike områder i analysen som har opplevd kjøretidsbesparelser. Kjøretidsbesparelsene er beregnet både relativt og absolutt utfra reiselengde og tidsbesparelse. Utfra dette er det beregnet estimert passasjervekst ved bruk av denne oppgavens beregnede elastisiteter og en tidligere relativ elastisitet for arbeidsreiser på 0,651 nevnt i Hamre (2002). Funnene fra testen indikerte at oppgavens elastisiteter gav høyere beregnet passasjervekst enn ved bruk av elastisiteten på 0,651. Det indikerer at en relativ elastisitet vedrørende reisetid for oppgaven ligger høyere enn 0,651, og at markedet dermed er også generelt mer sensitivt for reisetidsendringer enn hva som er estimert tidligere. Også her er det verdt å nevne at utelatte variabler, her hovedsakelig nye tog og mulighet for å tilby flere seter, har vært sterkt korrelert til reisetidsbesparelser og kan ha påvirket de høye reisetidselastisitetene som er fremkommet. Resultatene fra sammenlikningstesten er ikke presentert da disse kan benyttes til å finne estimatet på verdien B.

### 6.1.3 Pris

Ingen av analysene finner signifikante sammenhenger mellom prisreduksjoner og passasjervekst. Alle analysene gir høye p- verdier for elasticiteten. At en prisreduksjon ikke er funnet å ha signifikant påvirkning for antall reisende er noe overraskende. Prisutviklingen er sendt inn som et prissjokk i analysen. Generell økonomisk logikk forventer derfor et hopp i passasjertallet. Figur 30 undersøker hvorfor regresjonsanalysene ikke finner en signifikant sammenheng. Figuren gir oversikt over passasjerutviklingen for områdene utsatt for prisendringer. Vertikal akse viser vekst i reisende for et område fra året før innføring av nye priser til påfølgende år, altså 2011 til 2012. Horisontal akse viser prisreduksjon for området i absolutt krone verdi. Tallene ved de forskjellige punktene viser prosentandelen de enkelte observasjonene utgjør av totaltrafikken på Østlandet, som definerer påvirkningen punktene har for regresjonsanalysen (ved bruk av vektete observasjoner).



Figur 30: Pris-passasjervekst for ruterområdene 2011-2012

Figur 30 viser spredning i veksten for områdene. Tre av områdene har hatt negativ passasjerutvikling mellom 2011-2012. Resten ligger på positiv side. Noe av denne veksten vurderer regresjonsanalysen som en vekst forårsaket av andre faktorer enn et prissjokk, hovedsakelig befolkningsvekst mellom 2011 til 2012 da tilbudsendingene ikke kom før senere. Disse årsakene, samt at plottet ikke avdekker en tydelig trend mot

en generelt høyere vekst i reisende ved større prisreduksjoner gjør det vanskelig å estimere en signifikant sammenheng mellom pris og passasjerutvikling. Dette er trolig årsaken til at effekten av prisutviklingen ikke har fått et forventet estimat. Resultatene underbygger funnene fra Fouquet (2012) som forklarer at priselastisiteten er blitt lavere over tid og er avhengig av til hvilken grad et land er industrialisert. Paulley et al. (2006) nevner at priselastisiteten er dynamisk og får større påvirkning over tid. Dette kan til en viss grad kjennetegnes også i dette tilfellet hvor priskoeffisienten stiger under antakelsen av forsinket effektrealisering, men signifikante verdier uteblir. Paulley et al. (2006) nevner at arbeidsreiser er mindre prissensitive enn fritidsreiser. Denne analysen baseres på passasjerdata fra hverdager og dermed hovedsakelig arbeidsreiser. Dette kan også ha gitt utslag på resultatene for variabelen. Resultatene stemmer også noe overens med Wijeweera & Charles (2013) som viser til et lite prissensitivt togmarked i Australia. Hamre (2002) bemerker estimerer på økning i etterspørselen grunnet prisreduksjoner. Også Fearnley et al. (2012) nevner pris som en viktig etterspørselsdriver. Dette gjenspeiles ikke i resultatene som er fremkommet i dette tilfellet. Det er verdt å nevne at prisreduksjonen i togtilbudet har vært tilsvarende for busstilbudet. Dette betyr at togtilbudet ikke er blitt mer attraktivt sammenliknet med buss grunnet nytt sonesystem. Sammenliknet med biltrafikk har derimot togtilbudet blitt det gjennom prisreduksjonen, men analysen har ekskludert hvordan for eksempel bensinpriser, bomstasjoner og andel bileiere i befolkningen har utviklet seg. Om resultatene som er kommet frem er grunnet usikkerheter i rådataen og/eller om togmarkedet på Østlandet ikke er noe spesielt prissensitivt er ikke fullstendig avklart. Resultatene gir likevel ingen grunn til å tro at markedet er spesielt prissensitivt.

#### **6.1.4 Befolkningsvekst**

For alle analysetilfellene, med unntak av forsinkelsestilfelle 1 og 2 med ikke - vektete observasjoner, er det estimert signifikante sammenhenger mellom befolkningsvekst og passasjerutvikling. Elastisitetsverdiene varierer mellom D til D - 0,936 for analysene med vektete observasjoner, og mellom D - 1,788 til D - 2,298 for analysene med ikke - vektete observasjoner. Det er med andre ord vesentlige forskjeller mellom estimatene av elastisitetene ved bruk av vektete og ikke - vektete observasjoner. Ved å undersøke utviklingsplottene som er oppsummert i figur 28 og 29 er det tydelig at elastisitetene som er fremkommet ved ikke - vektete observasjoner virker å gjenspeile virkeligheten best da estimert utvikling og faktisk utvikling i forkant av ruteendringen er forholdvis lik

her. For analysene med bruk av vektete observasjoner er det tydelig å se en overestimert passasjervekst i forkant av ruteendringen ved studering av figur 28. Elastisitetsverdiene fremkommet ved vektete observasjoner kan også høres unaturlig høye ut, noe som virker å stemme med hva som fremkommer i utviklingsplottene. Disse verdiene forteller at en befolkningsvekst på ett prosent vil gi mer en ett prosent passasjervekst. En måte å forsvare elastisitetene på er likevel ved å tolke de som at det er  $D - 0,936$  til  $D$  ganger så sannsynlig at en tilflytter i et område benytter seg av toget enn en som alt er boende. I så tilfelle er trenden at tilflyttere benytter seg hyppigere av tog enn tidligere beboere. Om andelen av befolkningen som i utgangspunktet bruker tog er lavt trenger ikke en spesielt høy andel av tilflytterne å benytte toget for at tallene skal stemme. Alle elastisitetene, uavhengig av om observasjonene er vektet eller ikke, er likevel høye sammenliknet med elastisiteten fra Wijeweera & Charles (2013) liggende rundt 0,5.

#### **6.1.5 Modellene som helhet**

Regresjonsmodellen har gjennom alle analysene funnet signifikante sammenhenger for avgangsfrekvens og reisetid. Variabelen befolkningsvekst er funnet signifikant i fire av seks analyser. Pris er ikke blitt signifikant i analysene, noe som er forsøkt forklart i figur 30. Basert på endringene i forklaringsvariablene beskriver regresjonsmodellene mellom 80 – 90 % av observert endring i passasjerutviklingen. På generell basis er elastisitetstallene til de estimerte forklaringsvariablene høye sammenliknet med tidligere studier fra andre markeder. Dette kan være tilfellet, men også konsekvens av at effekten fra utelatte variabler blir innbakt i de estimerte koeffisientene. Ved å studere totalutviklingsplottene (figur 28 og 29) kan man se en forholdsvis jevn stigning i passasjerutviklingen grunnet befolkningsvekst. Dette indikerer at befolkningsveksten har utviklet seg i jevn takt i områdene inkludert i analysen. Effekten av utelatte variabler som også har hatt en jevn vekst kan dermed lettere bli en innbakt effekt i befolkningsveksten. Dette skjer da regresjonsanalysen ikke klarer å skille mellom effektene. Ved å inkludere variabler som dette er det sannsynlig å observere høy korrelasjon mellom variablene med lik utvikling og modellen ville ikke beskrevet situasjonen merkbart bedre. Bruttonasjonalprodukt og lønnsvekst kan være eksempler på slike variabler, med en trolig relativt jevn utvikling i analysens tidsrom. Dette kan også forklare befolkningsvekstens noe høye elastisiteter sammenliknet med rundt 0,5 fra Wijeweera & Charles (2013). Den store forskjellen mellom regresjonsmodellene med



vektede og ikke – vektede observasjoner er at befolkningsveksten virker overestimert for de vektede analysene, mens den er vesentlig lavere i de ikke - vektede analysene. I tillegg er elasticiteten for avgangsfrekvens høyere for analysene med ikke – vektede observasjoner. Totalutviklingsplottene viser også at analysene med ikke vektede observasjoner gjenspeiler virkeligheten best.

Utviklingen grunnet befolkningsvekst virker å være overestimert i tilfellene med vektede observasjoner. Årsaken til dette kan være at enkelte områder med stor trafikkandel har opplevd stor vekst uten tilbudsendringer og dermed påfører overestimerte elasticiteter for befolkningsvekst for resten av Østlandet som ikke har opplevd en tilsvarende vekst. Dersom det er mye støy i dataobservasjonene til de større områdene kan dette ha bidratt til overestimerte elasticiteter for befolkningsutviklingen, noe som er uheldig. Dette viser i så tilfelle at det er en risiko knyttet til å tildele enkelte observasjoner stor innflytelse på bekostning av andre, da man i praksis baserer analysen på færre observasjoner som virkelig "teller". Resultatene som er presentert er de som egner seg best med bakgrunn i datagrunnlaget som er levert inn. I totalutviklingsplottene kan man se at faktisk passasjerutvikling stiger relativt jevnt og bratt for årene etter 2012, i motsetning til den estimerte total utviklingen som er mer rykkete (dette gjelder spesielt ved tilfellene med umiddelbar effektrealisering). Estimert utvikling antar raskere effektrealisering ved endringer knyttet til tilbudsendringer enn hva som faktisk er observert. Dette forårsaker rykkene i veksten sammenliknet med faktisk vekst. I situasjonene som antar forsinkelser i effektrealiseringen er disse rykkene vesentlig mindre. I virkeligheten realiseres effekten gradvis over tid, derav en mer jevn stigning. Antakelsen om at kun halvparten av effekten realiseres innen første året (tilfelle 2) treffer best. Dette kan gjenspeiles i funnene til Fearnley & Bekken (2005) som finner at langtidselasticitetene er 1,84 ganger høyere enn kortidselasticitetene. Utfra dette kan man undre seg på om all effekt fra fullstendig implementert modell i 2015 er realisert. Når passasjerstatistikk for kommende år tilgjengeliggjøres er det spennende å se om grafen over faktisk utvikling viser en lik utvikling som vist årene etter 2012 eller om den flater mer ut.

Ekstrapoleringen av grafene for estimert utvikling viser antatt videre utvikling ved realisering av all resterende effekt. Dersom faktisk utvikling de kommende årene flates ut, slik som for årene før 2012, indikerer dette at all effekt av nytt tilbud er realisert.

## 6.2 Sammenlikning mot prognoser

Delkapittel 6.2 sammenlikner de beregnede effekter for de ulike analysetilfellene med prognosene for ruteendringen. Tabell 21 gir en oversikt over prognoser, resultater og avvik for tilfellene med vektete observasjoner. Tabell 22 gir tilsvarende informasjon ved bruk av ikke – vektete observasjoner. Relativt avvik er beregnet som resultatets relative forskjell til prognosen. Et negativt avvik betyr underestimert trafikkvekst. For tilfellene hvor det er antatt en forsinket effektrealisering er ikke all effekt fremkommet i 2013 og 2015. Verdiene med stjerne (\*) viser beregnede effekter ved bruk av elastisiteter som er estimert under analysene med forsinket effekt, men ved antakelse av at all effekt slår inn umiddelbart. Disse effektberegningene blir mye høyere og skulle ideelt blitt beregnet sammen med faktiske passasjertellinger for årene i etterkant, når all effekt er antatt være realisert. Beregningene er basert på likning 4.6. Det betyr at under beregning av effektene markert med stjerne (\*) blir både teller høyere og nevner lavere for likning 4.6, noe som gir store positive utslag i estimert effekt. Derfor gir ikke tallene et helt realistisk bilde da de ikke tar høyde for videre vekst uten tilbudsendringer.

Tabell 21: Sammenlikning av resultater og prognoser (vektet)

År	Verdi	Umiddelbar Realisering	Forsinkelses tilfelle 1	Forsinkelses tilfelle 2
2013	<i>Prognose</i>	5,2 %	5,2 %	5,2 %
	<i>Resultat</i>	8,7 %	8,5 % / 12,9 % *	7,8 % / 15,5%*
	<i>Avvik (absolutt)</i>	-3,5 %	-3,3 %	-2,6 %
	<i>Avvik (relativt)</i>	-67,3 %	-63,5 %	-50 %
2015	<i>Prognose</i>	20,9 %	20,9 %	20,9 %
	<i>Resultat</i>	17,5 %	21,8 % / 27,5 % *	23,4 % / 34,2 %*
	<i>Avvik (absolutt)</i>	3,4 %	-0,9 %	-2,5 %
	<i>Avvik (relativt)</i>	16,3 %	-4,3 %	-12,0 %

Tabell 22: Sammenlikning av resultater og prognoser (ikke – vektet)

År	Verdi	Umiddelbar Realisering	Forsinkelses tilfelle 1	Forsinkelses tilfelle 2
2013	<i>Prognose</i>	5,2 %	5,2 %	5,2 %
	<i>Resultat</i>	16,0 %	13,5 % / 21,5%*	11,4 % / 24,3 %*
	<i>Avvik (absolutt)</i>	-10,8 %	-8,3 %	-6,2 %
	<i>Avvik (relativt)</i>	- 207,7 %	- 159,6 %	-119,2 %
2015	<i>Prognose</i>	20,9 %	20,9 %	20,9 %
	<i>Resultat</i>	39,4 %	40,6 %/ 55,7%*	39,3 % /64,8 % *
	<i>Avvik (absolutt)</i>	-18,5 %	-19,7 %	-18,4 %
	<i>Avvik (relativt)</i>	-88,5 %	-94,3 %	-88,0 %

Både tabell 21 og 22 viser at ruteplanens effekt er underestimert for alle tilfellene i 2013. Det samme gjelder for 2015, med unntak av analysen med antakelse om umiddelbar effektrealisering og bruk av vektete observasjoner. Det er verdt å merke seg at prognosene har tatt utgangspunkt i umiddelbar effektrealisering. Det er vanskelig å avgjøre hvilket av analysetilfellene som er mest sammenliknbart. Tilfellet med umiddelbar effektrealisering er lettest da både prognosene og analysen da har like utgangspunkt.

Uavhengig av sammenlikningsgrunnlag viser alle tilfellene med bruk av vektete observasjoner små absolutte avvik og forholdsvis like beregnede effekter for de ulike årene liggende mellom 7-9 % for 2013 og 17-24 % for 2015. Dette er relativt overensstemmende med prognosenes beregnede effekt av ruteendringen, men tar ikke høyde for noe ytterligere effektrealisering i etterkant av 2015 for forsinkelsestilfellene. Ser vi på tallene med stjerne ser vi at avvikene blir vesentlig større.

For analysetilfellene ved bruk av ikke – vektete observasjoner er det derimot større avvik. De beregnede effekter av ruteplanendringen er større som følge av høyere elasticiteter for variablene knyttet til tilbudsendingene og ligger her rundt 11-16 % for

2013 og rundt 40 % for 2015. Dette gir store avvik mellom prognoser og resultater og enda større ved tallene med stjerne.

