

Exempel på nya lösningar från ABS

Officiell statistik i en digital värld – att gå från data till statistik

- SCB och Örebro universitets sommarskola i statistik 2021



Dr Anders Holmberg
Chief Methodologist-General Manager
Australian Bureau of Statistics

Tack till kollegor på ABS för original till några illustrationer

Australian Bureau of Statistics
Informing Australia's important decisions



Innehåll

- ▶ Nätverksanalys av handelsstatistik med GLIDE
 - Andra exempel, sårbarhet i leveranskedjor, företagsregisterkvalitet, transportanalys, arbetsmarknadsanalyser
- ▶ Visualisering
- ▶ Indikatorer och snabbare statistik

Vi utvecklar en miljö för att kunna tackla en komplex verklighet med data.



• **GLIDE** - (**G**raphically **L**inked **I**nformation **D**iscovery **E**nvironment)
En prototypplattform för att länka, utforska och analysera komplexa data från olika källor.



- Utvecklat av metodavdelningen; Ric Clarke m.fl.
 - Intresset från ämnesavdelningar och IT ökar gradvis
 - Externt intresse finns
-
- Prototypen levererar redan, målet är att bygga ut för att kunna göra analyser regelbundet.

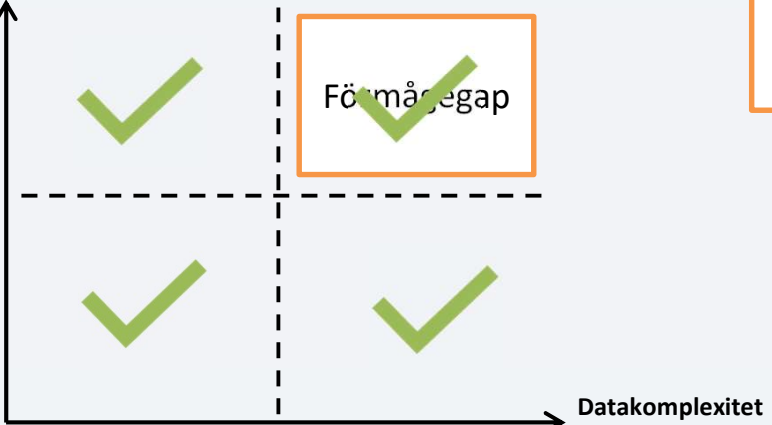
Fyller ett gap och är nyttskapande

Bättre att hantera okonventionella data källor (e.g. transaktioner, sensordata, etc.)

Komplex...

- Flera **sammanlänkade problem** att utforska ('wicked' problems)
- Vi **vet inte** alla **frågorna** på förhand

Komplexitet i analys och policy sfären



Enklare och snabbare att jobba multi-source, diversifierade data

Komplex...

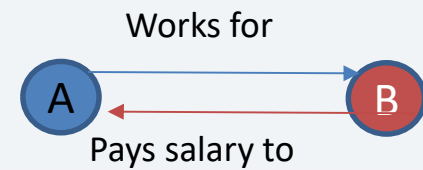
- **Stora** och ev. **semi-strukturerade** eller **ostrukturerade data**
- **Multipla enhetstyper** (t.ex. personer oc företag) med viktiga **relationer och samband dem emellan**
- **Multiple data sources** (t.ex. Transaktion händelser, surveyer, admin, sensorer)

How is GLIDE different to existing solutions?

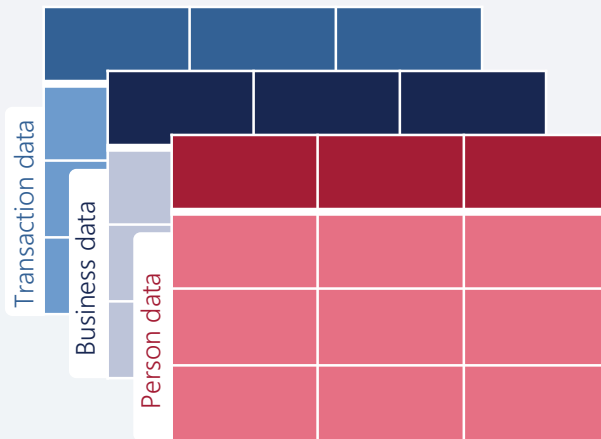
The use of graph database technology is the key differentiator of this new solution

