

# A MAP for the future of price indexes at Stats NZ

Matthew Stansfield  
Frances Krsinich

Ottawa Group 2022



## Introduction

- There is a need for production processes around multilateral indexes
- After 20 years Stats NZ is building MAP (Multilateral Application Pipeline) to generalise these
- Multiple benefits from generalisation
- R package *multilateral* for index estimation is now available on CRAN
- Other production processes will be made available as R packages

## Multilateral price indexes

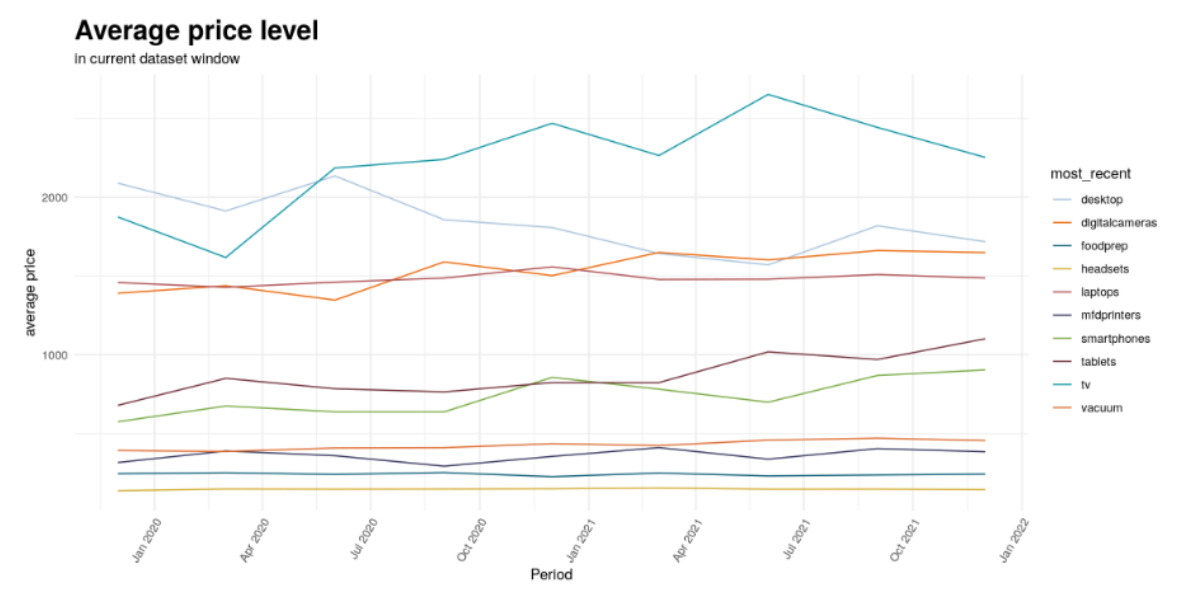
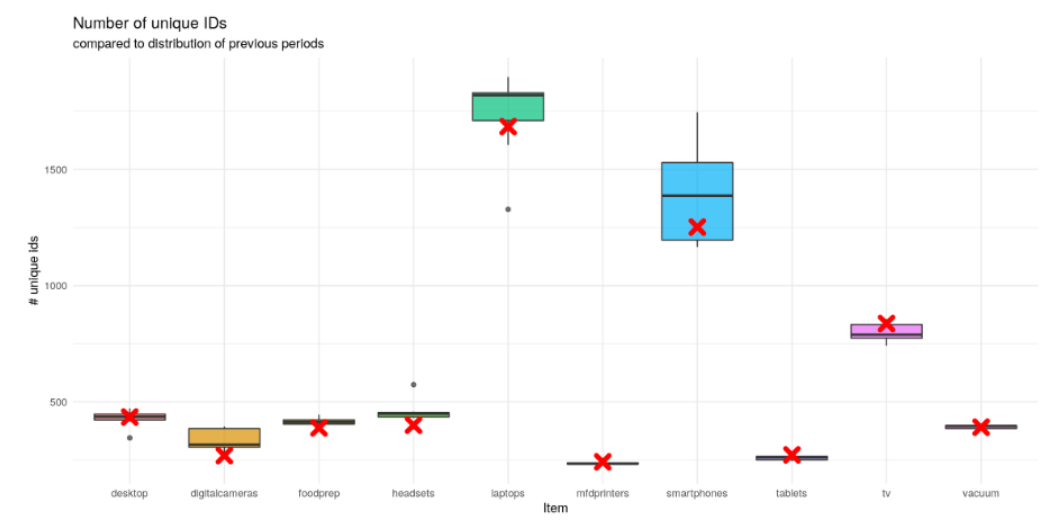
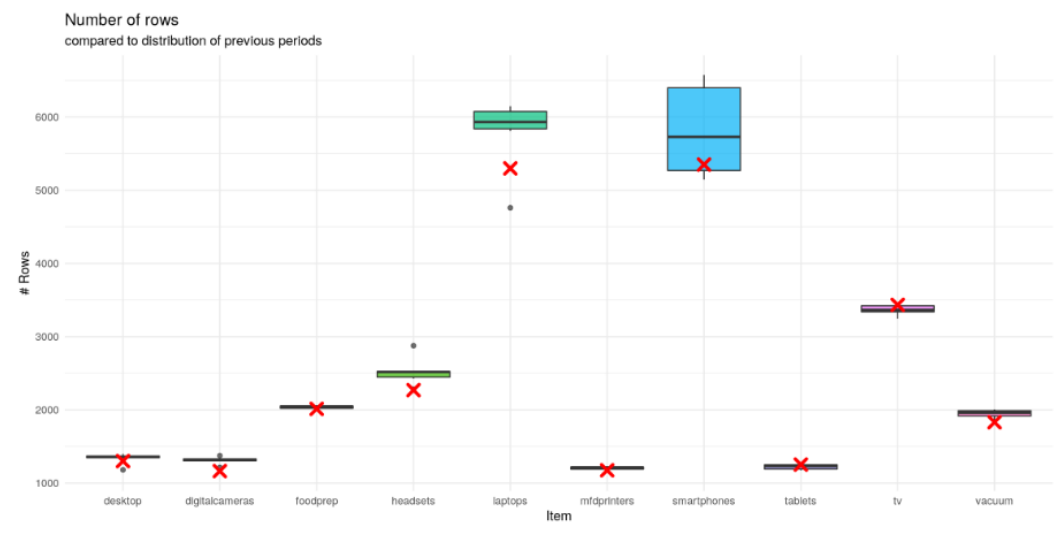
- Traditional methods don't work well with alternative data
  - chain drift (asymmetrical price/quantities due to sales)
  - implicit price movements associated with new products
- Over the last 20 years, significant research on multilateral methods
  - TDH, GEKS, TPD, GK, ITRYGEKS
- Stats NZ has adopted multilateral methods in production since 2001
  - used cars (2001, TDH), consumer electronics (2014, ITRYGEKS), rents (2019, TPD), overseas trade index (2013, 2020, TPD)
- 2019 internal review recommended consolidation of processes for both production and R&D

