

Oslobysykkel: utforskende og prediktiv analyse

Abay M. Kassa



Innhold

1 Motivasjon

2 Datainnsamling

3 Påvirkningen av vær

4 Prediksjon

5 Innsikt

6 Oppsummering



Oslo Bysykkel

Motivasjon



Motivasjon



- Noen stasjoner har vært brukt som **STARTE** mens andre som **RETUR** per timer



Datainnsamling



Datainnsamling

- Nummeret av turer fra hvert stasjoner

Datainnsamling

- Nummeret av turer fra hvert stasjoner
 - ▶ 2019: April – Desember
 - ▶ 2020, 2021: Januar – Desember
 - ▶ 2022: Januar - Mars

Datainnsamling

- Nummeret av turer fra hvert stasjoner
 - ▶ 2019: April – Desember
 - ▶ 2020, 2021: Januar – Desember
 - ▶ 2022: Januar - Mars
- Sanntidsdata
 - ▶ Oppdatere av hvor mange sykler tilgjengelig på hverstasjon

Datainnsamling



- Nummeret av turer fra hvert stasjoner
 - ▶ 2019: April – Desember
 - ▶ 2020, 2021: Januar – Desember
 - ▶ 2022: Januar - Mars
- Sanntidsdata
 - ▶ Oppdatere av hvor mange sykler tilgjengelig på hverstasjon



Datainnsamling



- Nummeret av turer fra hvert stasjoner
 - ▶ 2019: April – Desember
 - ▶ 2020, 2021: Januar – Desember
 - ▶ 2022: Januar - Mars
- Sanntidsdata
 - ▶ Oppdatere av hvor mange sykler tilgjengelig på hverstasjon
- Temperatur
- Vindhastighet

Datainnsamling



- Nummeret av turer fra hvert stasjoner
 - ▶ 2019: April – Desember
 - ▶ 2020, 2021: Januar – Desember
 - ▶ 2022: Januar - Mars
- Sanntidsdata
 - ▶ Oppdatere av hvor mange sykler tilgjengelig på hverstasjon



- Temperatur
- Vindhastighet



Datainnsamling



- Nummeret av turer fra hvert stasjoner
 - ▶ 2019: April – Desember
 - ▶ 2020, 2021: Januar – Desember
 - ▶ 2022: Januar - Mars
- Sanntidsdata
 - ▶ Oppdatere av hvor mange sykler tilgjengelig på hverstasjon



- Temperatur
- Vindhastighet
- Nedbør

Datainnsamling



- Nummeret av turer fra hvert stasjoner
 - ▶ 2019: April – Desember
 - ▶ 2020, 2021: Januar – Desember
 - ▶ 2022: Januar - Mars
- Sanntidsdata
 - ▶ Oppdatere av hvor mange sykler tilgjengelig på hverstasjon



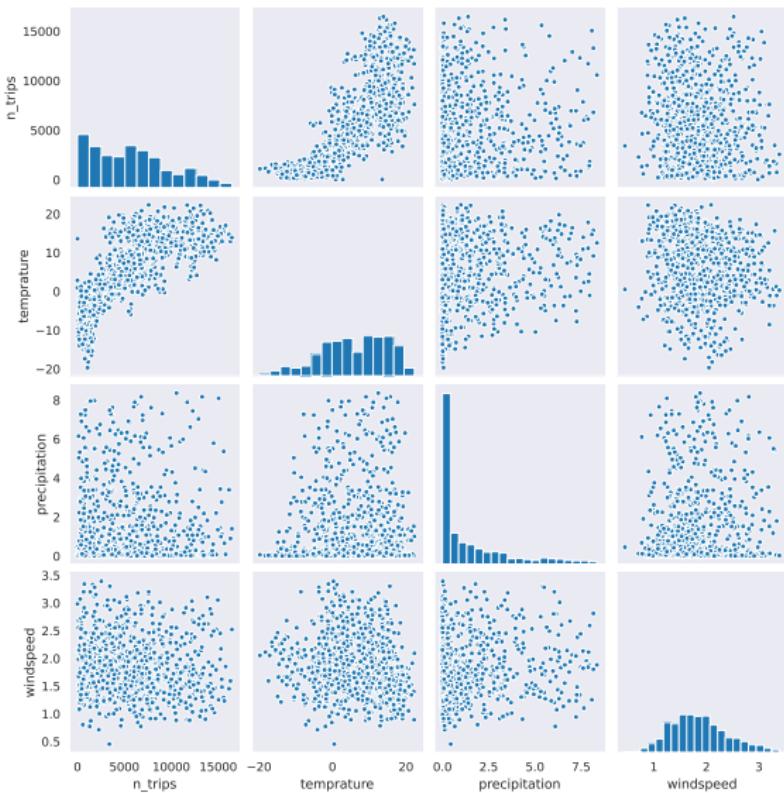
- Temperatur
- Vindhastighet



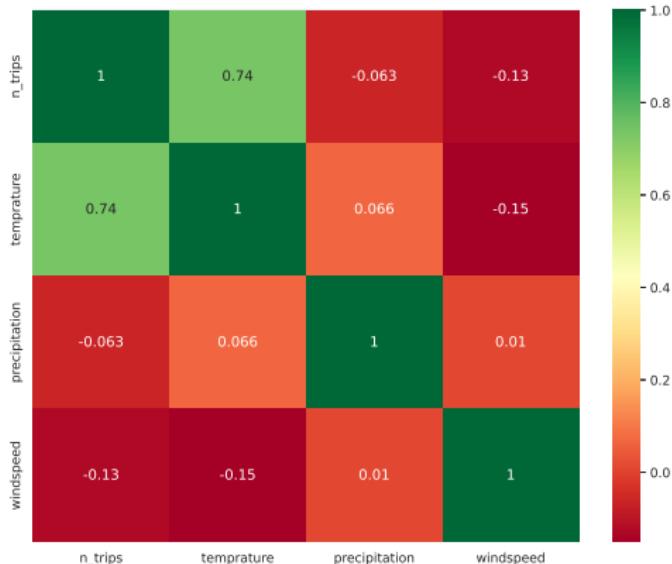
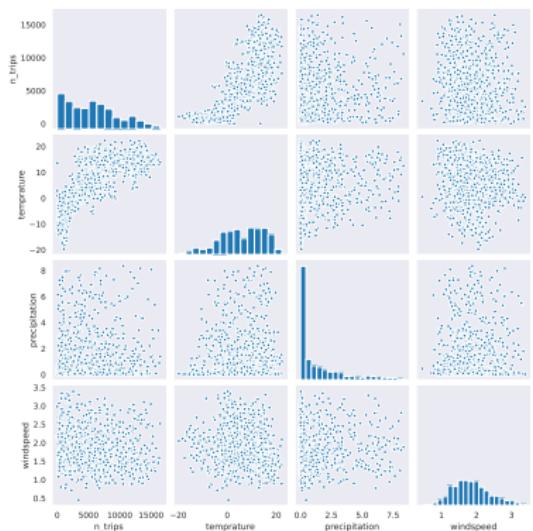
- Nedbør



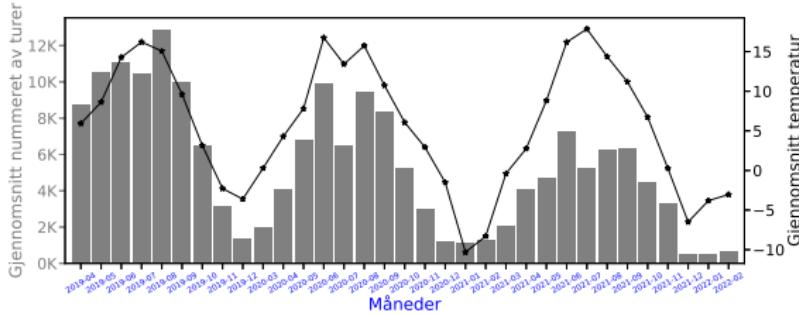
Oversikt av data



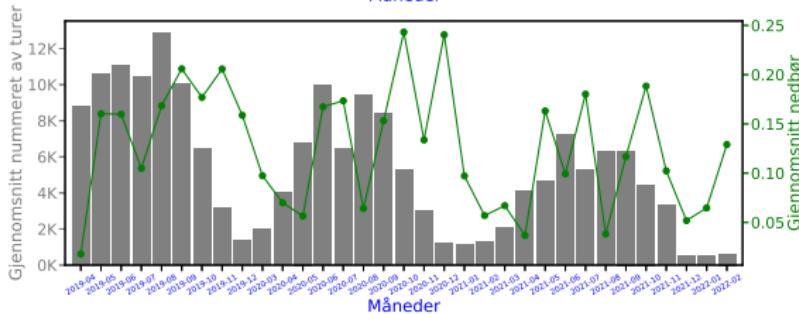
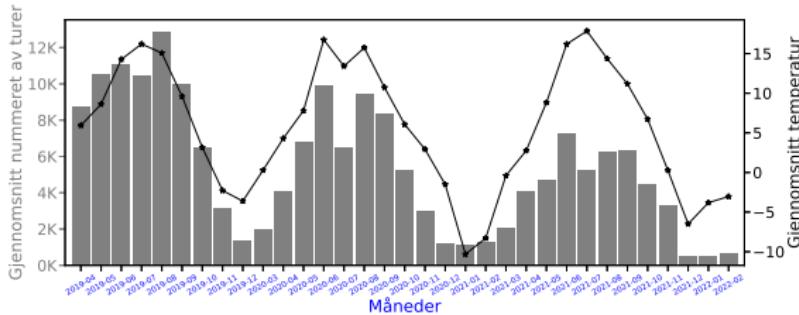
Oversikt av data



