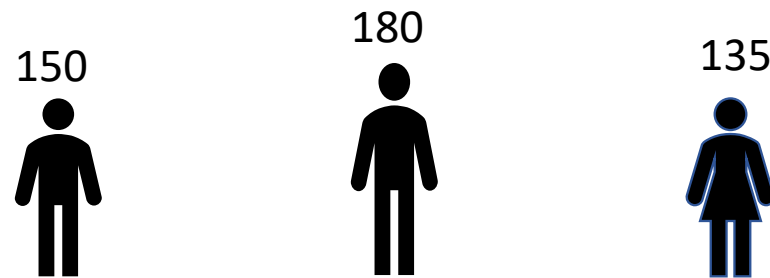


Relevant stats

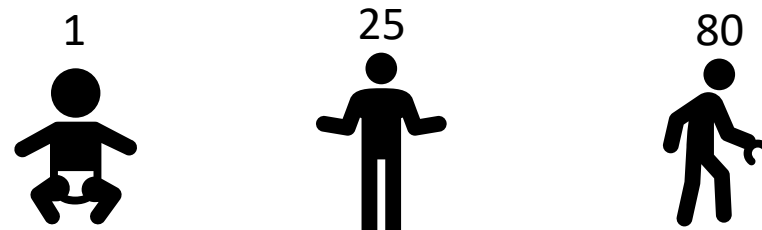
Variable

- Attribute of an entity(person, place, object, ...).
- Value of a variable can change.

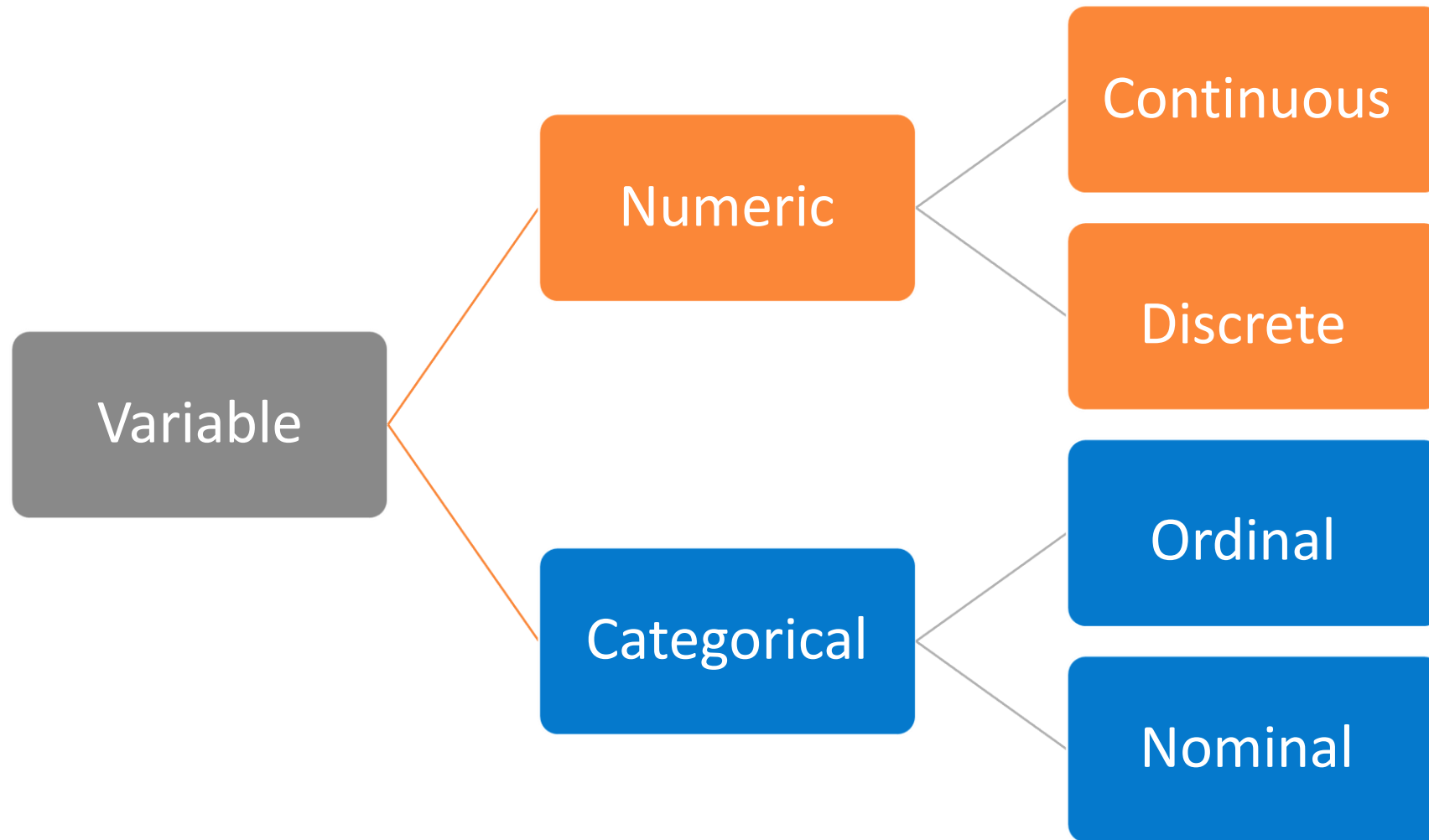
Height(cm)



Age(years)



Time

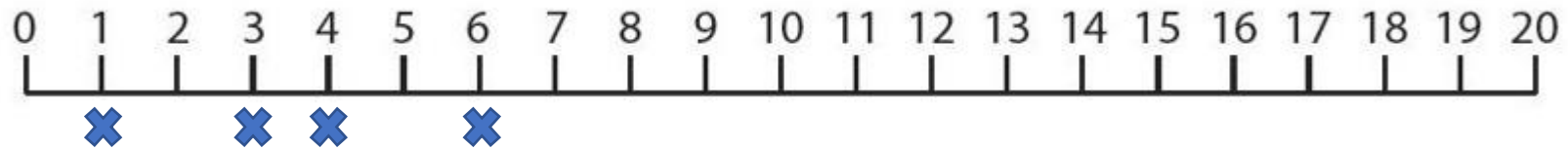


Measures of central tendency

Mean, Median, and Mode

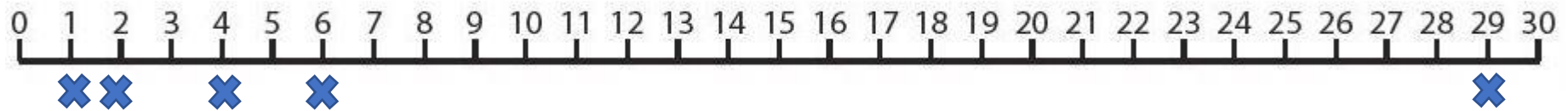
Mean

6, 1, 4, 3



Median

2, 29, 1, 6, 4



Mode

3, 5, 3, 7, 6, 15, 7, 3

1, 2, 1, 5, 23, 2

Measures of dispersion

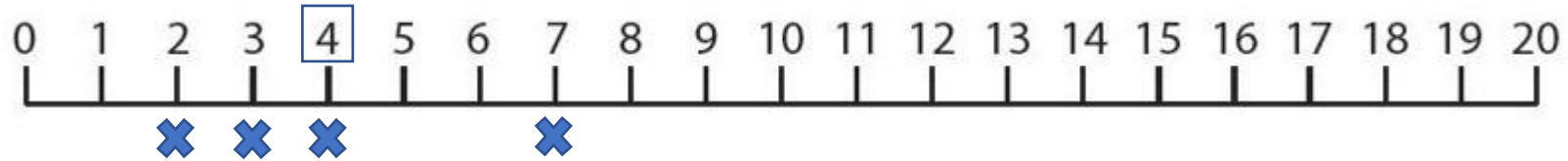
Range, Variance, Standard Deviation, and Quartiles