## Production processes

Production processes are needed in addition to the index estimation itself:

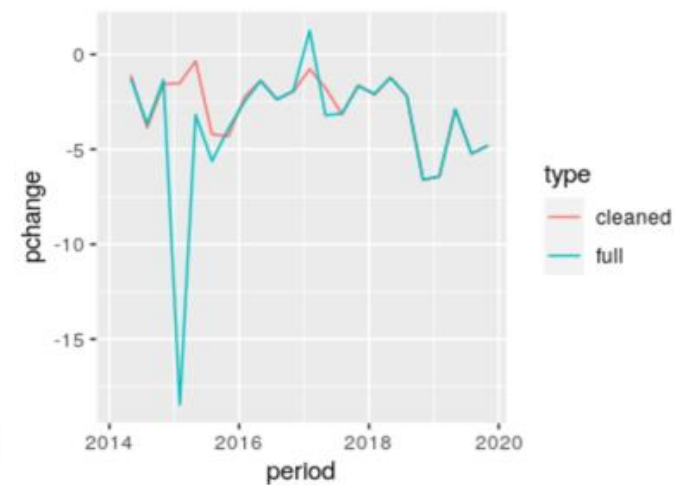
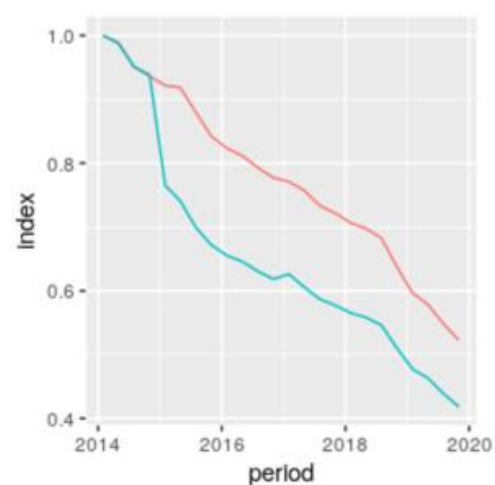
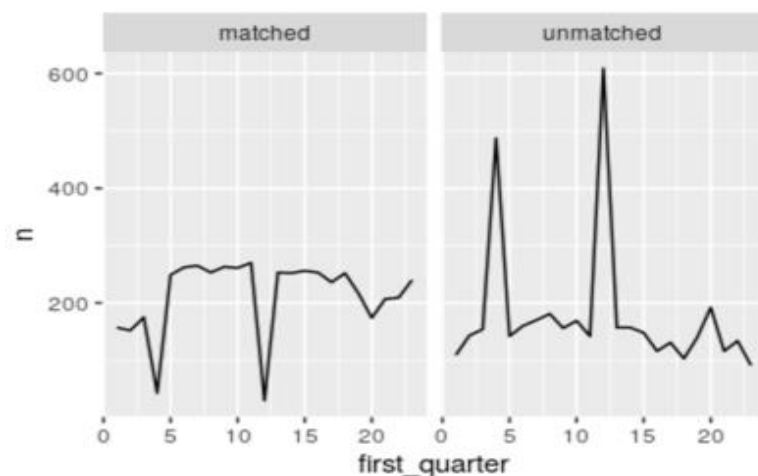
- **input diagnostics** to explore and validate source data
- **output diagnostics** to validate indexes, and compare them to previous production runs, effect of splicing on most recent movement
- **analytical measures** such as decomposition
- processes to identify and **deal with changes** – e.g. to coding of characteristics

