sesongjusteringskurs – påbygging 2024

FOR LANG TIDSSERIER?,
OM COVID-HÅNDTERING

REVISJONS-/«REFRESH-POLICY»,

OG EVT. LITT OM JWSACRUNCHER

DIREKTE VS. INDIREKTE SESONGJUSTERING

AV

JØRN IVAR HARME (S811)



For lange tidsserier?

COVID-håndtering

Revisjons-/«refresh-policy», og evt. om JWSACruncher

Direkte vs. indirekte sesongjustering





ESS guidelines on seasonal adjustment



- ... for å **fremme beste praksis**:
- oppnå harmonisering på tvers av nasjonale prosesser
- forbedre sammenlignbarheten mellom resultatene
- Øke robustheten til europeiske aggregater
 - A) Beste praksis (BP)
 - B) Akseptabelt
 - C) Bør unngås

ESS guideline, kap. 8.4, (for) lang tidsserier?

- Suboptimale resultater mot slutten av tidsserien...
- Gjennomsnittlige kalendereffekter mindre representative nå
- Over 20 år med data normalt for langt...
- Ingen generell fasit
- Glidende vindu på rundt 15 år, helst fra *januar.(YYYY-15)*
- Evt. pre-påske effekter krever lange TS for å få variasjon i antall dager mars/april i ulike påske-variable

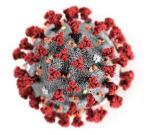
BP i ESS Guideline ang. outliere på slutten (kap.8.2)

 Outliers på slutten modelleres basert på statistiske kriterier og økonomisk informasjon

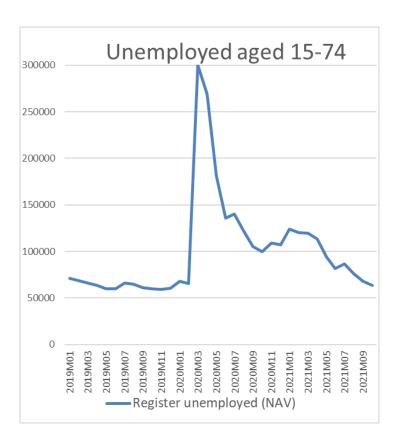
(spesielt ved sterke økonomiske endringer eller kriser)



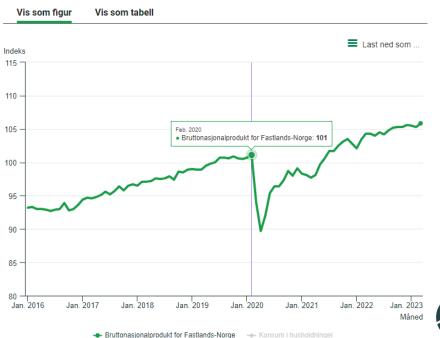
Om COVID-håndtering ved sesongjustering



- Er observasjonene under krisen representative...
- Bør obs. inngå i beregning av sesongkomponentene?



Figur 2. Bruttonasjonalprodukt og konsum i husholdninger. Månedlig. Sesongjustert. Volumindekser. 2019=100