Graph databases are useful for applications where we are:

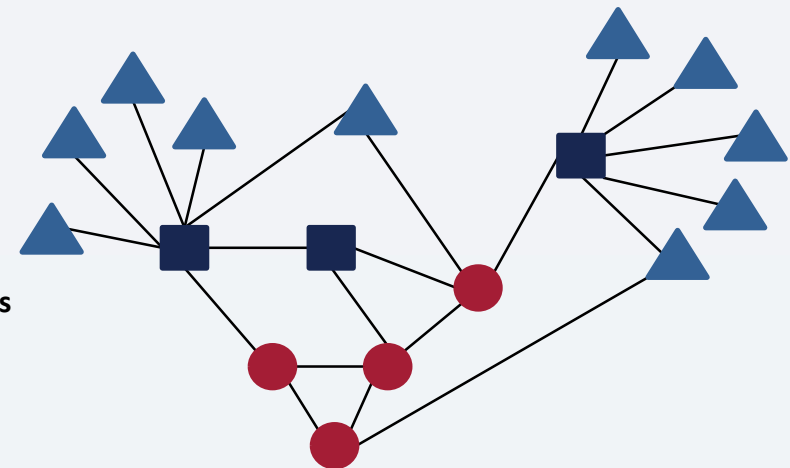
- working with multiple entities and 'many-to-many' relationships
- interested in relationships between data elements as well as the data elements themselves
- working with large and big data sets