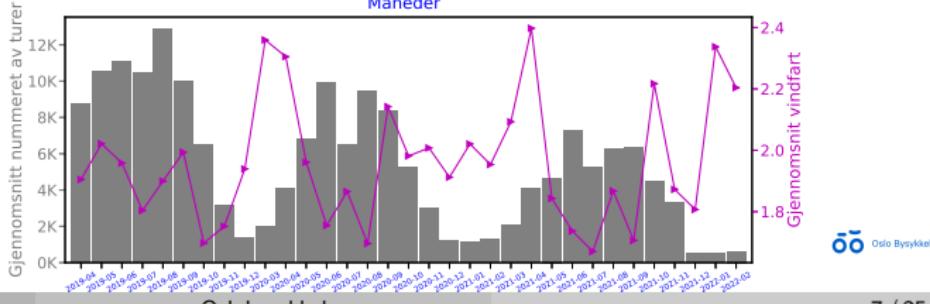
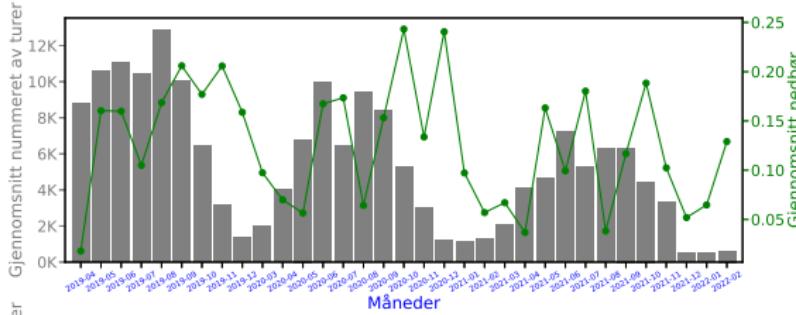
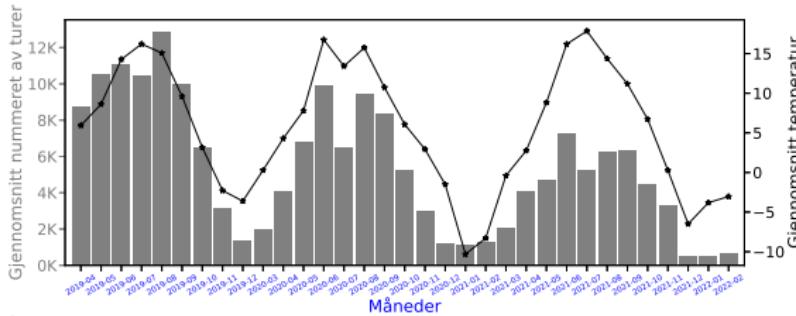
Påvirkningen av været



