

Arbejdsløshed - Ugeeksamen i Forecasting

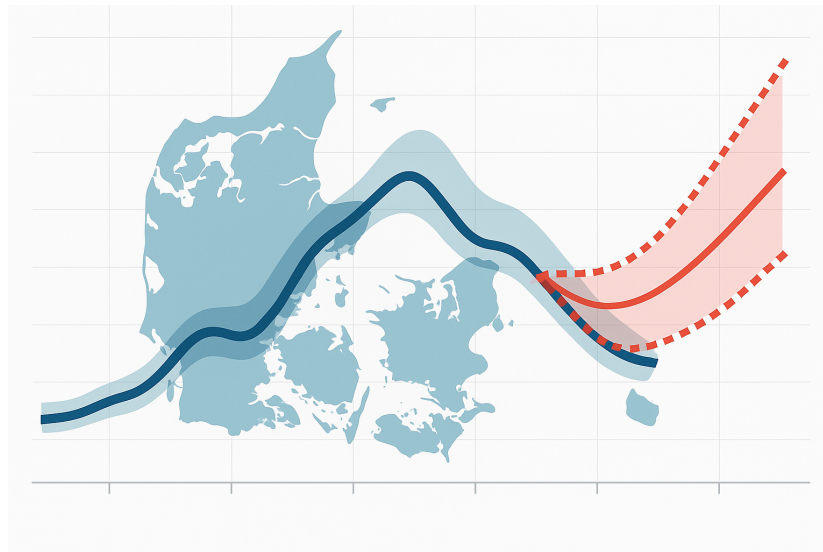
Christine Hegelund

Jing Wei

Marcus Nielsen

2025-06-13

Forecasting Eksamen



Antal tegn (inkl. mellemrum): 32.000

Vejledere:

Bjarne Taulo Sørensen

Indholdsfortegnelse

1	Introduktion	2
2	Problemformulering	2
2.1	Afgrænsning	2
2.1.1	Anvendelse af AI-værktøj	3
2.2	Definitioner/forkortelser	3
2.3	Struktur	3
3	Data og forberedelse	3
4	Eksplorativ dataanalyse (EDA)	4
4.1	Visualisering af udvikling og mønstre	4
4.1.1	Udvikling over tid	4
4.1.2	Sæsonmønstre	6
4.2	Deskriptive statistikker	8
4.2.1	Features	8
4.3	STL-dekomposition	10
5	Modelvalg	12
5.1	Valgte modeltyper	12
6	Modelvalidering	13
6.1	Splitning af datasættet	13
6.2	Time-series Cross-validation	13
6.2.1	Valg af bedste model for hver tidsserie, baseret på CV . .	14
6.3	Træning af vindermodel på hele datasættet og test på 2019-data	15
6.4	Sammenligning af trænings- og testscore	15
6.4.1	Analyse af residualer og autokorrelation	16
6.4.2	Ljung-tests for tidsserier	19
7	Sammenligning af 2019 forudsigelser og virkeligheden	20
8	Forecasting af 2020 for alle tidsserier	21

9	Konklusion	23
10	Kildeliste	24
11	Bilagsoversigt	25

1 Introduktion

Arbejdsløshedstal er en central indikator for et lands økonomiske tilstand og udvikling. De påvirker både den enkelte borger og den nationale økonomi og indgår som en væsentlig faktor i politiske og økonomiske beslutningsprocesser. I takt med øget datatilgængelighed og forbedrede statistiske værktøjer er det blevet muligt at analysere og forudsige sådanne udviklingstræk med større præcision.

Denne opgave undersøger udviklingen i arbejdsløsheden i Danmark i perioden 2007 til 2019 – opdelt på køn og region. Datagrundlaget består af ti tidsserier, én for hver kombination af fem regioner og to køn. Det giver mulighed for at analysere både regionale forskelle og kønsspecifikke mønstre i arbejdsløsheden.

Med udgangspunkt i klassiske tidsseriemodeller, ARIMA, ETS og en simpel benchmarkmodel (Seasonal Naive), undersøges, hvordan arbejdsløsheden har udviklet sig, og hvordan den kan fremskrives. Undervejs vurderes modellernes egnethed med fokus på præcision og residualstruktur, og der anvendes teknikker som STL-dekomposition og transformation af data. Målet er ikke blot at forudsige udviklingen i 2020, men også at vurdere, hvor godt klassiske modeller formår at håndtere forskellene mellem regioner og køn.

2 Problemformulering

Hvordan kan klassiske tidsseriemodeller anvendes til at analysere og forudsige udviklingen i arbejdsløsheden i Danmark, fordelt på region og køn?

For at besvare dette spørgsmål undersøges følgende delspørgsmål:

1. Hvordan varierer arbejdsløshedens udvikling og sæsonmønstre på tværs af regioner og køn, og hvordan kan disse identificeres gennem eksplorativ dataanalyse og dekomposition?
2. Hvordan performer modellerne ARIMA, ETS og Seasonal Naive i forhold til hinanden, når det gælder præcision, residualstruktur og prognoseegenskaber?
3. Hvilke modeller er bedst egnede til at fremskrive arbejdsløshedstallene for 2020, og hvordan varierer usikkerheden på tværs af serier?

2.1 Afgrænsning

Analysen baserer sig udelukkende på de udleverede arbejdsløshedstal for perioden januar 2007 til december 2019 og anvender ingen eksterne forklaringsvariable såsom COVID-19, økonomiske indikatorer eller politiske reformer. Dette valg er truffet for at isolere og vurdere forecasting-modellernes metodiske egenskaber.

Fokus er dermed på reproducerbar, statistisk modellering – ikke på årsagsforklaring eller samfundsøkonomisk fortolkning.

2.1.1 Anvendelse af AI-værktøj

ChatGPT (GPT-4o) er anvendt som støtteværktøj i forbindelse med idéudvikling, sproglig formulering og udformning af enkelte kdestumper. Værktøjet er kun brugt til tekniske og sproglige formål – al analyse, fortolkning og konklusion er udarbejdet selvstændigt af gruppens medlemmer.

2.2 Definitioner/forkortelser

Nedenfor er en oversigt over centrale begreber og forkortelser, som anvendes i opgaven:

- **ARIMA**: AutoRegressive Integrated Moving Average
- **ETS**: Exponential Smoothing State Space Model
- **SNAÏVE**: Seasonal Naive
- **RMSE**: Root Mean Squared Error
- **MAPE**: Mean Absolute Percentage Error
- **STL**: Seasonal-Trend decomposition using Loess
- **CV**: Cross-validation

2.3 Struktur

Opgavens struktur følger en klassisk tilgang til tidsserieanalyse og er inddelt i fem faser. Først gennemføres en eksplorativ dataanalyse med fokus på mønstre og variationer i arbejdsløsheden fordelt på region og køn. Dernæst estimeres tre modeller (ARIMA, ETS og Seasonal Naive) for hver tidsserie. I tredje fase valideres modellerne gennem krydsvalidering og residualanalyse. Herefter foretages fremskrivninger for 2020, og til sidst sammenlignes og konkluderes der på tværs af modeller og dataserier.