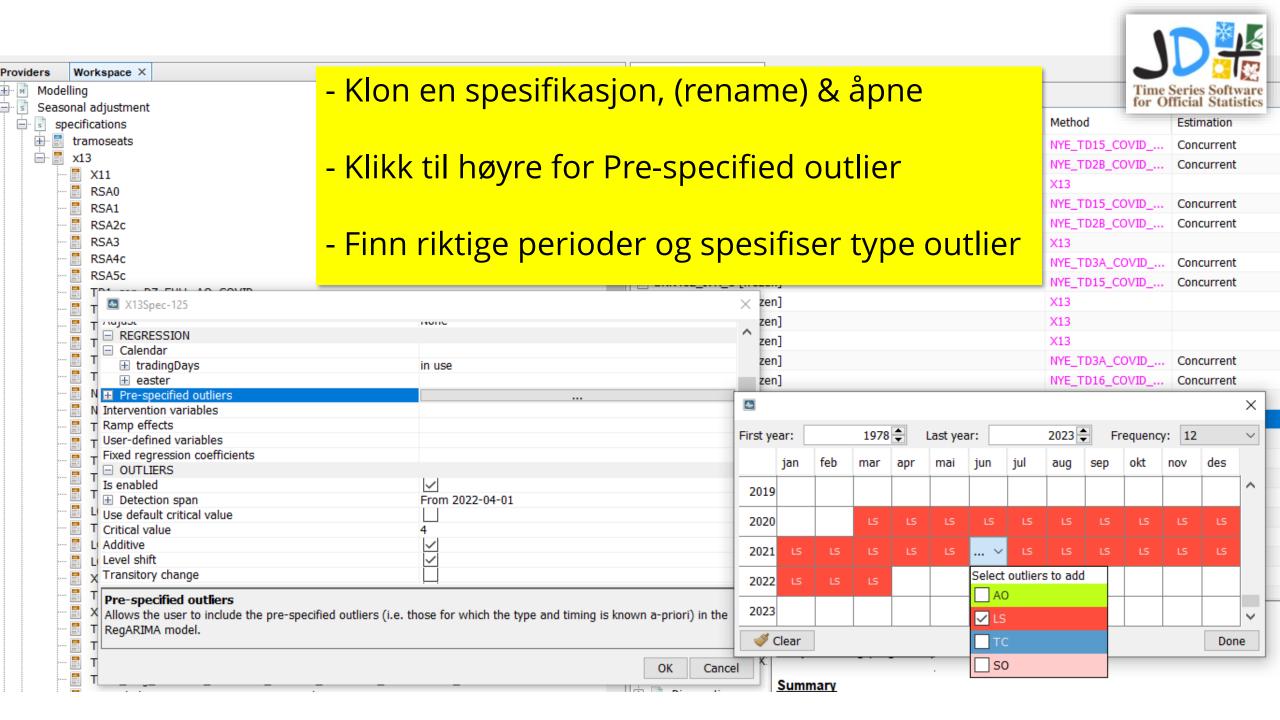


Om COVID-håndtering ved sesongjusterii

- Nasjonale tiltak (i større eller mindre grad) 2020M3 2022M2
 - Omicron-varianten => 2. periode med strenge COVID-19-tiltak: 2021M12-2022M2
- S811-anbefaling under og etter pandemien
 - Eksternt: https://github.com/statisticsnorway/ssb-seasonaladjustment-corona
 - Internt: https://statistics-norway.atlassian.net/wiki/spaces/s880/pages/3907584702/Sesongjustering+under+Koronakrisen

Sekvens av nivåskift 2020M3 - 2022M3 (evt. 2020Q1-2022Q1)

- Alt. <u>AO-sekvens</u> 2020.3-2022.2 & <u>LS2022.3</u> (om trend-tall *ikke* publiseres
 => trenden upåvirket til 2022M3...?
- Antok da at sesongmønsteret var som før (vanskelig å spå...)
- => Number of outliers (severe) i JD+ kan ignoreres



Om Korona-anbefalingen

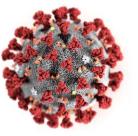
- Enkel, grei, krever ikke gjennomgang av alle seriene
- Vet at noe har skjedd: T-verdiene for LSS ikke avgjørende
- Potensielt to typer feil,
 - Spesifiserte ikke-ekstremverdier (taper primært frihetsgrader)
 - Uspesifiserte ekstremverdier (jamfør evt. kap.7 i Eurostat's
 - (egentlig en 3. type feil: feil type spesifisering)



LSS => trend ≈ sesongjusterte tall

(lite ny info med volatil trend, evt. 3MMA eller 5MMA i etterkant, evt. droppe publ.)

Evt. alternativ COVID-håndtering...

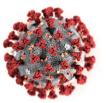


- Primært for korte serier (ikke råd til å «miste» så mange observasjoner...)
- Tidkrevende, trenger emne-kunnskap
- For 2020M3 2022M2, delperiode som er normal (UPÅVIRKET) for din serie?
 - Evt. opphold uten korona-nivåskift (separate sammenhengende COVID-per.)
 - Utfordrende og tidkrevende

• Automatisk outlier-søk i Korona-perioden med lavere kritisk verdi...



