**Hands-On Exercise 2**

ggplot2 themes: <https://ggplot2.tidyverse.org/reference/index.html#themes>

ggplot2 extra geoms and scales: <https://yutannihilation.github.io/allYourFigureAreBelongToUs/ggthemes/>

Hrbrthems package provides a base theme that focuses on typographic elements (typeface, hierarchy, contrast, consistency, alignment, white space, and colour), including where various labels are placed as well as the fonts that are used.

More production workflow: <https://hrbrmstr.github.io/hrbrthemes/articles/why_hrbrthemes.html>

A screenshot of a phone

Description automatically generated

grid.arrange() <https://cran.r-project.org/web/packages/gridExtra/vignettes/arrangeGrob.html> (gridExtra package)

plot\_grid() <https://wilkelab.org/cowplot/reference/plot_grid.html> (cowplot package)

patchwork <https://patchwork.data-imaginist.com/> (combining separate ggplot2 graphs in a single figure)

<https://patchwork.data-imaginist.com/articles/guides/assembly.html>

**A white background with black text

Description automatically generated** A close up of text

Description automatically generated

**LESSON TWO**

A screenshot of a computer

Description automatically generated

DO NOT give the chart on the left as the data is not continuous. The derived information is useless and might even give the wrong information. Change it to categorical instead as it should be and get frequencies instead.

A close-up of a graph

Description automatically generated

Using the chart on the left it might seem like the company is doing well but using the chart on the right you’ll realise they are not doing well. Every month losing a percentage of new customers and revenue. Customer loss decreasing more in the beginning then become closer towards the end. Neet to get the complete story of the data – reflect more insightful data (right) even though both are valid. Can consider showing the left chart first (make customers happy) then tell them the surprising information (right).

A chart with colorful circles and numbers

Description automatically generated

Show the appropriate image. The rest are country or sub-country (e.g. Hongkong) level, but it includes EU. Should not lump together as EU because it consists of very small trading partners (e.g. Turkey, that might have lower trade than Thailand) while also including very big trading partners e.g. France, Germany.

A graph and diagram of a graph

Description automatically generated with medium confidence

Historical data is important, snapshots can be misleading.

A close-up of a graph

Description automatically generated

Should see the full trend, if only using the chart on the left it might look like the numbers are growing. Seeing the full picture on the right you’ll realise that the number of China tourists are actually dropping.

* Avoid using points (single dots) to represent time series. If want to show change over time, should use other charts e.g. line charts
* Avoid using points to represent discrete values, not obvious to see. Can use bar charts etc. instead
* Bar charts MUST start from zero, cannot start from a value.
* Avoid using pie charts as the human eye is not very good at assessing areas.
* Avoid unnecessary colours and flashy backgrounds for your charts, make them very difficult to see
* Avoid using 3D graphics in graphs
* Avoid using artistic designs that make data difficult to see

A screenshot of a graph

Description automatically generated

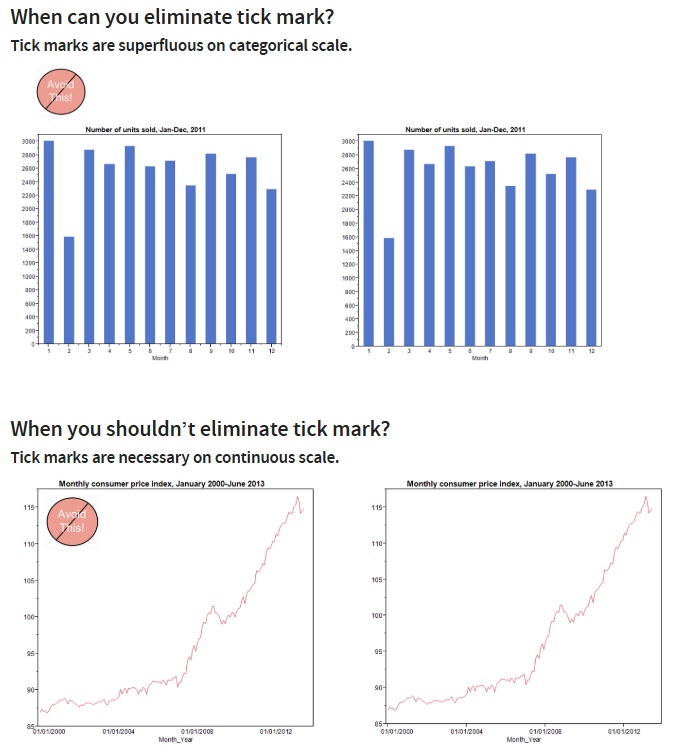
Chart on the right might be able to reflect the differences between charts easily but if want to show the change over time, not so useful. Line charts, however, can show rate of change as well (how steep line is)

A graph of potatoes and a chart of sales

Description automatically generated

If you want to use more interesting representations, make sure that it is accurate e.g. if 1 potato is 100, then 500 should have 5 potatoes.

* If you want different objects of the same colour in a graph to look the same, make sure that the background- the colour that surrounds them – is consistent.
* If you want objects in a graph to be easily seen, use a background colour that contrast sufficiently with the object.
* Use colour only when needed to serve a particular communication goal.
* Use different colours when they correspond to differences of meaning in the data. E.g. in bar charts use same colour since all the bars are representing the same thing.
* Use soft, natural colours to display most information and bright and/or dark colours to highlight information that requires greater attention.
* When using colour to encode a sequential range of quantitative values, stick with a single hue (or a small set of closely related hues) and vary intensity from pale colours for low values to increasingly darker and brighter colours for high values.
* Non-data components of a graph should be displayed just visibly enough to perform their role, but not more so, for excessive salience could cause them to distract attention from the data.
* To guarantee that most people who are colourblind can distinguish groups of data that are colour coded, avoid using a combination of red and green in the same display.



Depending on the nature of the chart, choose number of ticks and whether need to have ticks or not. Longer scale lines should also have more ticks

A graph with colored lines and text

Description automatically generated with medium confidence

Chart on the left is okay, standard way of doing things but if possible, the chart on the right is better.

A graph of a person and person

Description automatically generated with medium confidence

This chart is good – clear title, clear subtitle telling people what the chart measured by (% points, by income decile), axis clearly labelled (1 = lowest paid, 10 = highest paid)

A screenshot of a graph

Description automatically generated

Use annotations to describe massive charts so that people understand what is being shown

Weekday and weekend coloured clearly; grey shadings explained

A graph showing the growth of the company's sales

Description automatically generated

When something increases, another also increases DOES NOT MEAN CORRELATION. When trying to link two things together, do think about whether it is logical first

A graph showing countries/regions with different flags

Description automatically generated

Showing correlation between chocolate consumption and Nobel prize winners – might cause people to eat more chocolates in an attempt to be smarter. However, it might be that countries that can consume more chocolate = wealthier/well-to-do = greater access to education = higher chance of winning the Nobel prize

**In Class Exercise 2**

**A screenshot of a computer

Description automatically generated**



Since it is profit, it will be more insightful to show profits and losses. Reference line showing average might not be that useful, rather knowing what is above and below 0 would be better. Change the chart to a bar chart and increase the size of the line to become a bar.

**A screenshot of a computer

Description automatically generated**



Change to stepped colour to show positive and negative more obviously

To remove aggregation: Analysis > Aggregate measures (tick or untick)

A screenshot of a computer

Description automatically generated Drag state and drop over “Detail” and it will aggregate by state instead



A screenshot of a graph

Description automatically generated

Multiple in-built functions but we use in Ex2:

Quick table calculation > Percentile

Quick table calculation > Compute using > State