3 Data og forberedelse

Analysen bygger på månedlige arbejdsløshedstal fra Danmark i perioden januar 2007 til december 2019. Data er opdelt efter køn og region, hvilket giver ti separate tidsserier. Datasættet er udleveret i et forbehandlet tsibble-format med tydeligt definerede indeks- og nøglevariabler, og nedenfor indlæses datasættet:

```
# A tsibble: 6 x 4 [1M]
# Key:      kon, region [1]
  kon      region      yearmonth svalue
  <fct>    <fct>          <mth>  <dbl>
1 Kvinder Region Hovedstaden 2007 Jan   2.26
2 Kvinder Region Hovedstaden 2007 Feb   2.19
3 Kvinder Region Hovedstaden 2007 Mar   2.09
4 Kvinder Region Hovedstaden 2007 Apr   1.97
5 Kvinder Region Hovedstaden 2007 May   1.99
6 Kvinder Region Hovedstaden 2007 Jun   1.93
```

4 Eksplorativ dataanalyse (EDA)

Formålet med dette afsnit er at skabe et overblik over datasættets struktur forud for modelleringen. Ved hjælp af visualiseringer, deskriptive statistikker og STL-dekomposition undersøges arbejdsløshedens udvikling på tværs af regioner og køn i perioden 2007 til 2019. Analysen afdækker overordnede tendenser, sæsonmønstre og forskelle i niveau og variation. Resultaterne danner det metodiske fundament for det videre modelarbejde.

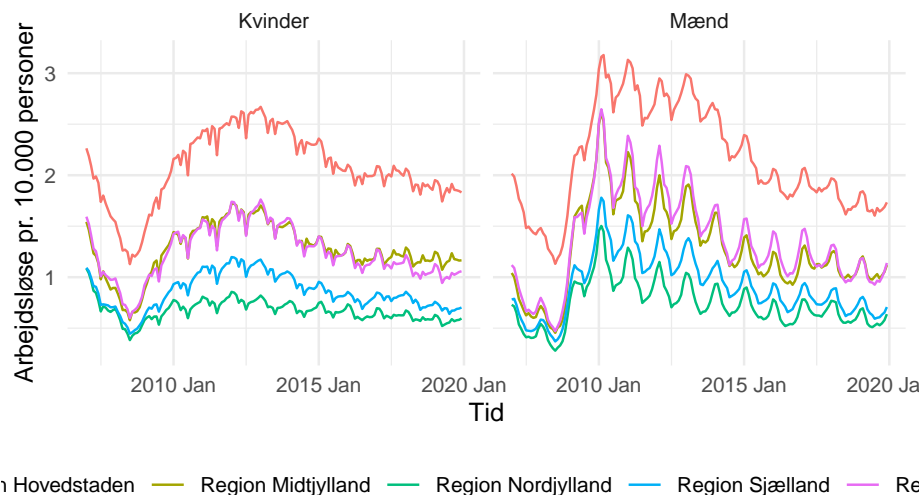
4.1 Visualisering af udvikling og mønstre

I dette afsnit undersøges, hvordan arbejdsløsheden har udviklet sig over tid – med fokus på tendenser, sæsonvariationer og forskelle mellem regioner og køn. Formålet er at identificere mønstre, som kan guide valg af modeller i den videre analyse.

4.1.1 Udvikling over tid

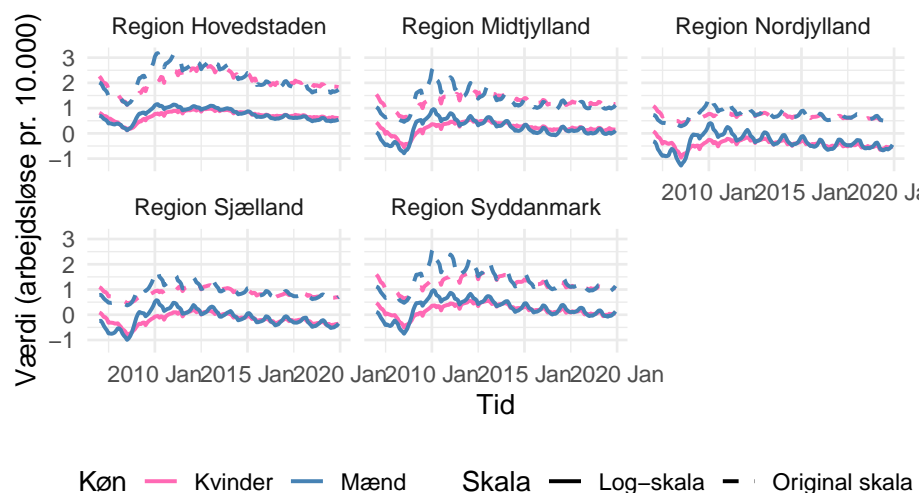
I de følgende figurer undersøges udviklingen i arbejdsløsheden over tid. Til at begynde med anvendes den oprindelige skala for at give et lettilgængeligt overblik (Figur 1). Fra Figur 2 og frem benyttes en log-transformation for at stabilisere variationen i serierne, især blandt mænd, og for at sikre sammenlignelighed i videre modellering og forecast.

Figur 1: Udvikling i arbejdsløsheden, fordelt på region og køn



Ovenstående figur viser den månedlige arbejdsløshed blandt mænd og kvinder i de fem danske regioner fra 2007 til 2019. Der fremgår et tydeligt sæsonmønster med højere ledighed i vintermånederne og lavere i sommerperioden. Niveauet varierer regionalt, hvor Region Hovedstaden typisk ligger højest og Nordjylland lavest. Samlet ses en faldende tendens efter finanskrisen.

Figur 2: Sammenligning af original og log-skala



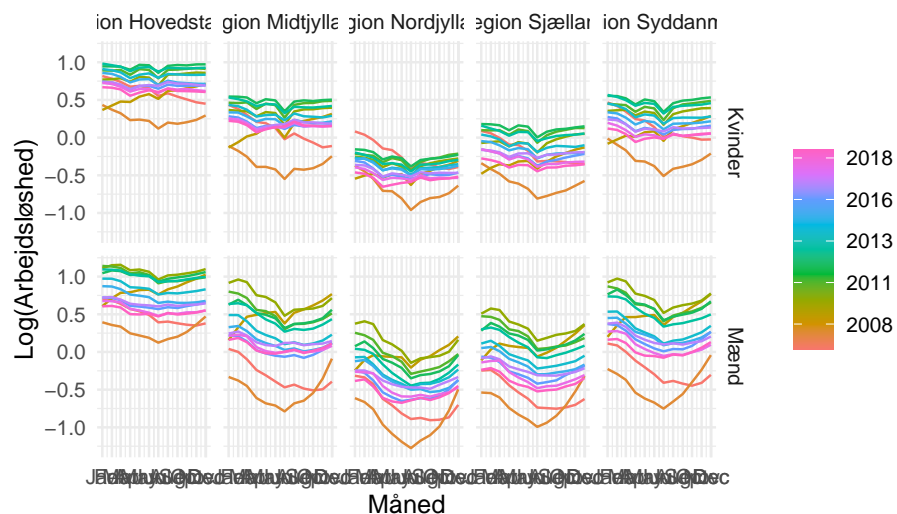
Figur 2 viser, hvordan log-transformeringen dæmper udsvingene, særligt i serier med højt niveau og stor variation, som det ses blandt mænd. Denne transforma-

tion gør det lettere at sammenligne udviklingen på tværs af regioner og køn og sikrer en mere stabil varians i det videre analysearbejde. Derfor anvendes den log-transformerede version fremadrettet i analysen.