Tabell 23 viser forskjellen mellom observert totalvekst for Østlandet og antatt totalvekst siden 2010 for de ulike årene. Observert totalvekst er basert på passasjertellingene og viser en vekst på 20,5 % og 50,7 % for 2013 og 2015. Trafikkprognosene hadde ingen referanse i totale reiser for 2010, men for 2008. Figur 8 viser at prognosene har estimert en totalvekst på 6,1 % mellom 2008-2013 uten tilbudsendringer. Dette gir en årlig vekst på 1,2 %<sup>3</sup>. Med bakgrunn i dette er det beregnet et referansetall for antall reisende i 2010, med utgangspunkt i reisetallet for 2008. Dette er brukt til å beregne prognosenes antatte vekst for årene 2013 og 2015 siden 2010.

Tabell 23: Observert totalvekst mot antatt totalvekst fra prognoser

År	2013	2015
Observert vekst	20,5 %	50,7 %
Prognose vekst	9,0 %	29,6 %
Avvik (absolutt)	-11,5 %	-21,1 %
Avvik (relativt)	-127,0 %	-71,3 %

Tabell 23 viser at prognosenes totale passasjervekst siden 2010 er langt lavere enn observert vekst. Hva som er årsaken til den ytterligere veksten vil forbli noe usikkert. Med utgangspunkt i en antakelse av at den store veksten som er observert i etterkant av 2012 kun er forårsaket av ruteendringen, er resultatene beregnet med ikke – vektete observasjoner forholdsvis overensstemmende med resultatene fra tabell 23. Utviklingsplottene i figur 28 underbygger også denne antakelsen og indikerer at ruteendringen har gitt en langt høyere effekt enn antatt. Tilbudsforbedringen fra ruteendringen må da trolig inkludere flere faktorer enn redusert reisetid og økt avgangsfrekvens - her stort sett nyere tog og økt setekapasitet.

<sup>3</sup> Årlig vekst =  $\sqrt[5]{1 + 0,061} = 1,2 \%$

Med utgangspunkt i at de beregnede effekter med vektete observasjoner stemmer må den store veksten etter 2012 også delvis kunne forklares av andre faktorer som ikke går under ruteendringen, altså ytre faktorer som for denne oppgaven hovedsakelig går under befolkningsvekst og pris.

Flyvbjerg et al. (2005) nevner i sitt studie at ni av ti trafikkprognoser i jernbaneprosjekter er overestimerte og at trafikkprognoser ikke er blitt bedre over et tidsintervall på 30 år. Også Pickrell (1990) trekker frem observasjoner av store overestimeringer i trafikkprognoser for jernbaneprosjekter. Resultatene som er fremkommet i tabell 21,22 og 23 viser ingen underbyggelse av dette. En fellesnevner for resultatene i oppgaven, med unntak av tilfellet med umiddelbar effektrealisering og vektete observasjoner, er at ruteendringens effekter er høyere enn antatt. For analysene med ikke vektete observasjoner er det estimert at faktisk effekt er godt over 100 % og 80 % høyere enn antatt effekt for henholdsvis 2013 og 2015. For de vektete analysene er disse avvikene vesentlig lavere, men effekten er likevel beregnet å være høyere enn antatt. Derfor er det vanskelig å si noe eksakt om prognosenes treffsikkerhet da resultatene viser store variasjoner i avvikene. Alle resultatene er avvikende fra Flyvbjerg et al. (2005) sin konklusjon om en gjennomsnittlig overestimering på 106 %. Resultatene med vektete observasjoner kan gjenkjennes i funn fra Button (2009) og Nicolaisen & Driscoll (2014) som viser indikasjoner på lavere overestimeringer og mer treffsikre prognoser.

Det er verdt å nevne at Flyvbjerg et al. (2005) ser på omfattende infrastrukturprosjekter som gir store endringer i trafikktilbudet. Høyhastighetsjernbane som blir studert av Beria & Grimaldi (2011) og Givoni & Dobruszkes (2013) er andre enorme infrastrukturprosjekter og kan defineres som helt nye trafikkprodukter. Denne oppgaven studerer mindre endringer. Dette gir grunnlag til å forvente lavere usikkerheter i prognosene da man bør forvente at større prosjekter også medfører større usikkerhet. I tillegg bør det igjen presiseres at nevnte studier tar direkte utgangspunkt i selve infrastrukturendringen, mens denne oppgaven ser på resultatet av infrastrukturendringen – ruteplanendringen. Pickrell (1990) sammenlikner prosjekters resultater mot prognoser gjort før endelig prosjektalternativ er valgt, og stiller spørsmål ved om overestimering av trafikkprognoser er bevisst for å få gjennomslag på ønskede prosjekter. For denne oppgaven er derimot prognosene gjort i etterkant av

prosjektvalget for å gi NSB realistiske forventninger rundt trafikkvekst. Prognosene er derfor gjort uten interesser om å få gjennomslag for infrastrukturprosjektet. Med bakgrunn i tidligere litteratur sin poengtering av trafikkprognosers store usikkerheter og tydelige trend til å overestimere fremtidig trafikk, kan det generelt oppsummeres med at oppgavens resultater avviker spesielt fra antakelsen om overestimering. De beregnede effekter viser seg å være høyere enn antatt med varierende avvik til prognosene. Prognosens treffsikkerhet er vanskelig å avgjøre med bakgrunn i resultatene, men prognoseavvikene fra de vektete analysene er tilsynelatende lave i forhold til eldre studier (Pickrell 1990; Flyvbjerg et al. 2005) og følger forbedringstrenden som poengteres i nyere studier (Button 2009; Nicolaisen & Driscoll 2014). For analysene med ikke – vektete obserasjoner viser resultatene ikke tilsynelatende forbedringer i avvikene mellom prognoser og resultater. Disse analysene er også mest overlappende med virkelig utvikling ved å studere totalutviklingsplottene.

### **6.3 Et kritisk blikk på oppgaven**

Alle oppgaver har svakheter og begrensninger ved seg. Denne oppgaven har gjort antakelser og generaliseringer som har vært nødvendige for gjennomføring av oppgaven, men ikke nødvendigvis gjenspeiler virkeligheten. Nevneverdige kritiske antakelser og generaliseringer blir beskrevet her. Det er viktig å være klar over begrensninger ved egen oppgave, men viktigere at det er nevnt hvilke antakelser og vilkår resultatene er bygget på ved fremtidig bruk.

Først er det viktig å bemerke at regresjonsmodellen i oppgaven har gjort en generalisering av totaltrafikken på Østlandet med rushtrafikk og grunnrutetrafikk sett under ett. Det er ikke gjort analyser som har studert utviklingen i ulike trafikktyper, slik det er gjort i trafikkprognosene av Dybdahl et al. (2011). Resultatene kan derfor ikke knyttes opp mot lokale områder. De estimerte elastisitetene passer best for totalutviklingen på Østlandet, men ikke nødvendigvis for totalutviklingen i de enkelte områder. Det er heller ikke skilt mellom elastisiteter for arbeids- og fritidsreiser, eller for lengden på reisene som tidligere litteratur skiller på (Jernbaneverket 2015; Fearnley et al. 2012; Paulley et al. 2006; Hamre 2002). Regresjonsanalysene gjennomføres under forutsetning om at alle områder som inngår responderer likt på endringer i togtilbudet. For eksempel antas en dobling av avgangsfrekvens mellom Lillestrøm og Oslo S å ha lik påvirkning som en tilsvarende endring mellom Drammen og Kongsberg.

Regresjonsmodellenes R – kvadrat på over 0,80 for alle tilfeller virker å være godt med tanke på rådataens usikkerhet. Det er viktig å være klar over at modellen gjør en forenkling av en sammensatt og komplisert virkelighet. Det er valgt å inkludere forklaringsvariabler som lettest kan knyttes til lokale observasjoner. Derfor utelukker modellen enkelte variabler som har betydning. Noen av disse variablene er beskrevet i 4.1.2 og er hovedsakelig utelatt da de er vanskelige å skalere mot ulike områder/stasjoner. Metodehåndboken til Jernbaneverket (2015) nevner faktorer som tilbringer tid til stasjon og antall stopp. Andre faktorer som markedsføring for togtrafikk og miljøengasjement har trolig også en påvirkning i folks reisevalg. Likevel er disse faktorene utelatt fra analysen. Det er også verdt å nevne at analysen ikke tar høyde for hvordan fordelingen av togavgangene for de ulike områdene er gjort over timen. En mer attraktiv avgangsfordeling blir derfor ikke inkludert, kun relativ vekst i antall avganger. Det er også kun tatt høyde for endring i antall reiser, ikke antall personkilometer. Det er nødvendig å være klar over at prognosene baserer sine tall på økning i totaltrafikk for alle dager i et helt år. Det betyr at det tas høyde for helgereiser og reiser i feriene. Denne oppgaven tar utgangspunkt i økninger basert på tall fra hverdagstrafikken. For at prognosene og oppgavens resultater skal være ideelt sammenlignbare må det derfor antas at ruteendringen har medført tilsvarende relativ vekst i helge- og feriereiser som det er funnet for hverdagsreiser i denne oppgaven. Den grove skaleringen og usikkerheten i rådataen bør igjen nevnes. Et bredere datagrunnlag ville gitt en mer robust analyse med flere observasjoner, samt et bedre overblikk over trafikkutviklingen med mindre sårbarhet for variasjoner. Olsson et al. (2015) er en av flere som nevner at prosjekters effekt ofte realiseres over tid. Forsinkelsestilfellene er brukt som et forsøk på å se hvordan resultatene påvirkes ved en effektrealisering over tid. Det er brukt forsinkelseskoeffisienter med verdier basert på antakelser, men Fearnley & Bekken (2005) gir en pekepinn på hvor dynamiske priselastisitetene er ved å nevne et forholdstall mellom langtidselastisiteter og korttidselastisiteter på 1,84. Et bedre skalert datagrunnlag ville gitt mer rom for å kunne forsøke å visualisere den faktiske forsinkelsen i realiseringen av effekter på et mer detaljert nivå gjennom totalutviklingsplottene. Resultatene som er fremkommet for de ulike forsinkelsestilfellene bygger også på forutsetningen om at det ikke har vært større endringer i forkant av 2010, hvor effekten av disse realiseres i passasjertallene etter 2010.

## 6.4 Generell diskusjon

Denne oppgaven kan sees på som en ex post evaluering av ruteplanendringens effekt på en viktig prestasjonsparameter, derav en prestasjonsmålt evaluering (Olsson et al. 2010). Trafikkprognosene kan tolkes som en ex ante evaluering og er brukt som referansegrunnlag for resultatene av ex post evalueringen. Olsson et al. (2010) nevner viktigheten ved innsamling av data på samme tidspunkt av året for å unngå endringer grunnet sesongvariasjoner. Dette er hensynstatt i oppgaven da alle tellinger som er brukt er gjort på høsten. Olsson et al. (2015) og Olsson et al. (2010) diskuterer bruken av referansepunkter og referansegrunnlag for evalueringsprosesser. Denne evalueringen kan sies å ha to ulike referansegrunnlag. Resultatene av ruteendringen er først presentert opp mot hvordan utviklingen ville vært uten endringer i togtilbudet (tabell 8,11,14,16,18,20). Deretter sammenliknes resultatene opp mot forventede resultater (tabell 21 og 22).

Med utgangspunkt i førstnevnte referansegrunnlag viser resultatene en positiv totalendring grunnet tilbudsendringene for alle år etter 2012. Med utgangspunkt i dette skulle man kunne konkludere med et vellykket prosjekt med solid passasjervekst. Også med bakgrunn i referansegrunnlaget hvor resultatene sammenliknes med forventede resultater kan det generelt konkluderes med at prosjektet har vært vellykket. Likevel viser også oppgavens resultater hvordan evalueringens resultater lett kan påvirkes ved valg av referansegrunnlag og referansetidspunkt. Dersom evalueringen kun hadde blitt gjort for året 2015 med bruk av vektete observasjoner og antakelse om umiddelbar effektrealisering ville resultatet være noe skuffende i forhold til hva som var forventet utfall. Utførelsen av flere analyser har gitt ulike resultater. Dette viser hvor sårbare og påvirkelige evalueringer er for forutsetningene som legges til grunn, noe som samsvarer med hva som poengteres i Olsson et al. (2015) og Olsson et al. (2010). Dette kan trolig også overføres til arbeidet med trafikkprognoser. Store prognoseusikkerheter kan med sannsynlighet begrunnes i at beregningene er basert på feil antakelser og forutsetninger, til tross for at analysemodellene er gode og er beregnet av dyktige analytikere. Dette poengteres også av Sanko et al. (2013) og Aarhaug et al. (2013) som viser til at usikkerheter i trafikkprognoser i stor grad kan relateres til usikkerheter ved forutsetningene og grunnlaget til en trafikkberegningmodell.

Mer empirisk kunnskap rundt tematikken kan forhåpentligvis gi fremtidige prognoser bedre indikasjoner på hva som er de rette forutsetningene for trafikkberegninger.



Den generelle trenden i faktisk passasjerutvikling er en brattere passasjervekst etter 2012 enn i forkant. Regresjonsmodellen har gitt sterke elastisiteter for variablene knyttet til en tilbudsending, altså avgangsfrekvens og reisetid, sammenliknet med elastisiteter funnet i tidligere litteratur. Oppgavens resultater viser at total trafikkvekst på Østlandet har vist seg å være langt sterkere enn hva som ble antatt i prognosene. Fearnley et al. (2012) nevner i sin analyse at det største potensialet i togmarkedet på Østlandet kan realiseres ved reisetidsforbedringer og hyppigere avganger rundt de største byene. Dette er gjort gjennom ruteendringen og oppgavens resultater virker å være i tråd med Fearnley et al. (2012) sine anbefalinger.

Kort oppsummert er det påvist god trafikkvekst på Østlandet, både som følge av tilbudsendinger og andre ytre faktorer. Likevel er de oppløftende resultatene ingen garanti for tilsvarende resultater i fremtiden. Logikken bak dette er at dersom det går lang tid mellom større endringer i togtilbudet vil etterspørselen etter et forbedret togtilbud gradvis bygges opp i takt med befolkningsveksten, uten at noe blir gjort. Dette kan kalles en undertrykt etterspørsel, eller et uutløst markedspotensial. Dersom dette har vært tilfelle på Østlandet før 2012 vil resultatene ved et endelig forbedret tilbud bli en suksess som ikke nødvendigvis gjentas i samme grad senere så sant etterspørselen ikke bygges opp i samme grad på ny.

## **6.5 Videre arbeid**

Gjennom oppgavens tidsperiode er det kommet frem ideer til videre arbeid rundt temaet. Først må det nevnes at denne analysen kun har grunnlag i et lite innblikk av togtrafikken på Østlandet over flere år, og med bakgrunn i dette forsøker å trekke store slutninger. En mer detaljert analyse av ruteendringens effekt som både kan inkludere flere forklaringsvariabler og et bredere datagrunnlag over flere år i etterkant av endringen er ønskelig. Når automatiske tellinger blir kvalitetssikret vil man kunne opparbeide et langt bredere datagrunnlag som bør benyttes for å evaluere effekten av fremtidige infrastruktur prosjekter på Østlandet, samt undersøke usikkerheten ved de manuelle tellingene. Fremgangsmåten som er brukt i denne analysen vil fortsatt være relativt overførbar til fremtidige analyser med liknende fokus.

Denne oppgaven har studert trafikkutviklingen forårsaket av ruteendringen i forbindelse med forbedret infrastruktur, og sammenliknet med trafikkprognoser for

selve ruteendringen. Et interessant videre arbeid er å ta mer direkte utgangspunkt i trafikkprognosene i forkant av infrastrukturutbyggelsen mellom Asker - Lysaker, slik Pickrell (1990) og Flyvbjerg et al. (2005) har gjort i sine studier. Givoni & Dobruszkes (2013) studerer hva som er overført trafikk og induisert trafikk ved innføring av høyhastighetsjernbane. Et tilsvarende studie for denne situasjonen, med en kartlegging av hva som er overført og induisert trafikk som følge av ruteendringen hadde vært av interesse.