Range

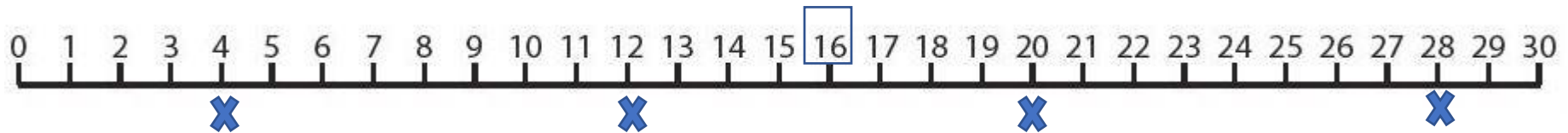
7, 9, 12, 23, 4

Variance & Standard Deviation

7, 2, 4, 3



4, 12, 20, 28



Quartiles and IQR

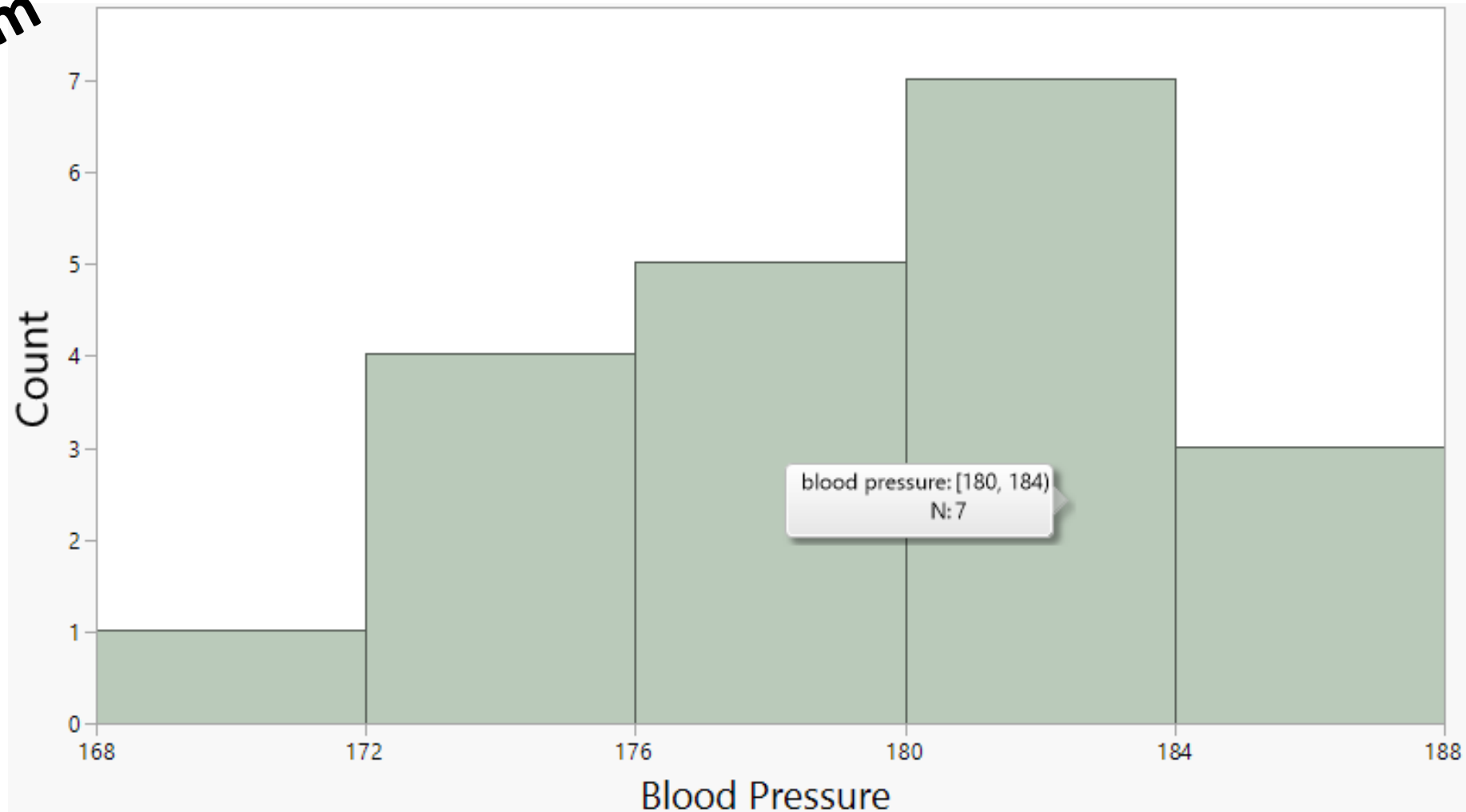
2, 4, 5, 6, 12, 15, 20, 22, 28

2, 4, 5, 6, 12, 15, 20, 22

Data Pattern/Distribution

183, **173**, 181, **181**, 184, **171**, 175, **179**, 178, **181**, 182, **178**, 186, **175**, 181, **181**, 179, **177**, 176, **184**