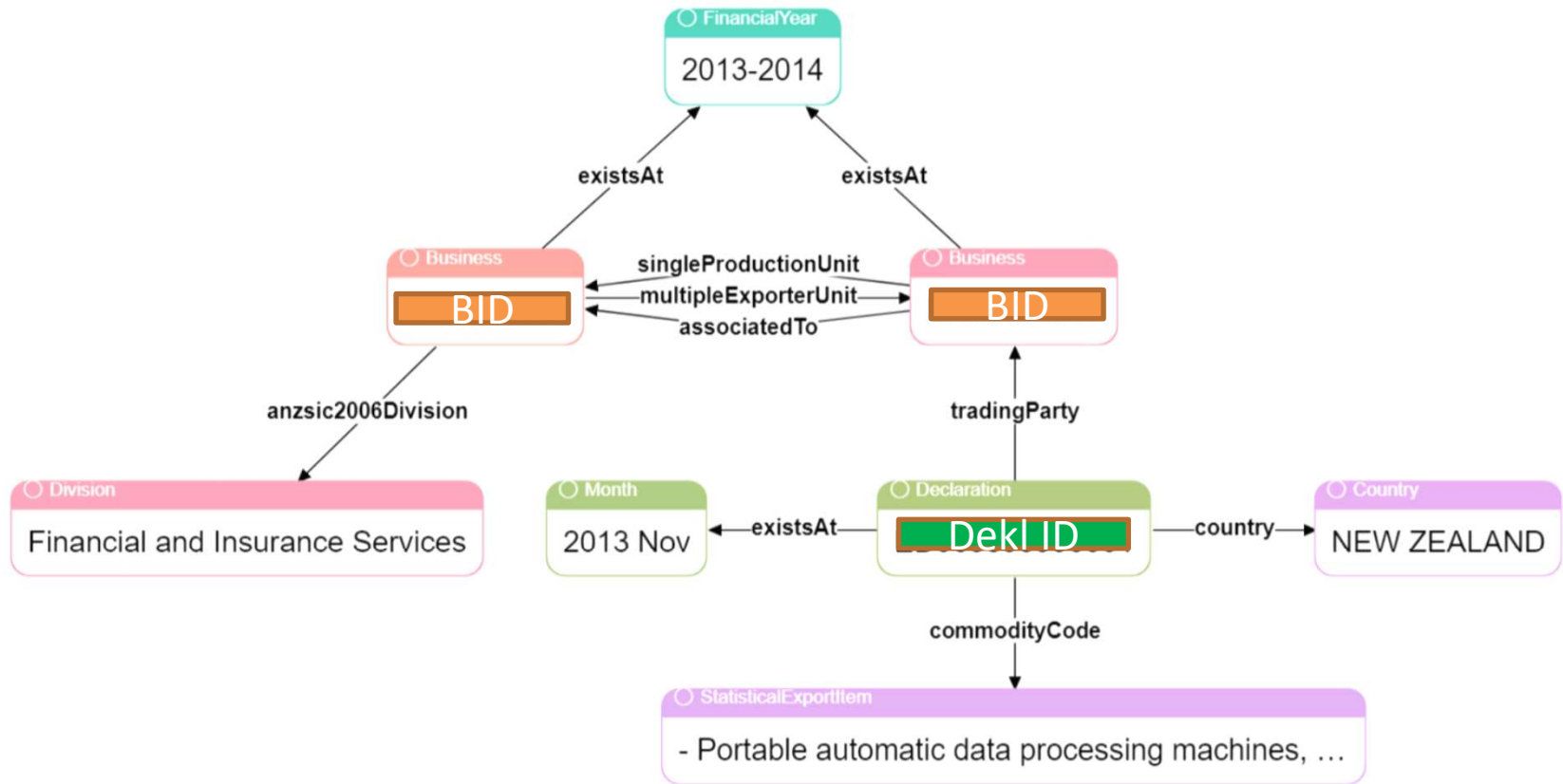


Relational databases



Graph databases





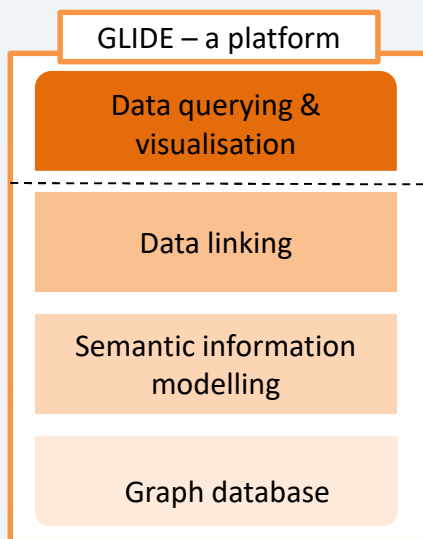
What is GLIDE?

...what the user sees and interacts with...

...under the hood...



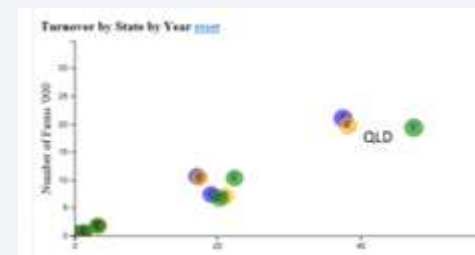
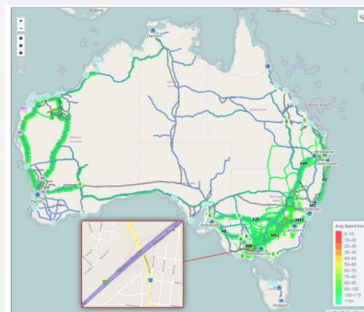
This is where the magic happens!



'Point and click' and a coding interface



- Summary statistics
- Modelling outputs
- Network analysis
- Interactive tables, graphs, maps and other data visualisations



Visualisation

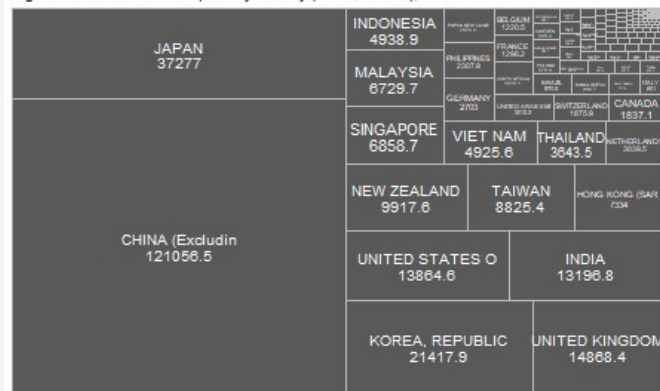
Current use case – Trade data clearance

- Collaboration with ABS International trade team
- Analysis of Imports and Exports for data clearance and briefing minutes
- Connecting to R for Visualisations for briefing minutes and releases

What are the top 10 import product types, and does this change over time?