# Example: input diagnostics



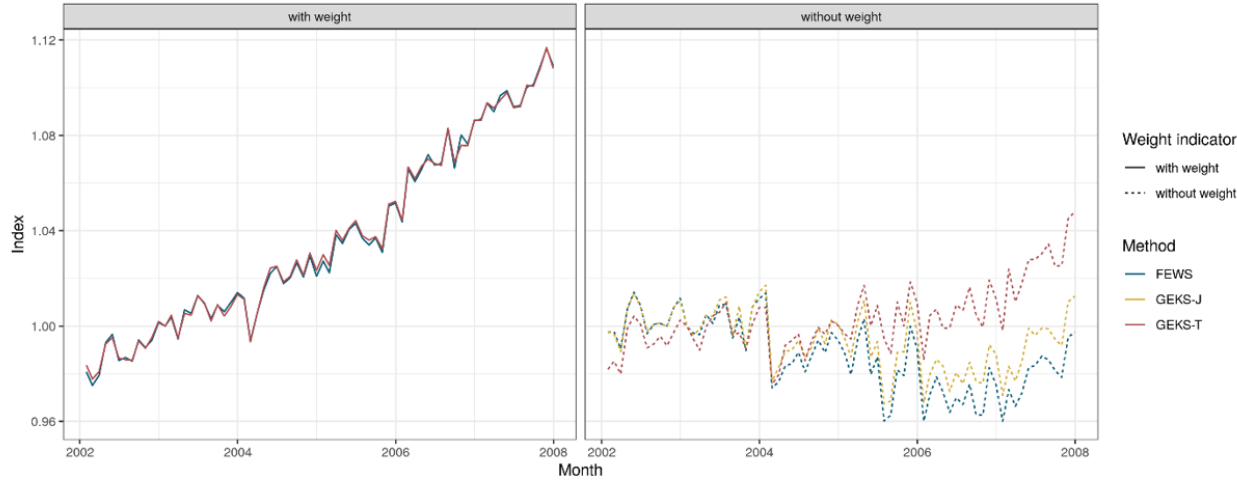
## Some production processes are non-trivial

- Index methodology only one part of production process
- Example – inconsistent coding of characteristics
  - Y|N|n.a. to Yes|No|N.A.