4.1.2 Sæsonmønstre

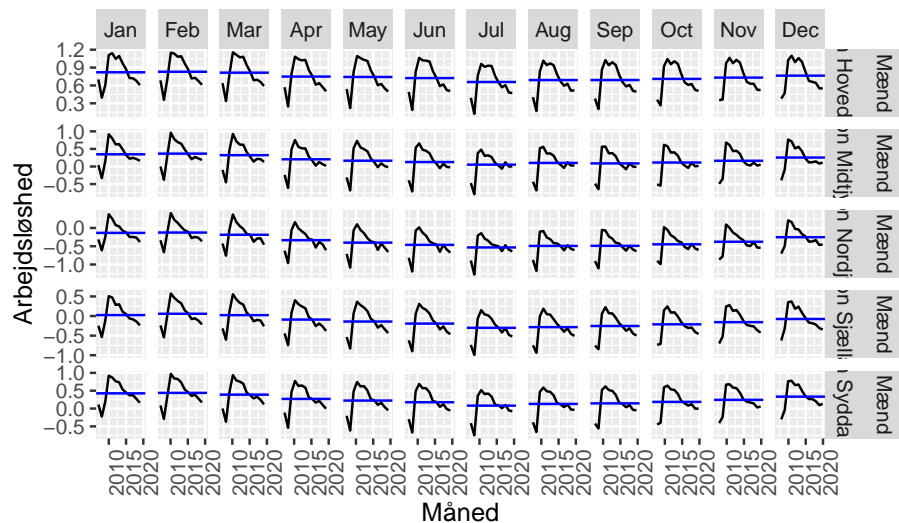
Med log-transformerede data som fundament undersøges nu sæsonmønstrene i arbejdsløsheden nærmere. Gennem sæsonplots og subserieplots vurderes, hvordan arbejdsløsheden typisk varierer over året – og hvordan dette adskiller sig mellem mænd og kvinder samt på tværs af regioner.

Figur 3: Sæsonmønstre i arbejdsløsheden pr. region og køn



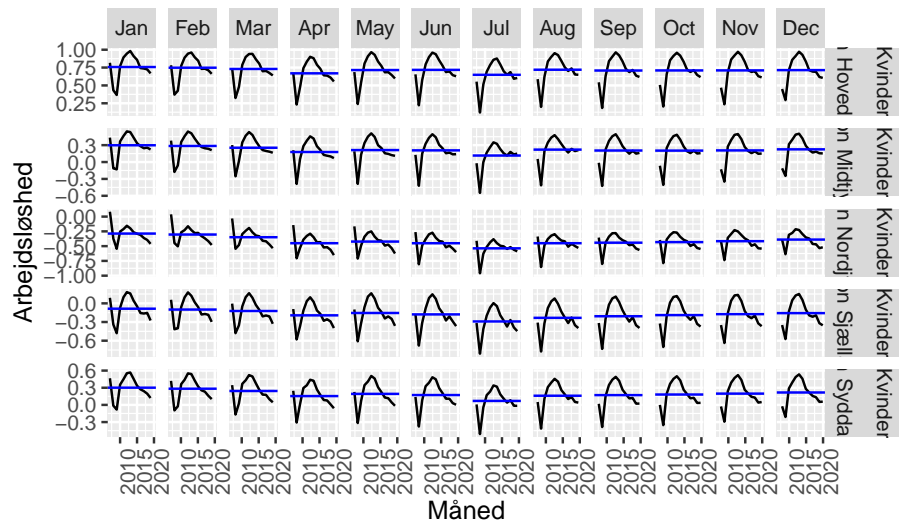
Figur 3 viser sæsonmønstre i arbejdsløsheden pr. region og køn (log-transformeret). Hver farve repræsenterer et år, og linjerne viser en tydelig sæsonrytme med høj ledighed i årets begyndelse og lav i sommermånederne. Mænd udviser generelt større sæsonudsving end kvinder. Niveaue er højest i Region Hovedstaden og lavest i Nordjylland og Sjælland. Negative log-værdier forekommer, hvor arbejdsløsheden ligger under 1 pr. 10.000.

Figur 4: Subserieplot for mænd pr. måned og region



Figur 4 viser subserieplottet for arbejdsløshedens udvikling for mænd pr. måned og region. Hver celle viser forløbet for en specifik måned over årene, med sorte linjer for de enkelte år og blå linjer for gennemsnit. Ledigheden toppe typisk i årets første måneder og falder hen over sommeren. Mænd udviser generelt større sæsonudsving og højere variation mellem år, især i vinterhalvåret.

Figur 5: Subserieplot for kvinder pr. måned og region



Sammenlignet med mænd viser kvinder tilsvarende sæsonmønstre med højere ledighed i årets begyndelse og lavere i sommermånederne. Udsvingene er dog

mindre markante, og variationen mellem år er mindre tydelig. Mønstret fremstår mere stabilt og jævnt fordelt på tværs af regioner, hvilket tyder på lavere volatilitet i kvindernes arbejdsløshed.

Udover sæsonmæssige udsving varierer arbejdsløsheden betydeligt på tværs af både geografi og køn. Dette undersøges nærmere i de følgende visualiseringer.

For at underbygge disse visuelle indsigter med konkrete tal, præsenteres i næste afsnit deskriptive statistikker. De giver en numerisk opsummering af centrale mål som gennemsnit, variation og ekstreme værdier og bidrager til at identificere særligt stabile eller volatile grupper, der kan kræve særlig opmærksomhed i den videre analyse.

4.2 Deskriptive statistikker

De foregående visualiseringer viste klare forskelle i arbejdsløshedens niveau, variation og sæsonmønstre på tværs af regioner og køn (Figur 1–5). For at supplere disse observationer med en kvantitativ opsummering præsenteres i Tabel 1 centrale deskriptive mål: gennemsnit, standardafvigelse, minimum og maksimum. Tabellen giver et hurtigt overblik over niveau og variation og fremhæver for eksempel Nordjylland som en stabil region, særligt blandt kvinder, mens andre kombinationer viser højere udsving.