Histogram



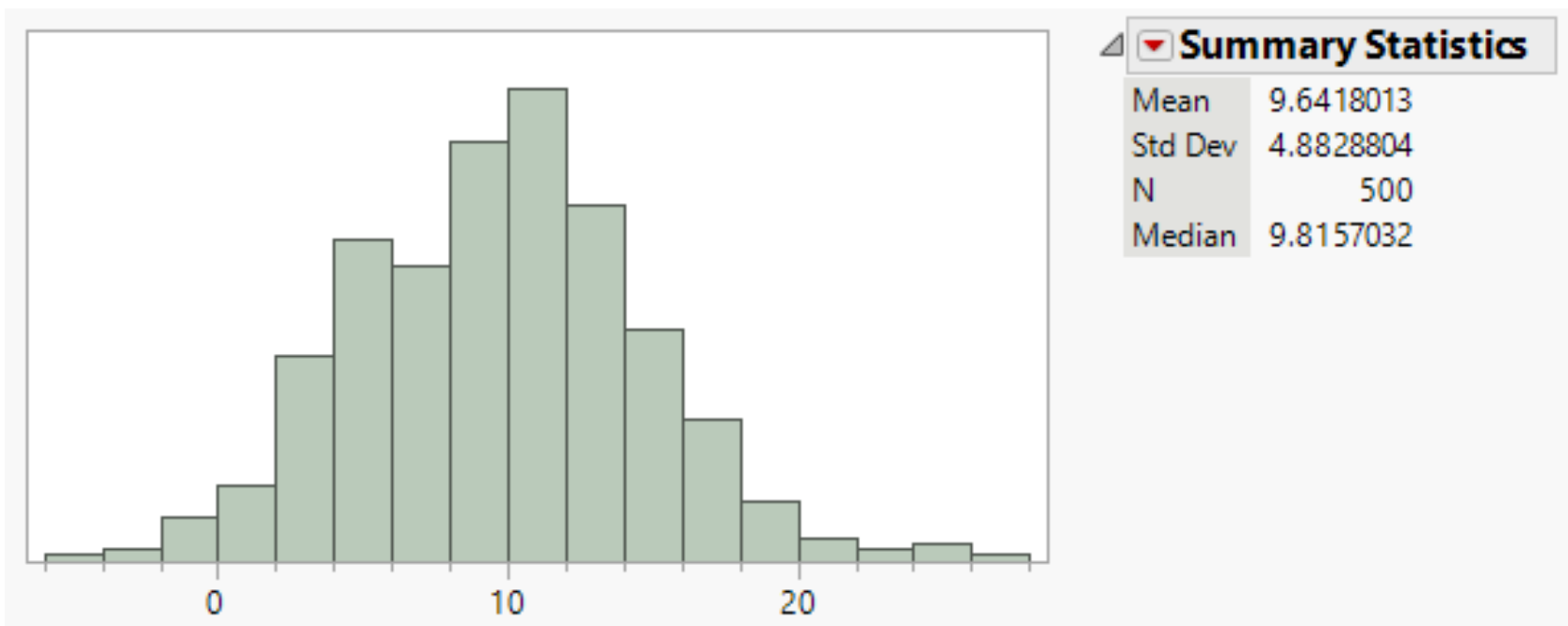
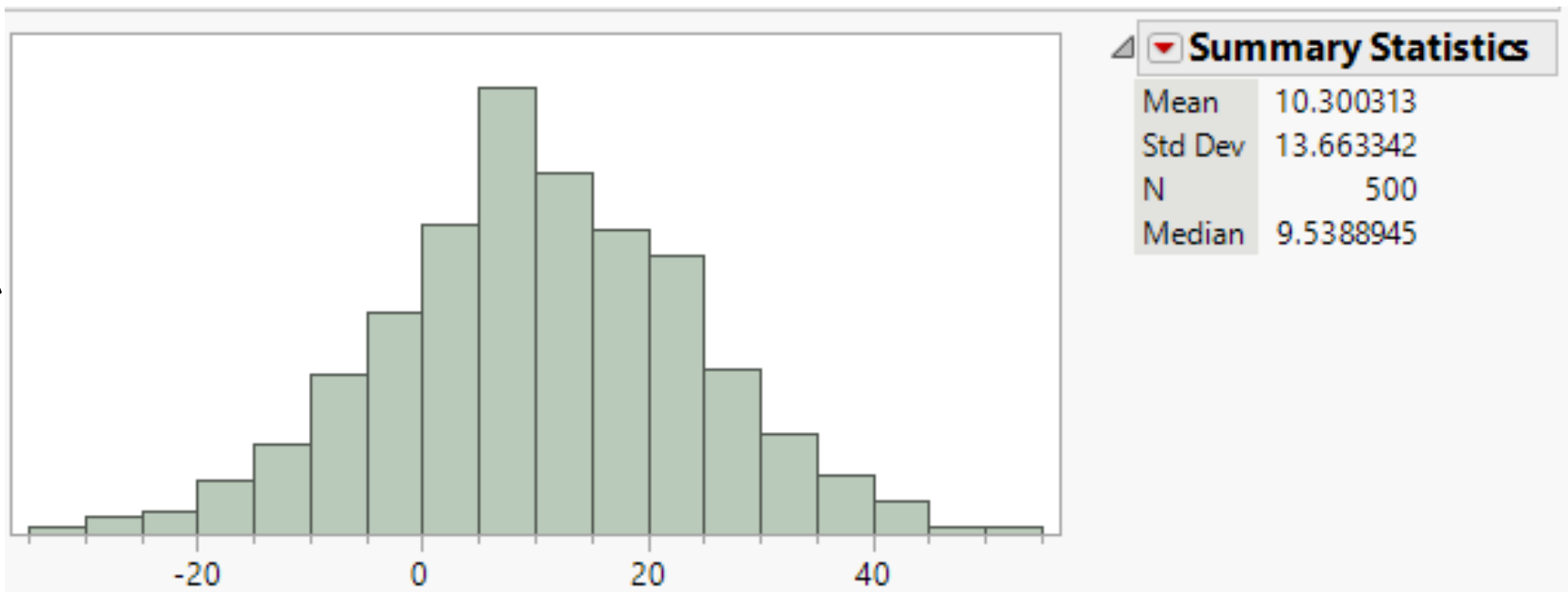
Data pattern

1)Center

2)Spread

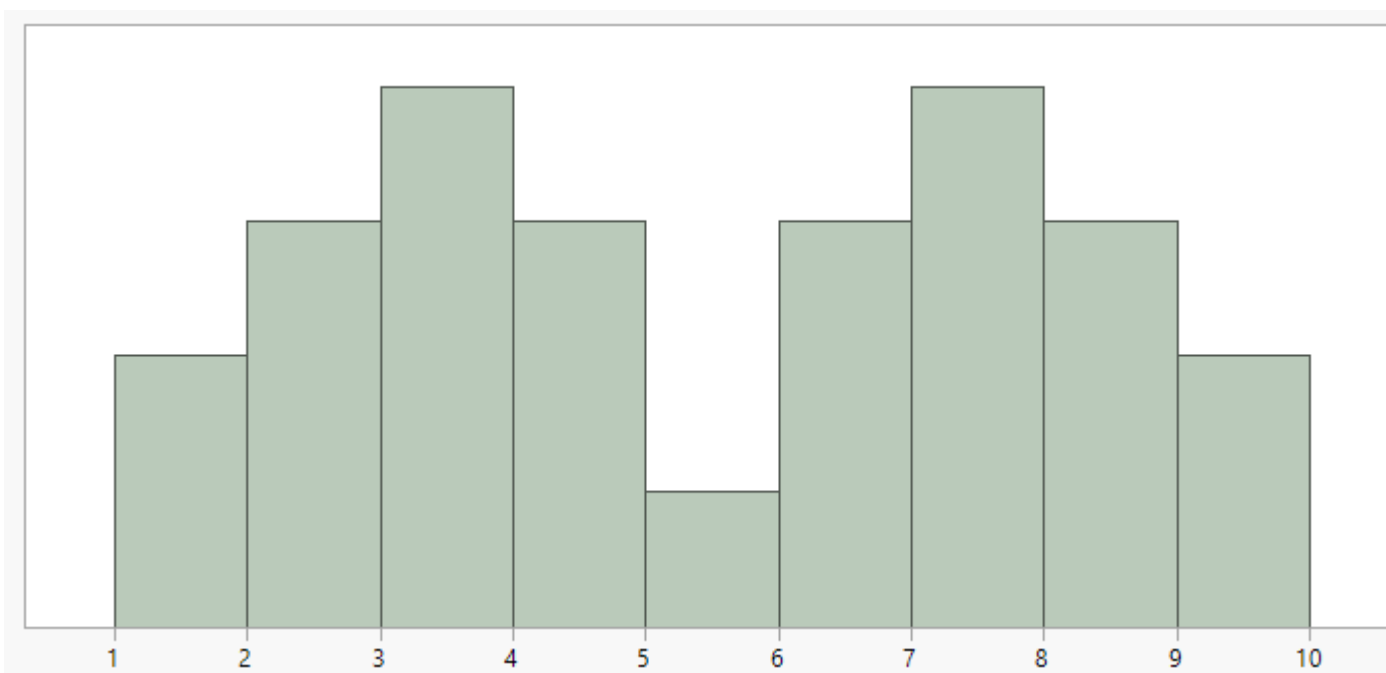
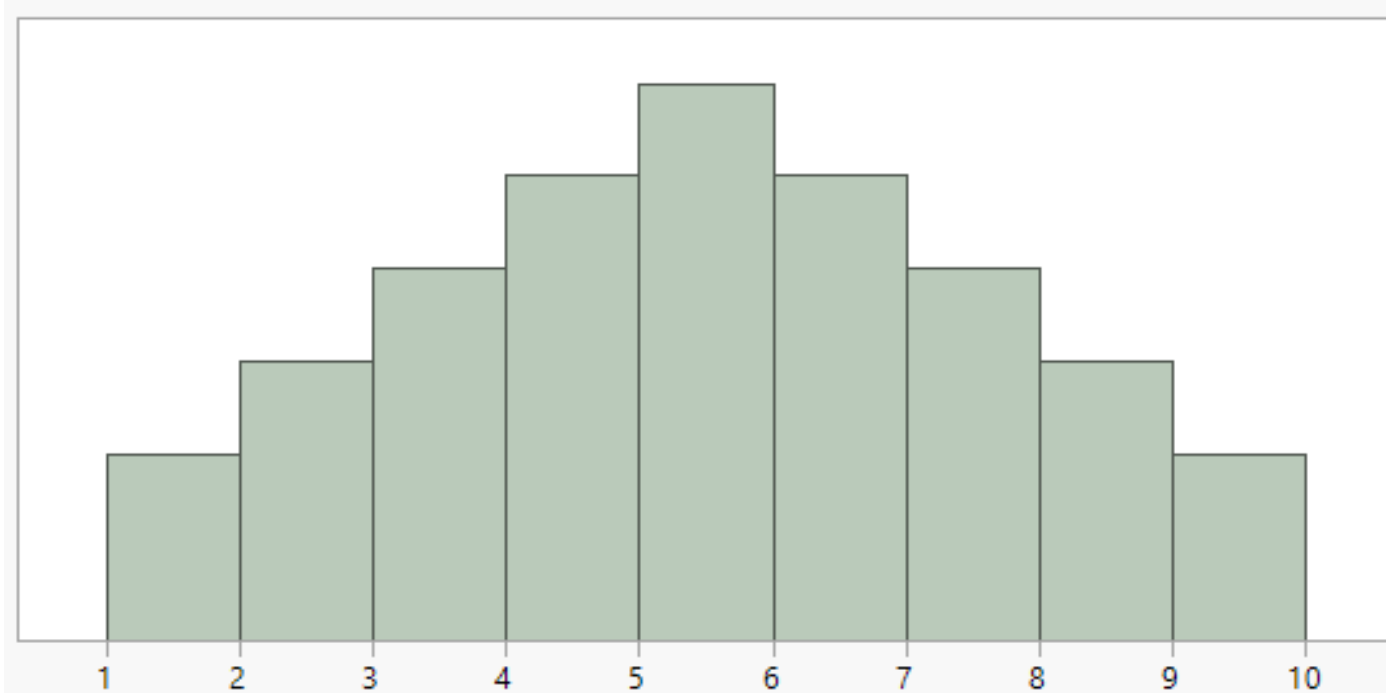
3)Shape