I forbindelse med denne oppgaven benytter jeg anledningen, likt tidligere studier, til å nevne fordelene med å danne et enda større forskningsgrunnlag rundt ex post analyser av jernbaneprosjekter, med spesielt fokus på hvordan trafikkutviklingen har utviklet seg i forhold til antatt. Med bakgrunn i oppgavens togmarked på Østlandet oppfordres det til å gjøre lignende evalueringer i etterkant av Follobanens ferdigstillelse, samt andre kommende endringer i togtilbudet. Denne oppgaven har sett på en viktig prestasjonsparameter, passasjerutviklingen. Andre prestasjonsmålte ex post evalueringer som tar for seg andre temaer ved ruteendringen er også nyttig for å få et mer helhetlig bilde av endringen. Eksempler på temaer kan være materiellutnyttelse, personellplanlegging og utvikling i kundetilfredshet.

## 7. Konklusjon

Gjennom oppgavens analyser og diskusjon er det fremkommet tilstrekkelig informasjon til å kunne besvare problemstillingen og de tilhørende forskningsspørsmål som har vært gjenstand for diskusjon. Oppgavens resultater er basert på et datagrunnlag som er sårbart og inneholder usikkerhet. I tillegg er det gjort antakelser under arbeidet som er beskrevet i oppgaven. Dette må hensynstas ved bruk av oppgavens funn.

Tilbudsendringene for 2013 og 2015 har gitt et sterkt løft i den totale passasjerutviklingen på Østlandet. Det er utført totalt seks analyser med forskjellige antakelser om hvor raskt effekten ved endringer i togtilbudet realiseres, og ulike antakelser rundt om analysens observasjoner er tildelt et vektall eller ikke. De beregnede effekter har variert mye utifra hvilke antakelser analysen er basert på. Generelt har bruk av vektall til observasjonene gitt lavere beregnede effekter for ruteendringen. Analysene med vektall har beregnet at antall reisende er 7-9 % høyere enn det ville vært uten endringer for 2013 og 17-24 % for 2015. Ved å utelukke vektall har de beregnede effekter variert mellom 11 - 16 % for 2013 og ligget rundt 40 % for 2015. For analysene som inkluderer forsinkelser i effektrealiseringen er ikke all effekt realisert for årene 2013 og 2015. Endelig effekt av tilbudsendringene vil dermed i virkeligheten være ytterligere større enn disse tallene. Ved å sammenlikne grafer mellom estimert passasjerutvikling og faktisk passasjerutvikling for de ulike analysene virker analysene uten vektall for observasjonene å stemme best.

Variasjonen i resultatene viser også til hvor sårbare evalueringresultatene er for analysens forutsetninger, et funn som det kan trekkes paralleller til i trafikkberegninger. Utførelsen av flere analyser med ulike forutsetninger er gjort for å få et bredest mulig evaluering grunnlag.

Fra analysenes inkluderte variabler er det generelt funnet at passasjerutviklingen påvirkes signifikant av endringer i avgangsfrekvens, reisetid og befolkningsvekst. Disse variablene viser seg å ha sterk påvirkning for passasjerutviklingen og elastisitetene som er fremkommet virker å være høyere enn estimer funnet i tidligere litteratur og forskning (Hamre 2002; Paulley et al. 2006; FitzRoy & Smith 1995; Wijeweera & Charles 2013; Fearnley et al. 2012). Resultatene viser også til at elastisitetsverdiene stiger

vesentlig under antakelsen av en forsinkelse i realiseringen av endringenes effekter, noe som er i tråd med Fearnley & Bekken (2005) sine beregninger. Dette gir grunnlag for å tro at endringer i togtilbudet får større påvirkning på sikt. Oppgaven har ekskludert enkelte variabler som antas å ha betydning for passasjerutviklingen. Det har med stor sannsynlighet gitt utslag på elastisitetsverdiene som er estimert og de resultatene som er fremkommet. Trolig har for eksempel utelukkelsen av nytt togmateriell og muligheten for å tilby flere seter som følge av dette, vært med på å trekke elastisiteten for avgangsfrekvens og reisetid opp. Det er ikke fremkommet noen signifikant sammenheng mellom prisreduksjoner og passasjerutvikling.

Samsvaret mellom oppgavens resultater og prognosene tilknyttet ruteendringen kan anses å være godt for analysene som har benyttet vektall og dermed stemme med nyere litteratur rundt prognoseusikkerhet (Button 2009; Nicolaisen & Driscoll 2014). For analysene hvor det er utelatt vektall er avvikene mellom resultatene og prognosene høye. Det er derfor vanskelig å si noe eksakt om prognosenes treffsikkerhet, men en rød tråd er at de beregnede effekter av ruteendringen er høyere enn forventet i prognosene. Dette er motstridene med den generelle trenden om overestimerte trafikkprognoser som er nevnt av tidligere studier (Flyvbjerg et al. 2005; Pickrell 1990; Givoni & Dobruszkes 2013; Beria & Grimaldi 2011). Avvikene på prognoseusikkerhetene varierer etter hvilke antakelser som er gitt rundt effektrealiseringen. For analysene hvor det antas en effektrealisering over tid er ikke fremtidig økning i etterkant av 2015 tatt høyde for i beregningene. Dette betyr at underestimeringene av ruteplanens effekt er noe høyere enn hva som er nevnt i oppgaven.

## 8. Litteraturliste

- Aarhaug, J., Caspersen, E., Fearnley, N., Farideh, R., Ranheim, P. & Steinsland, C. (2013). *Dokumentasjonsrapport: Inkrementell etterspørselsmodell*. Oslo: Transportøkonomisk institutt. s. 1-16. Hentet 25. Mars 2016 fra, <https://www.toi.no/getfile.php?mmfileid=35022>
- Austeng, K., Midtbø, J., Jordanger, I., Magnussen, O. & Torp, O. (2005). *Usikkerhetsanalyse Kontekst og grunnlag*. Trondheim: Concept Rapport (10). s. 60-75. Hentet 18. April 2016, fra <http://www.ntnu.no/documents/1261860271/1262010703/Concept%2010%20Usikkerhetsanalyse%20-%20Kontekst%20og%20grunnlag.pdf>
- Beria, P. & Grimaldi, R. (2011). *An Early Evaluation of Italian High Speed Projects*. Milano: *TeMa*. 4(3). s. 15- 26. Hentet 3. Mars 2016, fra [http://www.stagniweb.it/doc/Beria\\_AltaVelocita.pdf](http://www.stagniweb.it/doc/Beria_AltaVelocita.pdf)
- Brooks, C. (2008). *Introductory Econometrics for Finance*. 2. Utg. Cambridge: Cambridge University Press.
- Button, K. (2009). *Transit forecasting accuracy: Ridership forecasts and capital cost estimates*. Transportation and Economic Development Centre. s. 1-22. Hentet 20. Februar 2016, fra [https://www.researchgate.net/profile/Kenneth\\_Button/publication/268056254\\_Transit\\_Forecasting\\_Accuracy\\_Ridership\\_Forecasts\\_and\\_Capital\\_Cost\\_Estimates/links/54c668a80cf256ed5a9e63db.pdf](https://www.researchgate.net/profile/Kenneth_Button/publication/268056254_Transit_Forecasting_Accuracy_Ridership_Forecasts_and_Capital_Cost_Estimates/links/54c668a80cf256ed5a9e63db.pdf)
- Dybdahl, H., Dyb, V., Ekhaugen, T. & Homleid, T. (2011). *Trafikkberegninger, Østlandet 2012-2016*. Vista Analyse. (Epost til Torbjørn Stigen 22.01.2016)
- Fearnley, N., Bekken, J.T. (2005). *Etterspørselseffekter på kort og lang sikt: en litteraturstudie i etterspørselsdynamikk*. Oslo: Transportøkonomisk institutt. s. 27-33. Hentet 7. Mai 2016, fra <https://www.toi.no/getfile.php?mmfileid=1841>
- Fearnley, N., Aarhaug, J., Denstadli, J., Engebretsen, Ø. & Vågane, L. (2012). *Tilbuds og etterspørselssammenhenger i jernbanesektoren*. Oslo: Transportøkonomisk institutt. s. 1- 15. Hentet 25. Februar 2016, fra <https://www.toi.no/getfile.php?mmfileid=32204>
- FitzRoy, F. & Smith, I. (1995). *The demand for railtransport in European countries*. *Elsevier*. 2(3). s.153 – 158. Hentet 29. April 2016, fra <http://www.sciencedirect.com/science/article/pii/0967070X95967457>
- Flyvbjerg, B., Holm, M. & Buhl, S. (2005). *How (In)accurate are demand forecasts in public works projects*. Chicago: *American Planning Association*. 71(2). s.131 – 144. Hentet 1. Mars 2016, fra <http://flyvbjerg.plan.aau.dk/Traffic91PRINTJAPA.pdf>

- Fouquet, R. (2012). *Trends in income and price elasticities of transport demand (1850 - 2010)*. Elsevier. 50(1). s. 62-71. Hentet 2. Februar 2016, fra <http://www.sciencedirect.com/science/article/pii/S0301421512002066>
- Givoni, M. & Dobruszkes, F. (2013). *A Review of Ex-Post Evidence for Mode Substitution and Induced Demand Following the introduction of High speed Rail*. Taylor and Francis Online. 31(6). s. 1-24 s. Hentet 18 mars 2016, fra <http://www.tandfonline.com/doi/pdf/10.1080/01441647.2013.853707>
- Hamre, T. (2002). *NTM 5 – Den nasjonale persontransportmodellen – Versjon 5*. Oslo: Transportøkonomisk institutt. s. 21-26. Hentet 29. April 2016, fra <https://www.toi.no/getfile.php?mmfileid=7059>
- Jernbaneverket (2015). *Metodehåndbok: Samfunnsøkonomiske analyser for jernbanen 2015*. Oslo: Jernbaneverket. s. 56 – 68. Hentet 24. Januar 2016, fra <http://www.jernbaneverket.no/contentassets/a3e2c5d145d04eaa95c04e6ebe036e44/metodehandbok-2015.pdf>
- Litman, T. (2015). *Generated Traffic and Induced Travel. Implications for Transport Planning*. Victoria Transport Policy Institute. s.1- 4. Hentet 2. Februar 2016, fra <http://www.vtpi.org/gentraf.pdf>
- Løvås, G. (2004). *Statistikk for universiteter og høyskoler*. 2. Utg. Oslo: Universitetsforlaget.
- Nicolaisen, M. & Driscoll, P. (2014). *Ex- Post Evaluations of Demand Forecast Accuracy: A Literature Review*. Taylor and Francis Online. 34(4). s. 540 – 557. Hentet 16. Mars 2016, fra <http://dx.doi.org/10.1080/01441647.2014.926428>
- Nordli, O. (2015). *Dette er Høvik stasjon*. Hentet 15. Januar 2016, fra <http://www.jernbaneverket.no/Prosjekter/prosjekter/Hovik-stasjon/Prosjektartikler/Dette-er-Hovik-stasjon/>
- NSB (2007). *Markedsstrategisk produksjonsplan sluttrapport: Nytt rutekonsept for østlandet*. Oslo: NSB persontog og plan (Epost til Torbjørn Stigen 18.01.2016)
- NSB (2011). *Konsept og tilbudsbeskrivelse: Nytt rutetilbud for østlandet*. Oslo: NSB persontog og plan. (Epost til Torbjørn Stigen 15.01.2016)
- NSB (2016). *Kart over stasjoner i Norge: Rutekart lokaltog østlandet*. Hentet 17. Februar 2016, fra <https://www.nsb.no/reisemal/kart-over-togstasjoner-i-norge>
- Olsson, N. (2005). *Hvordan tror vi at det blir? Effektvurderinger av store offentlige prosjekter*. Concept Rapport (7). s. 13-16. Hentet 25. Januar 2016, fra <https://www.ntnu.no/documents/1261860271/1262010703/Concept%207%20Effektvurderinger.pdf>
- Olsson, N. & Veiseth, M. (2011). *Jernbanetrafikk*. Trondheim: Tapir akademiske forlag.

- Olsson, N., Krane, H., Veiseth, M. & Rolstadås, A. (2010). *Influence of reference points in ex post evaluations of rail infrastructure projects*. Elsevier. 17(4). s. 251-258. Hentet 23. februar 2016, fra <http://www.sciencedirect.com/science/article/pii/S0967070X1000020X>
- Olsson, N., Halse, A., Heggland, P., Killi, M., Kooji, R., Landmark, A.,... Østli, V. (2015). *Punktlighet i jernbanen -Hvert sekund teller*. Oslo: Sintef Akademiske forlag. s. 1-77s. Hentet 20. Februar 2016, fra <http://www.sintefbok.no/Product.aspx?sectionId=65&productId=1055&categoryId=39>
- Paulley, N., Balcombe, R., Mackett, R., Titheridge, H., Preston, J., Wardman, M.,...White, P. (2006). *The demand for public transport: The effects of fares, quality of service, income and car ownership*. Leeds: White Rose Research. s. 1-11. Hentet 13. Februar 2016, fra [http://eprints.whiterose.ac.uk/2034/1/ITS23\\_The\\_demand\\_for\\_public\\_transport\\_UPLOADABLE.pdf](http://eprints.whiterose.ac.uk/2034/1/ITS23_The_demand_for_public_transport_UPLOADABLE.pdf)
- Pickrell, D. (1990). *Urban Rail Transit Projects: Forecast versus actual ridership and Cost*. U.S. Department of Transportation. s. 1-15. Hentet 20 mars 2016, fra [http://people.plan.aau.dk/~mortenn/misc/alon/Pickrell%20\(1990\)%20Urban%20Rail%20Transit%20Projects%20Forecast%20Versus%20Actual%20Ridership%20and%20Cost.pdf](http://people.plan.aau.dk/~mortenn/misc/alon/Pickrell%20(1990)%20Urban%20Rail%20Transit%20Projects%20Forecast%20Versus%20Actual%20Ridership%20and%20Cost.pdf)
- Rencher, A. (2002). *Methods of multivariate analysis*. 2. Utg. John Wiley & Sons. Hentet 17. januar 2016, fra <http://tx.shu.edu.tw/~PurpleWoo/Literature/!DataAnalysis/Methods%20of%20Multivariate%20Analysis.pdf>
- Ruter (2016). *Sonekart*. Hentet 02. Mars 2016, fra <https://ruter.no/kjop-billett/billetter-og-priser/soner/>
- Samset, K. (2008). *Prosjekt i tidligfasen Valg av konsept*. 2. Utg. Trondheim: Tapir Akademiske forlag.
- Sanko, N., Morikawa, T. & Nagamatsu, Y. (2013). *Post-project evaluation of travel demand forecasts: Implications from the case of a Japanese railway*. Elsevier. 27(1). s. 209-217. Hentet 25. Mars 2016, fra <http://www.sciencedirect.com/science/article/pii/S0967070X13000279>
- Silverman, D. (2011). *Interpreting qualitative data*. 4 Utg. London: Sage Publications Ltd.
- SSB (2016). *Folkemengde og befolkningsendringar*. Hentet 15. Mars 2016, fra <https://www.ssb.no/statistikkbanken/SelectTable/hovedtabellHjem.asp?KortNavnWeb=folkemengde&CMSSubjectArea=befolkning&StatVariant=&PLanguage=0&checked=true>

Svingheim, N. (2008). *Historisk oversikt jernbanen i Norge*. Hentet 19. Januar 2016, fra <http://www.jernbaneverket.no/Jernbanen/Historie1/Historisk-oversikt-jernbanen-i-Norge/>

Svingheim, N. (2011). *Jernbanekart*. Hentet 19. Januar 2016, fra <http://www.jernbaneverket.no/Jernbanen/Jernbanekart/>

Wijeweera, A. & Charles, M. (2013). *An Empirical Analysis of the Determinants of Passenger Rail Demand in Melbourne, Australia*. *Elsevier*. 43(3). s. 249-264. Hentet 25. Mars 2016, fra <http://www.sciencedirect.com/science/article/pii/S0313592613500312>