Tabel 1: Deskriptiv statistik per region og køn

region	køn	Gennemsnit	Standardafvigelse	Minimum	Maksimum
Region Hovedstaden	Kvinder	2.07	0.36	1.13	2.67
Region Hovedstaden	Mænd	2.17	0.52	1.13	3.18
Region Midtjylland	Kvinder	1.27	0.25	0.58	1.72
Region Midtjylland	Mænd	1.29	0.44	0.45	2.62
Region Nordjylland	Kvinder	0.67	0.10	0.38	1.09
Region Nordjylland	Mænd	0.74	0.23	0.28	1.50
Region Sjælland	Kvinder	0.86	0.17	0.45	1.20
Region Sjælland	Mænd	0.92	0.30	0.37	1.78
Region Syddanmark	Kvinder	1.24	0.26	0.60	1.76
Region Syddanmark	Mænd	1.37	0.47	0.47	2.65

4.2.1 Features

Selvom deskriptive statistikker giver et første indtryk af variation og niveau, indfanger de ikke strukturelle træk som trend, sæsonmønstre og stationaritet. For at kvantificere disse egenskaber beregnes en række såkaldte features ved hjælp af ved hjælp af funktionen `features()` fra `feasts`-pakken (Hyndman & Athanasopoulos, 2021, kap. 4).

Tabel 2 præsenterer udvalgte mål, der beskriver centrale træk ved serierne og uddyber de visuelle indsigter fra Figur 1–5. Den bagvedliggende kode, herunder brugen af `features()`-funktionen og indlæsning af dataobjektet `feature_table.rds`, er dokumenteret i Bilag X.

Tabel 2: Deskriptiv statistik per region og køn

kon	region	acf1	seasonal_strength_year	var_tiled_mean	shift_level_max
Kvinder	Region Hovedstaden	0.97	0.65	0.97	0.73
Kvinder	Region Midtjylland	0.96	0.72	0.92	0.73
Kvinder	Region Nordjylland	0.89	0.87	0.73	0.73
Kvinder	Region Sjælland	0.96	0.81	0.92	0.73
Kvinder	Region Syddanmark	0.96	0.83	0.91	0.73
Mænd	Region Hovedstaden	0.98	0.81	0.98	0.73
Mænd	Region Midtjylland	0.98	0.84	0.93	0.73
Mænd	Region Nordjylland	0.96	0.91	0.80	0.73
Mænd	Region Sjælland	0.97	0.92	0.90	0.73
Mænd	Region Syddanmark	0.98	0.88	0.91	0.73

De kvantitative mål i Tabel 2 supplerer og styrker tolkningerne fra Figur 1 til 5. I det følgende fremhæves de mest centrale mønstre med fokus på fem analytiske temaer: autokorrelation, sæsonmønstre, variation og skift, stationaritet og trend.

Autokorrelation

Variablen `acf1` måler korttidsafhængighed, det vil sige, hvor stærkt observationer i serien er knyttet til deres nære fortid. Værdierne ligger generelt højt (typisk mellem 0,96 og 0,98), hvilket indikerer en glidende udvikling uden pludselige udsving. Dette stemmer overens med de jævne bevægelser og stabile tendenser, som blev observeret i Figur 1 og Figur 2.

Sæsonmønstre

Styrken af årlige sæsonmønstre, målt med `seasonal_strength_year`, varierer på tværs af grupper. Mænd udviser generelt mere udtalte sæsonudsving end kvinder. For eksempel har mænd i Nordjylland og Sjælland værdier på henholdsvis 0,91 og 0,92. Det svarer til de tydelige sæsonrytmer vist i Figur 3, Figur 4 og Figur 5. I modsætning hertil har kvinder i Region Hovedstaden en lavere værdi på 0,65, hvilket indikerer en mere stabil sæsonprofil.

Variation og skift

Målet `var_tiled_mean` afspejler graden af variation inden for afgrænsede perioder. Kvinder i Nordjylland har den laveste værdi (0,73), hvilket viser en forholdsvis stabil udvikling over tid. Dette understøttes af de glatte kurver i Figur 5. Variablen `shift_level_max` måler derimod pludselige niveauskift. Mænd i Midtjylland udviser den højeste værdi (1,05), hvilket sandsynligvis relaterer sig

til reaktioner på finanskrisen og kan ses i de kraftige fald i arbejdsløsheden i Figur 1 og Figur 2.

Stationaritet

Antallet af nødvendige differensieringer (ndiffs) angiver, hvorvidt en serie er stationær. De fleste serier kræver én differens, men enkelte, herunder kvinder i Nordjylland og mænd i både Midtjylland og Nordjylland, har værdi 0. Dette indikerer, at de allerede er stationære uden transformation, hvilket også kommer til udtryk i deres stabile forløb i Figur 3 og Figur 5.

Trend

Variablen `trend_strength` måler graden af trendstruktur i serien. Værdierne ligger generelt meget højt (typisk 0,97 til 0,99), hvilket bekræfter tydelige underliggende tendenser i arbejdsløsheden. Det gælder især Region Hovedstaden, hvor niveauet er højt, og udviklingen over tid er markant, som det fremgår af Figur 1 og Figur 2.

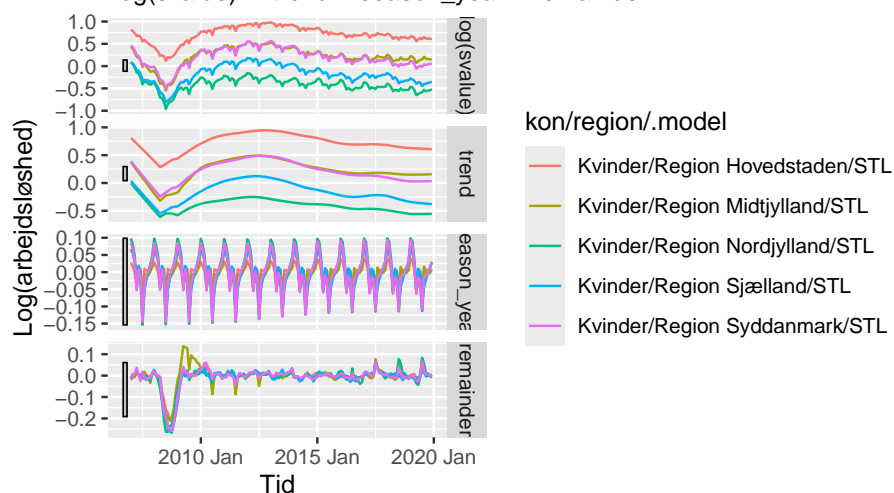
Samlet set bidrager funktionerne i Tabel 2 til at kvantificere og validere de mønstre, der tidligere blev identificeret visuelt. Autokorrelation, sæsonvariation og trendstruktur fremstår som gennemgående karakteristika i alle serier, men med klare forskelle mellem køn og regioner. Resultaterne danner et vigtigt grundlag for valg af modeller og videre analyse.

4.3 STL-dekomposition

STL-dekomposition anvendes for at undersøge de strukturelle komponenter i arbejdsløshedsserierne. Metoden opdeler tidsserierne i tre dele: trend, sæson og remainder, hvilket giver indblik i, hvor stor en del af variationen der skyldes henholdsvis langsigtet udvikling, tilbagevendende mønstre eller kortsigtede udsving (Hyndman og Athanasopoulos, 2021). Dekompositionen er udført separat for mænd og kvinder i hver region, og resultaterne vises i figur 6 og 7.