1) Center & Spread

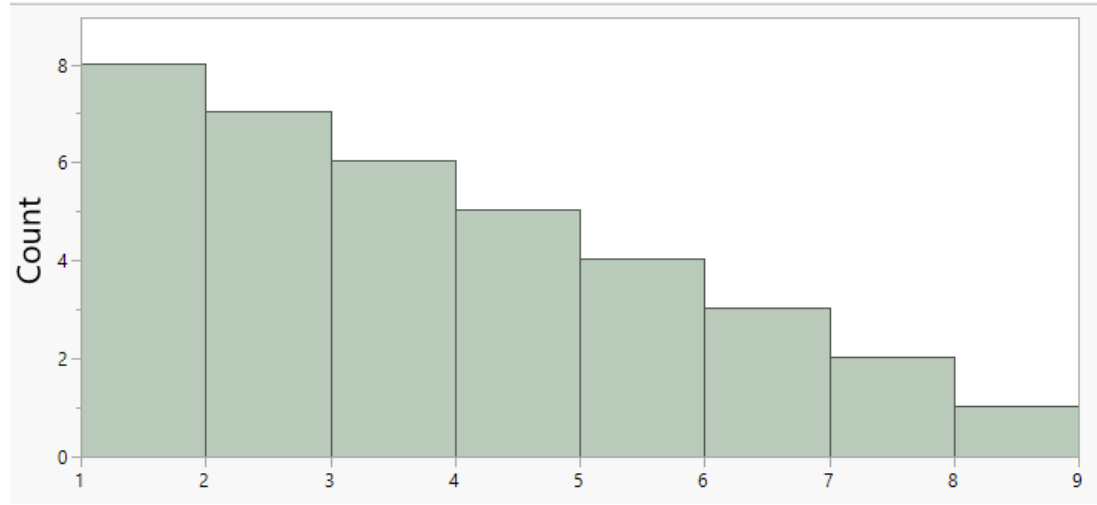


Shape

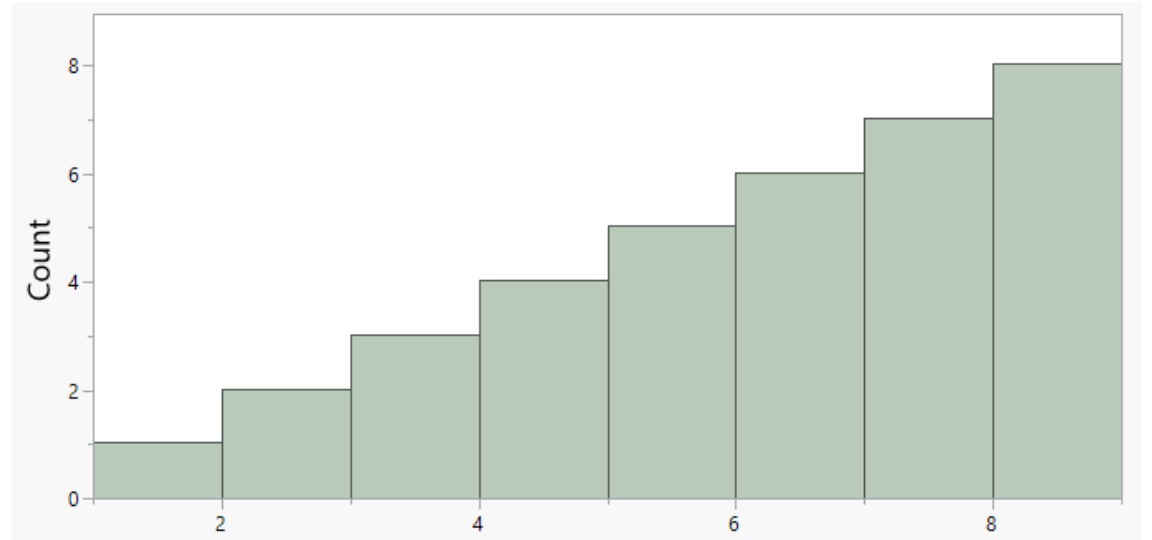
**Shape-
symmetry**



Shape- skewness

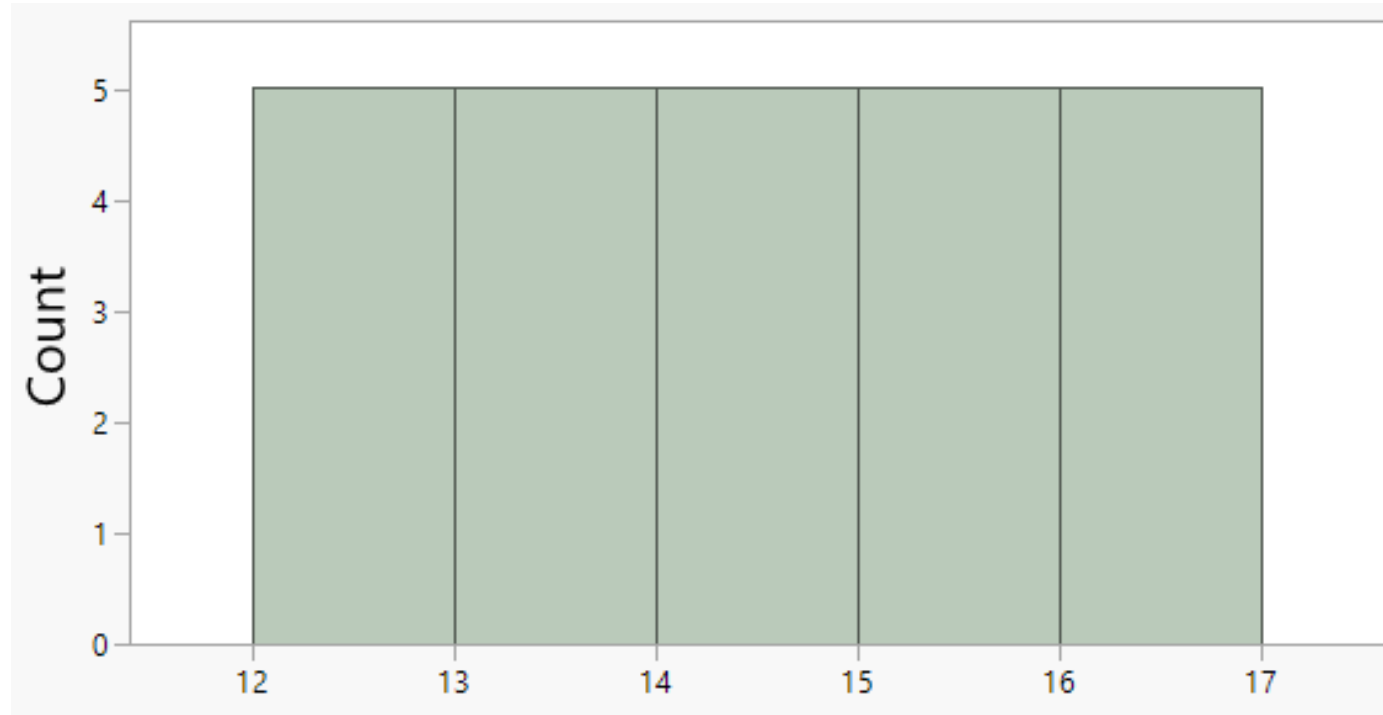


Points