How are exports spread across countries?

Figure X. Total value of exports by country (FOB \$m AUD), 2019



Source: Real GLIDE trade data, excludes passive claimants

• 2.9 Billion Triples

Merchandise exports data
Jul 2013 – Mar 2020

Merchandise Imports data
Jul 2014 – Mar 2020

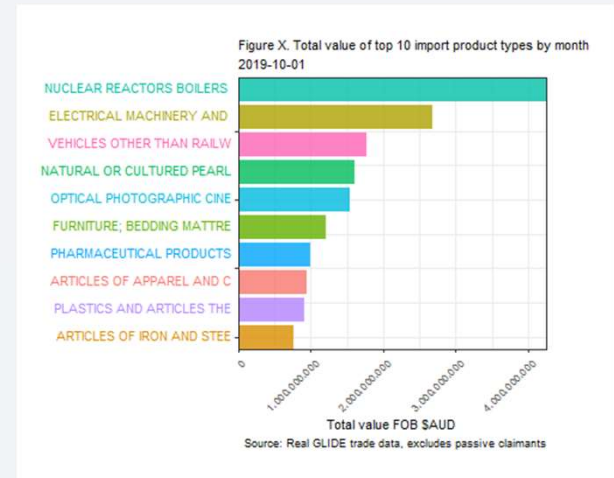


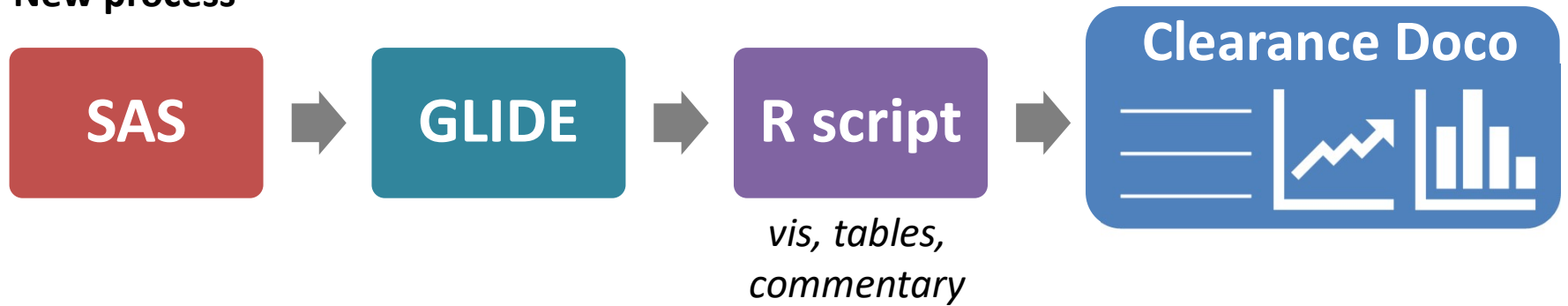
Figure X. Total value of top 10 import product types by month 2019-10-01