Figur 6: STL-dekomposition for kvinder – log-transformeret

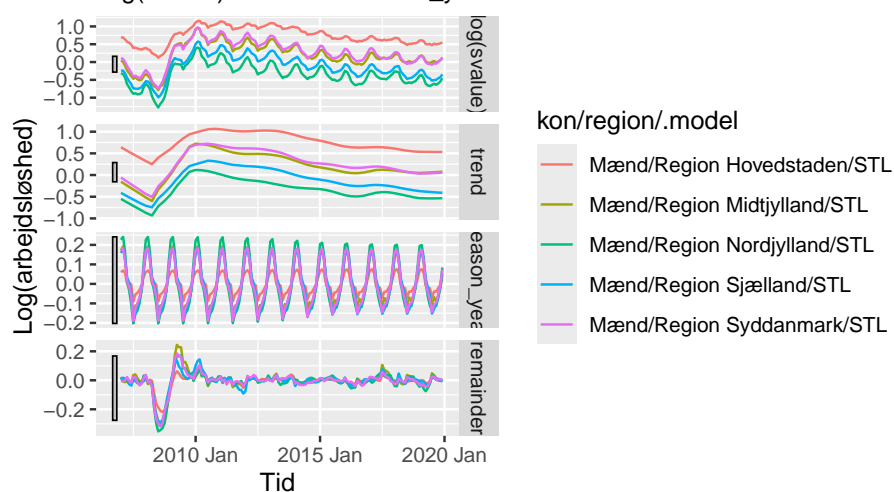
'log(svalue)' = trend + season_year + remainder



Figur 6 viser STL-dekompositionen for kvinder i de fem danske regioner. Der ses tydelige sæsonmønstre med tilbagevendende lavpunkter i sommermånederne og højere ledighed i vinterhalvåret. Trends varierer mellem regionerne: Region Hovedstaden udviser generelt højere ledighed, mens Nordjylland og Sjælland ligger lavere. Residualerne ligger stabilt omkring nul, hvilket indikerer, at modellen formår at fange de dominerende strukturer i data.

Figur 7: STL-dekomposition for mænd – log-transformeret

'log(svalue)' = trend + season_year + remainder



Figur 7 viser tilsvarende dekomposition for mænd. Her ses en tilsvarende stærk

sæsonkomponent, men med mere markante udsving og højere ledighedsniveauer – især efter finanskrisen. Trends følger samme overordnede forløb som for kvinder, men afvigelserne er større. Dette bekræfter tidligere observationer om højere volatilitet i mænds ledighed. Residualkomponenten er mere varierende, hvilket kan indikere, at nogle udsving ikke fanges fuldt ud af modellen.

Disse observationer fra STL-dekompositionen bekræfter, at både trend og sæsonmønstre varierer betydeligt på tværs af serierne. Det understreger behovet for modeller, der kan tilpasses hver series struktur, hvilket vil blive præsenteret i det følgende afsnit.

5 Modelvalg

I dette afsnit estimeres og sammenlignes tre klassiske modeltyper, ARIMA, ETS og benchmarkmodellen SNaive, for hver tidsserie ($\text{region} \times \text{køn}$). Hver modeltype har forskellige antagelser og fordele og kan fange forskellige karakteristika i data, såsom trend, sæson og autokorrelation. Modelvalget foretages automatisk for hver serie via fable-pakken, og resultaterne anvendes i den efterfølgende validering og forecast.

5.1 Valgte modeltyper

ARIMA (AutoRegressive Integrated Moving Average)

ARIMA-modeller anvendes til at modellere tidsserier med autokorrelation og ikke-stationaritet. De kombinerer autoregressive (AR) led, differens (I) for at opnå stationaritet, og glidende gennemsnit (MA) led til at modellere fejl. Ved at inkludere sæsonkomponenter (SARIMA) kan modellen håndtere årligt gentagende mønstre. ARIMA er særligt velegnet til serier med langsigtede trends og strukturelle skift, hvor tidligere værdier har stor indflydelse på den nuværende tilstand (Hyndman & Athanasopoulos, 2021, kap. 9).

ETS (Exponential Smoothing State Space Model)

ETS-modeller benytter eksponentiel glatning til at vægte nyere observationer højere og beskriver tidsserier ud fra tre elementer: fejl, trend og sæson. Hver komponent kan være additiv eller multiplicativ afhængigt af dataens karakter. ETS egner sig godt til serier med tydelige strukturer og sæsonrytmer, især når der ikke er behov for at modellere autokorrelation direkte (Hyndman & Athanasopoulos, 2021, kap. 8).

SNAÏVE (Seasonal Naive)

SNaive er en enkel benchmarkmodel, der baserer hvert forecast på den tilsvarende værdi fra samme sæson i den foregående periode. Modellen fanger sæsonmønstre, men tager ikke højde for trend eller afhængighed i data. Den anvendes til at vurdere, om mere komplekse modeller giver en mærkbart bedre forecast-præcision (Hyndman & Athanasopoulos, 2021, kap. 3 og 8).

6 Modelvalidering

Dette afsnit har til formål at sikre, at modellerne giver pålidelige forecasts og ikke blot overtilpasser sig historiske data. Gennem tidsserie-crossvalidering, fejlmålinger som RMSE og MAPE, samt residualanalyse identificeres den bedst egnede model for hver serie, baseret på forecastpræcision og modeladfærd.

6.1 Splitning af datasættet

Vi starter med at splitte datasættet op i et træningssæt og et testsæt. Test-sættet er dataen fra hele 2019, som er det sidste år i det samlede datasæt, og træningssættet er alt den forrige data fra 2007-2018.

```
train_data <- data |> filter_index(. ~ "2018 Dec")
test_data  <- data |> filter_index("2019 Jan" ~ "2019 Dec")
```

6.2 Time-series Cross-validation

Herunder udføres tidsserie-krydsvalidering udelukkende på træningsdataen med et rullende vindue på 60 måneder. Vi anvender `stretch_tsibble()` til at generere flere trænings- og test-split, som forskydes én måned ad gangen. Tre modeller evalueres: ARIMA, ETS og en sæsonbetonet naiv model (SNaive) – alle på log-transformerede data.