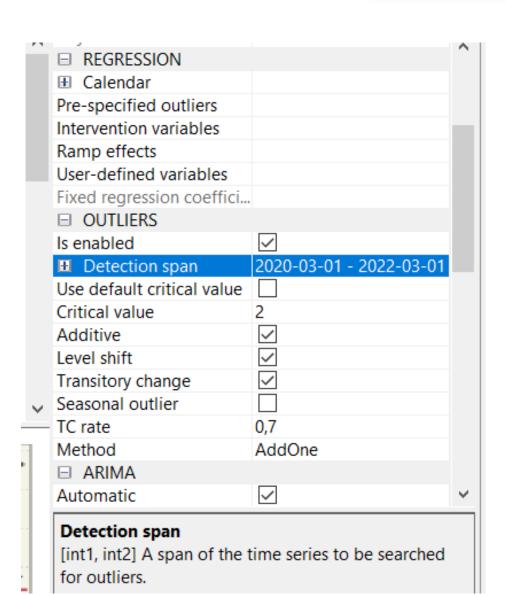
Evt. alternativ COVID-håndtering...





- Primært for kort serie (ikke råd til å «miste» så mange observasjoner...)
- Automatisk outlier søk
 - ∘ i Korona-perioden (2020M3-2022M3)
 - med lav kritisk verdi
- Ulik prekorrigering for seriene...
- Kan forandre seg over tid...
- Resultatet bør «hardkodes» («pre-specified» outlier)

Tidkrevende likevel...



Endret sesongmønster etter koronaperioden?

- Kan evt. forklare mer revisjon i en overgangsperiode
- Tar tid før mnd. med gammelt mønster ikke inngår i beregning av sesongfaktorene)
- X11/X12-metodene *håndterer gradvis endret sesongmønsteret*

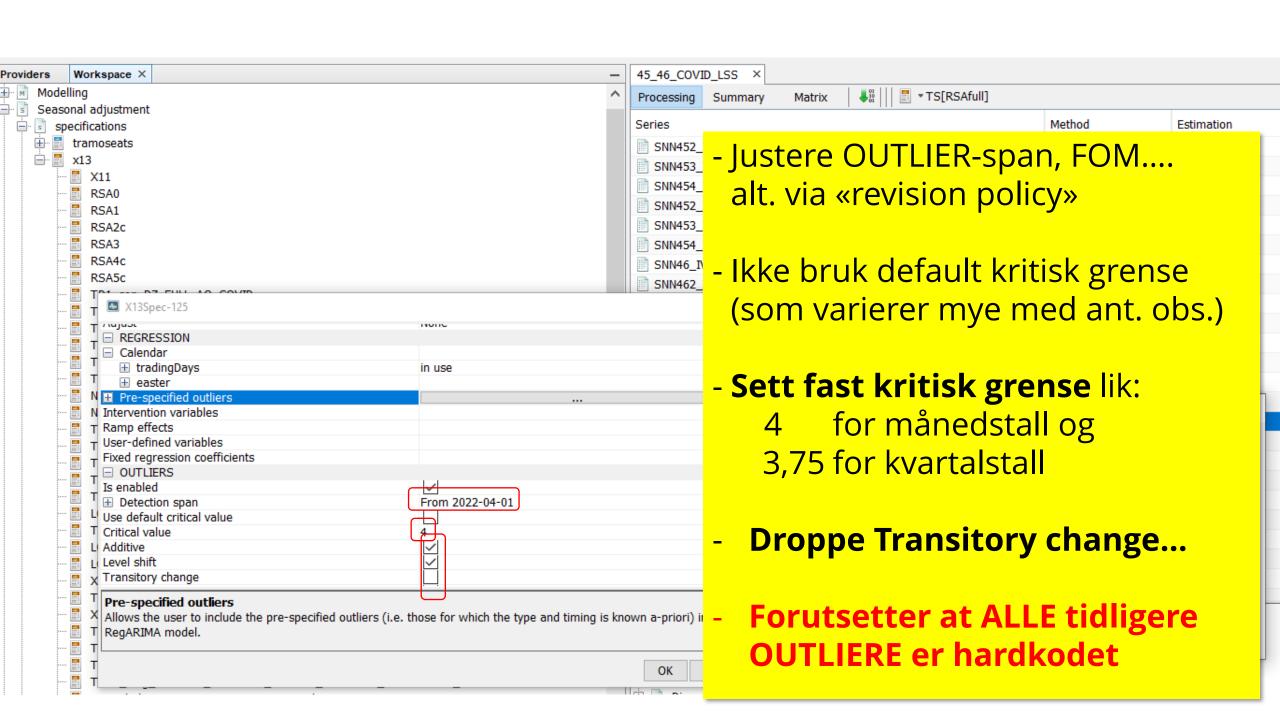
Når det har gått noen år...

