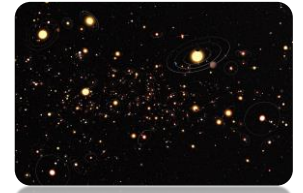
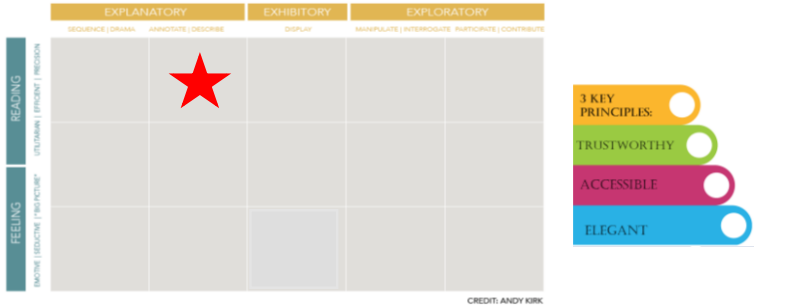




Visualization gives you answers to questions you didn't know you had.

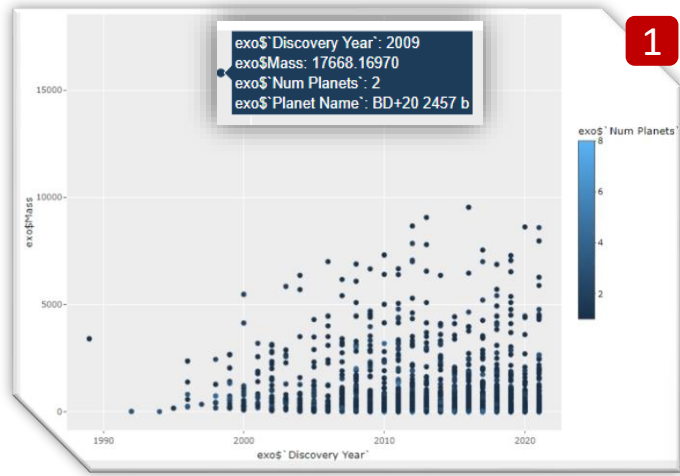


What is data visualisation and what makes a "good" data visualisation?
 As humans, we are better at processing information by looking at visuals. Hence creating an excellent visualisation allows greater and better interest. It has to be amazing, and visually appealing to develop attention. In specific, it concerns the framing of the visualisation and interactivity is considered too. A project map matrix can be used to place a type of visualisation. In this case, the visualisations will be for reading purposes to annotate/describe or display information about exoplanets.



Data Visualisation of Exoplanets using R Studio

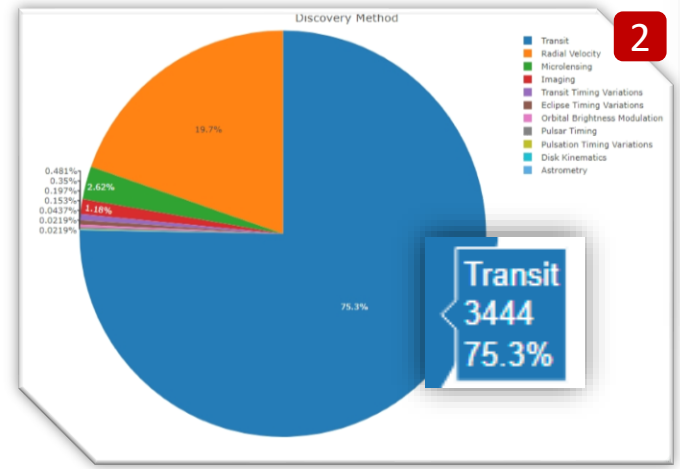
Tehræem Chaudhry Oxford Brookes University



The Exoplanet dataset consists of known exoplanets that were discovered using a variety of methods, and many have been detected using multiple methods. The data was produced by the NASA Exoplanet Archive online. The original dataset has **4575** rows and 23 columns "**dplyr**", "**ggplot2**", "**tidyverse**", "**plotly**" are the necessary packages needed to be installed to create these interactive styled visualisations which show data labels when hovered over.

Showing the range of masses of exoplanets discovered over the years, with added labels with the specific year, mass, exoplanets name, and number of planets orbited.

Showing the discovery method of exoplanets as a percentage with its quantity.



```
#Interactive Pie chart for Discovery Method using plotly
install.packages("plotly")
library(plotly)

exo1 <- all_exoplanets_2021
exo1 <- exo1 %>% group_by('Discovery Method')
exo1 <- exo1 %>% summarize(count = n())
vis <- exo1 %>% plot_ly(labels = ~`Discovery Method`, values = ~count)
vis <- vis %>% add_pie
vis <- vis %>% layout(title = "Discovery Method")
vis

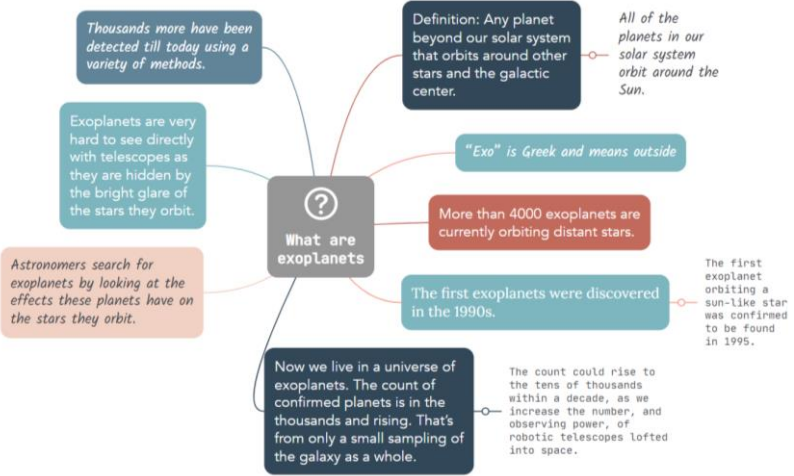
#Interactive Plot with multi labels using ggplot2 and plotly
install.packages("ggplot2")
library(ggplot2)

exo <- all_exoplanets_2021

exo_plot_hover <-
  ggplot(exo, aes(x =exo$`Discovery Year`,
                 y = exo$Mass,
                 colour = exo$`Num Planets`,
                 info = exo$`Planet Name`)) +
  geom_point()

ggplotly(exo_plot_hover)
```

- Key steps:**
- ✓ Import dataset into R studio (excel csv file)
 - ✓ Install necessary packages
 - ✓ Group columns selected
 - ✓ Redefine to simpler names for ease of use
 - ✓ Include all arguments in the function
 - ✓ Check there are no extra brackets or spaces



Concluding thought: Isn't Excel easier to use?
 Although time consuming, but when dealing with large datasets, it is incredibly useful and efficient to learn coding styles to develop graphical outputs. This is an example of what can be done using one kind software(R studio). The amazing thing about data visualisation is that there are many other softwares that create a wide range of advanced outputs.



"THE PURPOSE OF VISUALIZATION IS INSIGHT, NOT PICTURES."

"There is a magic in graphs. The profile of a curve reveals in a flash a whole situation."