Påvirkningen av været



Påvirkingen av været

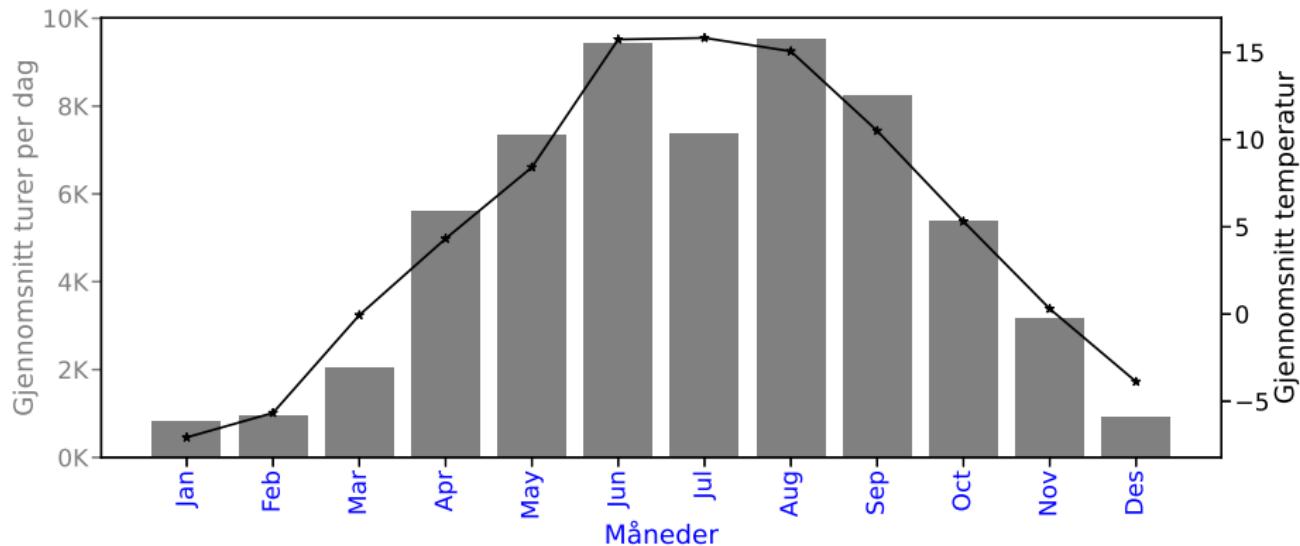


KPI
oo

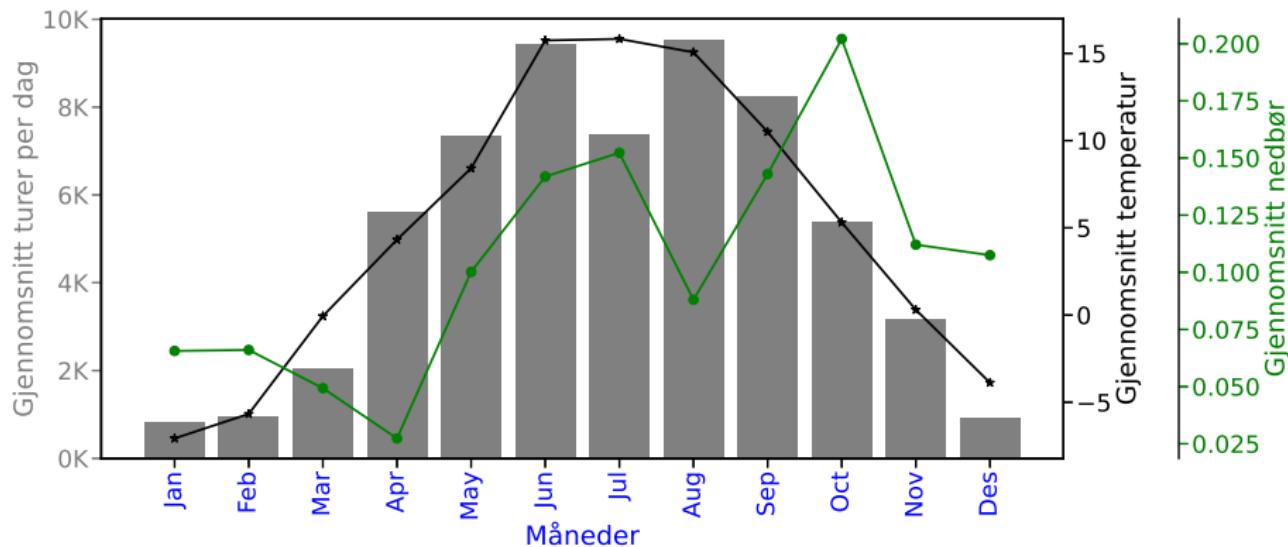
Påvirkningen av vær



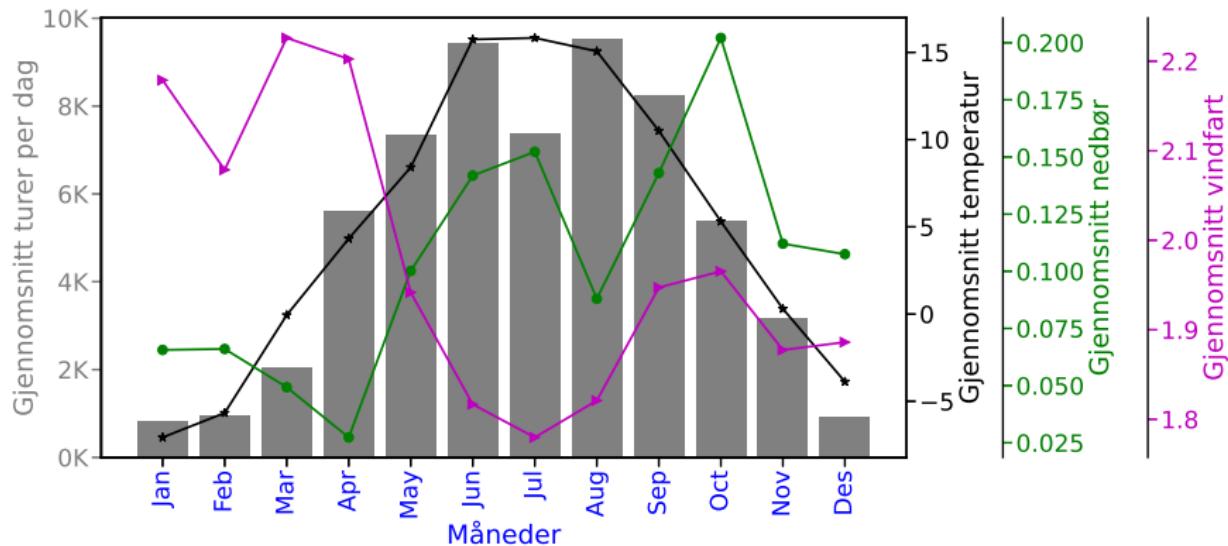
Påvirkningen av vær



Påvirkningen av vær

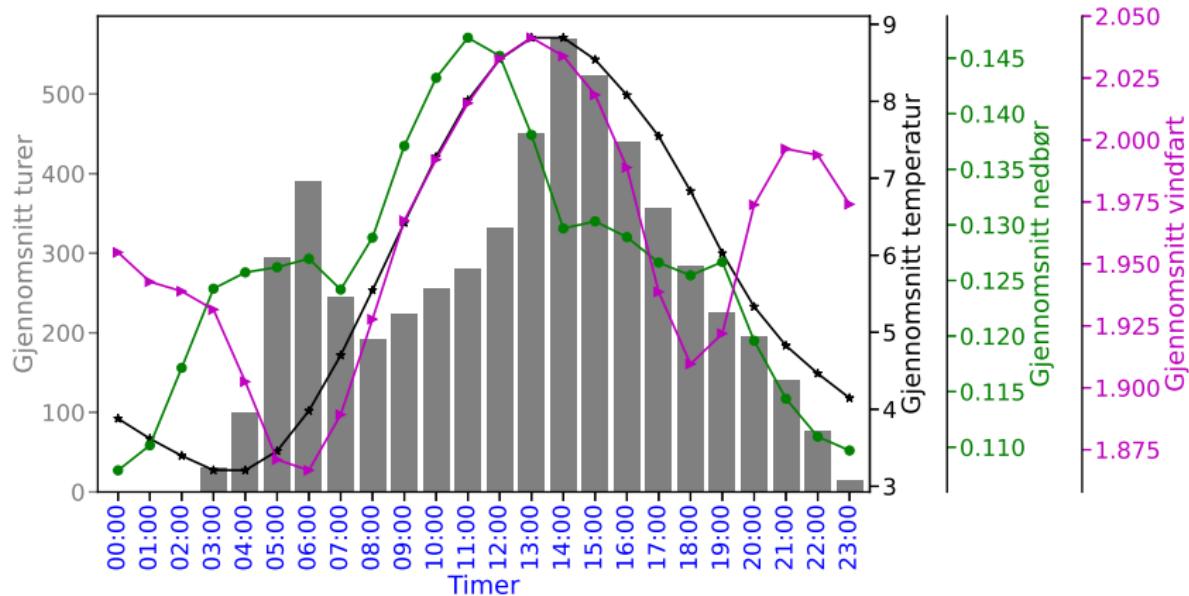


Påvirkningen av vær



Påvirkningen av vær

- Korrelasjon mellom være og nummeret av turer per time



Prediktiv modellering

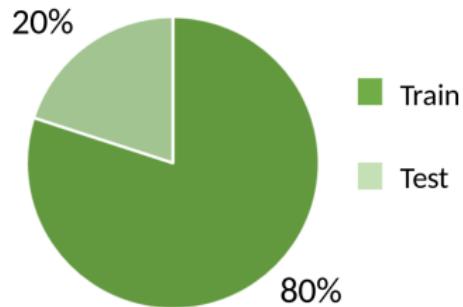
- Integrere Oslobysykle og værdata

Prediktiv modellering

- Integrere Oslobysykle og værdata
- Ren dataset

Prediktiv modellering

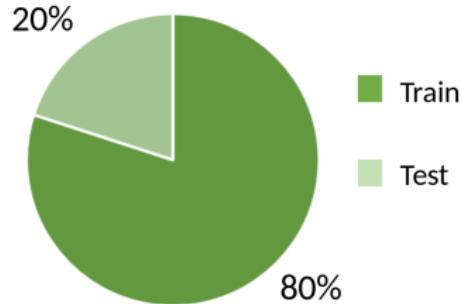
- Integrere Oslobysykle og værdata
- Ren dataset
- Del opp datasettet



- Design ML model

Prediktiv modellering

- Integrere Oslobysykle og værdata
- Ren dataset
- Del opp datasettet



- Design ML model

$$f \left(\text{Snowflake icon}, \text{Wind and leaves icon}, \text{Rain cloud icon} \right) = \text{oo} \text{ Oslo Bysykkel}$$

Forutsi det totale nummeret av turene

- Vi trener forskjellige ML modeller

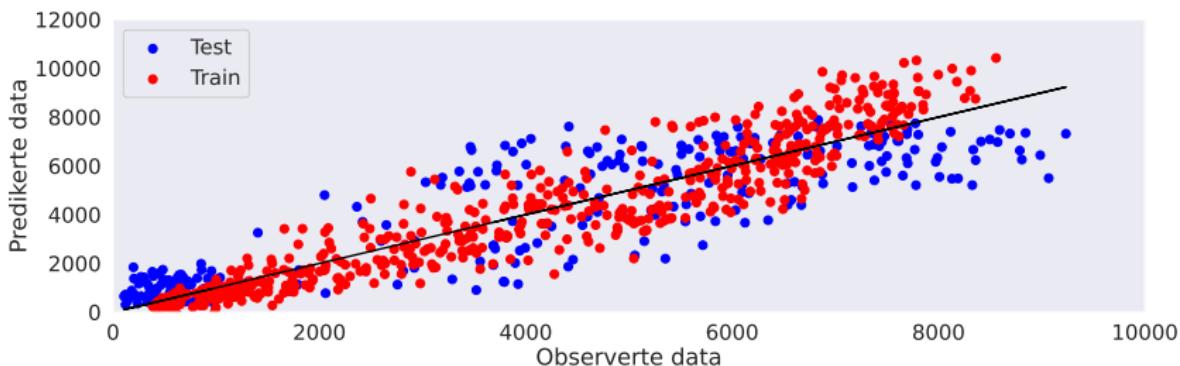
Forutsi det totale nummeret av turene

- Vi trener forskjellige ML modeller
 - ▶ RandomForest
 - ▶ KNN
 - ▶ XGBoost
 - ▶ Ensemble: Padding med RandomForest

Forutsi det totale nummeret av turene

- Vi trener forskjellige ML modeller
 - RandomForest
 - KNN
 - XGBoost
 - Ensemble: Padding med RandomForest

Model	R2 - Train	R2 - Test
XGBoost	72.1%	56.6%
KNN	75.1%	72.8%
Random Forest	94.4%	67.3%
Ensemble	88.5%	75.4%



Forutsi nummeret av turer på hver stasjon

- Startstasjon

$f\left(\text{temperatur}, \text{vær}, \text{nederbörd}\right) = \text{ø} \text{ Oslo Bysykkel}$ fra 

Forutsi nummeret av turer på hver stasjon

- Startstasjon

$$f\left(\text{Snowflake}, \text{Wind}, \text{Rain}\right) = \text{○○ Oslo Bysykkel fra } \text{[Image]}$$

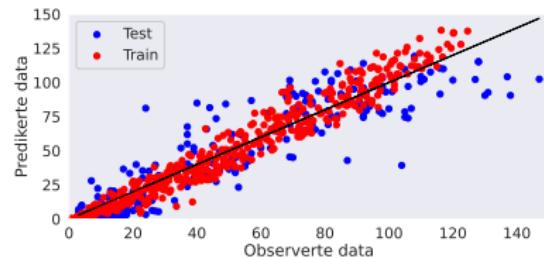
- Returstasjon

$$f\left(\text{Snowflake}, \text{Wind}, \text{Rain}\right) = \text{○○ Oslo Bysykkel til } \text{[Image]}$$

Forutsi nummeret av turer per stasjon

- Stasjon hvor turen startes
(494)

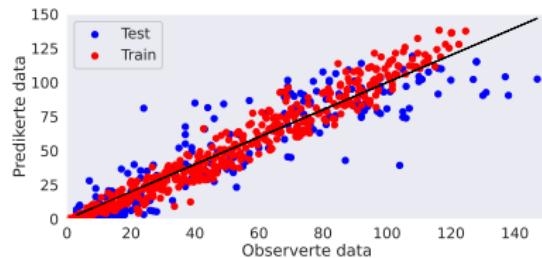
	R2 - Train	R2 - Test
Random Forest	93%	80%



Forutsi nummeret av turer per stasjon

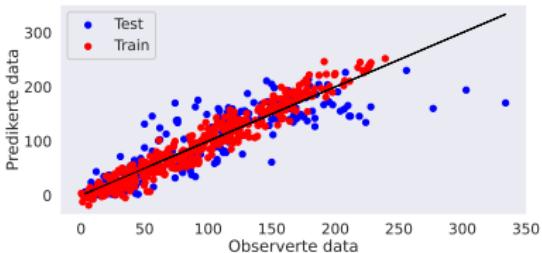
- Stasjon hvor turen startes
(494)

	R2 - Train	R2 - Test
Random Forest	93%	80%



- Stasjon hvor turen sluttet
(494)

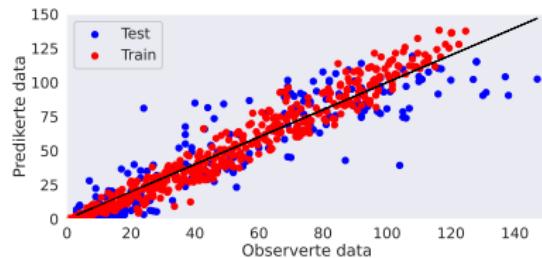
	R2 - Train	R2 - Test
Random Forest	93%	75%



Forutsi nummeret av turer per stasjon

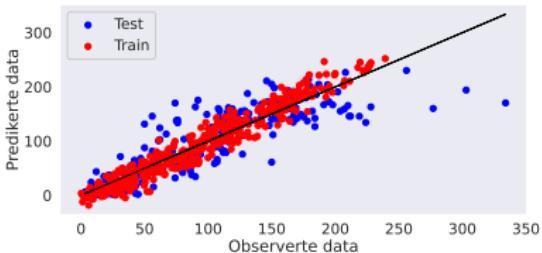
- Stasjon hvor turen startes
(494)