- Er sesongmønsteret **etter** «Korona-perioden» mer likt situasjonen **under** Korona-perioden **enn før**?
 - Hvis ja (varig endring), revurdere tidligere Korona-håndtering
 (få ned revisjonen raskt, men endret prekorrigering reviderer tidl. tall)
 - Hvis ja (varig endring fom. dato), evt. vurdere å modellere SO
 - **Hvis nei**, enklest å beholde tidligere metode (i hvert fall med lange T.S.)

S811-anbefaling også om outlier-håndtering generelt

Link: https://statistics-norway.atlassian.net/wiki/spaces/s880/pages/3907584702/Sesongjustering+under+Koronakrisen?preview=/3907584702/3907595080/Notat_om_Avvikling_av_korona_14mars2022.docx

- Beste praksis: fryse gamle outliere
- Kun søke etter nye på slutten av serien, det siste året,
 - ∘ OUTLIER-span
 - Alt. sette «revision policy» i JD+ (tema senere)
- Fast (høy) kritisk deteksjonsgrense
 - Unngår at outliere kommer med også kastes ut senere med tilhørende revisjon
- Jdemetra+ med defalult automatisk outlier-søk også etter TC...
- X-13ARIMA-SEATS kun defalult automatisk <u>outlier-søk etter AO og LS</u>
 - · Kun est. nivåskift (LS) legges tilbake i trend-komp., mao. «trend-skift», AO & TC tilordnes irregulær



Litt om ARIMA-modell seleksjon, REVISJON, ESS GUIDELINE, REVISJONSPOLICY & MULIGHERTER I JD+



Innledning: regARIMA-modellering sentral

- Framskriving for symmetriske filter/gjennomsnitt i X-12ARIMA
- Estimere Kalendereffekt- og outlier-modellering
- Automatisk valg av ARIMA-struktur (hvordan historien brukes, antall lag):
 - pickmdl{}: Første OK ARIMA-modell blant 5 pre-spesifiserte
 (i X-12ARIMA, mest brukt i SSB tidligere, OK for de fleste T.S., IKKE i JD+)
 Flere kombinasjonsmuligheter i r-pakken pickmdl (mer om dette onsdag)
 - automdl{} (TRAMO): «Beste» blant mange nyere, mer fleksibel: implementert i JD+!
- ESS Guideline anbefaler **automdl{}** (TRAMO), (ikke «ren» pickmdl{})

Fra ONS Guide to Seasonal Adjustment with X-12-ARIMA

Because of the effect of sampling variation or irregularity in the observed series, it is often possible to find different ARIMA models that will fit a given time series. One fundamental idea in choosing a model is to take the simplest model which will give a satisfactory fit, where simplest means having the smallest number of parameters. This principle is known as "parsimonious parameterisation", or parsimony for short. The process of searching for a model often involves adding parameters to avoid specification errors and then seeing which previously chosen parameters may be removed in the interests of parsimony. The

KILDE: https://ec.europa.eu/eurostat/cache/metadata/Annexes/lci_esqrs_uk_an_4.pdf



Selseksjonsprosedyren automdl{} (Beste blant mange)

Passende ARIMA-modeller også for tidsserier uten sesong

Endrer modellene oftere

• Seleksjon av ARIMA-modell **påvirket av outliere**...

Kan fort «ombestemme» seg... med nye data

Løpende identifisering av «beste» ARIMA-struktur...

- ... kombinert med automatisk outlier-søk kan gi mye endringer...
- Re-tolker historien: «føyer dataene best»
- Ny ARIMA-modell (marginalt bedre «in-sample» forecast)
 - •Nye outliere bak i tid...?
 - Nye estimater på kalender effekter



Revisjon (tidligere sj-tall endres)

- Når nye observasjoner og evt. reviderte rådata/vekter
- Revisjoner forvirrende for brukerne

Balansere presisjon og stabilitet...

(for mindre revisjon gjennom året)

• Bør ARIMA-modell re-identifiseres hver gang eller låses...?