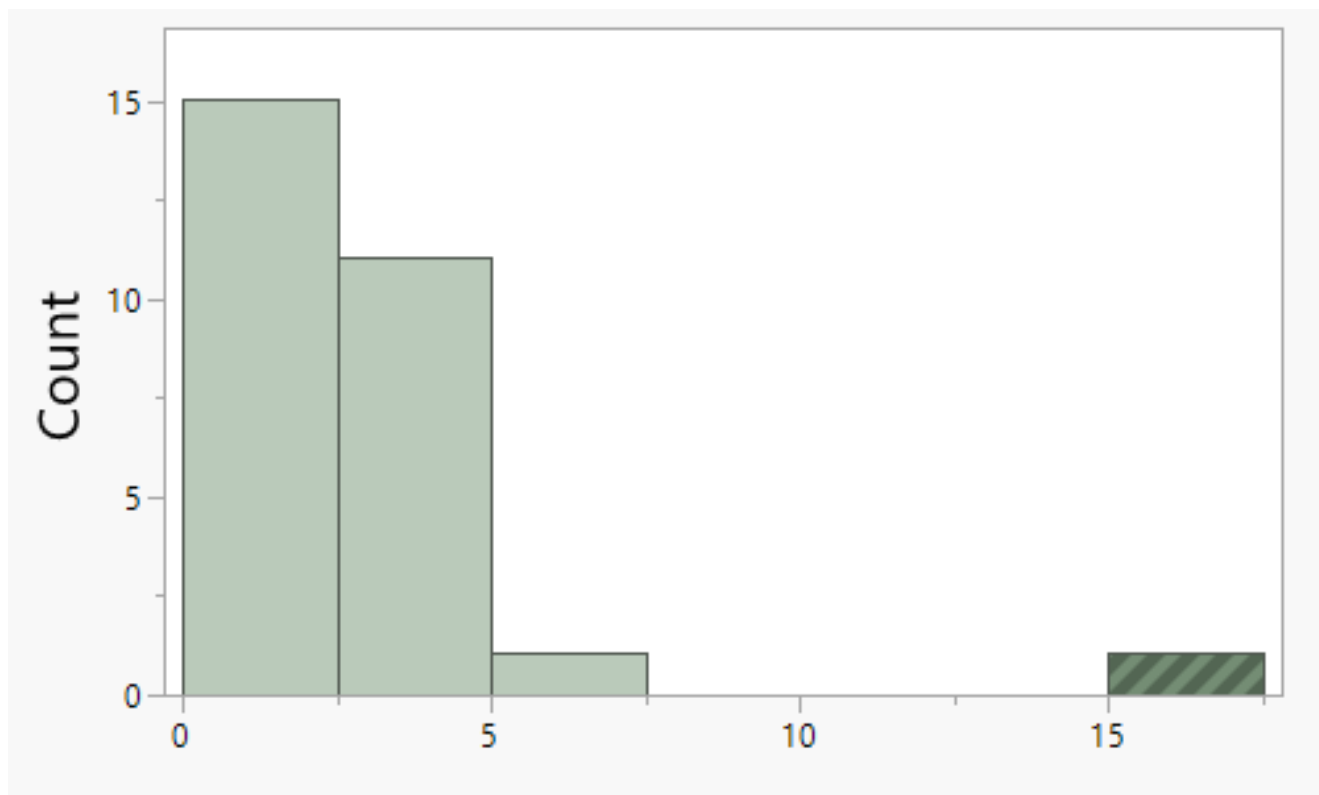
Points

**Shape -
Uniform**



Unusual features

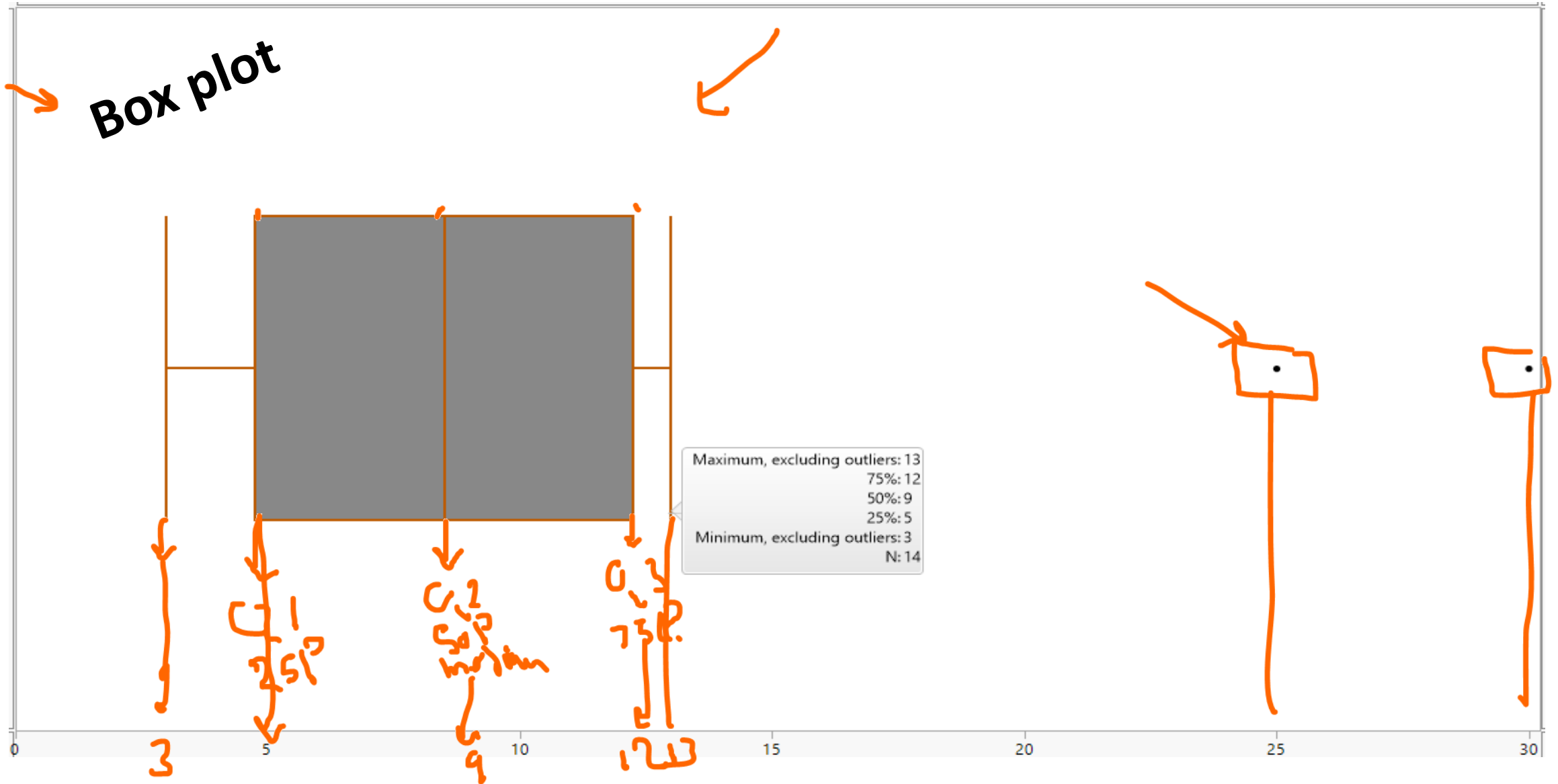
Gaps & Outliers



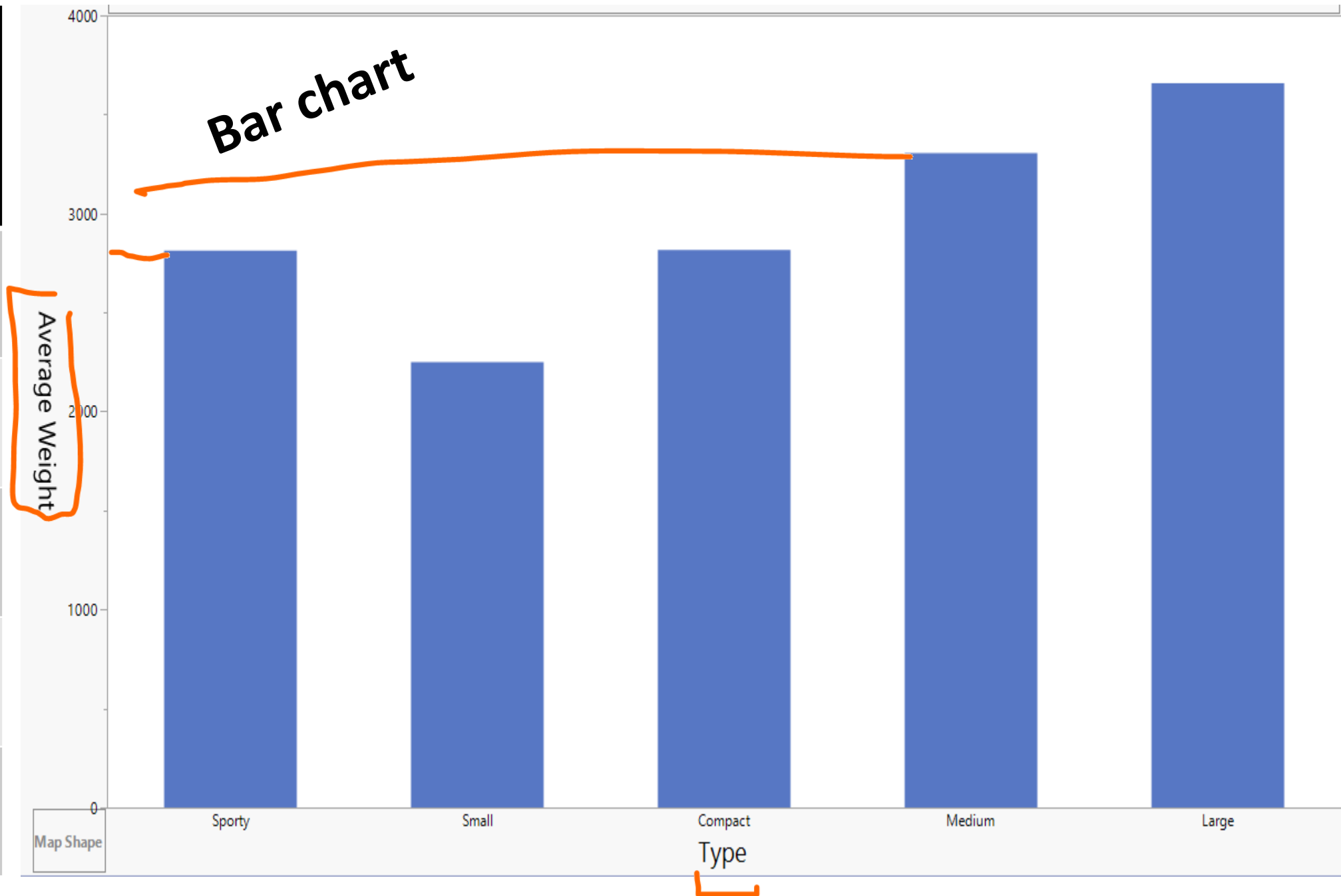
Graphs/Charts/Plots

6, 3, 10, 12, 4, 9, 6, 5, 8, 4, 13, 25, 30, 9

