GovTech Day 3

Hands-on Lab Outline

EXCEL LAB

Split-apply-combine with Excel

- Sort
- Filter
- Select columns
- Copy-and-paste
- Pivot tables
- Drop columns
- Deal with missing values

Merge sheets

Visualisation

- Bars
- Lines

Brainstorm further questions

TABLEAU LAB

Comparison with Excel

Tour of the Tableau Gallery

- Graphs
- Dashboards
- Maps
- Story boards

Tour of Tableau GUI to illustrate Tableau's dynamic functionality (Demo)

Basic Charts (to be handed out as instruction sheet)

- Read in data
- Show drag-and-drop tools into rows and columns

- Drag propertyType into Rows
- Drag Price into Columns
- Change Price to Average
- Change Price to Median
- Switch Axes
- Sort Ascending

- Sort Descending
- Change Colour
- Add Axes Labels
- Add Title
- Mouse over for Annotation
- Right-click to highlight
- Add note and more annotations, Keep Only, Exclude etc.
- Drag Type of Sale to columns
- Drag Price to rows
- Highlight outlier by clicking
- Exercise - which street has the highest price?
- Exercise - which marketSegment has the lowest price?

    - Show sub-categories
    - Show Datetime manipulation
- Drag market segment to columns
- Drag price to rows
- Convert contractDate to date
- Segment by contractDate Year
- Further sub-set by month
    - Graphs best practice
- Experiment with different ways of portraying bars using Show Me tab
- Show comparisons
- Show change

    - More tweaks
- Add filters
- Drag street to columns
- Drag Tenure to columns
- Drag price to rows
- Drag street to filters
- Drag tenure to filters
- Add Colours
- Add Price to Colours to highlight property prices
    - Take-away - rapidly generate graphs and explore data
    - *Try-your-own*
- HDB_2013 dataset
- Tasks
    - Which town has the largest elderly population?
    - Which town has the highest number of primary schools?
    - Which town is most expensive to live in?
    - Read in another data source -> mention joins

Maps (To be handed out as instruction sheet)

- Read KML shape files (hawker stall dataset, population dataset)

- Drag Longitude to Columns
- Drag Latitude to Rows
- Drag Geometry to Detail
- Drag Name to Detail
- Tour of Tableau Help looking at blending geographic data
- HDB_2013 map lat long for proportional map
- Access map layers from map to customise look
- Add filter to drill down to specific area (this allows us to create dashboard later)
    - Try your own
- Access data from govtech open data
- Download KML file
- Read into Tableau
    - As we break for lunch, think of own story

30mins

- Storytelling with dashboards
    - ❏ Illustration of a good story with hawker visualisation
    - ❏ (demo example with simple graphs)
- Report generation with dashboards (demo example with simple graphs)
- Dashboard best practices (go back and change example - add legends, colour, emphasis etc.)
- Try your own
- Share and discuss

30mins

- Interacting with a Tableau Viz with Tableau Server - subscribe, get updates, collaborate
- Sharing Vizes
- Connecting to other data sources like Google Big Query
- Administering Tableau Server

PYTHON

Behind the scenes on our datasets

- Wrangling JSON with Python (Jupyter Notebook)
- Some examples of wrangling with Pandas (Jupyter Notebook)

- Function and Method arguments allow more flexibility
- Libraries for almost anything you want to do eg. manipulate strings, make API queries, convert lat long data
    - Visualisation with Seaborn (Jupyter Notebook)

    - Modelling with SKLearn (Jupyter Notebook)
- HDB_2013 dataset too small
- Good for visualisation and storytelling, not so good for statistical modelling because sample size too small
- SG_Housing dataset better for statistical learning
- But still not big data
    - More Python fun - scraping Lee Hsien Loong's twitter feed.