	R2 - Train	R2 - Test
Random Forest	93%	80%



- Stasjon hvor turen sluttet
(494)

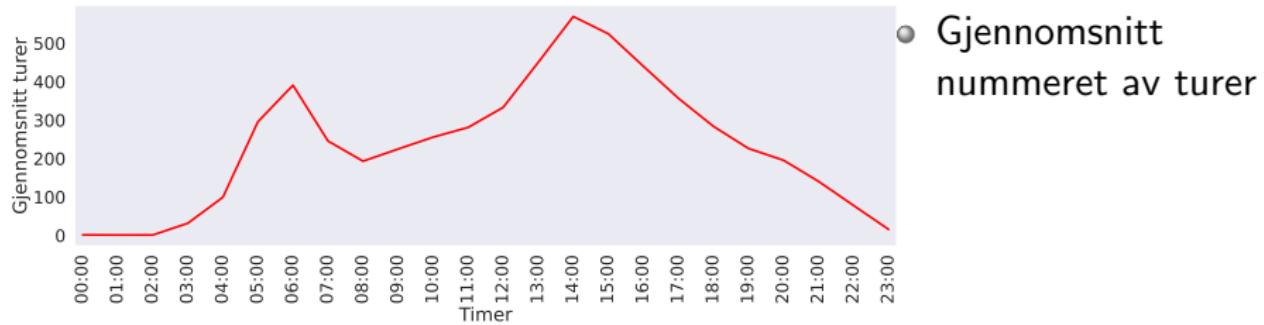
	R2 - Train	R2 - Test
Random Forest	93%	75%



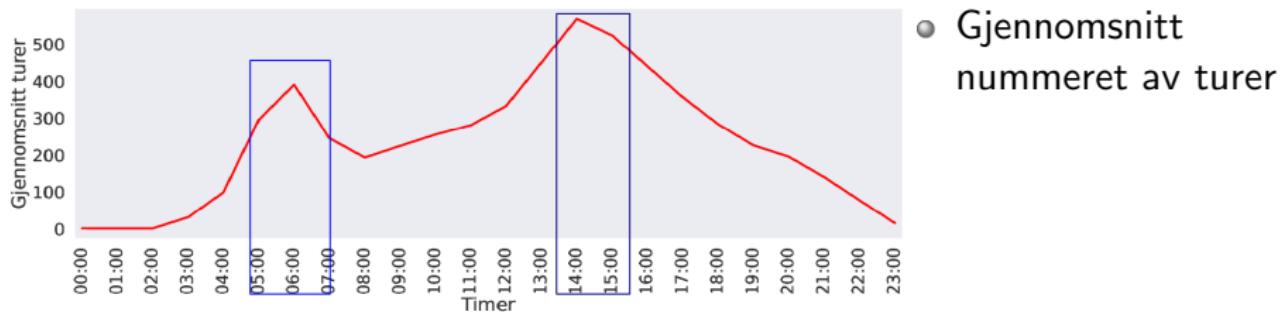
Motivasjon

- Nummeret av turer er korrelert med den vær
- Noen stasjoner har vært brukt som *STARTE* mens andre som *RETUR* per timer

Oslobysykkel Tur: Før og etter lunsj

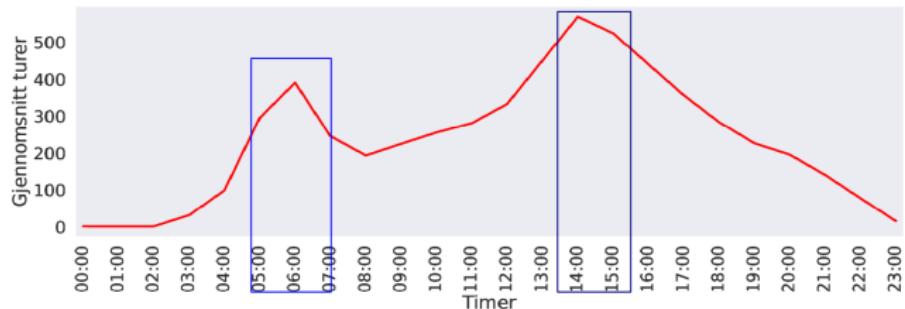


Oslobysykkel Tur: Før og etter lunsj



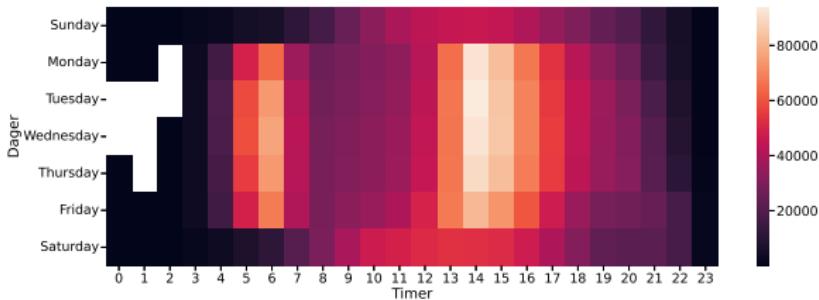
- Gjennomsnitt nummeret av turer

Oslobysykkel Tur: Før og etter lunsj



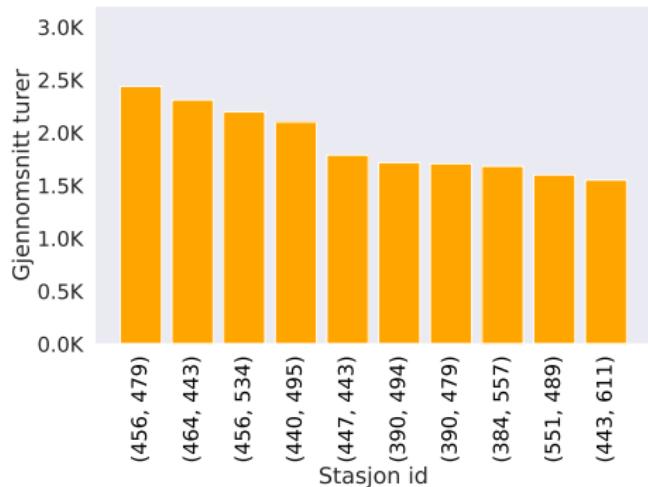
- Gjennomsnitt
nummeret av turer

- Totale nummer av
turer

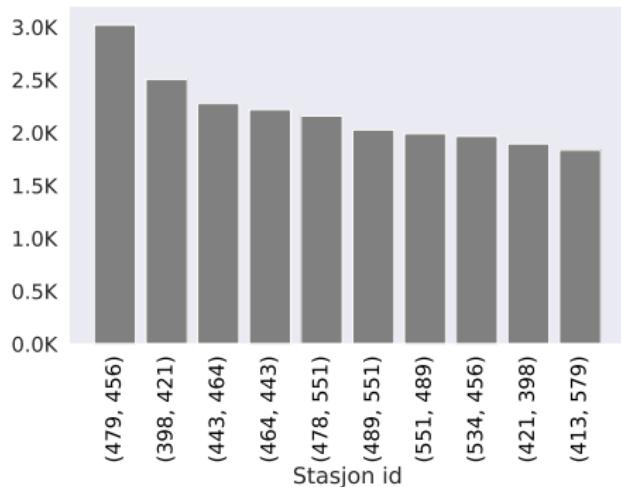


Oslobysykkel Tur: Før og etter lunsj

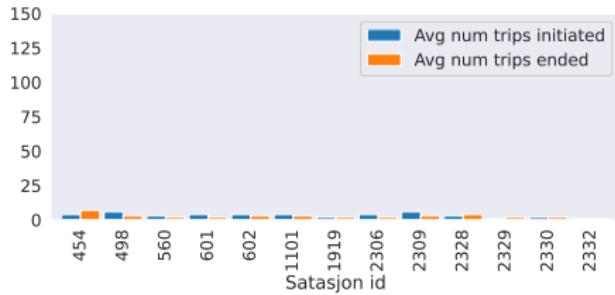
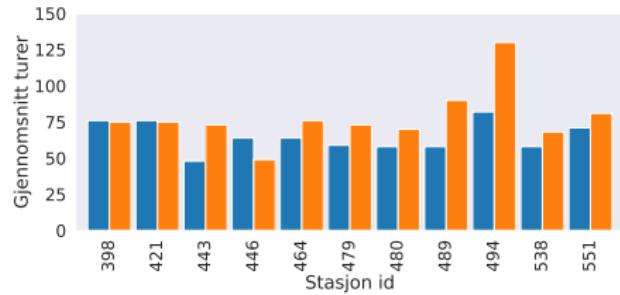
● Før lunsj