Source: Real GLIDE trade data, excludes passive claimants

Old process



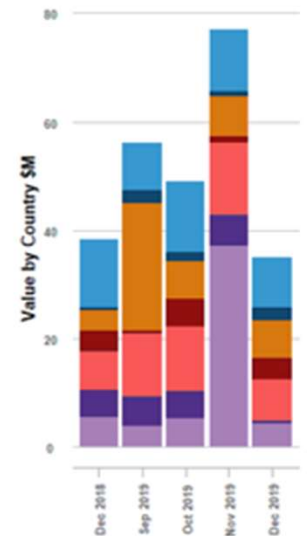
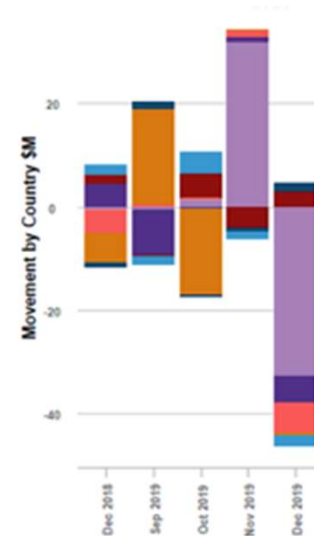
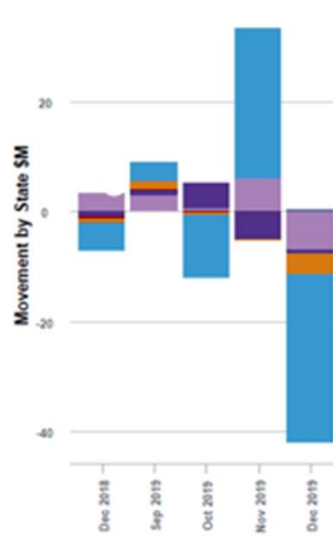
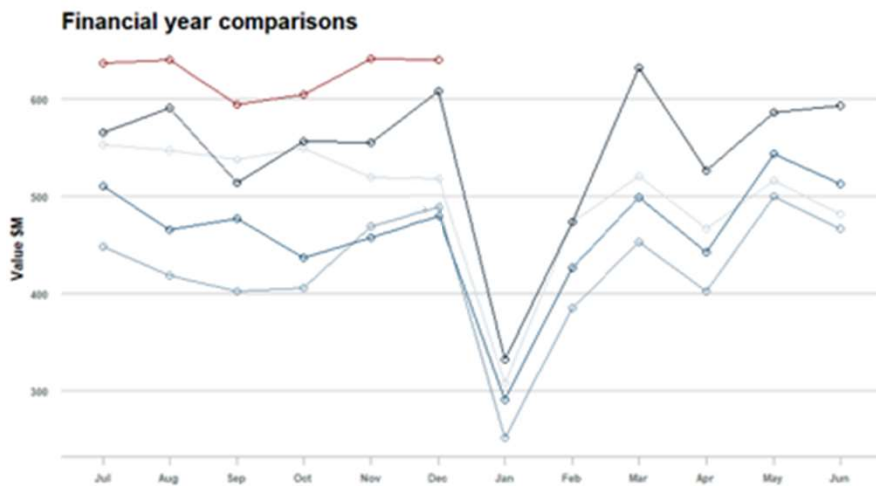
New process



The total value of imports **decreased** to \$24171M this month, driven by imports from **China** (down \$1416M).

By SITC, imports were **down** in **Telecommunications** (down \$423M) and **Miscellaneous manufactured articles** (down \$400M).

The main off setters of this move were imports from the **USA** (up \$1108M) with SITC increases to **Other transport equipment** (up \$816M) and **Petroleum & related materials** (up \$748M).

