**Hands-On Exercise 1 (ggplot2)**

A white background with black text

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a graph

Description automatically generated

A plot must have at least one geom; there is no upper limit. You can add a geom to a plot using the **+** operator. FULL LIST: <https://ggplot2.tidyverse.org/reference/#section-layer-geoms>

A screenshot of a computer code

Description automatically generated

Default bin is 30 for histogram

A screenshot of a graph

Description automatically generated

A screenshot of a graph

Description automatically generated

Geom\_point() used to create scatterplot

A screen shot of a graph

Description automatically generated

Adding a stat() can override the default geom, and adding a geom() can override a default stat

A screen shot of a graph

Description automatically generated A screen shot of a graph

Description automatically generated

A graph with black dots and numbers

Description automatically generated

A graph with black dots and white text

Description automatically generated

Full list of themes: <https://ggplot2.tidyverse.org/reference/ggtheme.html>

Themes control elements of the graph not related to data e.g. background colour, size of fonts, gridlines, colour of labels. Each theme element can be conceived of as either a line (e.g. x-axis), a rectangle (e.g. graph background), or text (e.g. axis title).

**LESSON ONE**

A graph with blue line and white text

Description automatically generated

Information said to be half yearly but visually graph is in quarters – 4.5% cannot be viewed via chart

A diagram of data exploration

Description automatically generated

Should put data into R, Tableau or Python to visualise first (data visualization)

* See how skewed data is (left/right, single/double hump) to clean data and know what model to use (select variables, feature engineering)
* Gain better understanding of data (exploration), remove outliers, trends
* View model output to make better models (model visualization)

**Visual Analytics** – always show them what they know first, then show them what they don’t (unexpected)

Need good visualisation tool that is interactive

**In Class Exercise 1**

**Data Type**

 Numerical type

* Character type

Date

Location

A screenshot of a computer screen

Description automatically generated

dragging another table will automatically allow for a relational join, can choose what you want to join by

A white speech bubble with black and grey graphics

Description automatically generated



bottom left corner > worksheet, dashboard (visualize and combining multiple worksheets), story (visualize and combine multiple dashboards, like a flipchart)

A screenshot of a computer

Description automatically generated A screenshot of a computer

Description automatically generated

Dragging region to row, order date to column and sales to the table itself

**To see sorted data table**

Go to top tab Analysis > View Data to see the data sorted by the variables you’ve chosen

Sorted by data in table first (sales) > column (order date) > row (region)

**Row/Column total**

Go to left tab Analytics and drag totals to table, hover over row or column to get respective totals

A screenshot of a computer

Description automatically generated

If drag new variable over SUM(Profit), it will replace the sales with profit

If drag new variable over table where sales is at, it will add to the table (showing both sales and profit)

A blue box with white text

Description automatically generated

If add more dates, it will split it into small segments, if only 1 then it will sort by biggest (year)

If add more variables to rows, the best suited graph visualization will automatically be chosen. Otherwise, can change under the top right tab “Show Me”

A screenshot of a computer

Description automatically generated A screenshot of a computer

Description automatically generated



Adding reference line to give value to the visualization, type “Avg =” and use “>” button to get <value>

**DASHBOARD**

Choose “Default”, under size choose “Automatic” so that it can be viewed correctly depending on the device that it is used e.g. phone, laptop, tablet etc.