Creating a digital Food Price Index from web-scraped data

Stats NZ
Tatauranga Aotearoa

1. Why the FPI?

- National interest
- Potential efficiencies
- Good online coverage
- Comparatively low churn rate (21%)
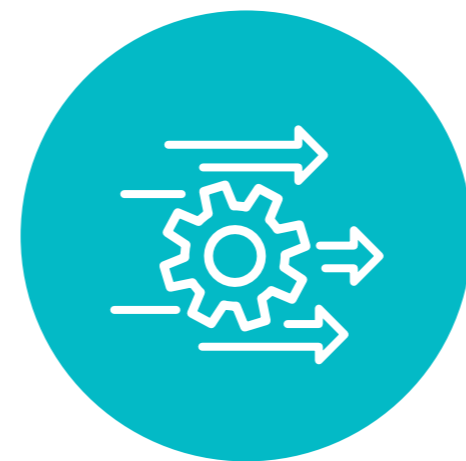
2. Data supply

- Pricestats data
- Daily web scraping of up to 30,000 items
- Two supermarkets (out of three)
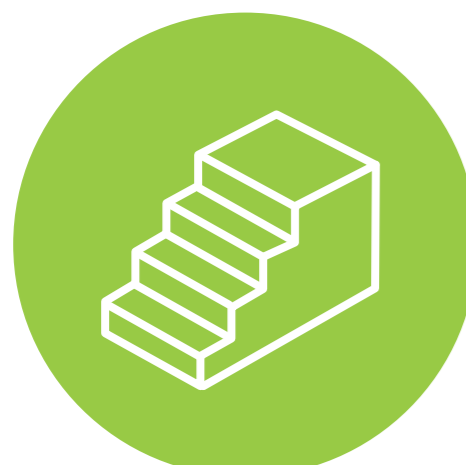- Metadata (retailer, date, product name, classification etc)

3. Methods

- Fixed Effects Window-Splice (FEWS) calculation
- Aggregation by COICOP method preferred so far

4. Considerations

- Splice position
- Window length
- Imputation
- Chain drift
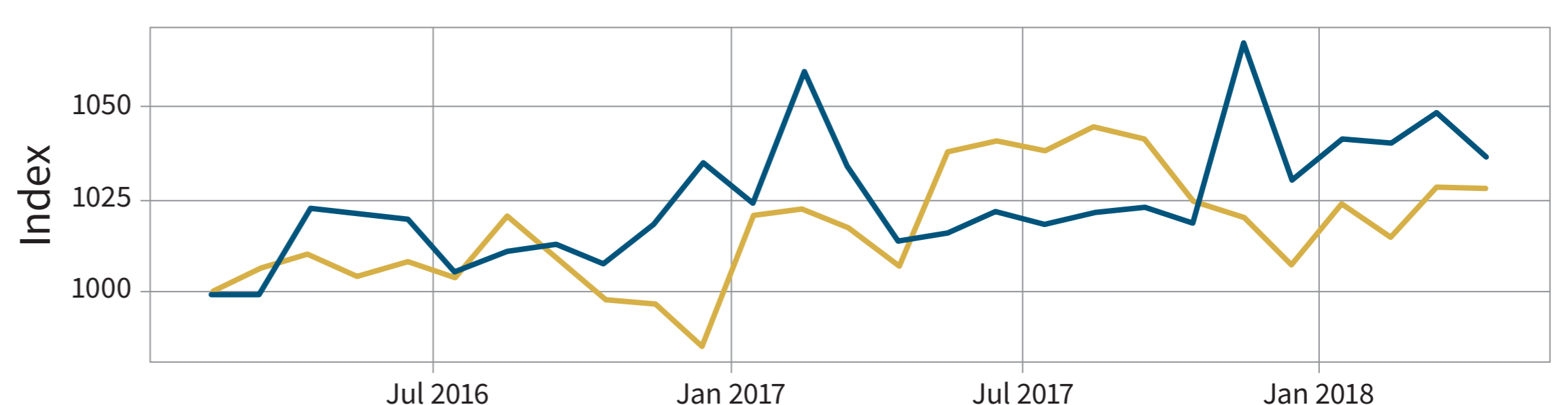- Coverage and data reliability

5. Obstacles

- Small number of online providers – need assurances or contingencies
- Online layout/structure
- Cost of purchasing data

Web-scraped data v published data
Base: Feb 2016 =1000

- Published FPI, without restaurant or takeaway meals
- Web-scraped data, FEWS, COICOP level 3

New Zealand Government