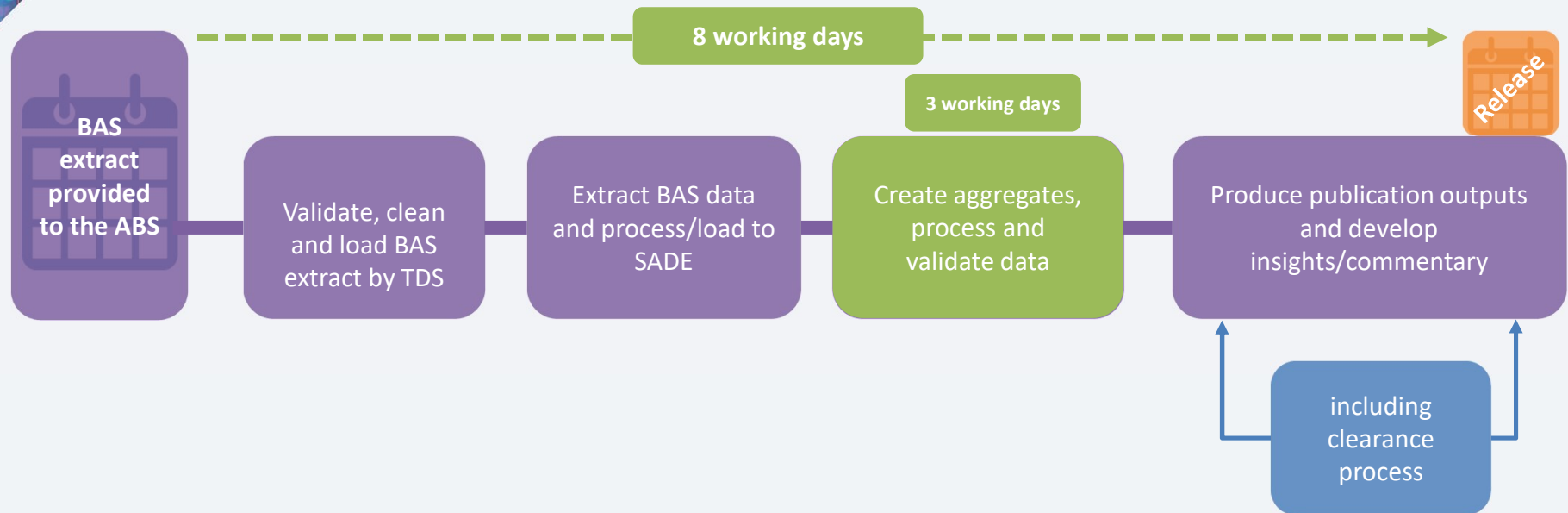


Exempel – Snabba Indikatorer med stora datamängder



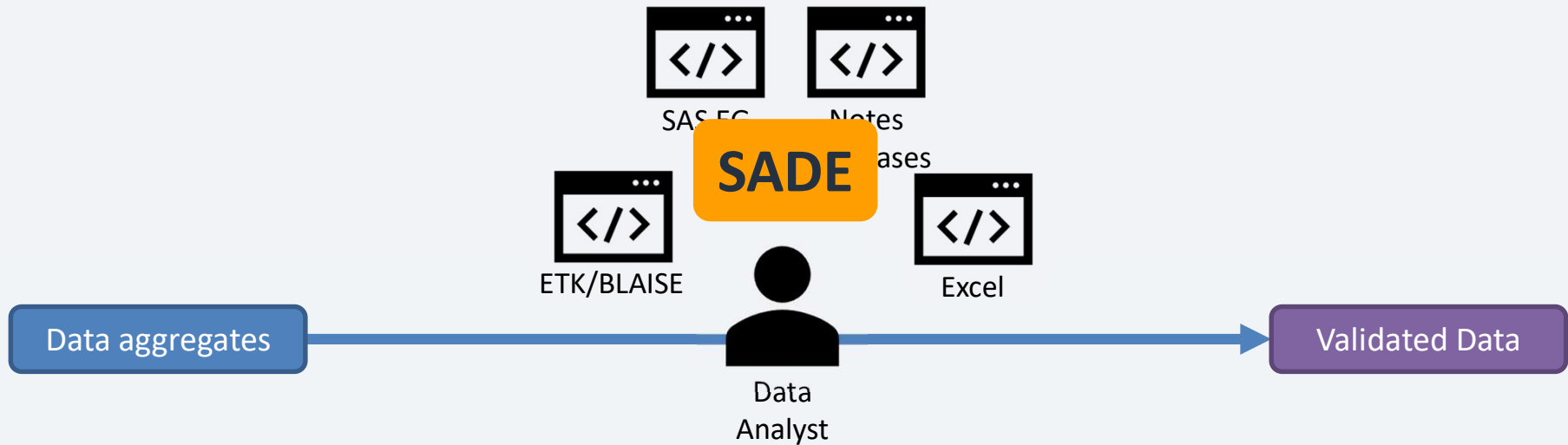
- ▶ **Big Data and Timely Indicators**
 - Ett projekt ambition att bygga ett produktionssystem med lösningar i molnmiljö för snabbare statistik med nya data.
 - Tre ekonomiska indikatorer, Compensation of employees, Household spending, Business turnover.
 - Har pågått ett par månader med den första indikatorn för omsättning.
- ▶ **Veckostatistik över lön och antal jobb**
 - Sattes först upp som indikator då pandemin slog till

