# Data Visualisation T2, 2025

## PBI for Front End

Sep 11, 2025

**Data Prep**

The data is drawn from the same 2 synthetic datasets as the Front End at time of writing:

* coles\_synthetic)dataset\_8\_weeks.csv
* woolworths\_cleaned.csv

Each has sufficiently similar fields to be combined, apart from the timeliness of their data: Coles has data for each of 8 weeks, noted as “Week 1”, “Week 2”, Week 3” etc.. whereas Woolworths has data for 5 unique, sporadically spaced dates.

For the sake of this exercise these two sets of dates were aligned arbitrarily in the “Supporting Data” spreadsheet (including one incorrectly formatted Woolworths date, formatted in US format).

**Data Model**

I leant into the metaphor of a stock market to visualise these data. The individual products could be considered as stocks and a representative basket of a typical shop could be an index, that snapshots the entire market / individual store.

After all, often consumers are looking for the best overall shop rather than the best deal on individual products.

I formed a representative basket based on this list built by Choice Magazine for their regular supermarket comparisons:

<https://www.choice.com.au/shopping/everyday-shopping/supermarkets/articles/how-we-surveyed-the-supermarkets>

To build the data model I generated a unique list of products by product code from the scraped data supplied. However there were some duplicate product codes (some products were being listed multiple times under different categories) so I arbitrarily removed dupes.

This is a key area of improvement for next semester!

Then I manually tagged products as belonging to the Choice Basket or not.

DAX Calculations identify the best available price for a representative Choice Basket at any given time at any given store.

You can read all other data model considerations in the actual PBIX file itself.

**Next semester**

Consider porting this view to Looker Studio.

Improve data scraping and formatting and link it to live data in Mongo DB

Toby 12 Sep