## **9. Liste over vedlegg**

**Vedlegg 1:** Residualplot og korrelasjonsmatriser for analysene

**Vedlegg 2:** Priser for ruterområdet

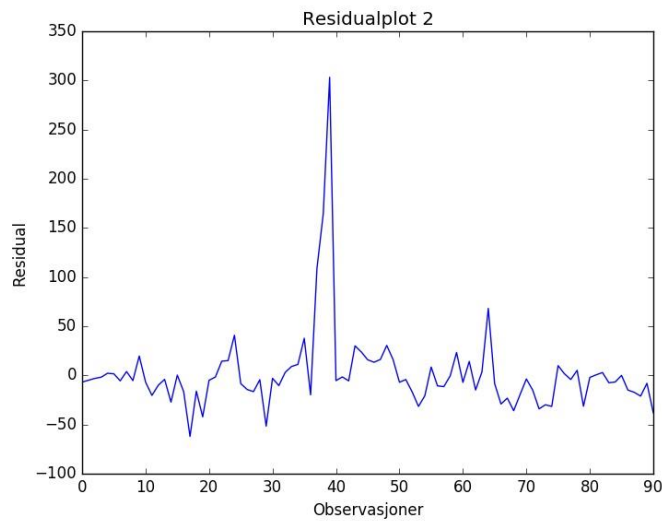
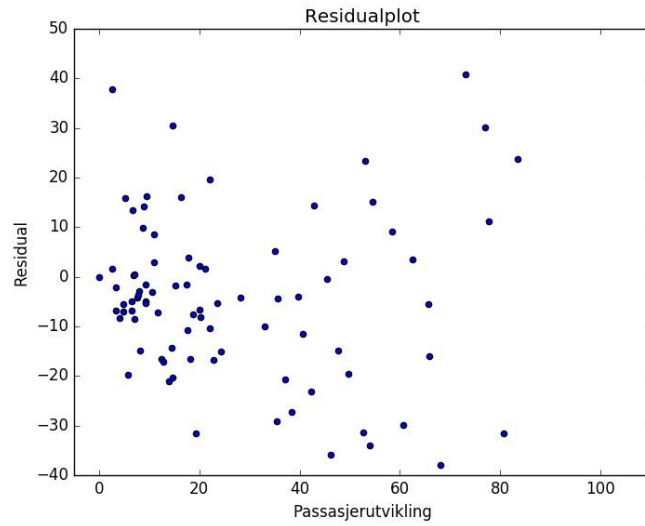
**Vedlegg 3:** Oversikt over rutetabeller for beregning av kjøretid

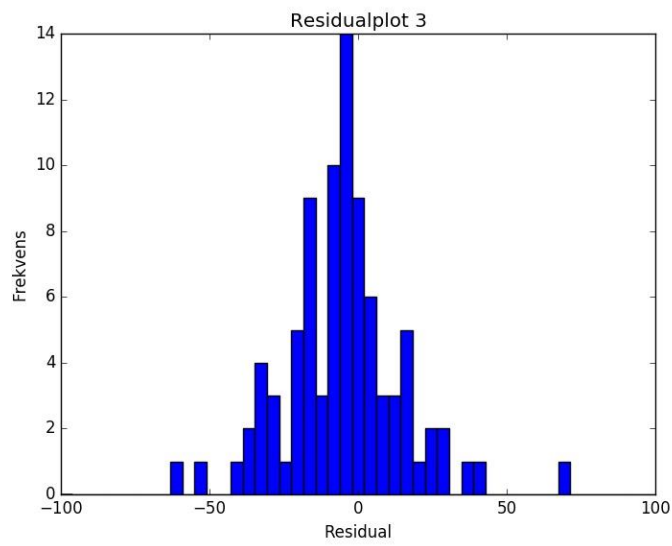
**Vedlegg 4:** Passasjertellingsprogram

**Vedlegg 5:** Analyseprogram

# Vedlegg 1 – Residualplot og korrelasjonsmatriser for analysene

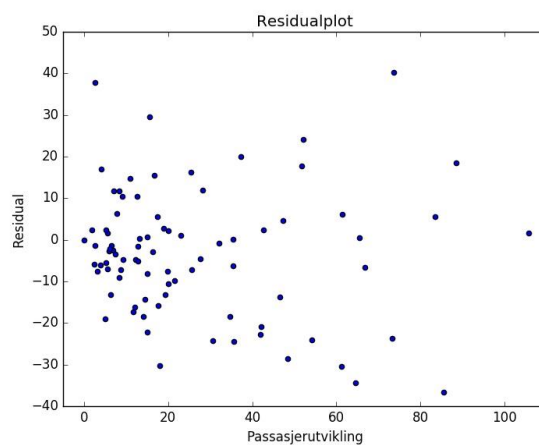
## Analyse med umiddelbar realisering av effekt (vektet)

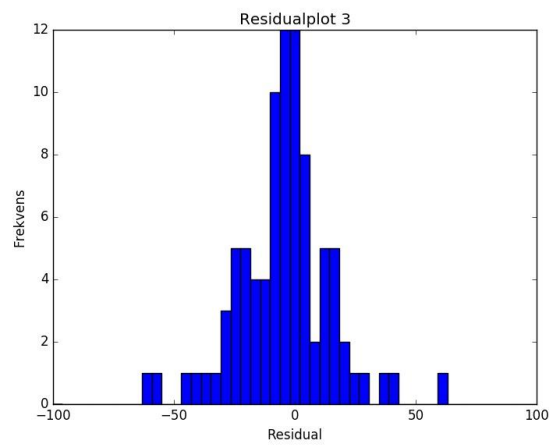
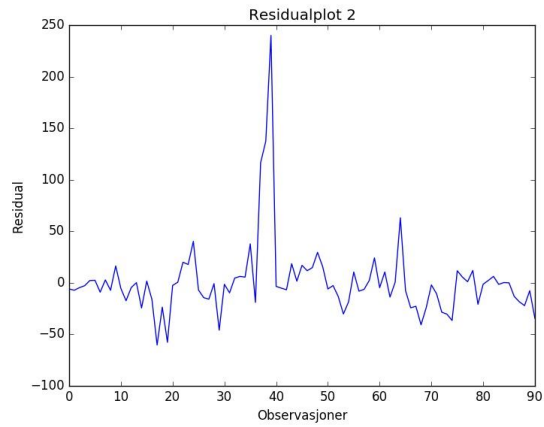




Variabler	Avgangsfrekvens	Reisetid	Pris	Befolkningsvekst
Avgangsfrekvens	1	0,203	0,331	0,608
Reisetid	0,203	1	0,0134	0,326
Pris	0,331	0,0134	1	0,674
Befolkningsvekst	0,608	0,326	0,674	1

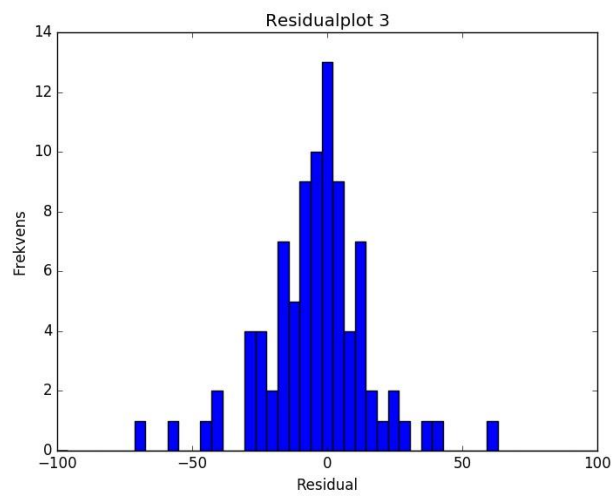
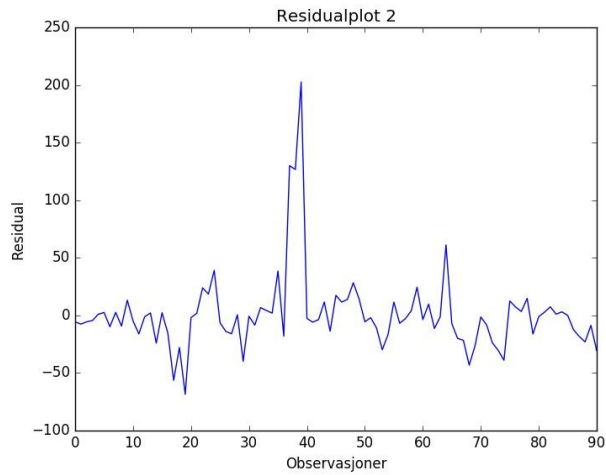
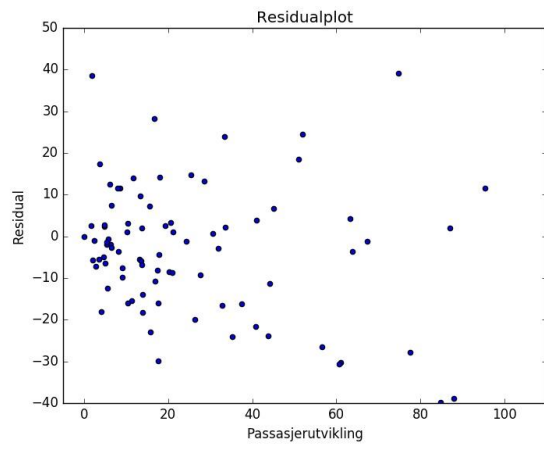
### Analyse med forsinkelsestilfelle 1 (vektet)





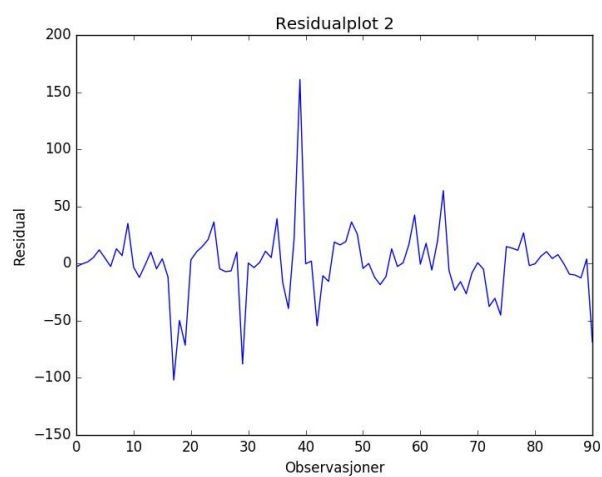
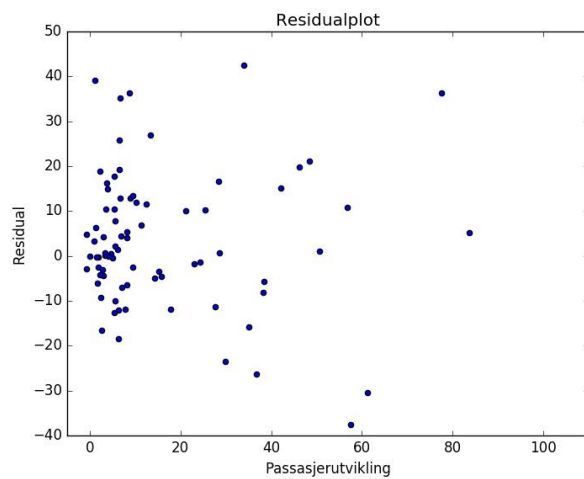
Variabler	Avgangsfrekvens	Reisetid	Pris	Befolkningsvekst
Avgangsfrekvens	1	0,227	0,379	0,605
Reisetid	0,227	1	0,0476	0,348
Pris	0,378	0,047	1	0,730
Befolkningsvekst	0,604	0,348	0,730	1

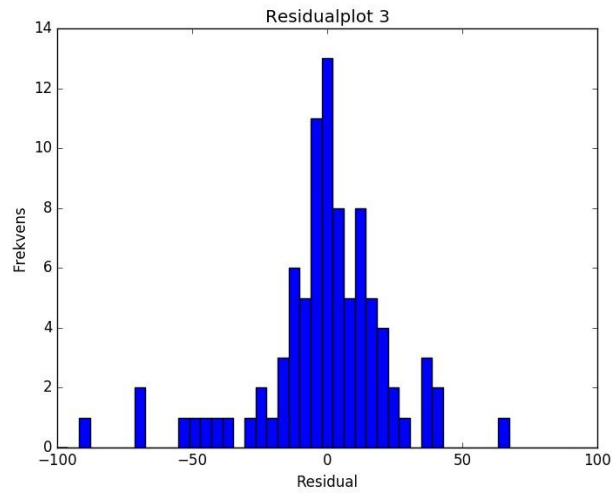
## Analyse med forsinkelsestilfelle 2 (vektet)



Variabler	Avgangsfrekvens	Reisetid	Pris	Befolkningsvekst
Avgangsfrekvens	1	0,239	0,390	0,593
Reisetid	0,238	1	0,068	0,356
Pris	0,390	0,068	1	0,749
Befolkningsvekst	0,593	0,356	0,749	1

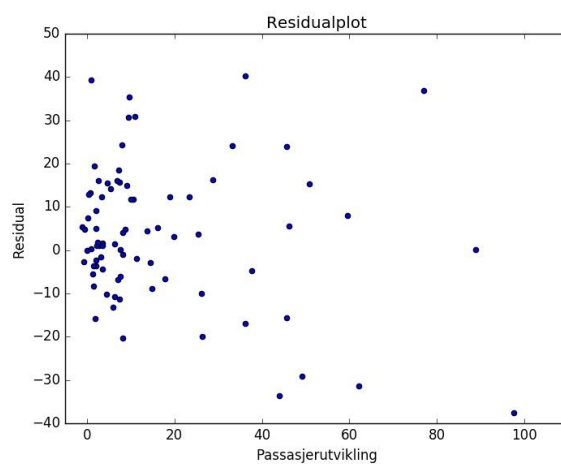
### Analyse med umiddelbar realisering av effekt (ikke - vektet)

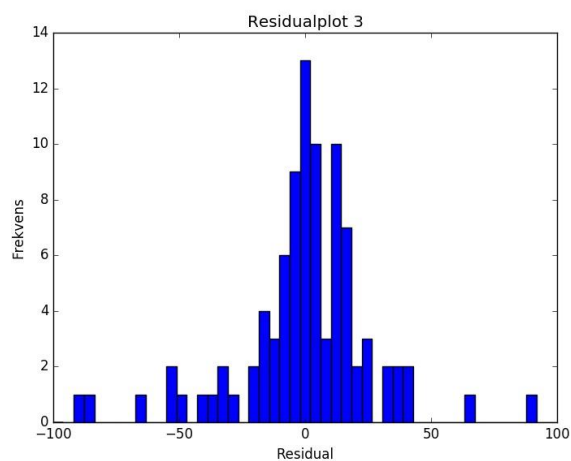
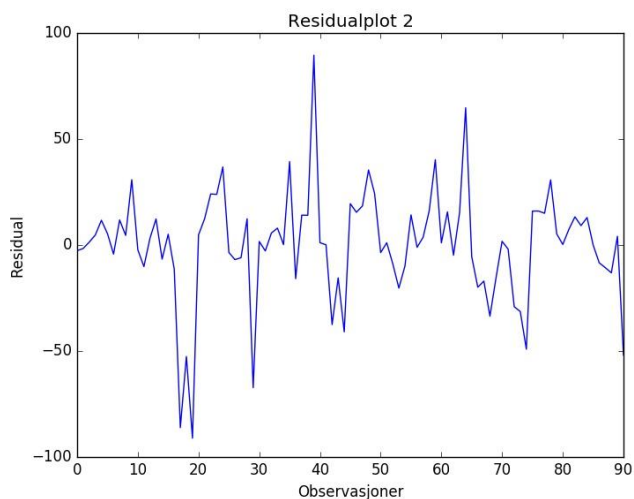




Variabler	Avgangsfrekvens	Reisetid	Pris	Befolkningsvekst
Avgangsfrekvens	1	0,202	0,331	0,608
Reisetid	0,202	1	0,0134	0,326
Pris	0,331	0,0134	1	0,674
Befolkningsvekst	0,608	0,326	0,674	1