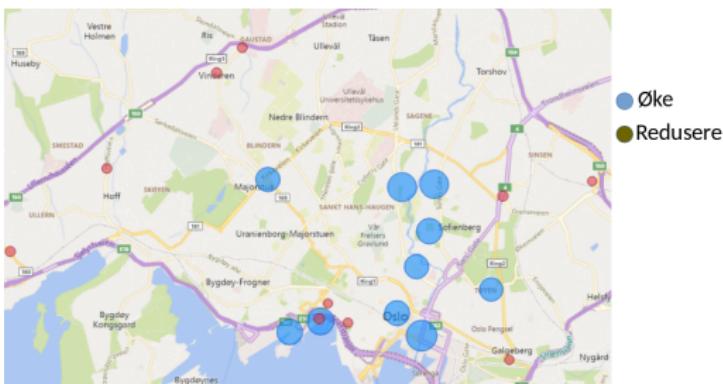
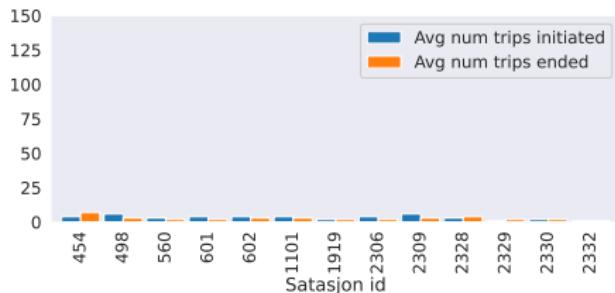
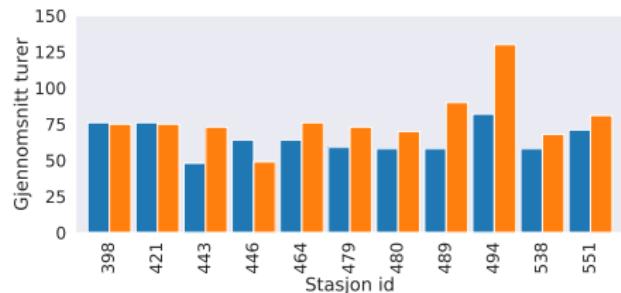
● Etter lunsj



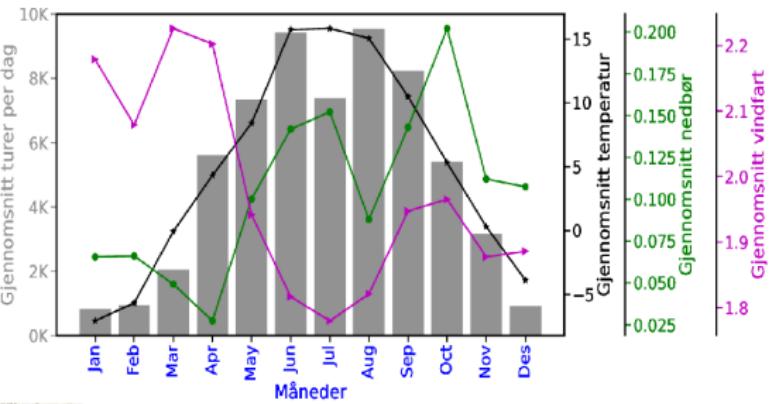
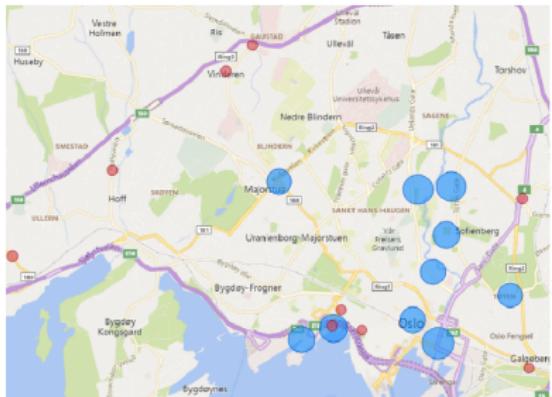
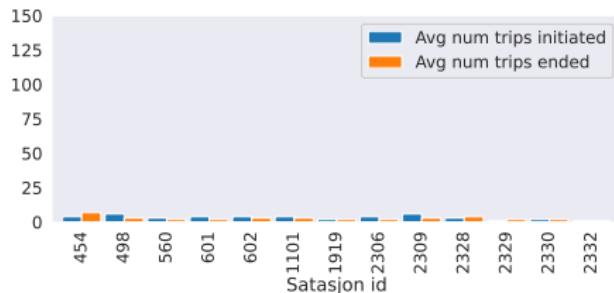
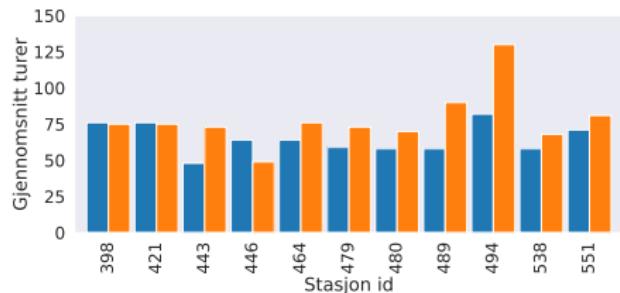
Øke kapasiteten av stasjonene



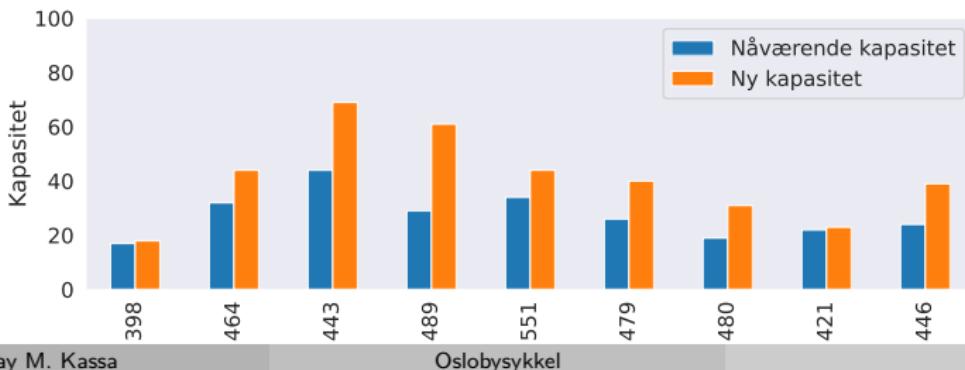
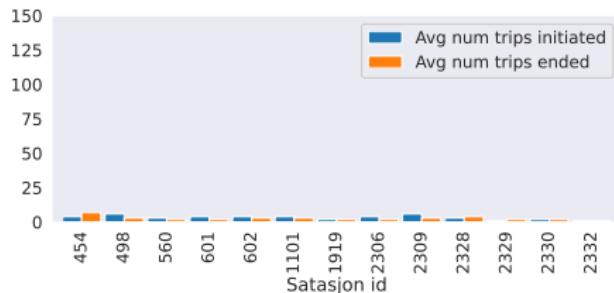
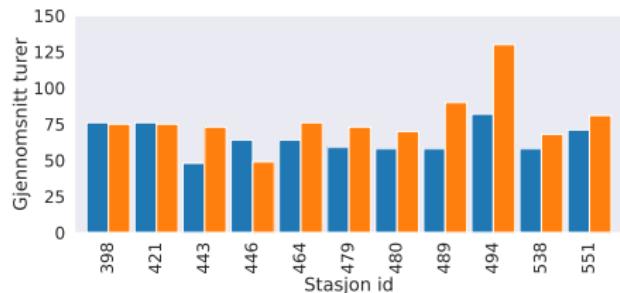
Øke kapasiteten av stasjonene



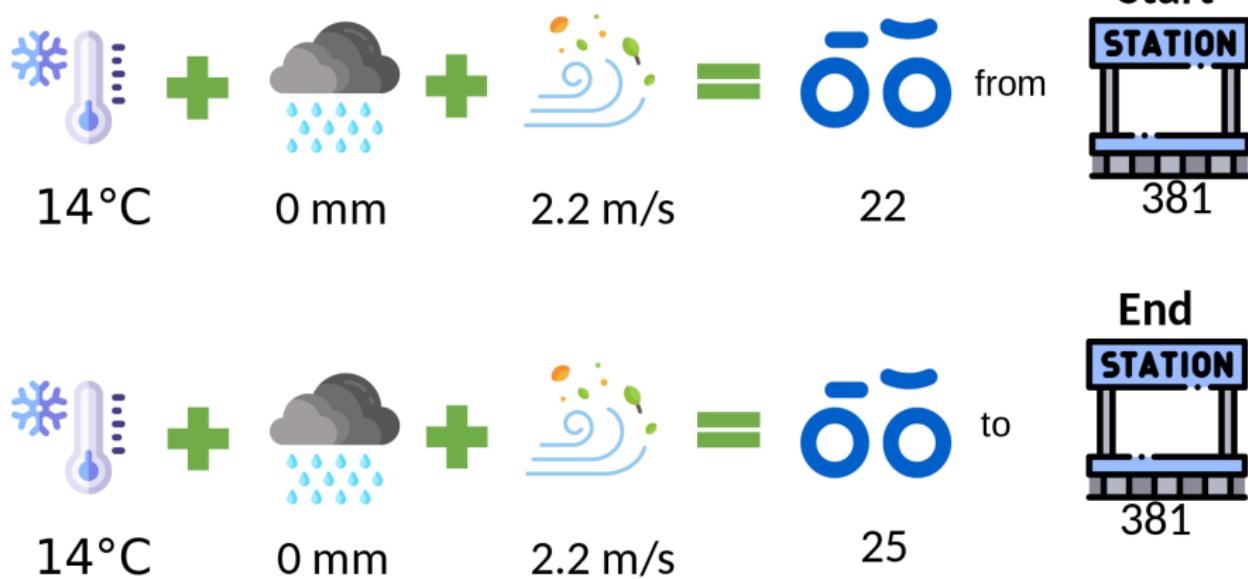
Øke kapasiteten av stasjonene



Øke kapasiteten av stasjonene

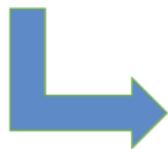


Oslobysykkel: Flyttestrategi



Oslobysykkel: Flyttestrategi

- $N = \text{UTE} - \text{INN}$
- Velg stasjonene hvor $N > 10\%$ av stasjonskapasitet