Forecastfejlene beregnes med `accuracy()`, og beregningerne køres parallelt for hurtigere eksekvering. Resultaterne gemmes med `write_rds()` for reproducérbarhed.

```
# plan(multisession)

# cv_data <- train_data |>
#   stretch_tsibble(.init = 60, .step = 1)

# cv_models <- cv_data |>
#   model(
#     ARIMA = ARIMA(log(svalue)),
#     ETS    = ETS(log(svalue)),
#     SNaive = SNAIVE(log(svalue))
#   )

# write_rds(cv_models, "cv_models.rds")

cv_models <- read_rds("data/cv_models.rds")
```

```
cv_accuracy <- cv_models |>
  accuracy()

# plan(sequential)
```

6.2.1 Valg af bedste model for hver tidsserie, baseret på CV

Her identificeres den bedste model for hver kombination af køn og region baseret på laveste RMSE. Ved at gruppere på køn og region og vælge modellen med mindst fejl (`slice_min()`), får vi én vinder pr. serie uden ligestilling mellem modeller (`with_ties = FALSE`). Dette gør det muligt senere at bruge den bedste model til forecast for hver delserie.

```
vinder_cv <- cv_accuracy |>
  group_by(køn, region) |>
  slice_min(RMSE, n = 1, with_ties = FALSE) |>
  ungroup()

vinder_cv |> select(køn, region, .model)
```

```
# A tibble: 10 x 3
   køn      region      .model
<fct> <fct>      <chr>
1 Kvinder Region Hovedstaden ETS
2 Kvinder Region Midtjylland ARIMA
3 Kvinder Region Nordjylland ARIMA
4 Kvinder Region Sjælland   ARIMA
5 Kvinder Region Syddanmark ARIMA
6 Mænd    Region Hovedstaden ARIMA
7 Mænd    Region Midtjylland ARIMA
8 Mænd    Region Nordjylland ARIMA
9 Mænd    Region Sjælland   ARIMA
10 Mænd    Region Syddanmark ARIMA
```

Resultaterne viser, at ARIMA-modellen klarer sig bedst i 9 ud af 10 tidsserier. Den eneste undtagelse er kvinder i Region Hovedstaden, hvor en ETS-model giver lavest RMSE. Dette tyder på, at ARIMA generelt formår at tilpasse sig strukturen i dataserierne bedre, hvilket ofte skyldes modellens fleksibilitet i forhold til både trend og sæson. Den ene ETS-vinder indikerer dog, at i nogle tilfælde kan eksponentiel glatning være mere passende – muligvis pga. mere stabil sæson uden kompleks autokorrelation.

6.3 Træning af vindermodel på hele datasættet og test på 2019-data

```
model_train <- train_data |>
  model(
    ARIMA = ARIMA(log(svalue)),
    ETS    = ETS(log(svalue)),
    SNaive = SNAIVE(log(svalue))
  )

train_accuracy <- model_train |>
  accuracy() |>
  select(kon, region, .model, RMSE_tr = RMSE, MAPE_tr = MAPE)

forecast_2019 <- model_train |>
  forecast(h = "12 months") |>
  inner_join(vinder_cv,
    by = c("kon", "region", ".model"))
```

Her gentrænes alle tre modeller på hele træningsperioden fra 2010 til 2018. Derefter laves et 12-måneders forecast for 2019, som svarer til testperioden. Ved at forecaste alle modeller og derefter filtrere til den tidligere udpegede vinder (`inner_join()` med `vinder_cv`), sikrer vi, at den bedste model anvendes til hver serie i den endelige evaluering.

Her sammenlignes trænings- og testfejl for hver serie ved at kombinere dem i én tabel (`train_vs_test`). Dette giver overblik over, hvorvidt modellerne generaliserer godt — altså om performance på testdata svarer til performance på træningsdata.

6.4 Sammenligning af trænings- og testscore

```
test_accuracy <- forecast_2019 |>
  accuracy(test_data) |>
  select(kon, region, .model, RMSE, MAPE)

train_vs_test <- left_join(train_accuracy, test_accuracy,
  by = c("kon", "region", ".model"))

train_vs_test <- train_vs_test |>
  filter(!is.na(RMSE)) |>
  relocate(RMSE, .after = RMSE_tr)
```

```
train_vs_test
```

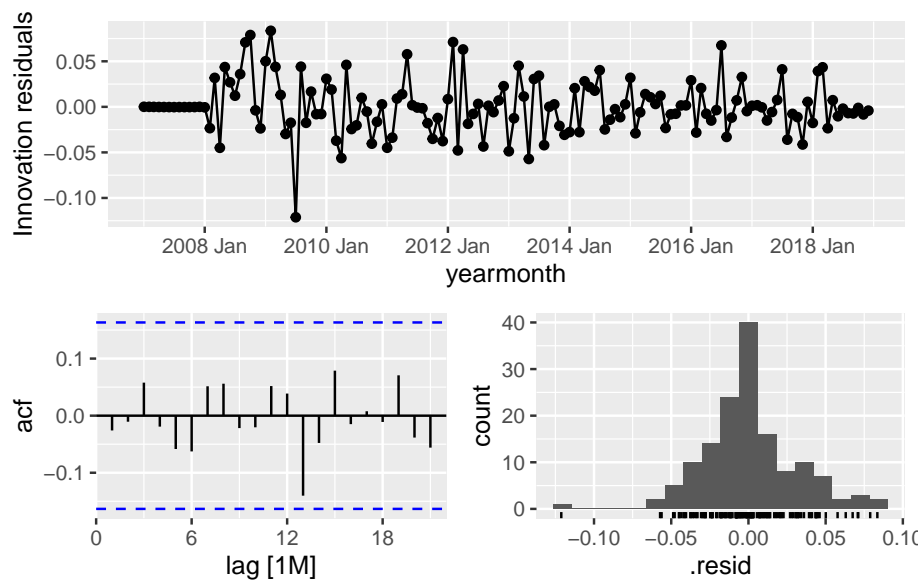
```
# A tibble: 10 x 7
  kon      region      .model RMSE_tr  RMSE MAPE_tr  MAPE
  <fct>   <fct>      <chr>   <dbl>  <dbl> <dbl> <dbl>
1 Kvinder Region Hovedstaden ETS      0.0434 0.0398    1.68  1.82
2 Kvinder Region Midtjylland ARIMA     0.0290 0.0436    1.70  3.43
3 Kvinder Region Nordjylland ARIMA     0.0133 0.0122    1.46  1.72
4 Kvinder Region Sjælland   ARIMA     0.0156 0.0280    1.28  2.88
5 Kvinder Region Syddanmark ARIMA     0.0238 0.0236    1.41  1.85
6 Mænd    Region Hovedstaden ARIMA     0.0383 0.0784    1.26  4.13
7 Mænd    Region Midtjylland ARIMA     0.0503 0.0193    2.41  1.52
8 Mænd    Region Nordjylland ARIMA     0.0296 0.0140    2.49  2.00
9 Mænd    Region Sjælland   ARIMA     0.0289 0.0153    1.94  1.85
10 Mænd   Region Syddanmark ARIMA     0.0464 0.0846    2.12  6.86
```

Sammenligningen af trænings- og testfejl viser, at modellerne generelt præsterer fornuftigt, men med variation. I flere serier er fejlene på testdata lavere end på træningsdata (f.eks. mænd i Region Midtjylland og Nordjylland). Omvendt ser vi enkelte serier med markant højere testfejl (f.eks. mænd i Region Hovedstaden og Syddanmark), hvilket kan indikere overfitting eller uforudsete udsving i 2019.