ESS guidelines on seasonal adjustment



- ... for å **fremme beste praksis**:
- oppnå harmonisering på tvers av nasjonale prosesser
- forbedre sammenlignbarheten mellom resultatene
- Øke robustheten til europeiske aggregater
 - A) Beste praksis (BP)
 - B) Akseptabelt
 - C) Bør unngås

ESS guideline om løpende eller faste valg (kap.6.2)

C. Løpende korrigering:

- Re-identifisering av ARIMA-struktur **hver gang**
- Estimering av alle parametere hver periode
- Max. presisjon (hver gang)
- Potensielt mer revisjon gjennom året
- Enkelt, tidligere mest normalt i SSB

«Rangering»: A = Beste praksis, B= akseptabelt/OK og **C = bør unngås**

ESS guideline om løpende eller faste valg (kap.6.2)

B. Fast korrigering:

Identifisering årlig

estimering kun årlig

<u>Fremskrevne sesong- og kalenderfaktorer anvendes gjennom året</u>

- Minimerer revisjon gjennom året
- Potensielt mye revisjon ved årsskifte (årlig gjennomgang)
- Ikke vanlig i SSB, tidligere vanskelig

Rangering: A = Beste praksis, **B= akseptabelt/OK** og C = bør unngås

ESS guideline om løpende eller faste valg (kap.6.2)

A. Delvis løpende korrigering (DLK):

- modeller identifiseres kun årlig (årlig/periodeisk gjennomgang)
- men<u>re-estimeres</u> hver periode
- Ideelt sett bør også <u>filter-lengdene</u> låses: (evt. mye jobb)
- Balansere presisjon og stabilitet
- Beste praksis!

Rangering: A = Beste praksis, B= akseptabelt/OK og C = bør unngås



Delvis løpende korrigering i JD+

Etter årlig gjennomgang:

- sørg for at modellene ikke re-identifiseres gjennom året!

(særlig viktig med automdl{}/TRAMO

med outlier-søk hele perioden)



Sette Refresh-polisy i JD+

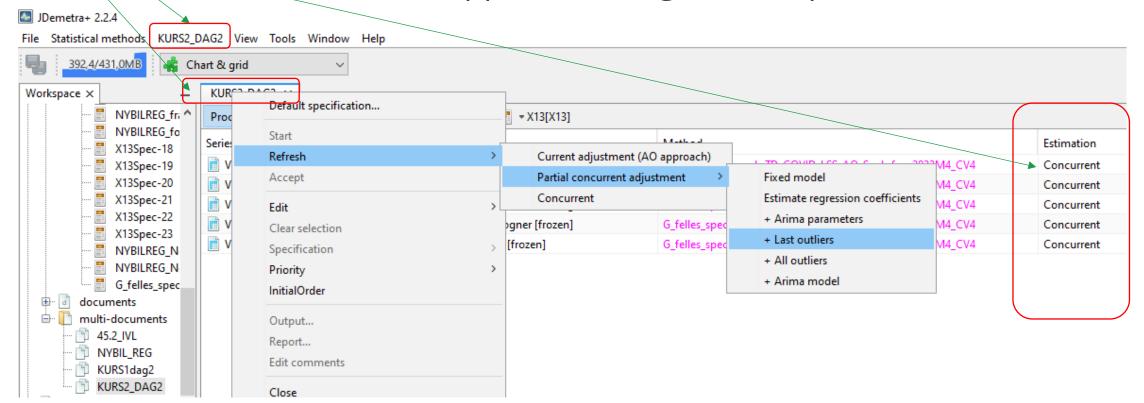


- Velg multi-dokumentet fra Workspace-vinduet og dobbeltklikk på det for å vise multidokument-menyen.
- Høyreklikk deretter, velg Refresh o Partial concurrent adjustment o og ett av alternativene...

S811 anbefaler: Partial concurrent adjustment + last outlier

• Sjekk at «Estimation» metoden oppdateres Lagre Workspace!

LAGRE!



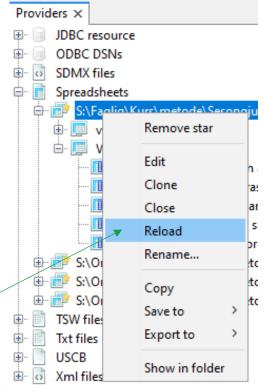


Delvis løpende korrigering i JD+

- Etter årlig gjennomgang, der refresh-policy er satt
- Åpne tidligere lagret worspace med «star-added» data)

Nye data tilgjengelig (i samme datafil): «Reload» data i Providers

• Kjør sesongjusteringen med tidligere identifiserte modeller:



Opsjoner i JD+ for delvis løpende korrigering



Kilde: Tabell fra
 JD+ User guide Version 2.2

 (User guide inkluderer et detaljert eksempel med side ved side sammenligning av initial SJ med de forskjellige refresh- alternativer)