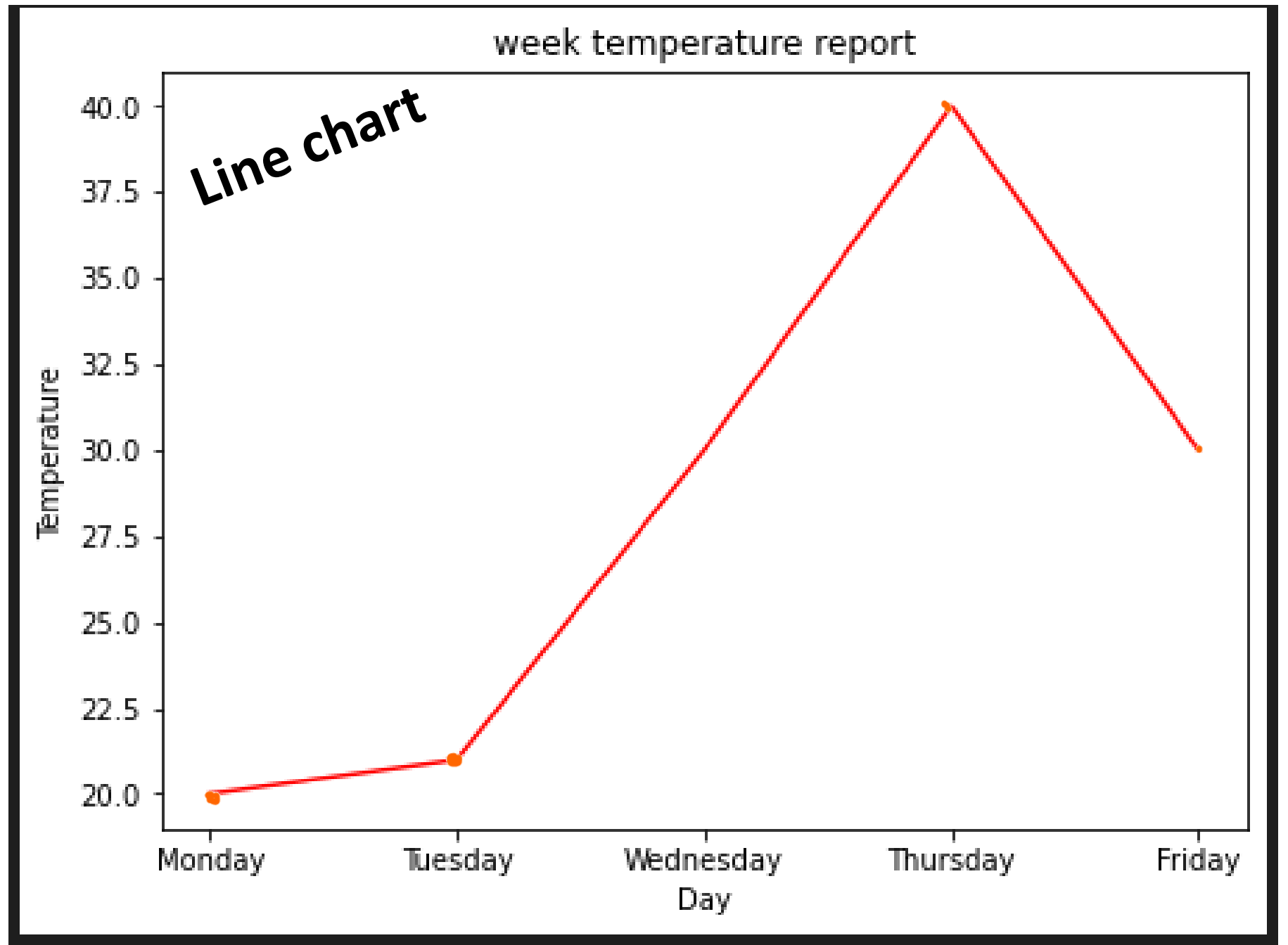
Box plot



| Car type | Avg weight(Kg) |
|----------|----------------|
| Sporty | 2813 |
| Small | 2250 |
| Compact | 2816 |
| Medium | 3305 |
| Large | 3657 |



