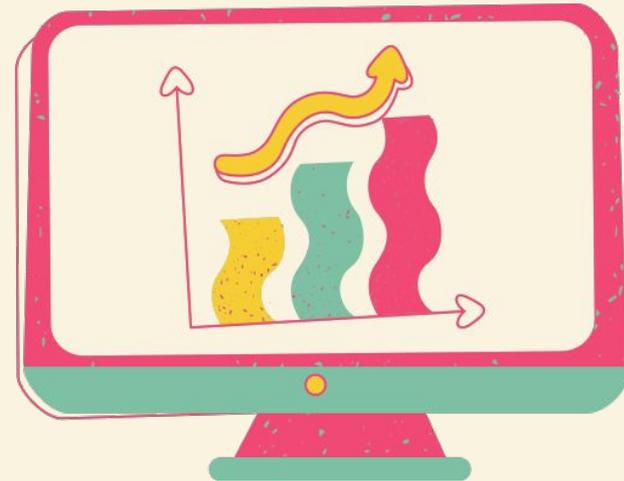


DATA SCIENCE

What is Data Science?

It's a field that uses machine learning, AI, processes, algorithms and other systems to **extract knowledge & insights** from large amounts of data and apply it.

Still unclear? Let's explore an example.





A group of 10 men (M) and 10 women (W) were selected for a test. The table below shows the results for each participant.

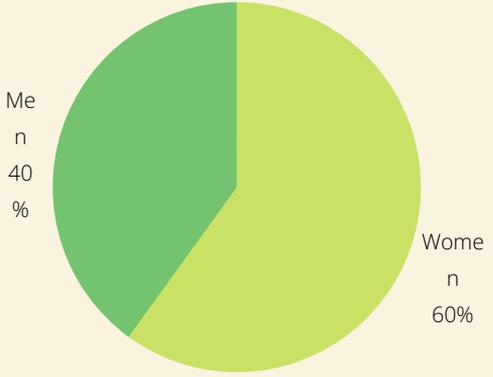
| | | | |
|-----------------------|------------|-------------|-------------|
| M 1 - Fail | W 1 - Pass | | |
| M 2 - Pass | W 2 - Pass | M 6 - Pass | W 6 - Pass |
| M 3 - Pass | W 3 - Fail | M 7 - Fail | W 7 - Fail |
| M 4 - Fail | W 4 - Pass | M 8 - Fail | W 8 - Pass |
| M 5 - Fail | W 5 - Pass | M 9 - Pass | W 9 - Fail |
| | | M 10 - Pass | W 10 - Pass |



While the raw data above doesn't really tell us anything, when ANALYSED, it reveals certain conclusions we can later use.



6 in 10 people who gave the test passed.