6.4.1 Analyse af residualer og autokorrelation

Der er herunder udvalgt 3 tidsserier med forskellige situationer ift. forskellen på trænings- og testscore, som der er blevet dykket længere ned i med `gg_tsresiduals()`.

```
model_train |>
  filter(kon == "Mænd", region == "Region Syddanmark") |>
  select(ARIMA) |>
  gg_tsresiduals()
```



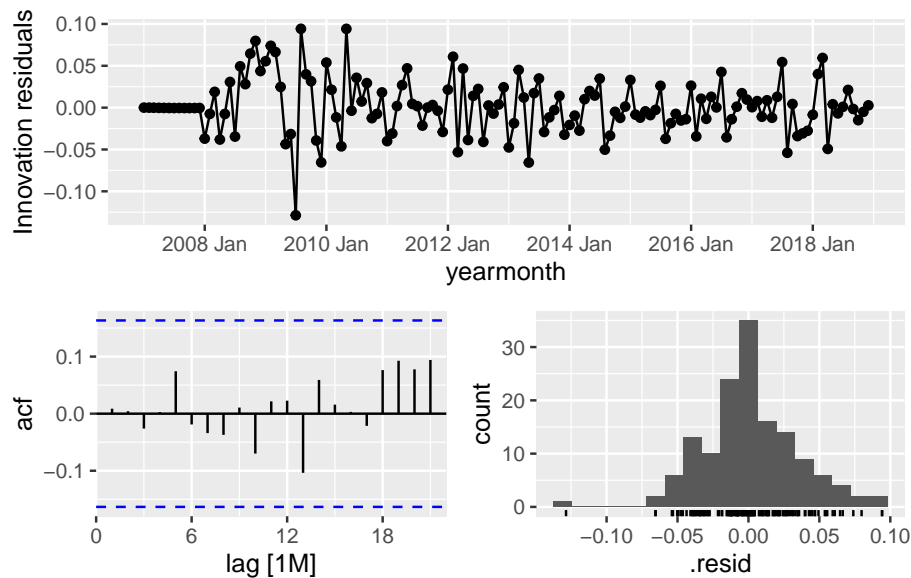
Der er kigget nærmere på tidsserien for mænd i region Syddanmark, da den har relativt høj test-RMSE ift. de andre og høj forskel på trænings- og testscore.

Residualerne ser ud til at være nogenlunde centreret omkring nul, men der er enkelte ekstreme udsving (outliers), særligt i 2009 og 2016.

ACF-plottet viser, at der ikke er signifikant autokorrelation i residualerne — alle ligger inden for konfidensgrænserne. Det er et tegn på, at modellen har fanget den systematiske struktur i data. Histogrammet indikerer en rimelig symmetrisk fordeling, men med lidt tungere haler end en ideel normalfordeling.

Samlet set tyder residualanalysen på, at modellen er acceptabel, men de ekstreme observationer og den relativt høje forecastfejl på testdata (RMSE = 0.085) antyder, at modellen kan være følsom over for enkelte udsving.

```
model_train |>
  filter(kon == "Mænd", region == "Region Midtjylland") |>
  select(ARIMA) |>
  gg_tsresiduals()
```



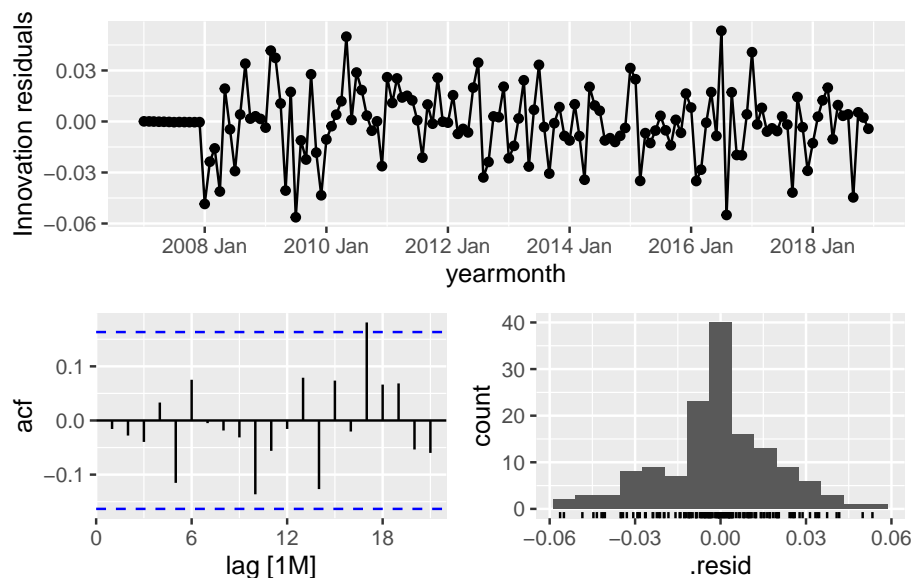
Der er kigget nærmere på tidsserien for mænd i region Midtjylland, da den har lav test-RMSE ift. trænings-RMSE.

Modellen ser ud til at have fanget strukturen i data rimeligt godt. Residualerne er generelt centreret omkring nul, men enkelte større afvigelser ses tidligt i serien.

ACF-plottet viser svag autokorrelation ved nogle højere lags, men ingen værdier ligger tydeligt uden for konfidensgrænserne. Histogrammet viser en nogenlunde symmetrisk fordeling med lette afvigelser fra normalitet.

Samlet set er der ikke tegn på systematiske fejl i residualerne. Det stemmer godt overens med den relativt lave RMSE på testdata ($\text{RMSE} = 0.019$) sammenlignet med en højere træningsfejl ($\text{RMSE} = 0.050$), hvilket tyder på, at modellen generaliserer bedre end forventet.

```
model_train |>
  filter(kon == "Kvinder", region == "Region Nordjylland") |>
  select(ARIMA) |>
  gg_tsresiduals()
```



Til sidst er der kigget nærmere på tidsserien for kvinder i region Nordjylland, da den har lav test-RMSE og lav trænings-RMSE.

Residualerne er pænt centreret omkring nul og uden tydelige mønstre over tid. Der ses enkelte udsving, men ingen systematiske afvigelser.

ACF-plottet viser lidt autokorrelation ved lag 17, men det virker rimelig tilfældigt og ellers ligger værdierne inden for konfidensgrænserne, hvilket tyder på, at modellen har fanget den væsentlige struktur i data. Histogrammet viser en forholdsvis symmetrisk fordeling af residualerne.

Samlet set understøtter residualanalysen, at modellen er velfungerende for denne serie. Det stemmer overens med en lav forecastfejl på testdata (RMSE = 0.012) og indikerer, at modellen generaliserer stabilt.

6.4.2 Ljung-tests for tidsserier

```
ljung_box_results <- model_train |>
  augment() |>
  features(.innov, ljung_box, lag = 24, dof = 3)

ljung_box_vindere <- ljung_box_results |>
  inner_join(vinder_cv, by = c("kon", "region", ".model")) |>
  select(1:5)
```

For hver vinder-model er der udført en Ljung-Box test med 24 lags og 3 frihedsgrader for at vurdere, om residualerne er serielt uafhængige.

Resultaterne viser, at 9 ud af 10 modeller har høje p-værdier, hvilket betyder, at vi ikke forkaster nulhypotesen om uafhængige residualer. Det tyder på, at modellerne har fanget den væsentlige struktur i data.

Dog er der en klar undtagelse: ETS-modellen for kvinder i Region Hovedstaden har en meget lav p-værdi, hvilket tyder på signifikant autokorrelation i residualerne. Det skaber tvivl om modellens gyldighed her og kunne indikere, at ARIMA muligvis havde været et bedre valg for denne serie.