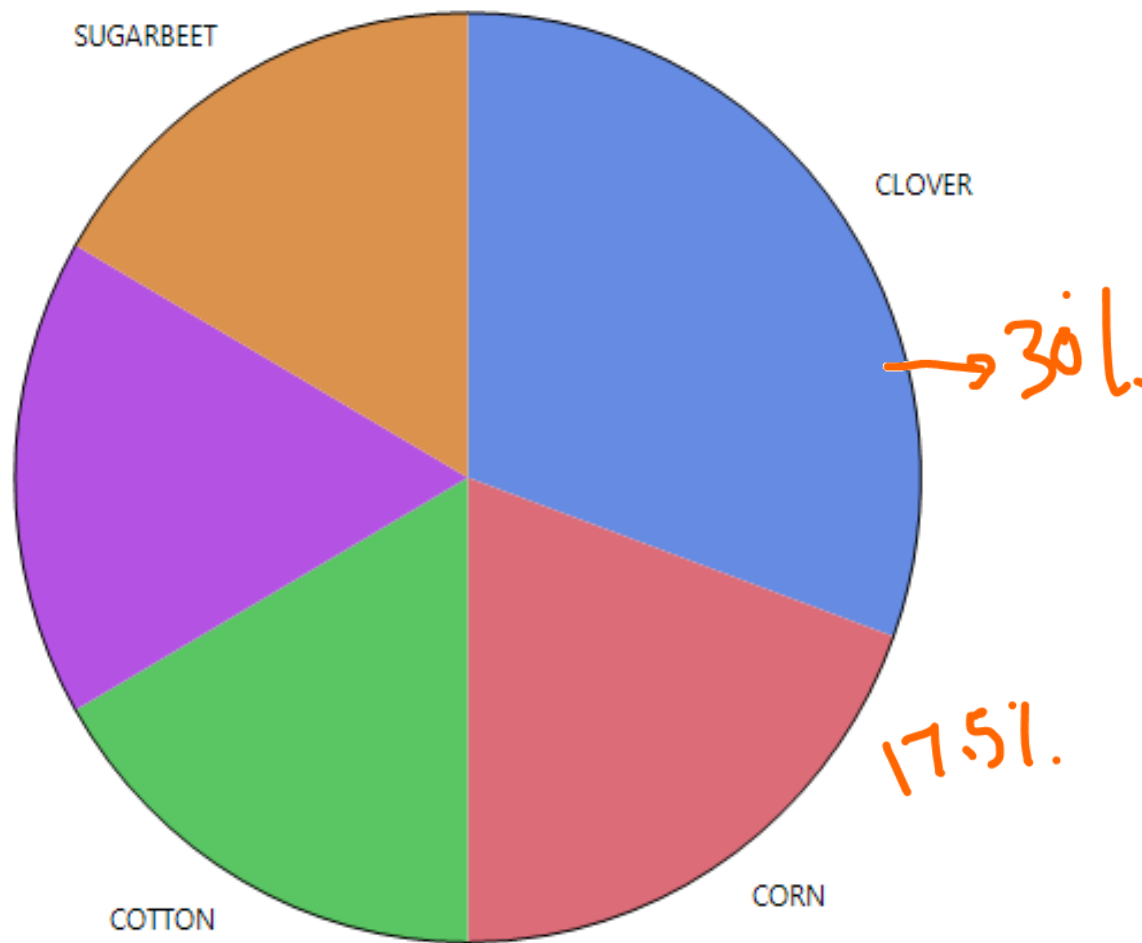
| Day | Temp |
|-----------|------|
| Monday | 20 |
| Tuesday | 21 |
| Wednesday | 30 |
| Thursday | 40 |
| Friday | 30 |



CLOVER
SOYBEAN
CORN
SUGARBEET
CLOVER
CLOVER
CLOVER
CORN
SOYBEAN
CORN
:
:
COTTON

pie chart

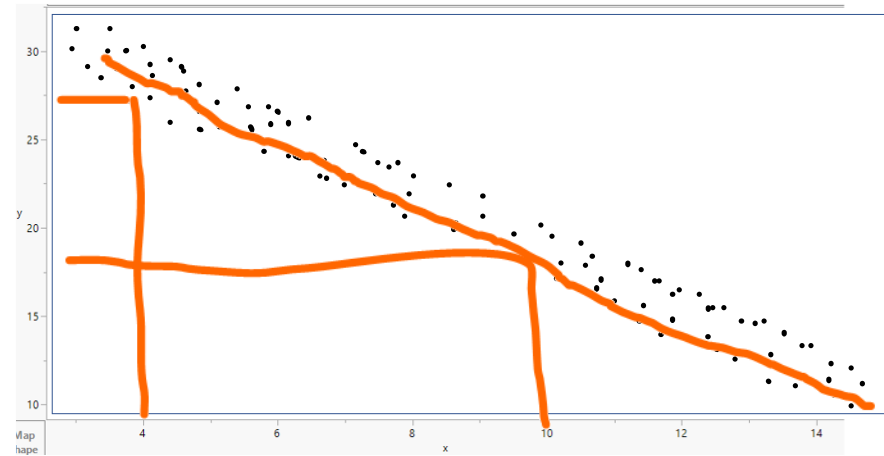
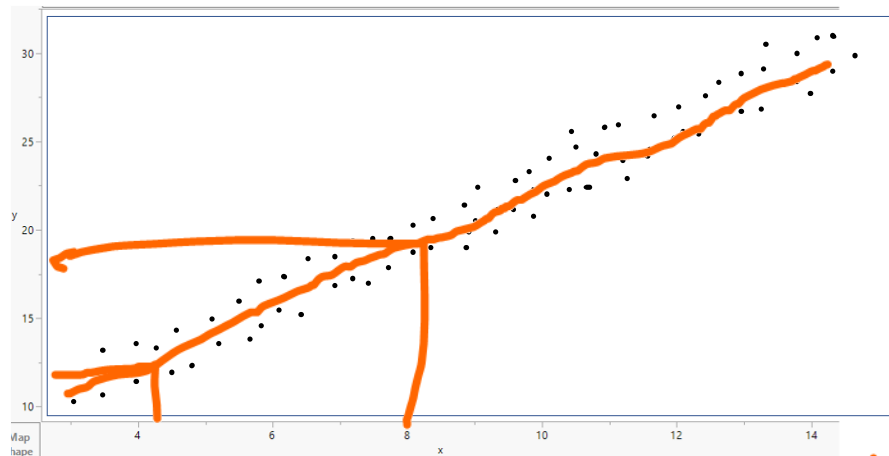
Clover 30%
CORN 17.5%
:
:
:



