

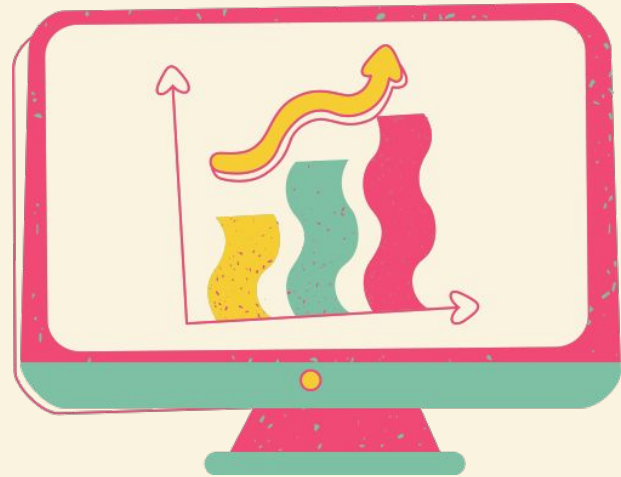


# DATA SCIENCE

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.

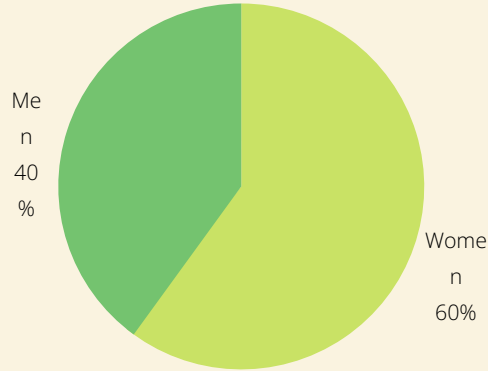
M 1 - Fail	W 1 - Pass		
M 2 - Pass	W 2 - Pass	M 6 - Pass	W 6 - Pass
<del>M 3 - Pass</del>	W 3 - Fail	M 7 - Fail	W 7 - Fail
M 4 - Fail	W 4 - Pass	M 8 - Fail	W 8 - Pass
M 5 - Fail	W 5 - Pass	M 9 - Pass	W 9 - Fail
		M 10 - Pass	W 10 - Pass



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.



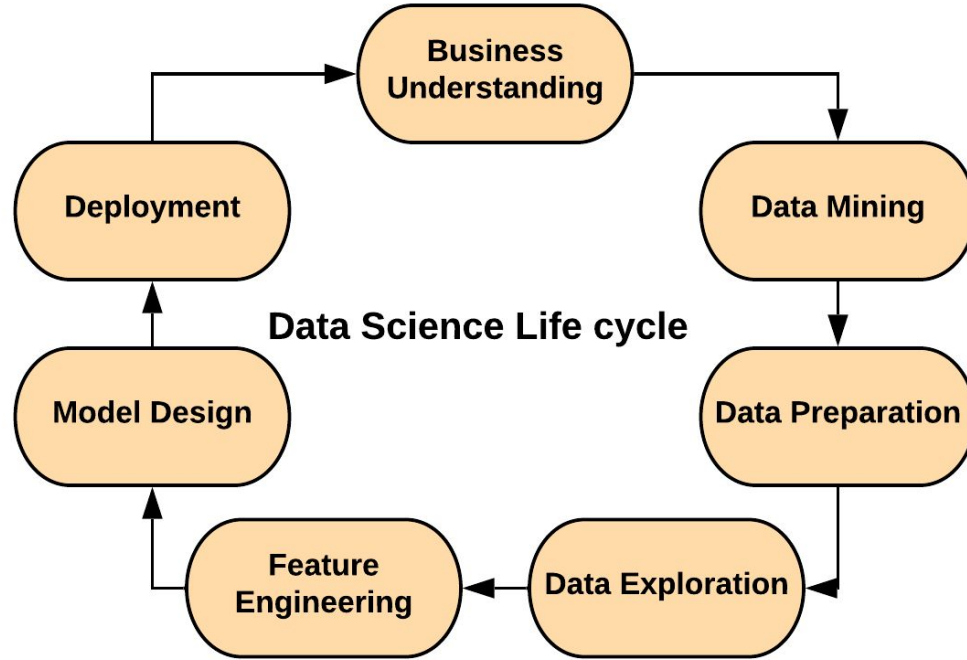
Women did 20% better than men in the test.

The data provided 2 conclusions when analysed.

This is the purpose of  
data science.



# The cycle of DATA SCIENCE





R



Python

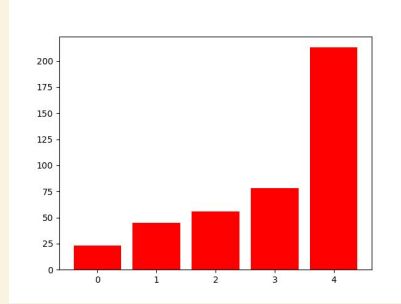


python<sup>TM</sup>

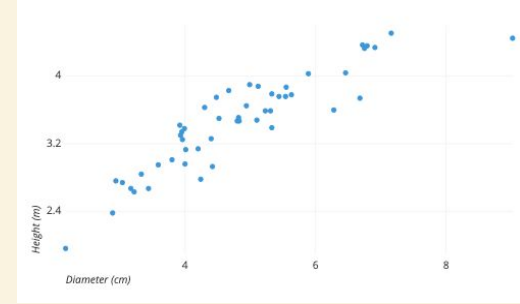
# Chart types



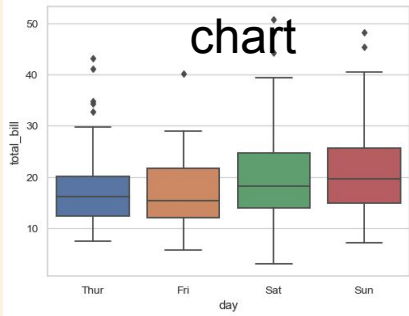
Pie



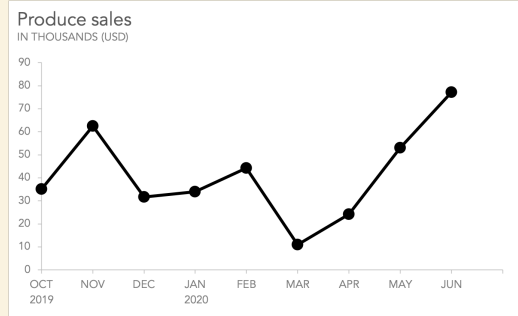
Bar chart



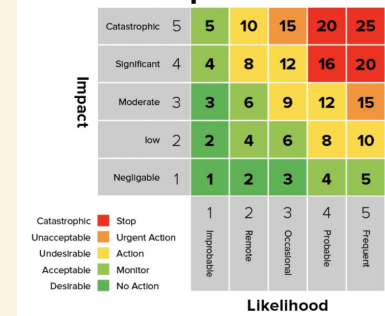
Scatterplot



Line graph



Boxplot



Heatmap



# Data Science with Google Sheets

Gender Deaths

File Edit View Insert Format Data Tools Extensions Help

Last edit was seconds ago

100% \$ % .0 .00 123 Calibri 10 B I S A

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30

Year	Gender	Value
2009	Females	1,940
2010	Females	2,129
2011	Females	2,096
2012	Females	2,250
2013	Females	2,388
2014	Females	2,417
2015	Females	2,555
2009	Males	5,849
2010	Males	5,285
2011	Males	5,254
2012	Males	5,452
2013	Males	5,627
2014	Males	5,843
2015	Males	6,200