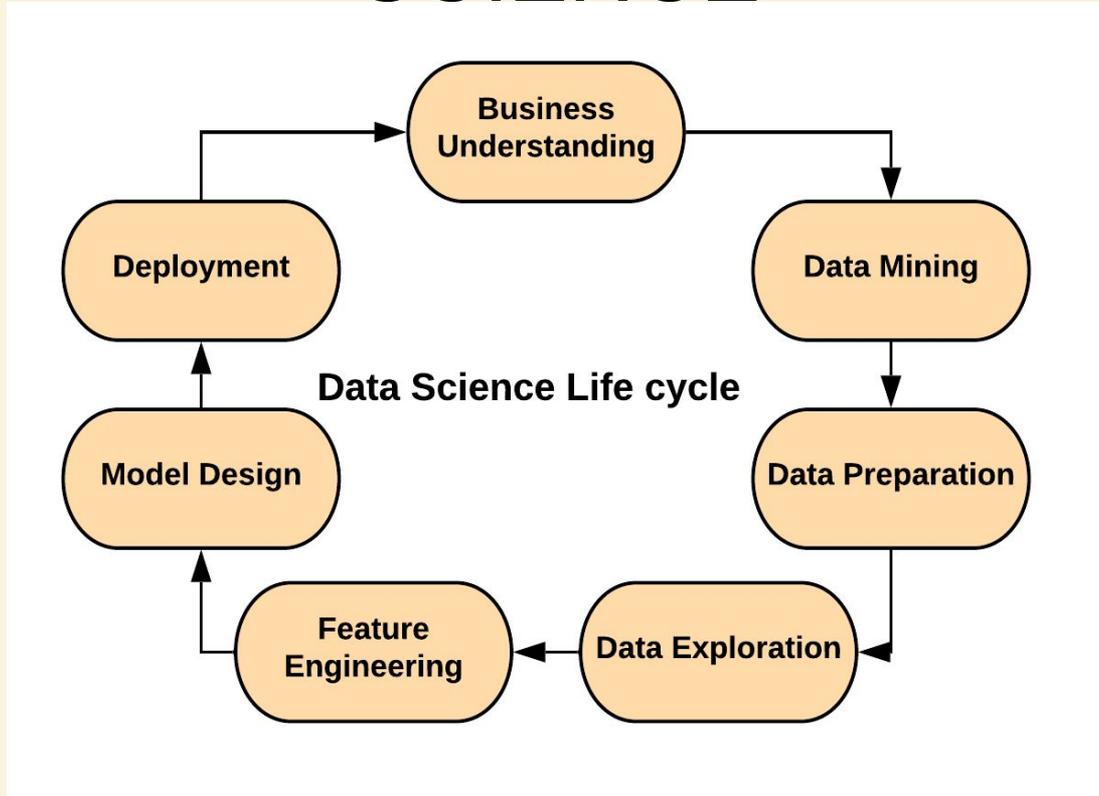


Women did 20% better than men in the test.

The data provided 2 conclusions when analysed.

This is the purpose of data science.

The cycle of DATA SCIENCE





R



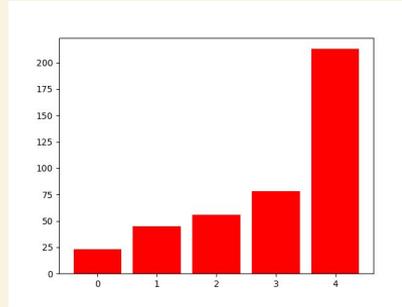
Python



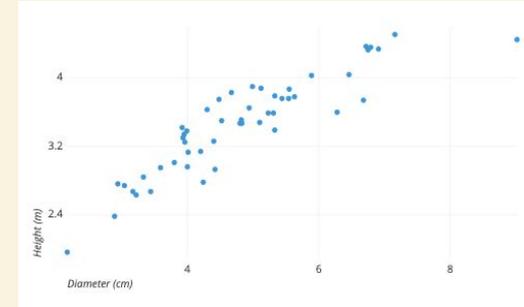
Chart types



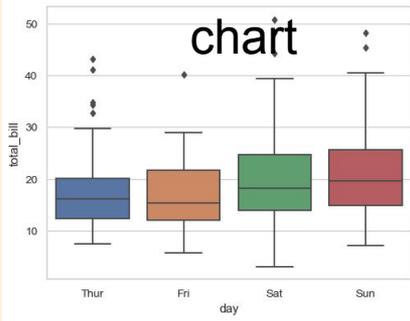
Pie



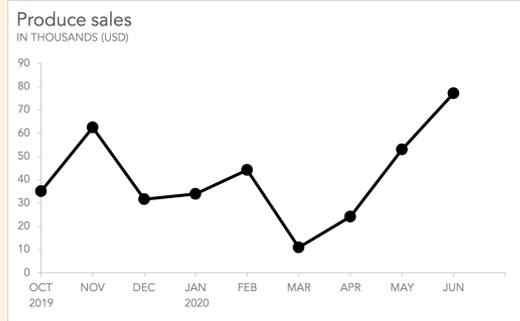
Bar chart



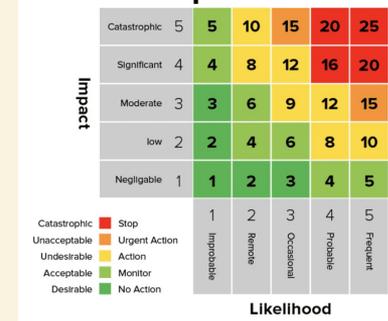
Scatterplot



Line graph



Boxplot



Heatmap