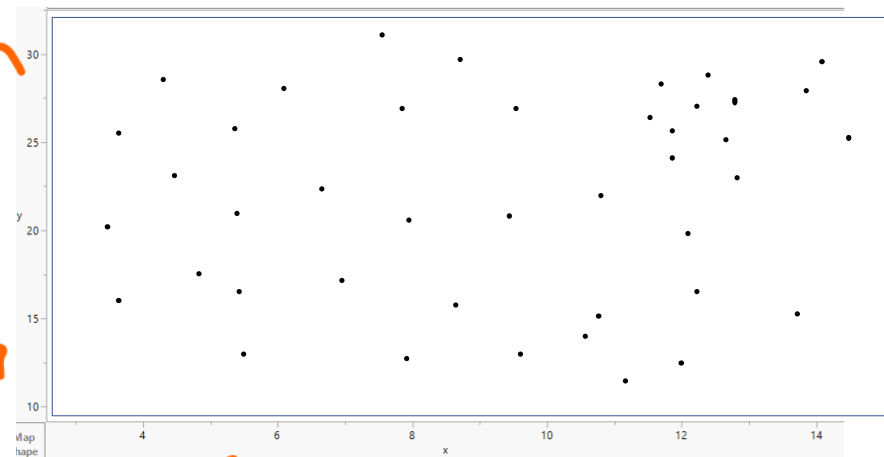
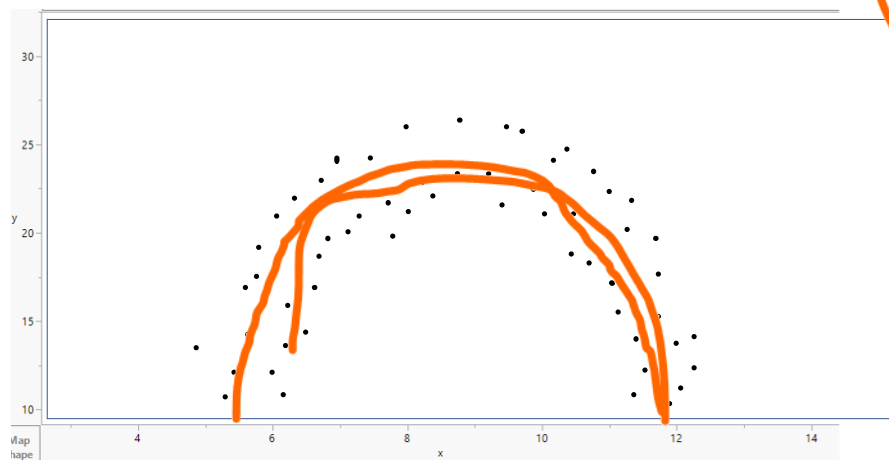
CROP

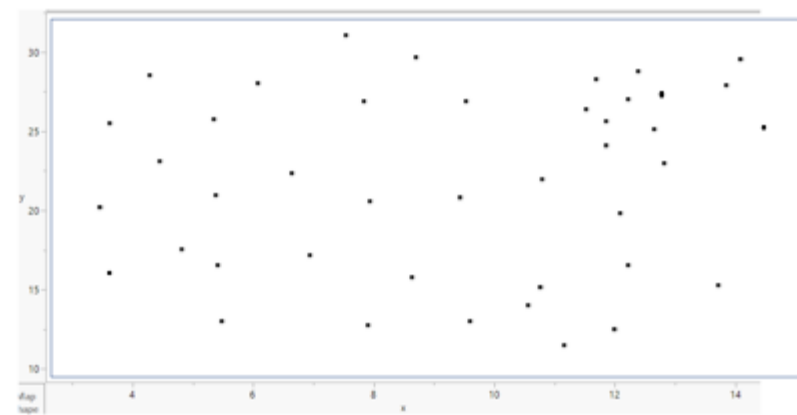
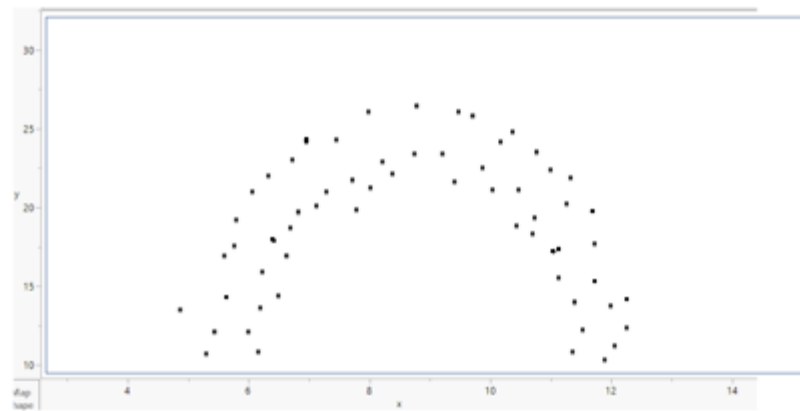
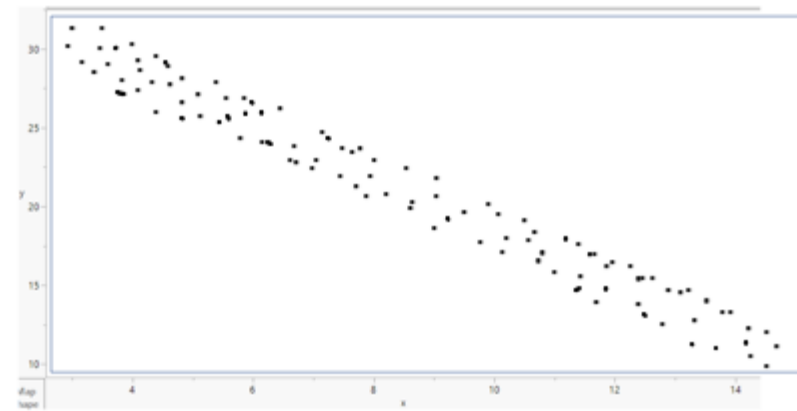
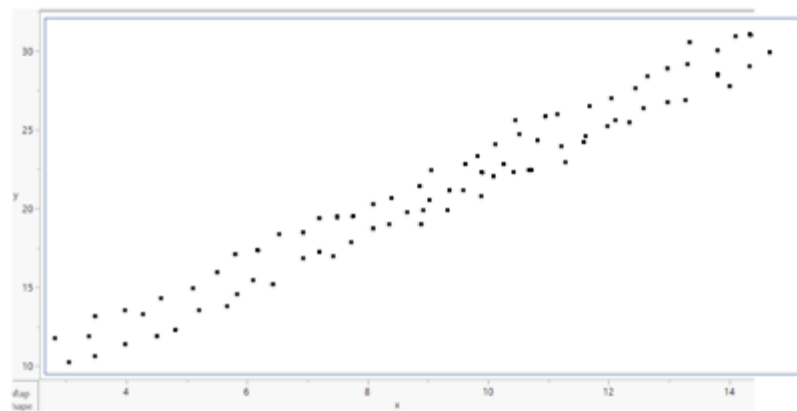
| X | Y |
|----|----|
| 1 | 54 |
| 5 | 22 |
| 7 | 40 |
| 9 | 12 |
| 15 | 29 |
| 21 | 11 |



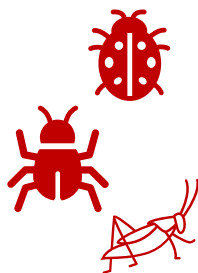
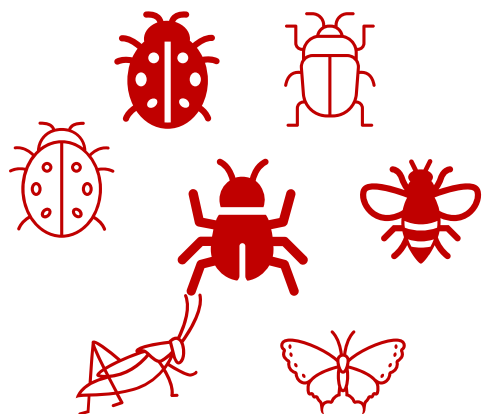
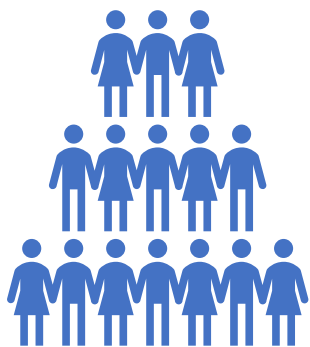


$y = mx + b$
 linearity
 slope
 Strength





Population vs Sample



Symbols

| | Population parameter | Sample statistics |
|--------------------|-----------------------------|--------------------------|
| Size | N | n |
| Mean | μ | \bar{x} |
| Variance | σ^2 | s^2 |
| Standard deviation | σ | s |

Equations

| | Population parameter | Sample statistics |
|--------------------|--|--|
| Size | N | n |
| Mean | $\mu = \frac{\sum_{i=1}^n x_i}{n}$ | $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ |
| Variance | $\sigma^2 = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N}$ | $S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$ |
| Standard deviation | $\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum_{i=1}^N (x_i - \mu)^2}{N}}$ | $S = \sqrt{S_{n-1}^2} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$ |