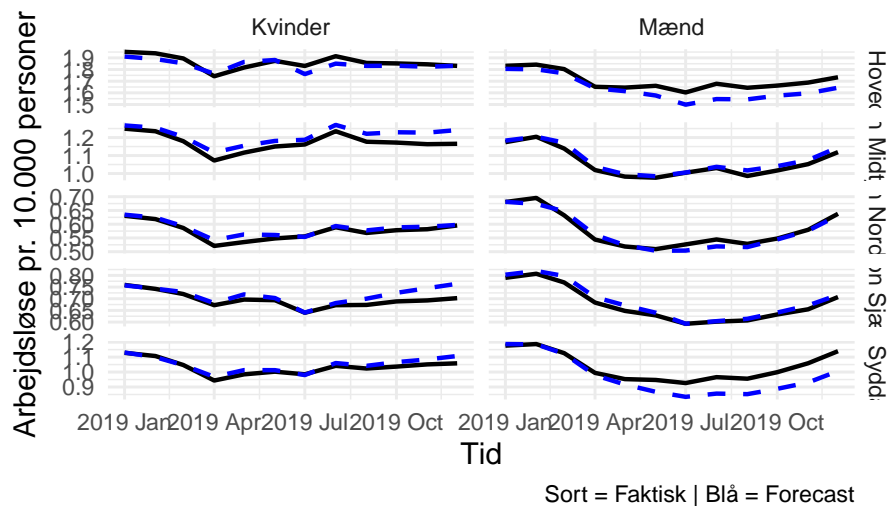
Samlet set tyder resultaterne på, at residualerne fra de fleste modeller kan betragtes som hvid støj, hvilket er et centralt krav i vurderingen af en god tidsseriemodel.

7 Sammenligning af 2019 forudsigelser og virkeligheden

```
forecast_vs_actual <- forecast_2019 |>
  left_join(
    test_data |> select(kon, region, yearmonth, actual = svalue),
    by = c("kon", "region", "yearmonth")
  )

forecast_vs_actual |>
  ggplot(aes(x = yearmonth)) +
  geom_line(aes(y = actual), color = "black", size = 0.9, linetype = "solid") +
  geom_line(aes(y = .mean), color = "blue", size = 0.9, linetype = "dashed") +
  facet_grid(region ~ kon, scales = "free_y") +
  labs(
    title = "Forecast vs. Faktiske værdier - 2019",
    y = "Arbejdsløse pr. 10.000 personer",
    x = "Tid",
    caption = "Sort = Faktisk | Blå = Forecast"
  ) +
  theme_minimal(base_size = 13)
```

Forecast vs. Faktiske værdier – 2019



Plottet viser en sammenligning mellem de faktiske observationer (sort) og modellernes punktprognoiser (blå) for 2019, opdelt på køn og region. Generelt følger prognoserne udviklingen i de faktiske data tæt, hvilket bekræfter, at modellerne har god forudsigelsesevne.

Særligt for mænd i Region Midtjylland og Nordjylland er der meget tæt overensstemmelse. Derimod ses systematiske afvigelser for mænd i Region Syddanmark og kvinder i Region Hovedstaden, hvor modellerne undervurderer eller overvurderer udviklingen. Det stemmer overens med de højere RMSE-værdier og residualanalyser for disse serier.

8 Forecasting af 2020 for alle tidsserier

```
data_train_final <- data

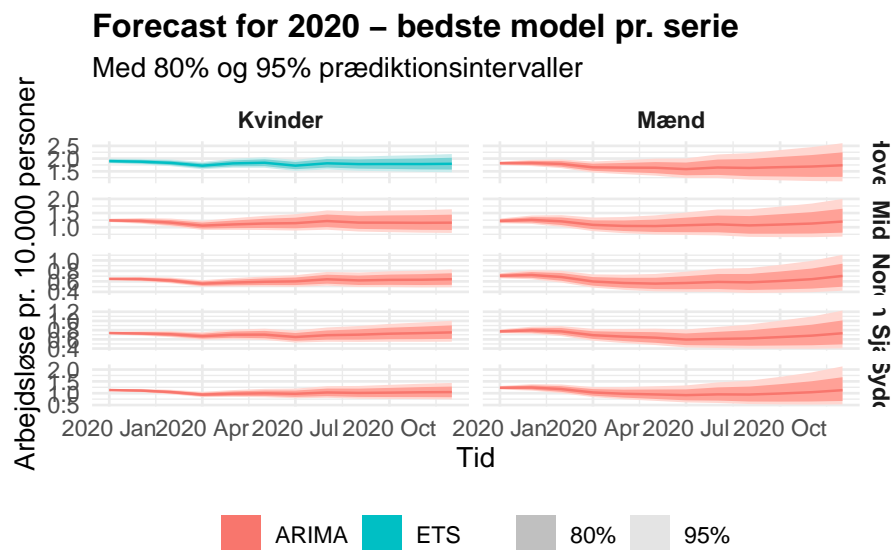
models_final <- data_train_final |>
  model(
    ARIMA = ARIMA(log(svalue)),
    ETS = ETS(log(svalue)),
    SNaive = SNAIVE(log(svalue))
  )

forecast_2020 <- models_final |>
  forecast(h = "12 months") |>
  inner_join(vinder_cv, by = c("kon", "region", ".model"))
```

```

forecast_2020 |>
  autoplot(level = c(80, 95)) +
  facet_grid(region ~ kon, scales = "free_y") +
  labs(
    title = "Forecast for 2020 - bedste model pr. serie",
    subtitle = "Med 80% og 95% prædiktionsintervaller",
    y = "Arbejdsløse pr. 10.000 personer",
    x = "Tid"
  ) +
  theme_minimal(base_size = 12) +
  theme(
    legend.position = "bottom",
    legend.title = element_blank(),
    plot.title = element_text(face = "bold"),
    strip.text = element_text(face = "bold")
  )

```



Plottet viser 12 måneders fremskrivning for 2020 baseret på den bedste model pr. serie, suppleret med 80% og 95% prædiktionsintervaller. For de fleste serier anvendes ARIMA (rød), mens ETS (turkis) kun bruges for kvinder i Region Hovedstaden.

De fleste forecasts viser relativt stabile niveauer i arbejdsløsheden, men for flere mandlige serier (især Region Syddanmark og Midtjylland) ses en tydelig stigning og stigende usikkerhed hen over året – hvilket afspejles i de bredere konfidensbånd.

Visualiseringen illustrerer både det forventede niveau og usikkerheden i prognoserne, og den bekræfter, at modellerne er i stand til at fange forskelle mellem regioner og køn.

9 Konklusion

10 Kildeliste

Hyndman, R. J., & Athanasopoulos, G. (2021). Forecasting: principles and practice (3rd ed.). OTexts. <https://otexts.com/fpp3/>

OpenAI. (2025). ChatGPT (v.4o) [Large language model]. <https://chat.openai.com/>

11 Bilagsoversigt

- Bilag 1: Features