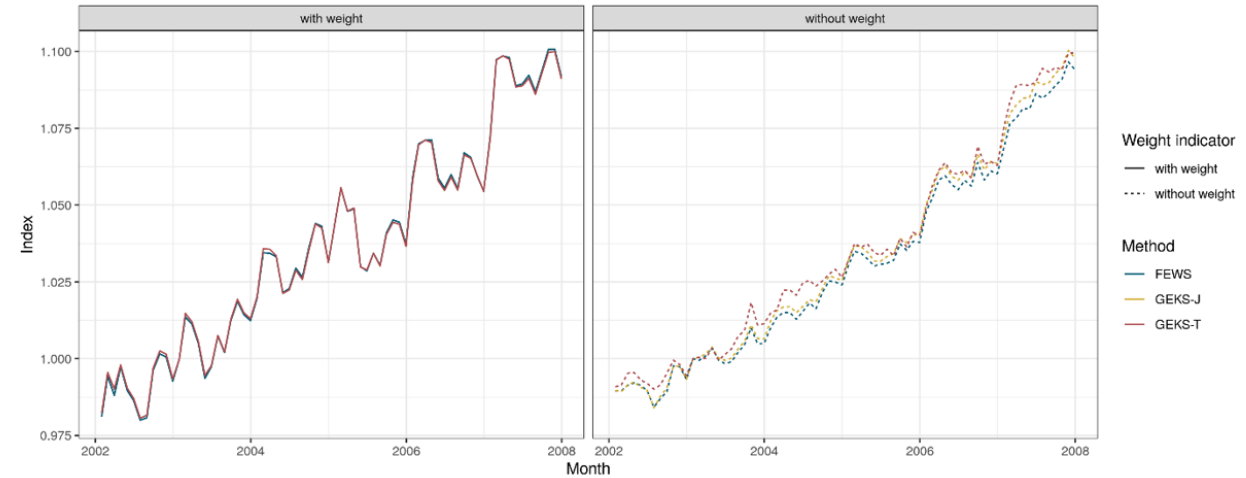


# Towards more automation of empirical testing **Stats NZ** Tatauranga Aotearoa

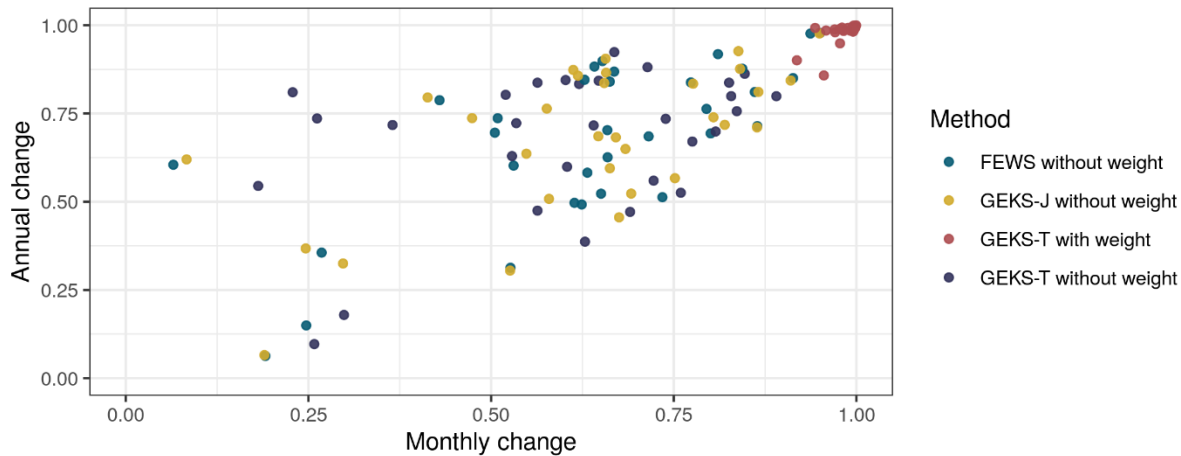
Blades: monthly price indexes (IRI supermarket scanner data)



Beer: monthly price indexes (IRI supermarket scanner data)



IRI Supermarket Products  
correlation against FEWS with weights



Enabling at-scale analysis with analytical interface:

- Multilateral methods and parameters (window length, splicing method) against historical benchmarks and one another
- Associations between results and characteristics such as churn, price change, technological change

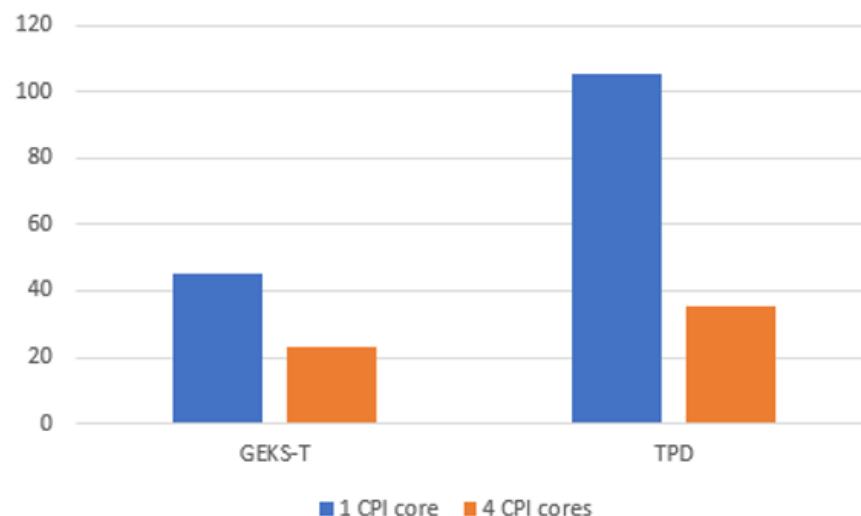
## Multiple benefits from generalisation