- **Mark disse stasjonene som mottakere**
- Flytte $N - \text{stasjonskapasitet} * 10\%$

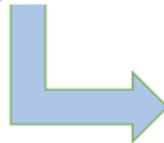
Oslobysykkel: Flyttestrategi

- N = UTE – INN
- Velg stasjonene hvor $N > 10\%$ av stasjonskapasitet



- Mark disse stasjonene som mottakere
- Flytte $N - \text{stasjonskapasitet} * 10\%$

Stasjoner og nummeret av
sykler å flytte fra

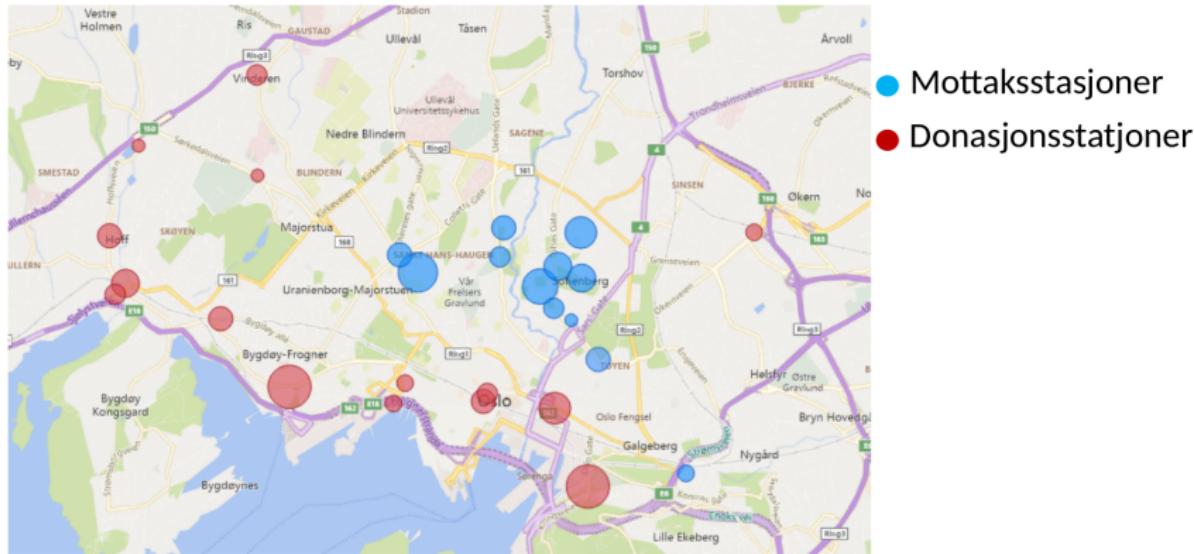


Velg $(\text{stasjonskapasitet}) * 30\%$ sykler
➢ Fra stasjoner med 80% av deres kapasitet full

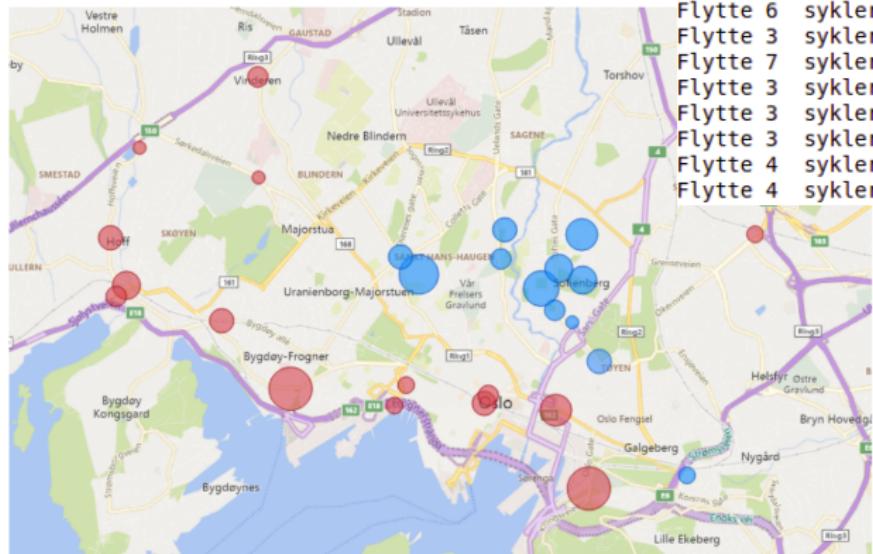
Oslobysykkel: Flyttestrategi

station_id	num_docks_available	num_bikes_available	%
2307	30	4	13.3
2306	18	2	11.1
2304	18	5	27.8
2270	19	3	15.8
614	24	11	45.8
618	42	5	11.9
621	27	4	14.8
613	27	8	29.6
620	40	12	30.0
596	18	4	22.2

Oslobysykkel: Flyttestrategi

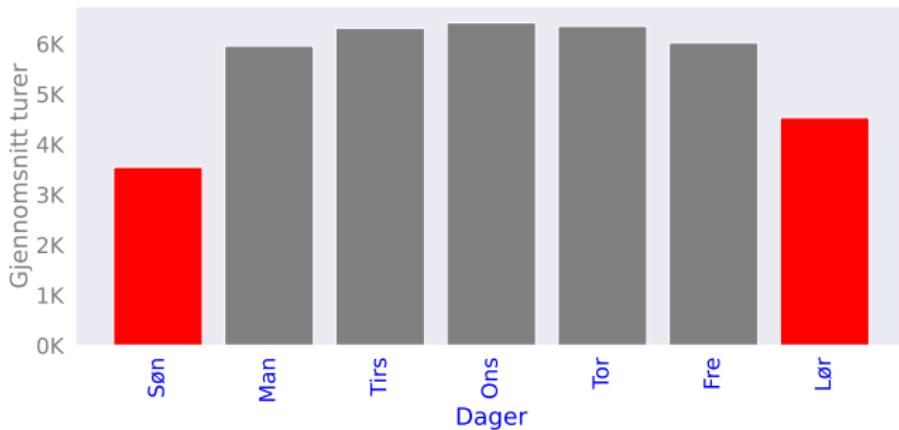


Oslobysykkel: Flyttestrategi

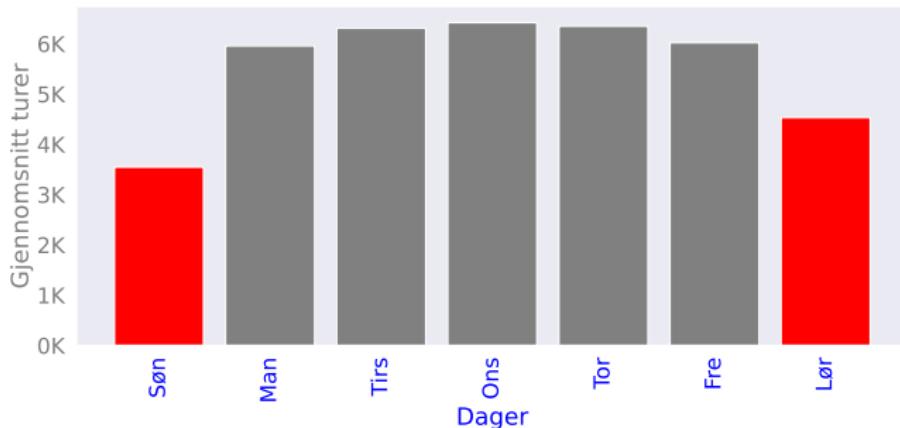


Flytte 2	sykler fra stasjon 491	til stasjon 2307
Flytte 4	sykler fra stasjon 557	til stasjon 2307
Flytte 5	sykler fra stasjon 523	til stasjon 2307
Flytte 5	sykler fra stasjon 465	til stasjon 2307
Flytte 7	sykler fra stasjon 611	til stasjon 2307
Flytte 3	sykler fra stasjon 489	til stasjon 2307
Flytte 3	sykler fra stasjon 557	til stasjon 2306
Flytte 4	sykler fra stasjon 523	til stasjon 2306
Flytte 6	sykler fra stasjon 611	til stasjon 2306
Flytte 3	sykler fra stasjon 489	til stasjon 2306
Flytte 7	sykler fra stasjon 430	til stasjon 2304
Flytte 3	sykler fra stasjon 391	til stasjon 2304
Flytte 3	sykler fra stasjon 495	til stasjon 2304
Flytte 3	sykler fra stasjon 480	til stasjon 2270
Flytte 4	sykler fra stasjon 396	til stasjon 2270
Flytte 4	sykler fra stasjon 437	til stasjon 2270

