Q1- How can a box plot be used to check if the data is normally distributed or skewed? How would you interpret in terms of skewness: if you see the median closer to 1st QR or 3rd QR in your boxplot?

We can describe a distribution as normal if there are equal proportions around the median in a box plot. When the median is closer to the 1st or 3rd quartile the box plot is skewed, if the median is in the 1st quartile the boxplot is positively skewed, and if its in the 3rd quartile the boxplot is negatively skewed.

Q2- Select a small dataset of your choice

Index	Feature 1	Feature 2	Feature 3	Target
1	4	11	5	A
2	7	2	9	A
3	1	9	1	B
4	3	10	4	B
5	5	1	2	B
6	7	0	1	A
7	9	2	0	B

a- Pick sample data for random forest (show in a table by respecting the recommendations from the inventors of the algorithm)

Index	Feature 1	Feature 2	Feature 3	Target
1	4	11	5	A
2	7	2	9	A
7	9	2	0	B
1	4	11	5	A
5	5	1	2	B
1	4	11	5	A
7	9	2	0	B

b- Pick sample data for Bagging (show in a table)

Index	Feature 1	Feature 2	Feature 3	Target
2	7	2	9	A
2	7	2	9	A
7	9	2	0	B
1	4	11	5	A
6	7	0	1	A
3	1	9	1	B
7	9	2	0	B