### Analyse med forsinkelsestilfelle 1 (ikke - vektet)

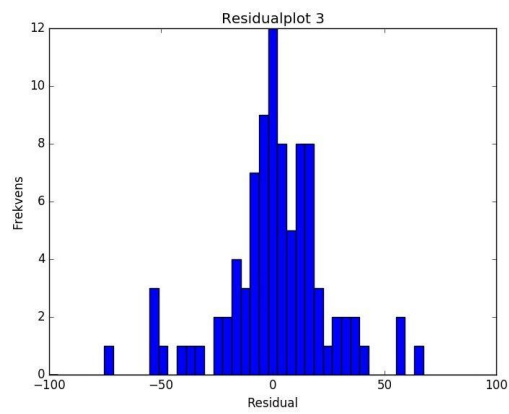
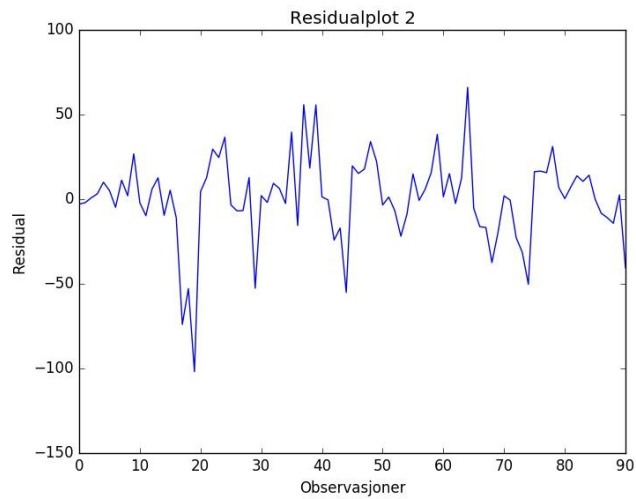
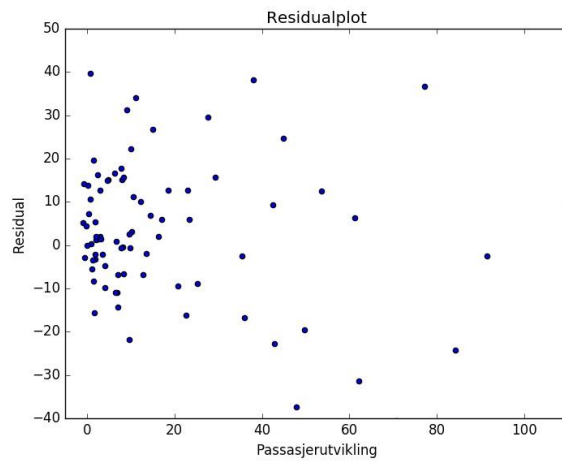




Variabler	Avgangsfrekvens	Reisetid	Pris	Befolkningsvekst
Avgangsfrekvens	1	0,227	0,379	0,605
Reisetid	0,227	1	0,048	0,348
Pris	0,379	0,048	1	0,730
Befolkningsvekst	0,605	0,348	0,730	1



## Analyse med forsinkelsestilfelle 2 (ikke - vektet)



<b>Variabler</b>	<b>Avgangsfrekvens</b>	<b>Reisetid</b>	<b>Pris</b>	<b>Befolkningsvekst</b>
Avgangsfrekvens	1	0,239	0,390	0,593
Reisetid	0,239	1	0,068	0,356
Pris	0,390	0,068	1	0,749
Befolkningsvekst	0,593	0,356	0,749	1

## Vedlegg 2: Priser i ruterområdet

### Priser etter ny sonestruktur

Billettslag Ruter	Priser 1. februar 2014				
	1 sone	2 soner	3 soner	4 soner	5+ soner
Enkeltbillett ombordkjøpt	50	70	90	110	130
	25	35	45	55	65
Enkeltbillett forhåndskjøpt	30	50	70	90	110
	15	25	35	45	55
24-timersbillett	90	140	190	190	190
	45	70	95	95	95
7-dagersbillett	230	420	610	610	610
	115	210	305	305	305
30-dagersbillett	650	1160	1670	1670	1670
30-dagersbillett B/U	325	425	525	525	525
30-dagersbillett H	325	580	835	835	835
30-dagers Studentbillett	390	700	1010	1010	1010
365-dagersbillett	6500	11600	16700	16700	16700
365-dagersbillett	3250	5800	8350	8350	8350

Billettslag Ruter	Priser 1. februar 2013				
	1 sone	2 soner	3 soner	4 soner	5+ soner
Enkeltbillett ombordkjøpt	50	70	90	110	130
	25	35	45	55	65
Enkeltbillett forhåndskjøpt	30	50	70	90	110
	15	25	35	45	55
24-timersbillett	80	130	180	180	180
	40	65	90	90	90
7-dagersbillett	220	400	580	580	580
	110	200	290	290	290
30-dagersbillett	630	1100	1570	1570	1570
30-dagersbillett B/U	320	390	460	460	460
Studentbillett	380	660	940	940	940
365-dagersbillett	6300	11000	15700	15700	15700

Billettslag Ruter	Priser 1. februar 2012				
	1 sone	2 soner	3 soner	4 soner	5+ soner
Enkeltbillett ombordkjøpt	50	70	90	110	130
	25	35	45	55	65
Enkeltbillett forhåndskjøpt	30	50	70	90	110
	15	25	35	45	55
Smartbillett	27	-	-	-	-
	13,50	-	-	-	-
Kupongkort (pris per reise)	27	45	63	81	99
	13,50	22,50	31,50	40,50	49,50
24-timersbillett	75	125	175	175	175
	38	63	88	88	88
7-dagersbillett	220	390	560	560	560
	110	195	280	280	280
30-dagersbillett	620	1070	1520	1520	1520
30-dagersbillett B/U/H	310	380	450	450	450
Studentbillett	380	650	920	920	920
365-dagersbillett	6200	10700	15200	15200	15200

Billettslag Ruter	Priser 2. oktober 2011				
	1 sone	2 soner	3 soner	4 soner	5+ soner
Enkeltbillett ombordkjøpt	44	64	84	104	124
	22	32	42	52	62
Enkeltbillett forhåndskjøpt	28	48	68	88	108
	14	24	34	44	54
Smartbillett	27	-	-	-	-
	13,50	-	-	-	-
Kupongkort (pris per reise)	27	45	63	81	99
	13,50	22,50	31,50	40,50	49,50
24-timersbillett	75	125	175	175	175
	38	63	88	88	88
7-dagersbillett	220	390	560	560	560
	110	195	280	280	280
30-dagersbillett	590	1040	1490	1490	1490
30-dagersbillett B/U/H	295	370	400	400	400
Studentbillett	355	625	895	895	895
365-dagersbillett	5900	10400	14900	14900	14900

## Priser ved gammel sonestruktur

### Månedskort, røde

Soner	2		3		4		5		6		7		8		9		10	
	U	P	U	P	U	P	U	P	U	P	U	P	U	P	U	P	U	P
1.2.2011	980	-	1160	-	1330	-	1440	-	1560	-	1680	-	1810	-	1930	-	2000	-
1.2.2010	930	-	1090	-	1250	-	1360	-	1470	-	1590	-	1710	-	1830	-	1940	-
1.3.2009	1065	905	1180	1065	1355	1225	1450	1315	1575	1430	1700	1540	1820	1650	1940	1760	2060	1870
1.8.2008	1065	905	1180	1065	1355	1225	1450	1315	1575	1430	1700	1540	1820	1650	1940	1760	2060	1870
1.1.2007*	1005	905	1180	1065	1355	1225	1450	1315	1575	1430	1700	1540	1820	1650	1940	1760	2060	1870
1.1.2006	-	-	1145	1025	1315	1185	1395	1265	1520	1365	1645	1480	1765	1595	1880	1710	1995	1810
1.1.2005	-	-	1105	990	1275	1150	1330	1205	1465	1315	1585	1430	1700	1550	1810	1660	1925	1735
1.1.2004	-	-	1070	965	1235	1120	1290	1175	1420	1280	1535	1395	1650	1510	1775	1620	1865	1690
1.1.2003	-	-	1050	945	1210	1095	1265	1150	1390	1255	1505	1370	1620	1480	1740	1585	1830	1655
1.1.2001	-	-	965	870	1125	1015	1180	1070	1300	1175	1415	1285	1525	1395	1645	1495	1730	1565
1.6.2000	-	-	915	835	1065	975	1120	1030	1230	1130	1340	1235	1445	1340	1555	1440	1635	1505
1.1.1999	-	-	-	715	-	835	-	880	-	965	-	1050	-	1130	-	1210	-	1265
1.1.1998	-	-	-	690	-	805	-	850	-	930	-	1015	-	1090	-	1170	-	1220
1.1.1997	-	-	-	675	-	790	-	835	-	910	-	995	-	1070	-	1145	-	1195
1.1.1996	-	-	-	655	-	765	-	810	-	885	-	965	-	1040	-	1110	-	1160

## **Vedlegg 3 – Oversikt over rutetabeller for beregning av kjøretid**

### **2010**

NSB Regiontog i Norge, gjelder i periode 13.06. – 11.12.10

NSB Lokaltog i Oslo-området, gjelder i perioden 13.12.09-12.06.10

### **2011**

NSB Regiontog i Norge, gjelder i periodene 12.06-25.06. og 08.08-10.12.11.

NSB Lokaltog i Oslo-området, gjelder i periodene 12.06-25.06. og 08.08.-10.12.11.

### **2012**

NSB Regiontog i Norge, gjelder i periodene 10.06-23.06. og 06.08-08.12.12.

NSB Lokaltog i Oslo-området, gjelder i periodene 10.06-23.06. og 06.08.-08.12.12.

### **2013**

NSB Østlandet, gjelder i periodene 09. juni 2013 – 23.juni 2013 og 19. august 2013 – 14. desember 2013

### **2014**

NSB Østlandet, gjelder i perioden 15. desember 2013 – 14. juni 2014

### **2015**

NSB Østlandet, gjelder i perioden 14. desember 2014 – 13. juni 2015

## Vedlegg 4: Passasjertellingsprogram i python

```
# -*- coding: cp1252 -*-
import xlswriter as xl
import xlrd
#Definerer områder, stasjoner og tognummer som er aktuelle for hvert enkelt område.
stasjoner = []
tittel = []
passasjerer = []
tognummerintervall = []
av_pa_totalutvikling = [] # 1 betyr område med, 0 betyr område ikke med.

tittel.append("Asker_Oslo_lokal")
stasjoner.append(["Asker", "Høn", "Vakås", "Hvalstad", "Billingstad", "Slependen"])
tognummerintervall.append([[2100,2290],[7100,7300]]) # 2200-2299 for 2013-14, 2100-2199
#2012, 2806 ekstra tog for noen år
av_pa_totalutvikling.append([1])

tittel.append("Asker_Oslo_knutepunkt")
stasjoner.append(["Asker", "Sandvika"])
tognummerintervall.append([[300,400],[500,600],[800,900],[1000,1200],[1600,1700],[1900,2000],[2290,2300]])
av_pa_totalutvikling.append([1])

tittel.append("Lillestrom_Oslo_lokal")
stasjoner.append(["Bryn", "Alna", "Nyland", "Grorud", "Haugenstua", "Høybråten", "Lørenskog", "Hanaborg", "Fjellhamar", "Strømmen", "Sagdalen", "Lillestrøm"]) # Tatt ut oslo s
tognummerintervall.append([[2100,2299],[1500,1520],[7100,7300]]) # 1500-1520 innsatstog
lokal
av_pa_totalutvikling.append([1])

tittel.append(u"Lillestrøm_Oslo_knutepunkt")
stasjoner.append(["Lillestrøm"])
tognummerintervall.append([[300,400],[500,600],[800,900],[1000,1100],[1600,1700]])
av_pa_totalutvikling.append([1])

tittel.append("Oslo_Ski_lokal")
stasjoner.append(["Nordstrand", "Ljan", "Hauketo", "Holmlia", "Rosenholm", "Kolbotn", "Solbråtan", "Myrvoll", "Greverud", "Oppegård", "Vevelstad", "Langhus"])
tognummerintervall.append([[2700,2899]])
av_pa_totalutvikling.append([1])

tittel.append("Ski_Sonsveien")
stasjoner.append(["Ås", "Vestby", "Sonsveien", "Kambo"])
tognummerintervall.append([[1100,1199],[101,102]])
av_pa_totalutvikling.append([1])

tittel.append("Ski_Mysen")
stasjoner.append(["Kråkstad", "Skotbu", "Tomter", "Knapstad", "Spydeberg", "Askim", "Slitu", "Mysen"])
tognummerintervall.append([[100,199],[1900,2000]]) # 100 - 199 før 2013, 1900-2000 fom.
2013
av_pa_totalutvikling.append([1])

tittel.append("Drammen_Kongsberg")
```

```
stasjoner.append(["Kongsberg","Darbu","Vestfossen","Mjøndalen","Hokksund","Gulskogen"])
tognummerintervall.append([[500,599],[1600,1649]])
av_pa_totalutvikling.append([1])
```

```
tittel.append("Drammen - Asker")
stasjoner.append(["Drammen","Lier","Brakerøya"])
tognummerintervall.append([[300,400],[800,900],[1600,1700],[2100,2300],[500,600],[1900,2000]])
av_pa_totalutvikling.append([1])
```

```
tittel.append("Drammen_Skien")
stasjoner.append(["Skien","Porsgrunn","Larvik","Sandefjord","Torp","Stokke","Tønsberg","Skoppum","Holmestrand","Sande"])
tognummerintervall.append([[300,352],[800,899]])
av_pa_totalutvikling.append([1])
```

```
tittel.append("Tangen_Lillehammer")
stasjoner.append(["Lillehammer","Moelv","Brumunddal","Hamar","Stange","Tangen"])
tognummerintervall.append([[300,355],[810,840]]) #810 -830 kom egentlig inn etter 2014.
av_pa_totalutvikling.append([1])
```

```
tittel.append("Skarnes_Kongsvinger")
stasjoner.append(["Kongsvinger","Skarnes"])
tognummerintervall.append([[1000,1099]])
av_pa_totalutvikling.append([1])
```

```
tittel.append("Asker_Spikkestad")
stasjoner.append(["Spikkestad","Røyken","Hallenskog","Heggedal","Gullhella","Bondivann"])
tognummerintervall.append([[2200,2300],[1100,1200],[2100,2180],[7100,7300]]) #2200-2300 gjelder for 2013-2015, 1100-1200 fra 2010-2012. 2104 er tillegg for 2011 og 2012. 2179 tillegg for 2012
av_pa_totalutvikling.append([1])
```

```
tittel.append(u"Lillestrøm_Eidsvoll")
stasjoner.append(["Eidsvoll","Eidsvoll Verk","Gardermoen"])
tognummerintervall.append([[300,380],[800,899],[500,599],[1601,1610]])
av_pa_totalutvikling.append([1])
```

```
tittel.append(u"Lillestrøm_Årnes")
stasjoner.append(["Årnes","Bodung","Haga","Auli","Rånåsfoss","Blaker","Sørumsand","Svingen","Fetsund","Nerdrum","Tuen"])
tognummerintervall.append([[1000,1099],[2100,2110]])
av_pa_totalutvikling.append([1])
```

```
tittel.append("Moss-Halden")
stasjoner.append(["Moss","Rygge","Råde","Fredrikstad","Sarpsborg","Halden"])
tognummerintervall.append([[100,200],[380,400]])
av_pa_totalutvikling.append([1])
```

```
tittel.append("Ski")
stasjoner.append(["Ski"])
tognummerintervall.append([[100,200],[380,400],[1100,1200],[1900,2000]])
av_pa_totalutvikling.append([1])
```

```
tittel.append(u"Lillestrøm_Dal")
stasjoner.append(["Leirsund","Lindeberg","Jessheim","Nordby","Hauerseter","Dal"])
```



```
tognummerintervall.append([[500,600],[1600,1700],[2180,2200]])
av_pa_totalutvikling.append([1])
```

```
#Åpner filer
```

```
#2015
```

```
femten = open('høst_2015.txt','r')
```

```
#2014
```

```
fjorten = open('høst2014.txt','r')
```

```
#2010
```

```
ti_av_vest = open('høst2010_av_vest.txt','r')
```

```
ti_pa_vest = open('høst2010_på_vest.txt','r')
```

```
ti_av_ost = open('høst2010_av_øst.txt','r')
```

```
ti_pa_ost = open('høst2010_på_øst.txt','r')
```

```
#2011
```

```
elleve_av_vest = open('høst2011_av_vest.txt','r')
```

```
elleve_pa_vest = open('høst2011_på_vest.txt','r')
```

```
elleve_av_ost = open('høst2011_av_øst.txt','r')
```

```
elleve_pa_ost = open('høst2011_på_øst.txt','r')
```

```
#2012
```

```
tolv_av_vest = open('høst2012_av_vest.txt','r')
```

```
tolv_pa_vest = open('høst2012_på_vest.txt','r')
```

```
tolv_av_ost = open('høst2012_av_øst.txt','r')
```

```
tolv_pa_ost = open('høst2012_på_øst.txt','r')
```

```
#2013
```

```
tretten_av_vest_L = open('høst2013_av_vest_L.txt','r')
```

```
tretten_pa_vest_L = open('høst2013_på_vest_L.txt','r')
```

```
tretten_av_ost_L = open('høst2013_av_øst_L.txt','r')
```

```
tretten_pa_ost_L = open('høst2013_på_øst_L.txt','r')
```

```
tretten_av_vest_R = open('høst2013_av_vest_R.txt','r')
```

```
tretten_pa_vest_R = open('høst2013_på_vest_R.txt','r')
```

```
tretten_av_ost_R = open('høst2013_av_øst_R.txt','r')
```

```
tretten_pa_ost_R = open('høst2013_på_øst_R.txt','r')
```

```
####FUNKSJONER FOR Å LESE TEKSTFORMATENE TIL DATA ###
```

```
def lesformat2(inputfile):
```

```
    stations = {}
```

```
    line = inputfile.readline()
```

```
    line = inputfile.readline()
```

```
    while line != "":
```

```
        parts = line.split("\t")
```

```
        if parts[3] not in stations:
```

```
            stations[parts[3]] = {}
```

```
            tognummerdict = stations[parts[3]]
```

```
            tognummerdict[parts[1]]=[int(parts[8].strip()),int(parts[9].strip())]
```

```
            line = inputfile.readline()
```

```
    return stations
```

```
def lesformat1(inputfile):
```

```
    trainnumber = {}
```

```
    line = inputfile.readline()
```

```
    parts = line.split("\t")
```

```
    for i in range(2,len(parts)):
```

```
        trainnumber[i]=parts[i].strip()
```

```

stations = {}
line = inputfile.readline()

while line != "":
    parts = line.split("\t")
    if parts[0] not in stations and parts[0].strip() != "":
        stations[parts[0]] = {}

    for i in range(2, len(parts)):

        if parts[i].strip() != "":
            tognummerdict = stations[parts[0]]
            tognummerdict[trainnumber[i]] = int(parts[i].strip())

    line = inputfile.readline()

return stations

### LESER FILER FOR ALLE ÅR ###

av_pa_femten = lesformat2(femten)

av_pa_fjorten = lesformat2(fjorten) #14

av_vest_ti = lesformat1(ti_av_vest) #10
pa_vest_ti = lesformat1(ti_pa_vest)

av_ost_ti = lesformat1(ti_av_ost)

pa_ost_ti = lesformat1(ti_pa_ost)

av_vest_elleve = lesformat1(elleve_av_vest) #11
pa_vest_elleve = lesformat1(elleve_pa_vest)

av_ost_elleve = lesformat1(elleve_av_ost)

pa_ost_elleve = lesformat1(elleve_pa_ost)

av_vest_tolv = lesformat1(tolv_av_vest) #12
pa_vest_tolv = lesformat1(tolv_pa_vest)

av_ost_tolv = lesformat1(tolv_av_ost)

pa_ost_tolv = lesformat1(tolv_pa_ost)

av_vest_tretten_L = lesformat1(tretten_av_vest_L) #13
pa_vest_tretten_L = lesformat1(tretten_pa_vest_L)

av_ost_tretten_L = lesformat1(tretten_av_ost_L)

pa_ost_tretten_L = lesformat1(tretten_pa_ost_L)

av_vest_tretten_R = lesformat1(tretten_av_vest_R)

pa_vest_tretten_R = lesformat1(tretten_pa_vest_R)