Ytterligere innsikt

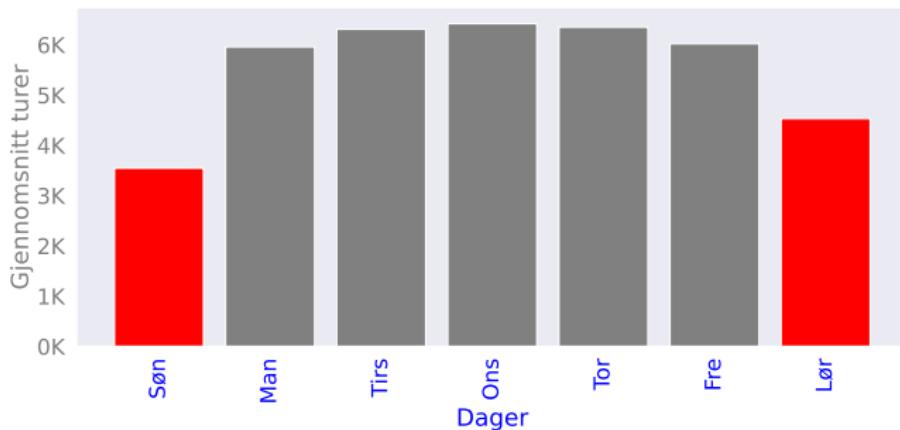


Ytterligere innsikt



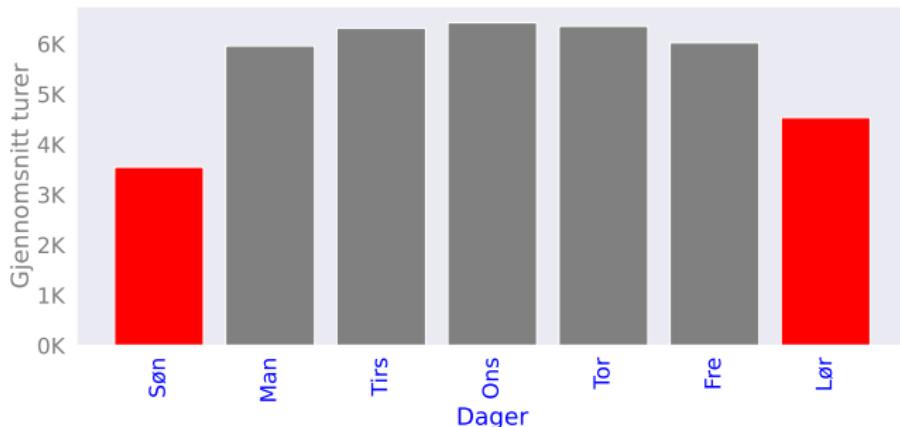
- Oppmuntre Oslofolk

Ytterligere innsikt



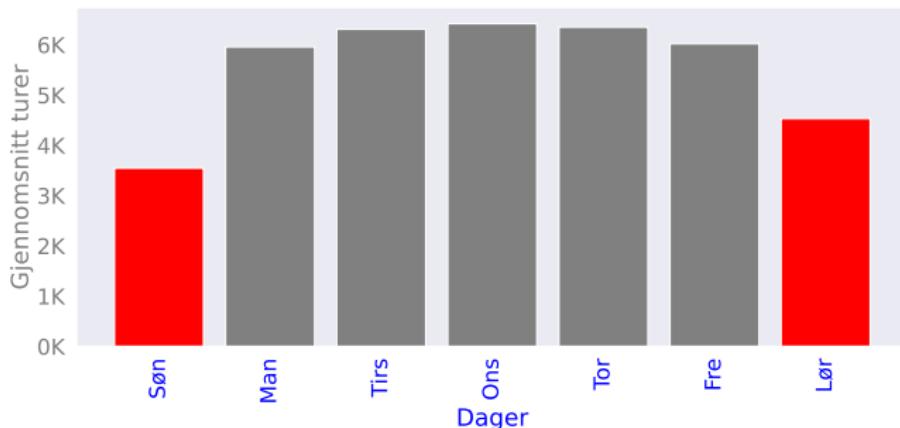
- Oppmuntre Oslofolk
 - ▶ Implementere spesialtilbud
 - ▶ Designpriser

Ytterligere innsikt



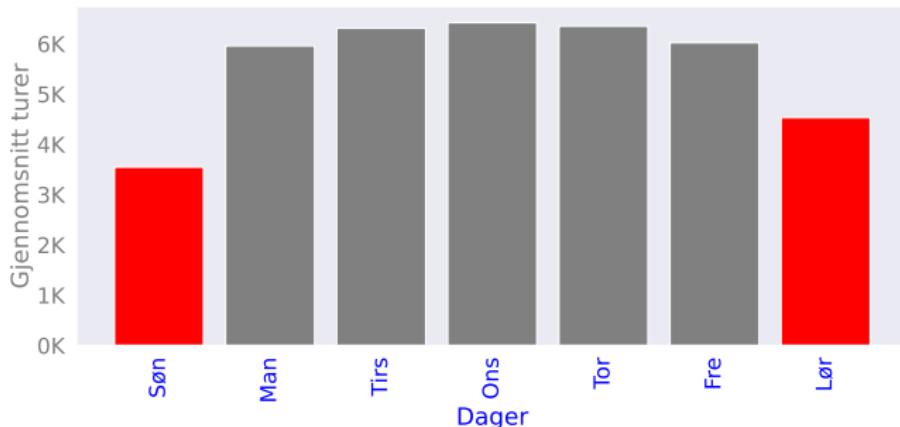
- Oppmuntre Oslofolk
 - ▶ Implementere spesialtilbud
 - ▶ Designpriser
- Jobber på barn

Ytterligere innsikt



- Oppmuntre Oslofolk
 - ▶ Implementere spesialtilbud
 - ▶ Designpriser
- Jobber på barn
- Sesong helgepass

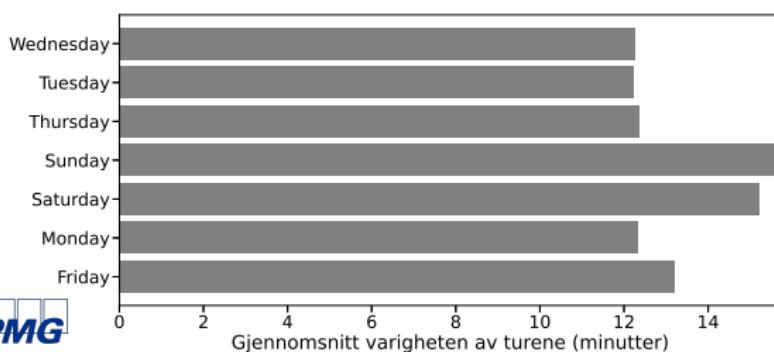
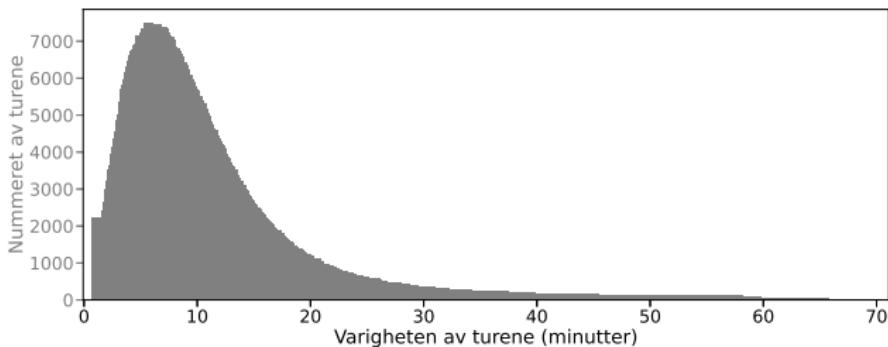
Ytterligere innsikt



- Oppmuntre Oslofolk
 - ▶ Implementere spesialtilbud
 - ▶ Designpriser
- Jobber på barn
- Sesong helgepass
- Lav helgeabonnement

Ytterligere innsikt

- Nummeret av turer

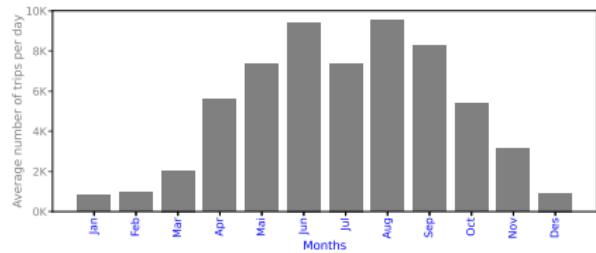


- Gjennomsnitt turer

Oppsummering

Øke kapasiteten av stasjoner

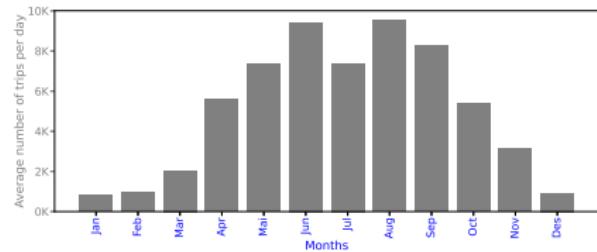
- Varmer månader



Oppsummering

Øke kapasiteten av stasjoner

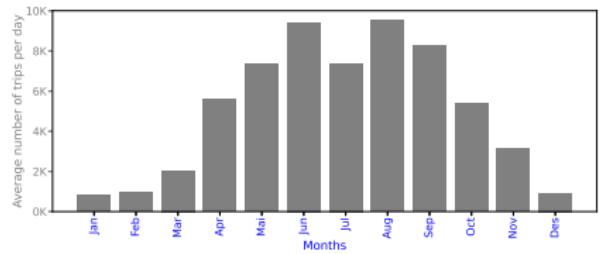
- Varmer månader
- Fra MAI til SEPTEMBER



Oppsummering

Øke kapasiteten av stasjoner

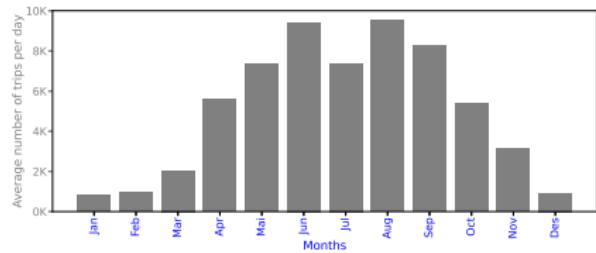
- Varmer månader
- Fra MAI til SEPTEMBER
- **MAI** er gode for reklamer