Indikatorer – Ett engagemang för ökad aktualitet



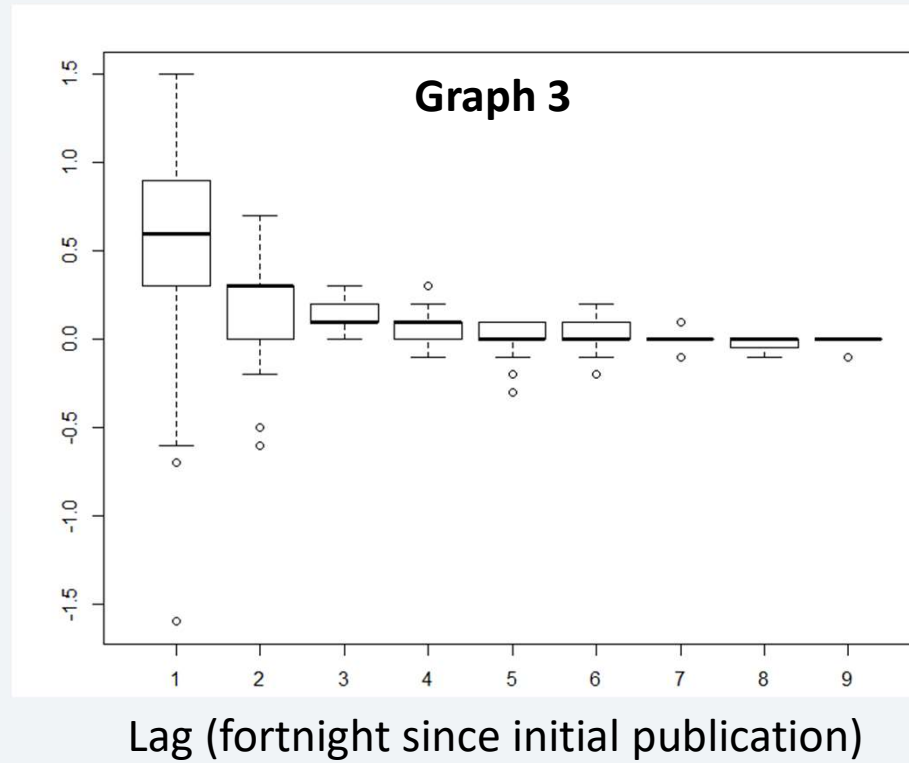
Epic 1a – Processing environment: current vs planned

New business statistics world



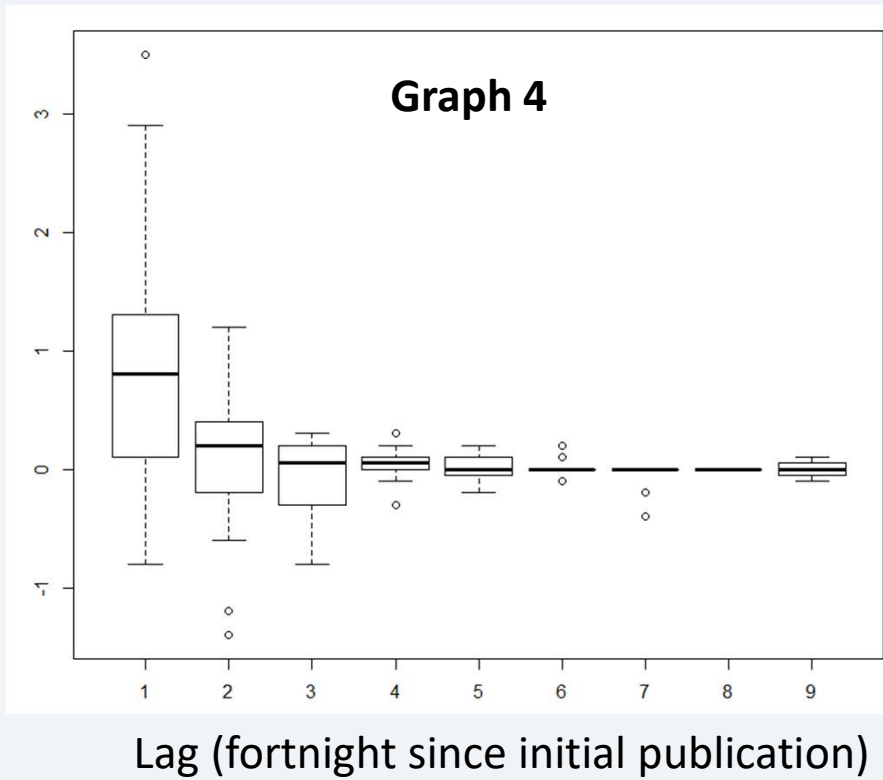
Size of revisions – (payroll) jobs

Size of revision (ppts)



Size of revisions – wages

Size of revision (ppts)



Tack så mycket!
Frågor?