- **automation** to reduce manual and error-prone processes
- **transparency**, with code open for review and reuse by others
- **diagnostics, monitoring and analysis** incorporated with index estimation
- *multilateral* R package is **optimised for speed** and has **hedonic** functionality for all the methods Stats NZ uses (i.e. TDH, ITRYGEKS)
- **consistent interface** across product types, data sources and methods
- can **link to documentation and training** from interface



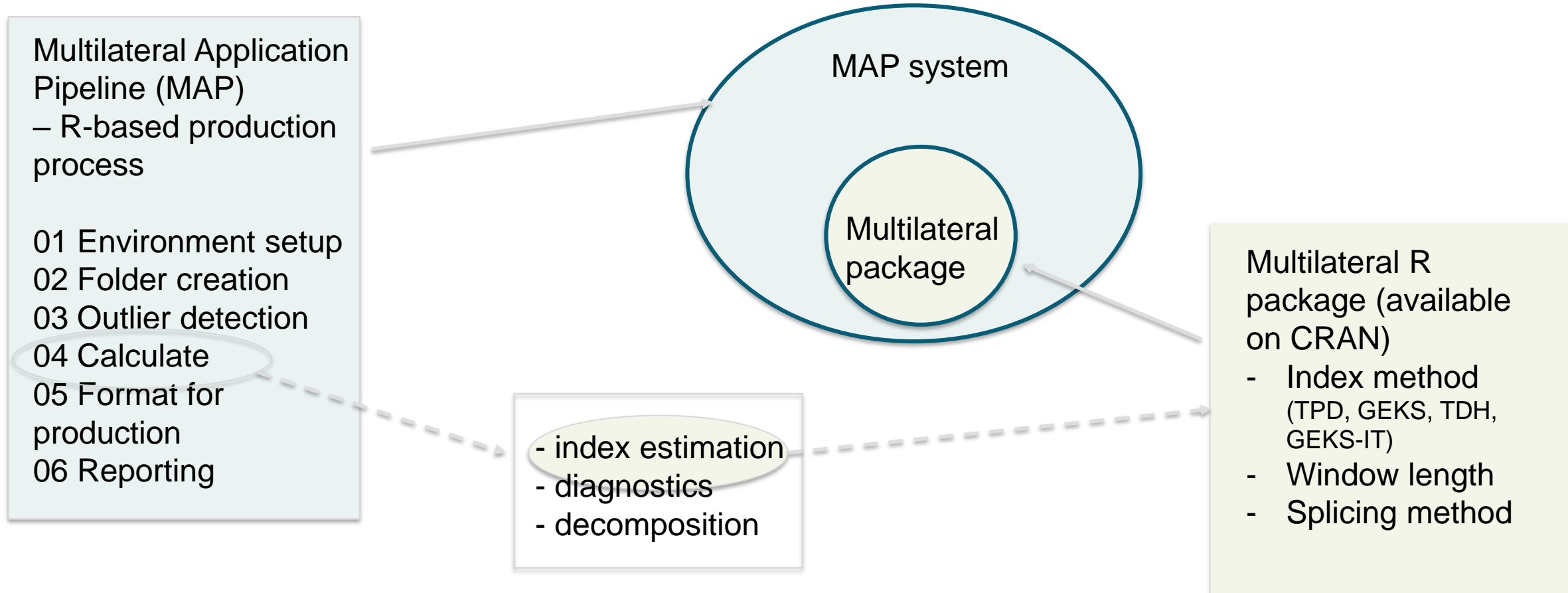
## Improved performance with parallel processing

- Two years of supermarket scanner data (not all retailers) – weekly data, 50 million observations
- Comparing performance for GEKS-T and TPD, window length 13 months and geomean splicing



*GEKS-T 45 min (1 core), 23 min (4 cores) TPD 105 min (1 core), 36 min (4 cores)*

# The multilateral application pipeline (MAP)



## Conclusion

- Production involves non-trivial processes other than index estimation
- Stats NZ is consolidating all our production processes
  - already migrated: used cars & consumer electronics
  - now migrating: rents and overseas trade indexes
- On the horizon: supermarkets, HPI prototype and in-house GS1
- *Multilateral* R package available on CRAN
- Expanding functionality / front-end for R&D
- We will also package up other MAP processes (e.g. diagnostics, points contributions, reporting) and make them available

**Thank you!**