Oppsummering

Øke kapasiteten av stasjoner

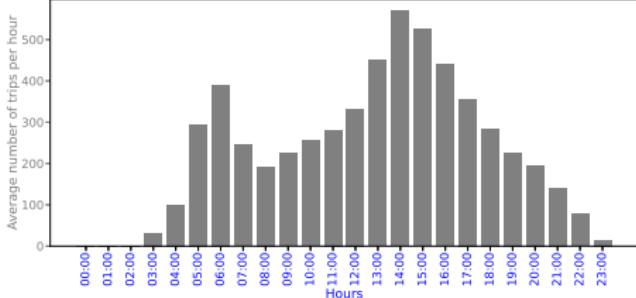
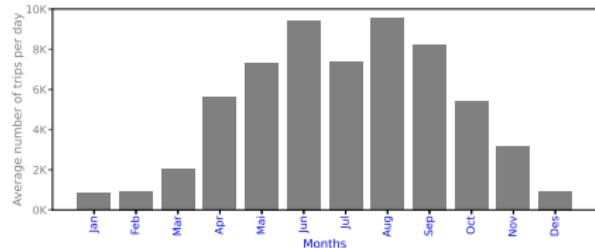
- Varmer månader
- Fra MAI til SEPTEMBER
- **MAI** er gode for reklamer



Oppsummering

Øke kapasiteten av stasjoner

- Varmer månader
- Fra MAI til SEPTEMBER
- **MAI** er gode for reklamer

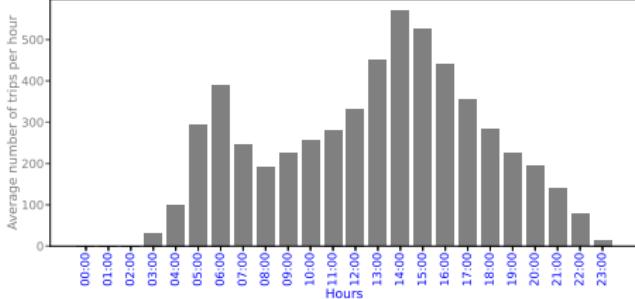
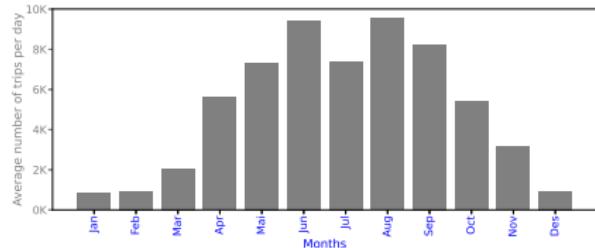


Flytte sykler på en optimal måte

Oppsummering

Øke kapasiteten av stasjoner

- Varmer månader
- Fra MAI til SEPTEMBER
- **MAI** er gode for reklamer



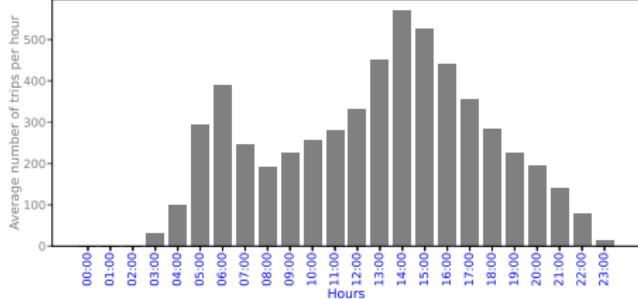
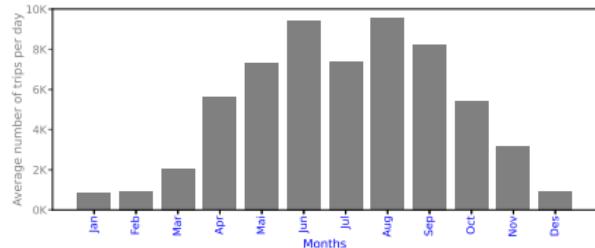
Flytte sykler på en optimal måte

- Fokusere på optattstasjonene

Oppsummering

Øke kapasiteten av stasjoner

- Varmer månader
- Fra MAI til SEPTEMBER
- **MAI** er gode for reklamer



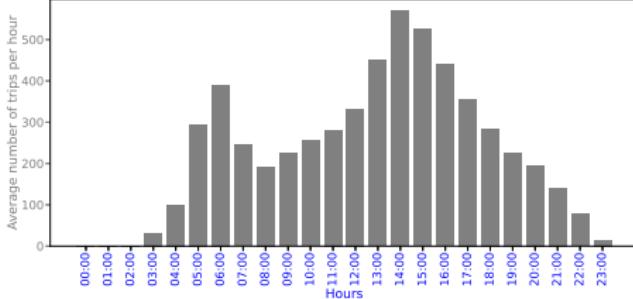
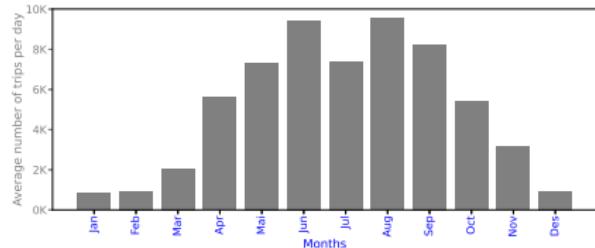
Flytte sykler på en optimal måte

- Fokusere på optattstasjonene
- Identifisere travle tider

Oppsummering

Øke kapasiteten av stasjoner

- Varmer månader
- Fra MAI til SEPTEMBER
- **MAI** er gode for reklamer



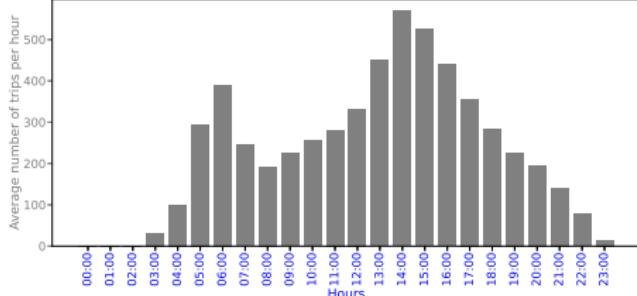
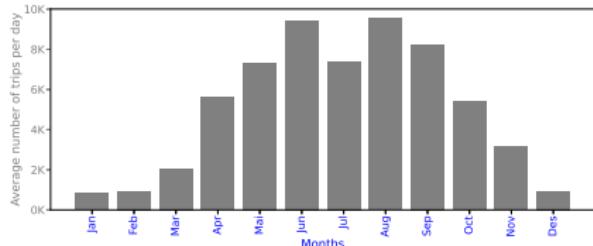
Flytte sykler på en optimal måte

- Fokusere på optattstasjonene
- Identifisere travle tider
 - ▶ Arbeidsuke

Oppsummering

Øke kapasiteten av stasjoner

- Varmer månader
- Fra MAI til SEPTEMBER
- **MAI** er gode for reklamer



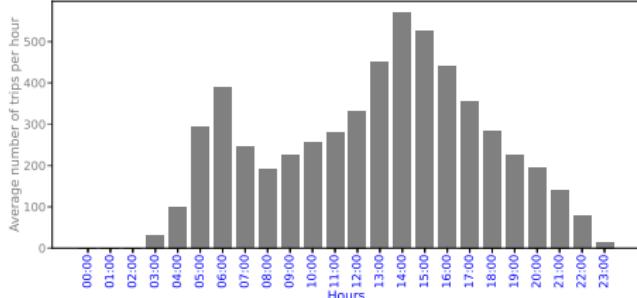
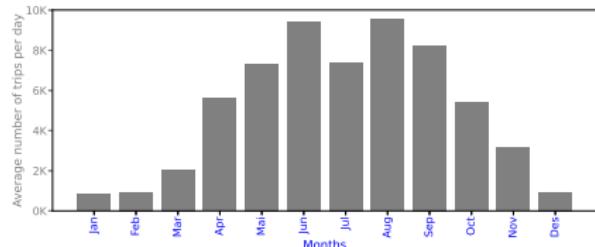
Flytte sykler på en optimal måte

- Fokusere på optattstasjoner
- Identifisere travle tider
 - ▶ Arbeidsuke
 - ▶ Til arbeidsplass og Hjem timer

Oppsummering

Øke kapasiteten av stasjoner

- Varmer månader
- Fra MAI til SEPTEMBER
- **MAI** er gode for reklamer



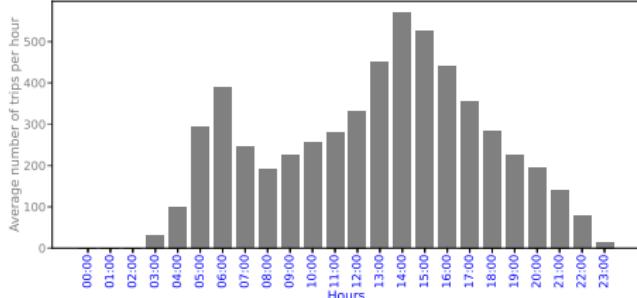
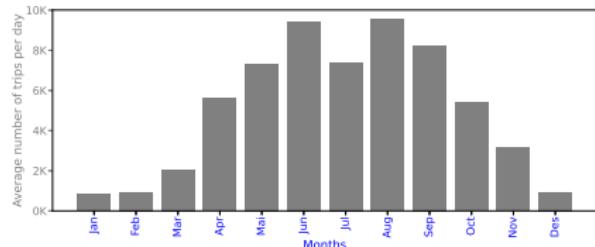
Jobber på Turist



Oppsummering

Øke kapasiteten av stasjoner

- Varmer månader
- Fra MAI til SEPTEMBER
- **MAI** er gode for reklamer



Jobber på Turist



Thank you!