```

```
av_ost_tretten_R = lesformat1(tretten_av_ost_R)
```

```
pa_ost_tretten_R = lesformat1(tretten_pa_ost_R)
```

```
### LAGER LØKKER SOM SLÅR SAMMEN ÅR 2010 - 2013 TIL SAMME FORMAT SOM  
AV PÅ 2014 ###
```

```
# 2010
```

```
av_pa_tjueti = {}
```

```
for key in av_vest_ti.keys():  
    if key not in av_pa_tjueti:  
        av_pa_tjueti[key]={}  
    for key2 in av_vest_ti[key].keys():  
        if key2 not in av_pa_tjueti[key]:  
            av_pa_tjueti[key][key2]=[-1,-1]  
            av_pa_tjueti[key][key2][1] = av_vest_ti[key][key2]
```

```
for key in pa_vest_ti.keys():  
    if key not in av_pa_tjueti:  
        av_pa_tjueti[key]={}  
    for key2 in pa_vest_ti[key].keys():  
        if key2 not in av_pa_tjueti[key]:  
            av_pa_tjueti[key][key2] = [-1,-1]  
            av_pa_tjueti[key][key2][0] = pa_vest_ti[key][key2]
```

```
for key in av_ost_ti.keys():  
    if key not in av_pa_tjueti:  
        av_pa_tjueti[key]={}  
    for key2 in av_ost_ti[key].keys():  
        if key2 not in av_pa_tjueti[key]:  
            av_pa_tjueti[key][key2] = [-1,-1]  
            av_pa_tjueti[key][key2][1] = av_ost_ti[key][key2]
```

```
for key in pa_ost_ti.keys():  
    if key not in av_pa_tjueti:  
        av_pa_tjueti[key]={}  
    for key2 in pa_ost_ti[key].keys():  
        if key2 not in av_pa_tjueti[key]:  
            av_pa_tjueti[key][key2] = [-1,-1]  
            av_pa_tjueti[key][key2][0] = pa_ost_ti[key][key2]
```

```
#2011
```

```
av_pa_tjueelleve = {}
```

```
for key in pa_vest_elleve.keys():  
    if key not in av_pa_tjueelleve:  
        av_pa_tjueelleve[key] = {}  
    for key2 in pa_vest_elleve[key].keys():  
        if key2 not in av_pa_tjueelleve[key]:  
            av_pa_tjueelleve[key][key2] = [-1,-1]  
            av_pa_tjueelleve[key][key2][0] = pa_vest_elleve[key][key2]
```

```

for key in av_vest_elleve.keys():
    if key not in av_pa_tjueelleve:
        av_pa_tjueelleve[key] = {}
    for key2 in av_vest_elleve[key].keys():
        if key2 not in av_pa_tjueelleve[key]:
            av_pa_tjueelleve[key][key2] = [-1,-1]
            av_pa_tjueelleve[key][key2][1] = av_vest_elleve[key][key2]

for key in av_ost_elleve.keys():
    if key not in av_pa_tjueelleve:
        av_pa_tjueelleve[key] = {}
    for key2 in av_ost_elleve[key].keys():
        if key2 not in av_pa_tjueelleve[key]:
            av_pa_tjueelleve[key][key2] = [-1,-1]
            av_pa_tjueelleve[key][key2][1] = av_ost_elleve[key][key2]

for key in pa_ost_elleve.keys():
    if key not in av_pa_tjueelleve:
        av_pa_tjueelleve[key] = {}
    for key2 in pa_ost_elleve[key].keys():
        if key2 not in av_pa_tjueelleve[key]:
            av_pa_tjueelleve[key][key2] = [-1,-1]
            av_pa_tjueelleve[key][key2][0] = pa_ost_elleve[key][key2]

```

#2012

```
av_pa_tjuetolv = {}
```

```

for key in av_vest_tolv.keys():
    if key not in av_pa_tjuetolv:
        av_pa_tjuetolv[key] = {}
    for key2 in av_vest_tolv[key].keys():
        if key2 not in av_pa_tjuetolv[key]:
            av_pa_tjuetolv[key][key2] = [-1,-1]
            av_pa_tjuetolv[key][key2][1] = av_vest_tolv[key][key2]

```

```

for key in pa_vest_tolv.keys():
    if key not in av_pa_tjuetolv:
        av_pa_tjuetolv[key] = {}
    for key2 in pa_vest_tolv[key].keys():
        if key2 not in av_pa_tjuetolv[key]:
            av_pa_tjuetolv[key][key2] = [-1,-1]
            av_pa_tjuetolv[key][key2][0] = pa_vest_tolv[key][key2]

```

```

for key in av_ost_tolv.keys():
    if key not in av_pa_tjuetolv:
        av_pa_tjuetolv[key] = {}
    for key2 in av_ost_tolv[key].keys():
        if key2 not in av_pa_tjuetolv[key]:
            av_pa_tjuetolv[key][key2] = [-1,-1]
            av_pa_tjuetolv[key][key2][1] = av_ost_tolv[key][key2]

```

```

for key in pa_ost_tolv.keys():
    if key not in av_pa_tjuetolv:
        av_pa_tjuetolv[key] = {}
    for key2 in pa_ost_tolv[key].keys():

```

```

    if key2 not in av_pa_tjuetolv[key]:
        av_pa_tjuetolv[key][key2] = [-1,-1]
    av_pa_tjuetolv[key][key2][0] = pa_ost_tolv[key][key2]

# 2013
av_pa_tjuetretten = {}

for key in av_vest_tretten_L.keys():
    if key not in av_pa_tjuetretten:
        av_pa_tjuetretten[key] = {}
    for key2 in av_vest_tretten_L[key].keys():
        if key2 not in av_pa_tjuetretten[key]:
            av_pa_tjuetretten[key][key2] = [-1,-1]
            av_pa_tjuetretten[key][key2][1] = av_vest_tretten_L[key][key2]

for key in pa_vest_tretten_L.keys():
    if key not in av_pa_tjuetretten:
        av_pa_tjuetretten[key] = {}
    for key2 in pa_vest_tretten_L[key].keys():
        if key2 not in av_pa_tjuetretten[key]:
            av_pa_tjuetretten[key][key2] = [-1,-1]
            av_pa_tjuetretten[key][key2][0] = pa_vest_tretten_L[key][key2]

for key in av_ost_tretten_L.keys():
    if key not in av_pa_tjuetretten:
        av_pa_tjuetretten[key] = {}
    for key2 in av_ost_tretten_L[key].keys():
        if key2 not in av_pa_tjuetretten[key]:
            av_pa_tjuetretten[key][key2] = [-1,-1]
            av_pa_tjuetretten[key][key2][1] = av_ost_tretten_L[key][key2]

for key in pa_ost_tretten_L.keys():
    if key not in av_pa_tjuetretten:
        av_pa_tjuetretten[key] = {}
    for key2 in pa_ost_tretten_L[key].keys():
        if key2 not in av_pa_tjuetretten[key]:
            av_pa_tjuetretten[key][key2] = [-1,-1]
            av_pa_tjuetretten[key][key2][0] = pa_ost_tretten_L[key][key2]

for key in av_vest_tretten_R.keys():
    if key not in av_pa_tjuetretten:
        av_pa_tjuetretten[key] = {}
    for key2 in av_vest_tretten_R[key].keys():
        if key2 not in av_pa_tjuetretten[key]:
            av_pa_tjuetretten[key][key2] = [-1,-1]
            av_pa_tjuetretten[key][key2][1] = av_vest_tretten_R[key][key2]

for key in pa_vest_tretten_R.keys():
    if key not in av_pa_tjuetretten:
        av_pa_tjuetretten[key] = {}
    for key2 in pa_vest_tretten_R[key].keys():
        if key2 not in av_pa_tjuetretten[key]:
            av_pa_tjuetretten[key][key2] = [-1,-1]
            av_pa_tjuetretten[key][key2][0] = pa_vest_tretten_R[key][key2]

for key in av_ost_tretten_R.keys():

```

```

if key not in av_pa_tjuetretten:
    av_pa_tjuetretten[key] = {}
for key2 in av_ost_tretten_R[key].keys():
    if key2 not in av_pa_tjuetretten[key]:
        av_pa_tjuetretten[key][key2] = [-1,-1]
        av_pa_tjuetretten[key][key2][1] = av_ost_tretten_R[key][key2]

for key in pa_ost_tretten_R.keys():
    if key not in av_pa_tjuetretten:
        av_pa_tjuetretten[key] = {}
    for key2 in pa_ost_tretten_R[key].keys():
        if key2 not in av_pa_tjuetretten[key]:
            av_pa_tjuetretten[key][key2] = [-1,-1]
            av_pa_tjuetretten[key][key2][0] = pa_ost_tretten_R[key][key2]

### FUNSKJON SOM SUMMERER OPP AVSTIGENDE/PÅSTIGENDE PER OMRÅDE ###

passasjerer = {} # Dictionary med områder --> år -- > av og på stigende
f = av_pa_tjueti
g = av_pa_tjueelleve
h = av_pa_tjuetolv
i = av_pa_tjuetretten
j = av_pa_fjorten
k = av_pa_femten
reisende_years = [f,g,h,i,j,k]
years = ["2010","2011","2012","2013","2014","2015"]

for omrader in tittel:
    passasjerer[omrader] = {}
    passasjerer[omrader]["2010"] = []
    passasjerer[omrader]["2011"] = []
    passasjerer[omrader]["2012"] = []
    passasjerer[omrader]["2013"] = []
    passasjerer[omrader]["2014"] = []
    passasjerer[omrader]["2015"] = []

for i_om, omrade in enumerate(tittel):
    for i_y, year in enumerate(years):

        pastigende = 0
        avstigende = 0
        y = reisende_years[i_y]
        for stasjon in stasjoner[i_om]:

            for tognummer in y[stasjon].keys():

                B_found = False

                for togintervall in tognummerintervall[i_om]:
                    if int(tognummer) >= togintervall[0] and int(tognummer) <= togintervall[1]:
                        B_found = True

                if B_found:
                    pastigende = pastigende + y[stasjon][tognummer][0]
                    avstigende = avstigende + y[stasjon][tognummer][1]

```

```

antall_reisende = pastigende + avstigende

passasjerer[omrade][year] = [antall_reisende]

### TELLESYSTEM FOR ANT TOG TELT PR OMRÅDE PR ÅR ###
tog_ant = {}
tognumme = {}

a = av_pa_tjueti
b = av_pa_tjueelleve
c = av_pa_tjuetolv
d = av_pa_tjuetretten
e = av_pa_fjorten
l = av_pa_femten
data_years = [a,b,c,d,e,l]
for omrader in tittel:
    tog_ant[omrader] = {}
    tognumme[omrader] = {}
    tog_ant[omrader]["2010"] = []
    tog_ant[omrader]["2011"] = []
    tog_ant[omrader]["2012"] = []
    tog_ant[omrader]["2013"] = []
    tog_ant[omrader]["2014"] = []
    tog_ant[omrader]["2015"] = []
    tognumme[omrader]["2010"] = []
    tognumme[omrader]["2011"] = []
    tognumme[omrader]["2012"] = []
    tognumme[omrader]["2013"] = []
    tognumme[omrader]["2014"] = []
    tognumme[omrader]["2015"] = []

for i_om, omrade in enumerate(tittel):
    for i_y, year in enumerate(years):

        tognummerliste = []
        x = data_years[i_y]

        for stasjon in stasjoner[i_om]:

            for tognummer in x[stasjon].keys():

                T_found = False

                for togintervall in tognummerintervall[i_om]:

                    if int(tognummer) >= togintervall[0] and int(tognummer) <= togintervall[1]:
                        T_found = True

                if T_found:
                    if tognummer not in tognummerliste:
                        tognummerliste.append(tognummer)
        tognumme[omrade][year] = tognummerliste
        antalltog = len(tognummerliste)
        tog_ant[omrade][year] = antalltog