3	Option	Meaning
varesties	Partial concurrent adjustment \rightarrow Fixed model	The ARIMA model, outliers and other regression
		parameters are not re-identified and the values of
		all parameters are fixed. The transformation type
		remains unchanged.
	Partial concurrent adjustment → Estimate regression coefficients	The ARIMA model, outliers and other regression
		parameters are not re-identified. The coefficients
		of the ARIMA model are fixed, other coefficients
		are re-estimated. The transformation type remains
		unchanged.
	Partial concurrent adjustment → Estimate regres- sion coefficients + Arima parameters	The ARIMA model, outliers and other regression
		parameters are not re-identified. All parameters of
		the RegARIMA model are re-estimated. The trans-
		formation type remains unchanged.
	Partial concurrent adjustment → Estimate regression coefficients + Last outliers	The ARIMA model, outliers (except from the out-
		liers in the last year of the sample) and other re-
		gression parameters are not re-identified. All pa-
		rameters of the RegARIMA model are re-esti-
		mated. The outliers in the last year of the sample
		are re-identified. The transformation type remains
		unchanged.
	Partial concurrent adjustment → Estimate regression coefficients + all outliers	The ARIMA model and regression parameters,
		except from outliers) are not re-identified. All pa-
		rameters of the RegARIMA model are re-esti-
		mated. All outliers are re-identified. The transfor-
		mation type remains unchanged.
	Partial concurrent adjustment → Estimate regres- sion coefficients + Arima model	Re-identification of the ARIMA model, outliers
		and regression variables, except from the calendar
		variables. The transformation type remains un-
		changed.



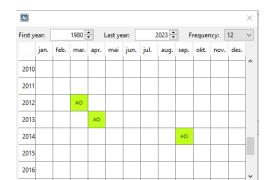
Delvis løpende korrigering + Last outlier i JD+

- Hva skjer egentlig; avhenger av outlier-opsjonene...
- Uten outlier-span:
 - Gamle automatisk identifiserte outliere (ikke siste år)

flyttes automatisk over til prespesified-outlier

PS. Sett først revision policy på slutten når du er ferdig med gjennomgangen...!

 Oppgave: for næring 47 eller 47.1, sjekk lista over Prespesified-outlier før/etter delvis løpende korrigering med Last outlier



JWSACruncher

EN JD+ PLUG-IN



JWSACruncher: en JD+ plug-in

- Kommandolinjeverktøy for batch-kjøring
- For masseproduksjon uten «klikking»
- Systematisk lagring av output-data i CSV-filer:
 - sesongjusterte data, trend, kalenderjusterte tall
 - Kvalitetsindikatorer

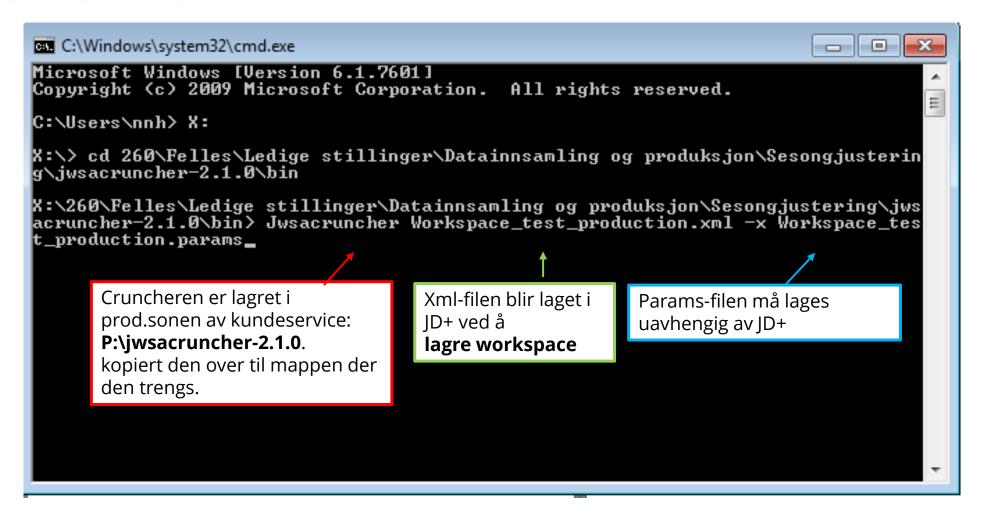
JWSACruncher: en JD+ plug-in

Om jwsacruncher:

