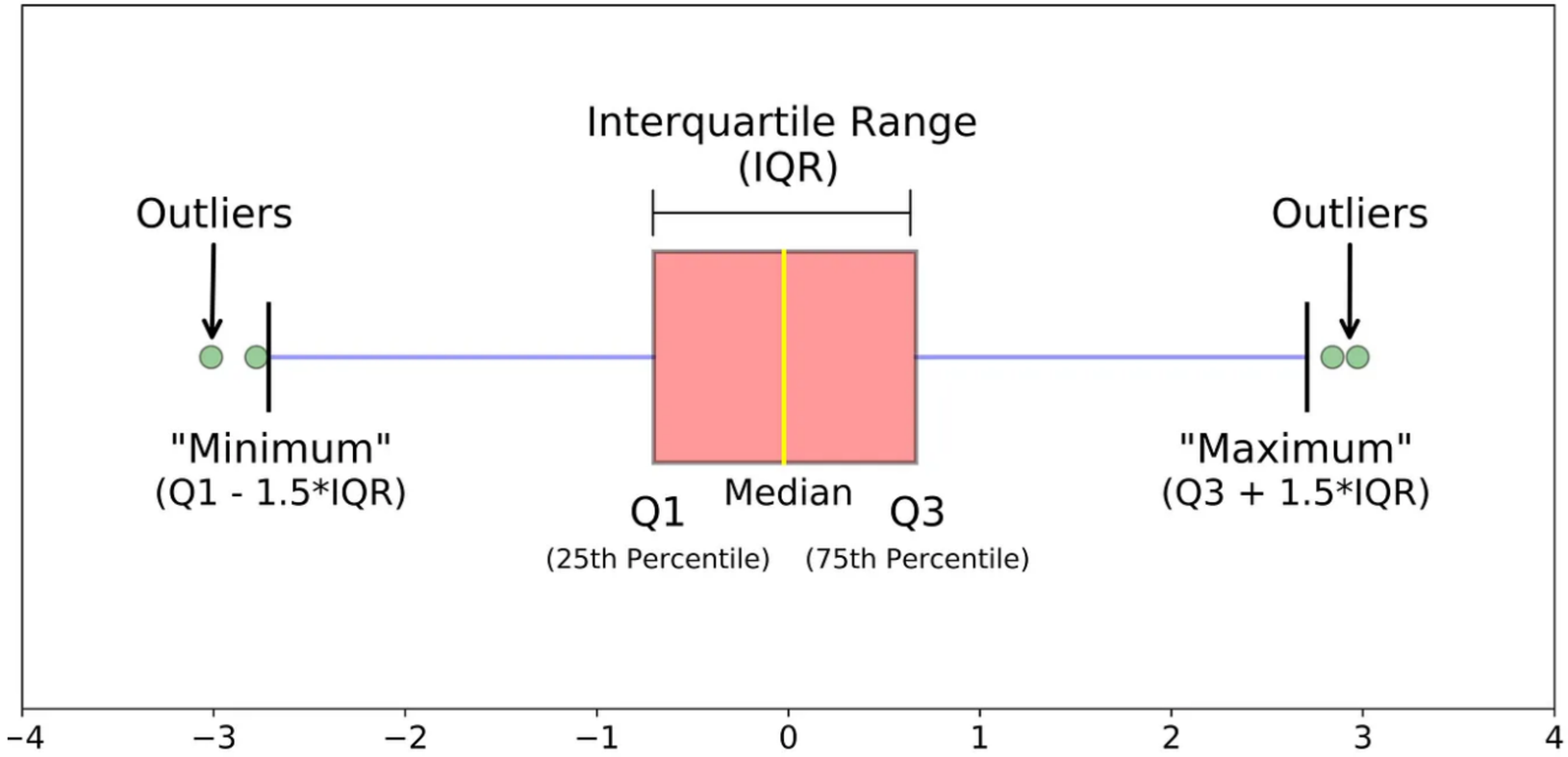
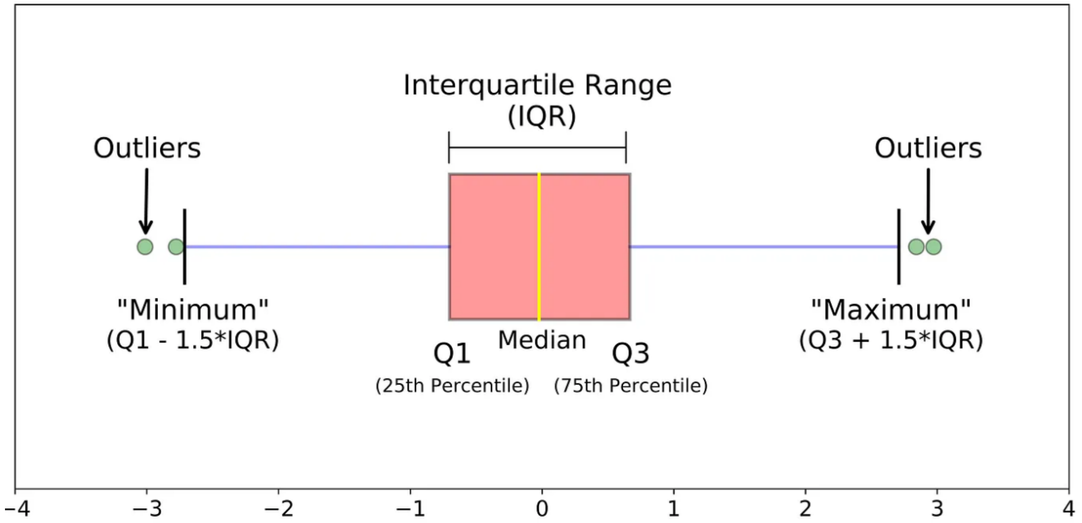
Recitation week 7

Boxplots and Outliers



Boxplots and Outliers

- To make a Boxplot in Google sheets we need to calculate:
  - Minimum value => =min(data range)
  - Maximum value => =max(data range)
  - Q1 => =quartile(data range, 1)
  - Q3 => =quartile(data range, 3)
  - IQR => Q3 - Q1
- A value is outlier when if it is
  - Higher than Q3 + 1.5 * IQR
  - Lower than Q1 - 1.5 * IQR



Boxplots and Outliers

1. Drawing Boxplots

- To draw a Boxplot for a variable in google sheets we need the information in the table
- Then, select this information and insert chart

Varaible	Min	Q1	Q3	Max
...	...	...	..	...

Boxplots and Outliers

Detecting outliers

- Unfortunately google sheets doesn’t show the outliers in Boxplots
- We need to find them by ourself
- Outliers are:
  - < Q1 - 1.5 * IQR (LB)
  - > Q3 + 1.5 * IQR (UB)

Varaible	Min	Q1	Q3	Max
...	...	...	..	...

Boxplots and Outliers

Detecting outliers

- Unfortunately google sheets doesn’t show the outliers in Boxplots
- We need to find them by ourself
- Outliers are values that are lower than the lower bound, and higher than the higher bound
- LB = Q1 - 1.5 * IQR
- UB = Q3 + 1.5 * IQR

IQR	LB	HB		
...	...	...	..	

Boxplots and Outliers

Detecting outliers

- After finding the LB and UB, we can filter our variable to be between LB and UB.
- Our variable has no outlier anymore!