```

```
### REGNER UT TOTAL PASSASJERTALL FOR ALLE ÅR ###
```

```
r = av_pa_tjueti  
t = av_pa_tjueelleve  
y = av_pa_tjuetolv  
p = av_pa_tjuetretten  
j = av_pa_fjorten  
l = av_pa_femten  
data_aar = [r,t,y,p,j,l]
```

```
totalt_antall_reiser = {} # År -> reiser, for områdene som er tatt med.
```

```
for i_y, year in enumerate(years):
```

```
    totalereiser = 0  
    x = data_aar[i_y]
```

```
    for i_o, omrade in enumerate(tittel):
```

```
        if av_pa_totalutvikling[i_o][0] == 1:
```

```
            totalereiser = totalereiser + passasjerer[omrade][year][0]
```

```
    totalt_antall_reiser[year] = totalereiser
```

```
# Beregner vekt hvert område har å si : Vekt = antall reisende i området/totalt antall reiser  
det året.
```

```
vekt = {}
```

```
for omrader in tittel:
```

```
    vekt[omrader] = {}  
    vekt[omrader]["2010"] = []  
    vekt[omrader]["2011"] = []  
    vekt[omrader]["2012"] = []  
    vekt[omrader]["2013"] = []  
    vekt[omrader]["2014"] = []  
    vekt[omrader]["2015"] = []
```

```
for i_om, omrade in enumerate(tittel):
```

```
    for i_y, year in enumerate(years):  
        andel = passasjerer[omrade][year][0]  
        totaldel = totalt_antall_reiser[year]
```

```
        weight = (float(andel)/float(totaldel))*100 #prosentandel  
        vekt[omrade][year] = weight
```

```
### PRODUSERER OUTPUTFIL ###
```

```
workbook = xl.Workbook('Resultatfil.xlsx')  
worksheet = workbook.add_worksheet()  
worksheet.set_column('A:F',25)  
bold = workbook.add_format({'bold':True})  
worksheet.write('A1','Omrade - Passasjerer',bold)  
worksheet.write('A25','Omrade - Antall togtellinger',bold)
```

```
## For passasjerer ##
```



```

for i_y, year in enumerate(years):
    i_col = 1 + i_y
    worksheet.write(0,i_col,year)

for i_o, omrader in enumerate(passasjerer.keys()):
    i_row = i_o + 1
    worksheet.write(i_row,0,omrader)

    for i_y, year in enumerate(years):

        i_col = i_y + 1
        tall = passasjerer[omrader][year]
        worksheet.write(i_row,i_col,tall[0])

### For togantall ##
for i_o, omrader in enumerate(tog_ant.keys()):
    i_row = i_o + 25
    worksheet.write(i_row,0,omrader)

    for i_y, year in enumerate(years):

        i_col = i_y + 1
        togantall = tog_ant[omrader][year]

        worksheet.write(i_row,i_col,togantall)

### For totale reiser ###

worksheet.write('A46','TOTAL PASSASJERUTVIKLING',bold)

for i_y, year in enumerate(years):

    i_col = i_y + 1
    i_row = 46
    inndata = totalt_antall_reiser[year]
    worksheet.write(i_row,i_col,inndata)

# For vekting
jahre = ["2010", "2011", "2012", "2013", "2014", "2015"]
worksheet.write('K1','Vekting for hvert omrade',bold)
for i_y, year in enumerate(years):
    i_col = i_y + 11
    worksheet.write(0,i_col,year)

for i_o, omrader in enumerate(vekt.keys()):
    row = i_o + 1
    col = 10
    worksheet.write(row,col,omrader)

    for i_y, year in enumerate(jahre):
        v_col = 11 + i_y
        x = vekt[omrader][year]

```

```

worksheet.write(row,v_col,x)

workbook.close()

### Leser inn outputfila og omformer til relativ passasjervekst per område og relativ vekst i
antall tog
filelocation = "C:/Users/tostigen/Desktop/Program_utgangspunkt/Resultatfil.xlsx"

workbook = xlrd.open_workbook(filelocation)

passasjerark = workbook.sheet_by_index(0)

outputfile = xl.Workbook("Relativeresultater.xlsx")
ark = outputfile.add_worksheet()
ark.write('A1','Omrade - Passasjerer')

aar = ["2010","2011","2012","2013","2014","2015"]

for i_y, years in enumerate(aar):
    i_col = 1 + i_y
    ark.write(0,i_col,years)

for i_o, omrader in enumerate(passasjerer.keys()):
    i_row = i_o + 1
    ark.write(i_row,0,omrader)
aar = ["2010","2011","2012","2013","2014","2015"]
a = len(passasjerer.keys())
b = len(aar)

for i in range(a):
    rad = i + 1
    ark.write(rad,1,0)

    for j in range(b-1):
        nykol = j + 2
        nyverdi = passasjerark.cell(rad,nykol).value
        orginalverdi = passasjerark.cell(rad,1).value
        orgv = int(orginalverdi)
        nyver = int(nyverdi)
        delta = nyver - orgv

        indeks = (delta/orginalverdi)*100

        ark.write(rad,nykol,indeks)

#For togantall
for i_o, omrader in enumerate(tog_ant.keys()):
    i_row = i_o + 25
    ark.write(i_row,0,omrader)
    ark.write(i_row,1,0)
    for j in range(b-1):
        nyk = j + 2
        nyv = passasjerark.cell(i_row,nyk).value
        orgv = passasjerark.cell(i_row,1).value
        delta = nyv - orgv
        ind = (delta/orgv)*100

```

```

ark.write(i_row,nyk,ind)

#For totalt antall reiser
aar = ["2010","2011","2012","2013","2014","2015"]
ark.write('A46','TOTAL PASSASJERUTVIKLING')

for i_y, ar in enumerate(aar):

    i_col = i_y + 1
    i_row = 46
    inndata = totalt_antall_reiser[year]
    ark.write(i_row,i_col,inndata)

# For vekting
jahre = ["2010","2011","2012","2013","2014","2015"]
ark.write('K1','Vekting for hvert område',bold)
for i_y, year in enumerate(jahre):
    i_col = i_y + 11
    ark.write(0,i_col,year)

for i_o, omrader in enumerate(vekt.keys()):
    row = i_o + 1
    col = 10
    ark.write(row,col,omrader)

    for i_y, year in enumerate(jahre):
        v_col = 11 + i_y
        x = vekt[omrader][year]
        ark.write(row,v_col,x)

outputfile.close()

aar = ["2010","2011","2012","2013","2014","2015"]

tognummer_sjekk = {}
for omr in tittel:
    tognummer_sjekk[omr] = {}

    for y in aar:
        tognummer_sjekk[omr][y] = {}
        for nummer in tognumme[omr][y]:
            tognummer_sjekk[omr][y][nummer]=[]

for i_o, omrader in enumerate(tittel):

    for togn in tognummer_sjekk[omrader]["2010"]:

        avstig = 0
        pastig = 0

        for stasjon in av_pa_tjueti.keys():

            if stasjon in stasjoner[i_o]:

                if togn in av_pa_tjueti[stasjon]:

```

```
avstig = avstig + av_pa_tjueti[stasjon][togn][1]
pastig = pastig + av_pa_tjueti[stasjon][togn][0]
```

```
tognummer_sjekk[omrader]["2010"][togn]= [pastig,avstig]
```

```
for i_o, omrader in enumerate(tittel):
```

```
for togn in tognummer_sjekk[omrader]["2011"]:
```

```
avstig = 0
pastig = 0
for stasjon in av_pa_tjueelleve.keys():
```

```
    if stasjon in stasjoner[i_o]:
```

```
        if togn in av_pa_tjueelleve[stasjon]:
            avstig = avstig + av_pa_tjueelleve[stasjon][togn][1]
            pastig = pastig + av_pa_tjueelleve[stasjon][togn][0]
```

```
tognummer_sjekk[omrader]["2011"][togn]= [pastig,avstig]
```

```
for i_o, omrader in enumerate(tittel):
```

```
for togn in tognummer_sjekk[omrader]["2012"]:
```

```
avstig = 0
pastig = 0
```

```
for stasjon in av_pa_tjuetolv.keys():
```

```
    if stasjon in stasjoner[i_o]:
```

```
        if togn in av_pa_tjueelleve[stasjon]:
            avstig = avstig + av_pa_tjuetolv[stasjon][togn][1]
            pastig = pastig + av_pa_tjuetolv[stasjon][togn][0]
```

```
tognummer_sjekk[omrader]["2012"][togn]= [pastig,avstig]
```

```
for i_o, omrader in enumerate(tittel):
```

```
for togn in tognummer_sjekk[omrader]["2013"]:
```

```
avstig = 0
pastig = 0
```

```
for stasjon in av_pa_tjuetolv.keys():
```

```
    if stasjon in stasjoner[i_o]:
```

```
        if togn in av_pa_tjuetretten[stasjon]:
            avstig = avstig + av_pa_tjuetretten[stasjon][togn][1]
            pastig = pastig + av_pa_tjuetretten[stasjon][togn][0]
```

```
tognummer_sjekk[omrader]["2013"][togn]= [pastig,avstig]
```

```

for i_o, omrader in enumerate(tittel):

    for togn in tognummer_sjekk[omrader]["2014"]:

        avstig = 0
        pastig = 0

        for stasjon in av_pa_fjorten.keys():

            if stasjon in stasjoner[i_o]:

                if togn in av_pa_fjorten[stasjon]:
                    avstig = avstig + av_pa_fjorten[stasjon][tog][1]
                    pastig = pastig + av_pa_fjorten[stasjon][tog][0]

                tognummer_sjekk[omrader]["2014"][tog]= [pastig,avstig]

for i_o, omrader in enumerate(tittel):

    for togn in tognummer_sjekk[omrader]["2015"]:

        avstig = 0
        pastig = 0

        for stasjon in av_pa_femten.keys():

            if stasjon in stasjoner[i_o]:

                if togn in av_pa_femten[stasjon]:
                    avstig = avstig + av_pa_femten[stasjon][tog][1]
                    pastig = pastig + av_pa_femten[stasjon][tog][0]

                tognummer_sjekk[omrader]["2015"][tog]= [pastig,avstig]

tognummerliste = {}

for omrader in tittel:
    tognummerliste[omrader] = []
    for y in aar:
        for tognummm in tognummer_sjekk[omrader][y].keys():
            if tognummm not in tognummerliste[omrader]:
                tognummerliste[omrader].append(tognummm)

for omrader in tittel:
    tognummerliste[omrader] = sorted(tognummerliste[omrader])

## Output fil ##
print tognummer_sjekk[u"Lillestrøm_Eidsvoll"]["2013"]

for omrader in tognummer_sjekk.keys():
    tittel = omrader + "test.xlsx"

```

```

output = xl.Workbook(tittel)
shhet = output.add_worksheet()
shhet.write('A1',omrader)
a = ["2010","2011","2012","2013","2014","2015"]
c = 1
for i in range(len(a)):
    shhet.write(0,c,a[i])
    c = c + 2

for i_n, nummer in enumerate(tognummerliste[omrader]):
    r = i_n + 1
    shhet.write(r,0,nummer)

    coll = 1
    a = ["2010","2011","2012","2013","2014","2015"]
    for years in a:
        if nummer in tognummer_sjekk[omrader][years]:
            pastigende = tognummer_sjekk[omrader][years][nummer][0]
            avstigende = tognummer_sjekk[omrader][years][nummer][1]
            soll = coll + 1
            shhet.write(r,coll,pastigende)
            shhet.write(r,soll,avstigende)
        else:
            shhet.write(r,coll + 1,0)
            shhet.write(r,coll,0)

    coll = coll + 2

output.close()

```

## Vedlegg 5: Analyseprogram i python

```
# -*- coding: cp1252 -*-
import xlswriter as xl
import xlrd
import numpy as np
import matplotlib.pyplot as plt
import statsmodels.api as sm
from statsmodels.sandbox.regression.predstd import wls_prediction_std
import math

# Lager oversikt over endelig output fil for analyser

fil = xl.Workbook('Analysefil.xlsx')
ark = fil.add_worksheet()

ark.set_column('A:K',25)
#bold = ark.add_format({'bold':True})
ark.write('A1','Observasjon')
ark.write('B1','Relativ passasjervekst')
ark.write('C1','Relativ togvekst')
ark.write('D1','Reisetid')
ark.write('E1','Pris')
ark.write('F1','Flirt, Ja =1, Nei = 0')
ark.write('G1','Befolkning')

### Leser inn resultatfil fra annet program. Legger til hvert passasjerer(år,område) som en
observasjon i fil
filelocation =
"C:/Users/tostigen/Desktop/Program_utgangspunkt/relative_resultater_0,5,0,75,1,1,1.xlsx"
workbook = xlrd.open_workbook(filelocation)
passasjerark = workbook.sheet_by_index(0)
liste = passasjerark.col_values(0)
lis = []
for i in range(len(liste)):
    if liste[i] is not "":
        if liste[i] not in lis:
            lis.append(liste[i])

a = len(lis)-2

for i in range(a):
    rows = i + 1
    row_ant = i + 25
    for j in range(6):
        cols = j + 1
        colvekt = j + 11
        if passasjerark.cell_value(rows,cols) is not "":
            observasjon = passasjerark.cell_value(rows,cols)
            togantall = passasjerark.cell_value(row_ant,cols)
            vekt = passasjerark.cell_value(rows,colvekt)
            indeks = j + i*6
            index = indeks + 1
            ob =int(observasjon)
```

```

    tog = int(togantall)
    ve = int(vekt)
    ark.write(index,1,observasjon)
    ark.write(index,2,togantall)
    ark.write(index,9,vekt)
    omrade = passasjerark.cell_value(rows,0)
    ark.write(index,0,omrade)
## Leser inn annen data for å legge til hvert område ###

fillokasjon =
"C:/Users/tostigen/Desktop/Program_utgangspunkt/input_forsinkelse_0,5,75,1,1,1.xlsx"

arbeidsbok = xlrd.open_workbook(fillokasjon)
inputark = arbeidsbok.sheet_by_index(0)

# Legger frekvens grunnrute

indeksliste = [0,0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0,
    2, 0, 0, 0, 0, 0, 0, 0, 3, 0, 0, 0, 0, 0, 0, 0, 0,
    4, 0, 0, 0, 0, 0, 0, 0, 5, 0, 0, 0, 0, 0, 0, 0, 0,
    6, 0, 0, 0, 0, 0, 0, 0, 7, 0, 0, 0, 0, 0, 0, 0, 0,
    8, 0, 0, 0, 0, 0, 0, 0, 9, 0, 0, 0, 0, 0, 0, 0, 0,
    10, 0, 0, 0, 0, 0, 0, 0, 11, 0, 0, 0, 0, 0, 0, 0, 0,
    12, 0, 0, 0, 0, 0, 0, 0, 13, 0, 0, 0, 0, 0, 0, 0, 0,
    14, 0, 0, 0, 0, 0, 0, 0, 15, 0, 0, 0, 0, 0, 0, 0, 0,
    16, 0, 0, 0, 0, 0, 0, 0, 17, 0, 0]

for i in range(1,139):
    r = i
    if inputark.cell_value(r,0) is not "":
        for q in range(6):

            colaar = q + 2

            f = inputark.cell_value(r,colaar)
            tid = inputark.cell_value(r+1,colaar)
            ruter = inputark.cell_value(r+2,colaar)
            flirt = inputark.cell_value(r+3,colaar)
            bv = inputark.cell_value(r+4,colaar)

            indekss = indeksliste[i]*6 + q +1
            #ark.write(indekss,2,f)
            ark.write(indekss,3,tid)
            ark.write(indekss,4,ruter)
            #ark.write(indekss,5,flirt)
            ark.write(indekss,5,bv)

fil.close()

# Leser inn analysefil på nytt

location = "C:/Users/tostigen/Desktop/Program_utgangspunkt/Analysefil.xlsx"
dataark = xlrd.open_workbook(location)
analyseark = dataark.sheet_by_index(0)
null_liste = [1,7,13,19,25,31,37,43,49,55,61,67,73,79,85,91,97]