https://github.com/jdemetra/jwsacruncher/wiki

- Øyvind Bruer-Skarsbø: Flytte sesongjustering med Jdemetra+ over til Dapla: Link: https://git-adm.ssb.no/pages/~OBR/oyvinds_blog/master/browse/posts/2022-01-28-sesongjusteringdapla/
- Flere linker fra sesongjusteringsforum hhv. 2022 & 2023 v/ Rakel Gading:
 - Sesongjusteringsforum-20220518_Hvordan_bygge_opp_et_produksjonsløp_for_sesongjustering_på_Dapla.mp4
 - Om sesongjustering på Dapla og hvordan ledige stillinger har gjort det i praksis Video-link: <u>SESONGJUSTERINGSFORUM-20230510_121243-Opptak av møte.**mp4**</u>

JWSA Cruncher





jwsa<u>cruncher</u> på:

- alle SAS compute-servere
- Dapla
- DaplaLab





Bruer-Skarsbø, Øyvind

Sep 1

jwsacruncher på Dapla Lab

jwsacruncher er installert i følgende tjenester på Dapla Lab:

- Jupyter
- Jupyter-playground
- Rstudio

Vi har foreløpig valgt å ikke installere jwsacruncher i Vscode-python for å holde antall avhengigheter i denne tjenesten til et minimum. Hvis noen har behov for dette så kan de bare melde fra til oss i denne kanalen og så vurderer vi det på nytt.

På Dapla Lab kan man jobbe med data i bøtter som et vanlig filsystem. En praktisk konsekvens av dette er at det blir mye lettere å kjøre sesongjustering med jwsacruncher på Dapla. Vi har lagt ut et eksempel-workspace i produktbøtta til Dapla Felles som alle har tilgang til og kan kjøre fra Dapla Lab tjenestene nevnt over. For å prøve det kan du gjøre følgende

- 1. Åpne en av tjenestene der jwsacruncher er installert i Dapla Lab
- 2. Under tjenestekonfigurasjonen velger du dapla-felles-developers og at bøtter skal monteres
- 3. Inne tjenesten åpner du en terminal
- 4. Kjør sesongjusteringen med denne kommandoen:



Direkte vs. indirekte sesongjustering



Direkte vs. indirekte sesongjustering (SJ)

- En (del)total kan SJ direkte eller indirekte
- Indirekte SJ: sesongjustere delseriene, så summe SJ-tall (evt. vekte, evt. diff.)
- <u>Direkte SJ:</u> summer delseriene, så sesongjuster (del)totalen
 - Anbefales i utgangspunktet i **ESS guideline** (kap .5.4) for enkelhet særlig om likt mønser
- Forskjellig resultat / uoverensstemmelser
- Fare for residual-sesong ved indirekte SJ SJEKK!



40

Direct vs. indirect SA...

Litt veiledning:

Indirekte SJ hvis:

- Ulikt sesongmønster for de ulike underseriene
- Bruker-krav: konsistente og sammenhengende TS, særlig additivt

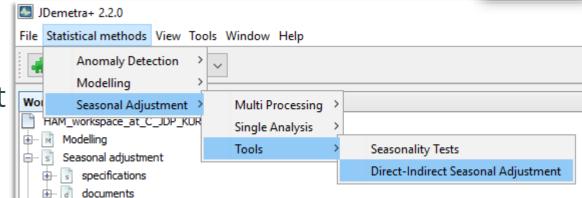
Sammenlign kvaliteten; residual-sesong?