```



```

#### Gjør analyser i python ####
# lager matriser for både forklaringsvariabler og responsvariabler med og uten vektning. Går
gjennom
# legger til i responsvariabel vektet og uvektet.
rv = np.empty(0)
rvv = np.empty(0)
responser = analyseark.col_values(1)
vekker = analyseark.col_values(9)
#vekker = np.ones(len(vekker))*10
for i in range(1,len(responser)):
    if i not in null_liste:
        rv = np.append(rv,responser[i])
        vektetrespons = responser[i]*vekker[i]
        rvv = np.append(rvv,vektetrespons)
responsvariabel = np.transpose(np.matrix(rv))
responsvariabel_vektet = np.transpose(np.matrix(rvv))
# Legger til forklaringsvariable vektet og uvektet.
fv = [[] for i in range(91)]
fvv = [[] for i in range(91)]
tog = analyseark.col_values(2)

tid = analyseark.col_values(3)

ruter = analyseark.col_values(4)

#flirt = analyseark.col_values(5)
befv = analyseark.col_values(5)

#for i in range(1,71):
el = 0
for i in range(1,109):

    if i not in null_liste:

        element = [tog[i],tid[i],ruter[i],befv[i]]
        vektet_element = [tog[i]*vekker[i],tid[i]*vekker[i],ruter[i]*vekker[i],befv[i]*vekker[i]]
        fv[el]= element
        fvv[el] = vektet_element
        el = el + 1
forklaringsvariabel = np.matrix(fv)
forklaringsvariabel_vektet = np.matrix(fvv)
model = sm.OLS(responsvariabel,forklaringsvariabel)
results = model.fit()
print results.summary()
ols_resid = sm.OLS(responsvariabel,forklaringsvariabel).fit().resid
new = results.get_robustcov_results(cov_type='HAC',maxlags=1)
print new.summary()
koeffsienter = results.params
togkoe = koeffsienter[0]
tidkoe = koeffsienter[1]
rutrkoe = koeffsienter[2]
bfvkoe = koeffsienter[3]

#### Lager residual plot ####

```

```

faktisk_observasjon = []

vektet_observasjon = []
for i in range(1,len(responser)):
    if i not in null_liste:
        faktisk_observasjon.append(responser[i])
        vektet_observasjon.append(responser[i]*vekter[i])

forventet_observasjon = []
forventet_vektet = []

for i in range(1,len(responser)):
    if i not in null_liste:

        verdi = togkoe*tog[i] + tidkoe*tid[i] + rutrkoe*ruter[i]+bfvkoe*befv[i]
        ver = togkoe*tog[i]*vekter[i] + tidkoe*tid[i]*vekter[i] +
rutrkoe*ruter[i]*vekter[i]+bfvkoe*befv[i]*vekter[i]
        forventet_observasjon.append(verdi)
        forventet_vektet.append(ver)

residualer = []
residualer_vekt = []
for i in range(len(forventet_observasjon)):

    res = faktisk_observasjon[i]-forventet_observasjon[i]
    residualer.append(res)
    res_ve = vektet_observasjon[i] - forventet_vektet[i]
    residualer_vekt.append(res_ve)

residualplot = plt.figure()
plt.title("Residualplot")
plt.xlabel("Passasjerutvikling")
plt.ylabel("Residual")
plt.scatter(forventet_observasjon,residualer)
plt.axis([-5,110,-40,50])
plt.show()
residualplot.savefig("Residualplot.jpg")
#
obs = []
for i in range(len(residualer)):
    obs.append(i)
observasjonplot = plt.figure()
plt.title("Residualplot 2")
plt.xlabel("Observasjoner")
plt.ylabel("Residual")

plt.plot(obs,residualer)
observasjonplot.savefig("Residualplot_2.jpg")
#
bins = np.linspace(-100,100,50)
histoplot = plt.figure()
plt.title("Residualplot 3")
plt.xlabel("Residual")
plt.ylabel("Frekvens")
plt.hist(residualer,bins)

```

```

histoplot.savefig("Residualplt_3.jpg")

corr_matrix = np.corrcoef(forklaringsvariabel_vektet, rowvar = 0)
sm.graphics.plot_corr(corr_matrix)
plt.show()
print corr_matrix

### Estimerer forbedring pga. ruteendring pr. område for alle år ###
def tilbudsforbedring(omrade):
    omrader = analyseark.col_values(0)
    tilbudsforbedring_aar = []
    for i in range(1,len(omrader)):
        if omrade == omrader[i]:
            tilbudsokning = togkoe*analyseark.cell_value(i,2) + tidkoe*analyseark.cell_value(i,3)
            tilbudsforbedring_aar.append(tilbudsokning)
    return tilbudsforbedring_aar

def annen_forbedring(omrade):
    omrader = analyseark.col_values(0)
    annenforbedring_aar = []
    for i in range(1,len(omrader)):
        if omrade == omrader[i]:
            annenokning = rutrkoe*analyseark.cell_value(i,4) + bfvkoe*analyseark.cell_value(i,5)
            annenforbedring_aar.append(annenokning)
    return annenforbedring_aar

def befolkning_forbedring(omrade):
    omrader = analyseark.col_values(0)
    befolkningforbedring_aar = []
    for i in range(1,len(omrader)):
        if omrade == omrader[i]:
            befokn = bfvkoe*analyseark.cell_value(i,5)
            befolkningforbedring_aar.append(befokn)
    return befolkningforbedring_aar
### Estimerer absolutt passasjervekst pr område forårsaket av endring for årene
2011,2012,2013,2014 ###

loc = "C:/Users/tostigen/Desktop/Program_utgangspunkt/Resultatfil.xlsx"
arket = xlrd.open_workbook(loc)
absoluttepassasjerer = arket.sheet_by_index(0)

estimert_befolkningvekst = {}
estimert_passasjervekst = {}
estimert_nivauten_endring = {}
omrader = absoluttepassasjerer.col_values(0)
opprinnelig_reisende = absoluttepassasjerer.col_values(1)
for i in range(1,20):
    if omrader[i] is not "":
        estimert_befolkning_elleve =
opprinnelig_reisende[i]*befolkning_forbedring(omrader[i])[1]/100
        estimert_befolkning_tolv =
opprinnelig_reisende[i]*befolkning_forbedring(omrader[i])[2]/100
        estimert_befolkning_tretten =
opprinnelig_reisende[i]*befolkning_forbedring(omrader[i])[3]/100

```

```

    estimert_befolkning_fjorten =
    opprinnelig_reisende[i]*befolkning_forbedring(omrader[i])[4]/100
    estimert_befolkning_femten =
    opprinnelig_reisende[i]*befolkning_forbedring(omrader[i])[5]/100
    estimert_befolkningvekst[omrader[i]] =
    [estimert_befolkning_elleve,estimert_befolkning_tolv,estimert_befolkning_tretten,estimert_befolkning_fjorten,estimert_befolkning_femten]

```

```

    estimert_vekst_elleve = opprinnelig_reisende[i]*tilbudsforbedring(omrader[i])[1]/100
    estimert_vekst_tolv = opprinnelig_reisende[i]*tilbudsforbedring(omrader[i])[2]/100
    estimert_vekst_tretten = opprinnelig_reisende[i]*tilbudsforbedring(omrader[i])[3]/100
    estimert_vekst_fjorten = opprinnelig_reisende[i]*tilbudsforbedring(omrader[i])[4]/100
    estimert_vekst_femten = opprinnelig_reisende[i]*tilbudsforbedring(omrader[i])[5]/100
    estimert_passasjervekst[omrader[i]] =
    [estimert_vekst_elleve,estimert_vekst_tolv,estimert_vekst_tretten,estimert_vekst_fjorten,estimert_vekst_femten]

```

```

    vekst_uten_elleve = opprinnelig_reisende[i]*annen_forbedring(omrader[i])[1]/100
    vekst_uten_tolv = opprinnelig_reisende[i]*annen_forbedring(omrader[i])[2]/100
    vekst_uten_tretten = opprinnelig_reisende[i]*annen_forbedring(omrader[i])[3]/100
    vekst_uten_fjorten = opprinnelig_reisende[i]*annen_forbedring(omrader[i])[4]/100
    vekst_uten_femten = opprinnelig_reisende[i]*annen_forbedring(omrader[i])[5]/100
    estimert_nivauten_endring[omrader[i]] =
    [vekst_uten_elleve,vekst_uten_tolv,vekst_uten_tretten,vekst_uten_fjorten,vekst_uten_femten]

```

```

totalvekst_elleve = 0
totalvekst_tolv = 0
totalvekst_tretten = 0 # Total passasjerøkning som følge av ny ruteplan 2013
totalvekst_fjorten = 0 # Total passasjerøkning som følge av ny ruteplan 2014
totalvekst_femten = 0
total_uten_elleve = 0
total_uten_tolv = 0
total_uten_tretten = 0 # Estimert økning uten ny innføring av ny ruteplan 2013
total_uten_fjorten = 0 # Estimert økning uten ny innføring av ny ruteplan 2014
total_uten_femten = 0
total_bef_ell = 0
total_bef_tolv = 0
total_bef_tret = 0
total_bef_fjor = 0
total_bef_femt = 0

```

for key in estimert\_passasjervekst.keys():

```

totalvekst_elleve = totalvekst_elleve + estimert_passasjervekst[key][0]
totalvekst_tolv = totalvekst_tolv + estimert_passasjervekst[key][1]
totalvekst_tretten = totalvekst_tretten + estimert_passasjervekst[key][2]
totalvekst_fjorten = totalvekst_fjorten + estimert_passasjervekst[key][3]
totalvekst_femten = totalvekst_femten + estimert_passasjervekst[key][4]

```

```

total_uten_elleve = total_uten_elleve + estimert_nivauten_endring[key][0]
total_uten_tolv = total_uten_tolv + estimert_nivauten_endring[key][1]
total_uten_tretten = total_uten_tretten + estimert_nivauten_endring[key][2]
total_uten_fjorten = total_uten_fjorten + estimert_nivauten_endring[key][3]
total_uten_femten = total_uten_femten + estimert_nivauten_endring[key][4]

```

```

total_bef_ell = total_bef_ell + estimert_befolkningvekst[key][0]
total_bef_tolv = total_bef_tolv + estimert_befolkningvekst[key][1]
total_bef_tret = total_bef_tret + estimert_befolkningvekst[key][2]
total_bef_fjor = total_bef_fjor + estimert_befolkningvekst[key][3]
total_bef_femt = total_bef_femt + estimert_befolkningvekst[key][4]

### lager plot som ser på faktisk vekst, estimert vekst (med og uten ruteendring) for hele
østlandet samlet. (i prosenter). Bruker fortsatt 2010 som referansepunkt.

totalpassasjerer_faktisk = []

for i in range(6):
    totalpassasjerer_faktisk.append(absoluttepassasjerer.cell_value(46,i+1))

passasjervekst_total = [] #passasjervekst_total = y punkter for virkelig utvikling
for i in range(6):
    passasjervekst_total.append(100*(totalpassasjerer_faktisk[i]-
totalpassasjerer_faktisk[0])/totalpassasjerer_faktisk[0])

# Estimerer vekst for alle år med kun tilbudsøkning
elleve = 100*(totalvekst_elleve)/totalpassasjerer_faktisk[0]
tolv = 100*(totalvekst_tolv)/totalpassasjerer_faktisk[0]
tretten = 100*(totalvekst_tretten)/totalpassasjerer_faktisk[0]
fjorten = 100*(totalvekst_fjorten)/totalpassasjerer_faktisk[0]
femten = 100*(totalvekst_femten)/totalpassasjerer_faktisk[0]

estimert_kun_okning = [0,elleve,tolv,tretten,fjorten,femten]

# Estimerer vekst for alle år med både tilbudsøkning og normal vekst.
ell = 100*(totalvekst_elleve + total_uten_elleve)/totalpassasjerer_faktisk[0]
tol = 100*(totalvekst_tolv + total_uten_tolv)/totalpassasjerer_faktisk[0]
tret = 100*(totalvekst_tretten + total_uten_tretten)/totalpassasjerer_faktisk[0]
fjrtt = 100*(totalvekst_fjorten + total_uten_fjorten)/totalpassasjerer_faktisk[0]
fme = 100*(totalvekst_femten + total_uten_femten)/totalpassasjerer_faktisk[0]

estimert_okning_alt = [0,ell,tol,tret,fjrtt,fme]

elleveb = 100*(total_bef_ell)/totalpassasjerer_faktisk[0]
tolvb = 100*(total_bef_tolv)/totalpassasjerer_faktisk[0]
trettb = 100*(total_bef_tret)/totalpassasjerer_faktisk[0]
fjortb = 100*(total_bef_fjor)/totalpassasjerer_faktisk[0]
femt b = 100*(total_bef_femt)/totalpassasjerer_faktisk[0]

estimert_befolkning = [0,elleveb,tolvb,trettb,fjortb,femt b]

x_ekstrapolert = [10,11,12,13,14,15,16,17,18,19]
tilbud_pol = [0, -0.7574490174973617, 1.418324094543317, 16.517488611615647,
18.810006109709253, 44.47252196412006,50.903,54.923,58.943,58.943]
total_pol = [0, 1.8234164181793893, 7.691796779232443, 24.918949640100436,
29.928496852918027, 57.9510596251615,64.38,70.816,77.246,79.659]
totalfigur = plt.figure()
x = [10,11,12,13,14,15]
plt.title("Totalutvikling siden 2010")

```

```

plt.xlabel("Years")
plt.ylabel("Passasjerutvikling (%)")

plt.axis([10,15,0,60])
plt.plot(x,passasjervekst_total, label = "Faktisk utvikling")

plt.plot(x,estimert_kun_okning, label = "Utvikling grunnet tilbudsvekst")

plt.plot(x,estimert_okning_alt, label = "Estimert total utvikling")

plt.plot(x,estimert_befolkning, label = "kun befolkning")
plt.legend(loc = 2,prop={'size':11})

plt.show()
totalfigur.savefig("Totalutvikling.jpg")

#### Ruter resultater er ulogiske. Plotter et diagram over innføringsåret (2012) med
passasjerutvikling = [1.545,6.27,2.58,10.911,5.19,3.55,-10.48,16.53,-4.748,-0.863,3.65]
prisreduksjon = [260,260,360,260,360,40,160,480,480,260,480]
vekt = [4.51,9.29,11.96,7.54,13.65,6.08,2.12,4.96,3.70,4.55,2.38]
vet = [5,9,12,8,14,6,2,5,4,5,2]

ruterplot = plt.figure()
plt.title("Vekst og pris utvikling 2011-2012")
plt.xlabel("Pris reduksjon 2011-2012 (kr)")
plt.ylabel("Vekst i reisende")
plt.scatter(prisreduksjon,passasjerutvikling)
for i, txt in enumerate(vet):
    plt.annotate(txt,(prisreduksjon[i],passasjerutvikling[i]))
ruterplot.savefig("Ruterplot.jpg")
plt.show()

# Beregner effekter av ruteendringen #

tjuetretten = absoluttepassasjerer.cell_value(46,4)
tjuefjorten = absoluttepassasjerer.cell_value(46,5)
tjuefemten = absoluttepassasjerer.cell_value(46,6)
refftretten = tjuetretten-totalvekst_tretten
reffjorten = tjuefjorten - totalvekst_fjorten
reffemten = tjuefemten - totalvekst_femten

a = tjuetretten - totalvekst_tretten
b = tjuefjorten - totalvekst_fjorten
c =tjuefemten - totalvekst_femten
print a,b,c

d = totalvekst_tretten totalvekst_tretten totalvekst_tretten - vekst_for_ruteendring
e = totalvekst_fjorten - vekst_for_ruteendring
f = totalvekst_femten - vekst_for_ruteendring

print "Effekt_tretten"
print "-----"
print (d/a)*100, "%"
print "-----"
print "Effekt_fjorten"

```

```
print "-----"  
print (e/b)*100, "%"  
print "-----"  
print "Effekt_femten"  
print "-----"  
print (f/c)*100, "%"  
print "-----"
```



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

Postboks 5003  
NO-1432 Ås  
Norway