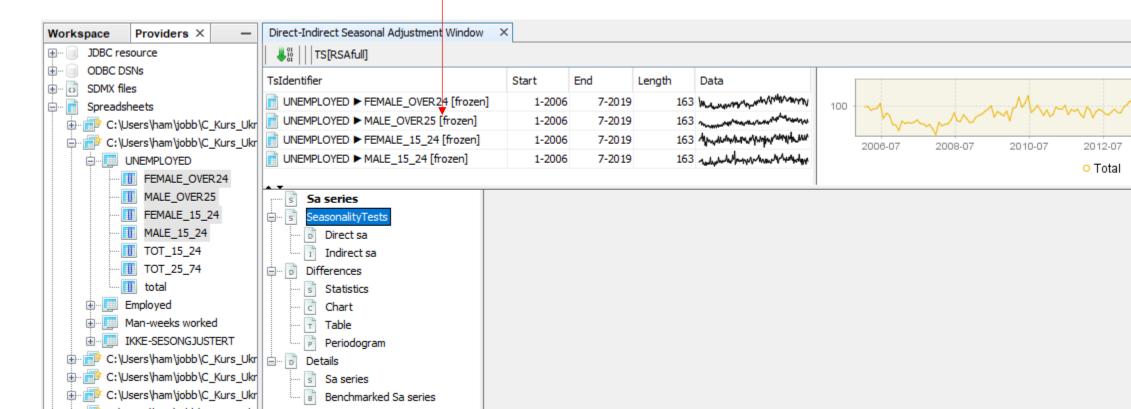
- Kanskje blandet indirekte tilnærming for KNR el VHI (Men ingen klare retningslinjer for valg)
- Detaljerte delserier fra utvalgsundersøkelser
 → høy standardfeil → kanskje problemer for indirekte SA...

Direct vs. indirect SA in JD+



- JD+ tilbyr funksjonalitet som letter sammenligningen
- Dra og slipp del-serier til panelet øverst til venstre (fra Providers)
- Panelet til høyre viser summen av utvalgte serier.



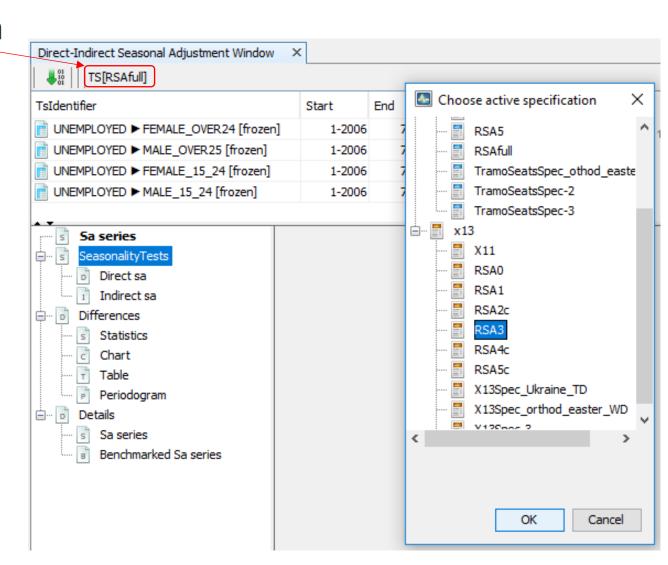


Direct vs. indirect SA in JD+



By default, the pre-defined TRAMO/SEATS specification is used [RSAfull]

To change it, click on the button

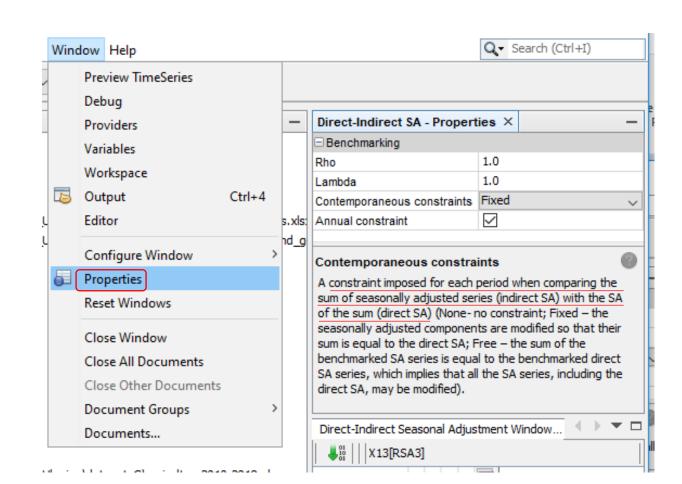


Direct vs. indirect SA in JD+



- Benchmarking mulig, men...
- Ulemper: skjevheter i SJ-tall,
- Ikke optimalt

- Diskutert i ESS Guideline kap. 3.3
- I hovedmenyen, klikk på:
 Window → Properties,
 en kan spesifisere benchmarking
 opsjoner for
 direct-indirect sammenligning



Direct vs. indirect SA: test-rsultater



- Next, run the process by clicking the button with a green arrow
- The bottom panel presents the detailed results.
- **Seasonality test** node presents the outcome Direct SA and Indirect SA
- Check for presence of residual seasonality and calendar effects especially in indirect SA

