Quartiles and Percentiles

Quartiles and Percentiles are measures of variation, which describes how spread out the data is.
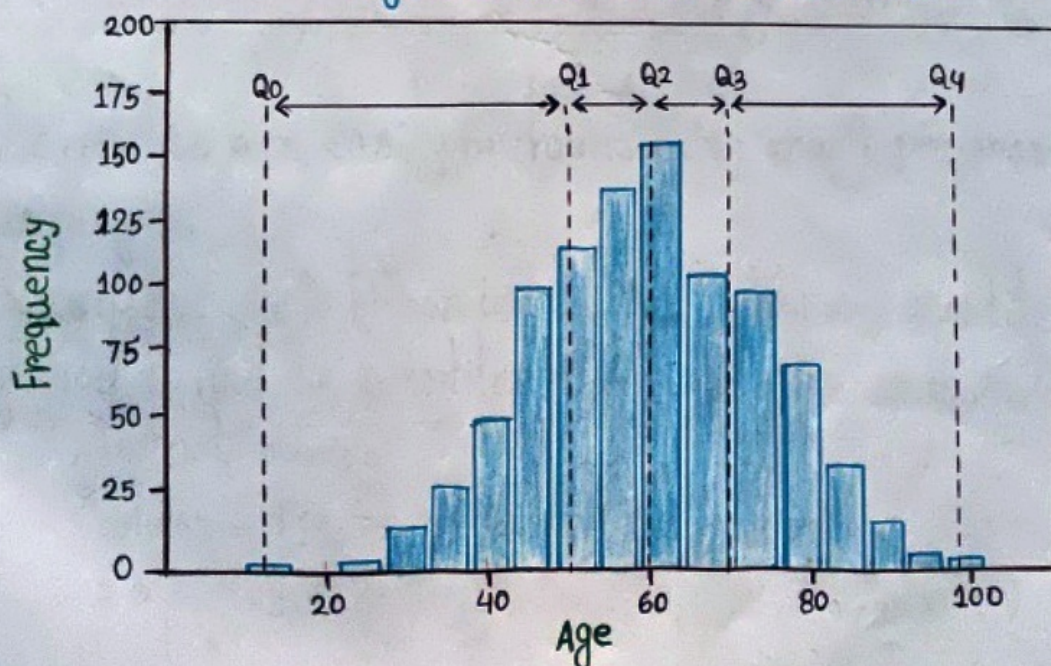
Quartiles and percentiles are both types of quantiles.

Quartiles

Quartiles are values that separate the data into four equal parts.

Here is a histogram of age of all 934 Bharat Ratna prize winners up to the year 2020, showing the quartiles: (It's only an example)

Age of Bharat Ratna winners

The quartiles (Q_0, Q_1, Q_2, Q_3, Q_4) are the values that separate each quarter.

Between Q_0 and Q_1 are the 25% lowest values in the data. Between Q_1 and Q_2 are the next 25%. And so on.

- Q_0 is the smallest value in the data.
- Q_1 is the value separating the first quater from second quater of the data.
- Q_2 is the middle value (median), separating the bottom from the top half.
- Q_3 is the value separating the third quater from the fourth quater.
- Q_4 is the largest value in the data.

Calculating Quartiles with Python

Quartiles can easily be found with many programming languages.

Example: With Python use the NumPy library quantile() method to find the quartiles of the values 13,21,21,40,42,48,55,72:

```
import numpy
Values = [13,21,21,40,42,48,55,72]
X = numpy.quantile (values,[0,0.25,0.5,0.75,1])
print (x)
```
Ans. [13. 21. 41. 49.75 72.]

Percentiles

Percentiles are values that separate the data into 100 equal parts. For example, The 95th percentile separates the lowest 95% of the values from the top 5%.

The 25th percentile ($P_{25\%}$) is the same as the first quartile (Q_1).

The 50th percentile ($P_{50\%}$) is the same as the second quartile (Q_2) and the median.

The 75th percentile ($P_{75\%}$) is the same as the third quartile (Q_3).

Example in Python

With Python use the NumPy library percentile() method to find the 65th percentile of the values 13, 21, 21, 40, 42, 48, 55, 72:

```python
import numpy
values = [13, 21, 21, 40, 42, 48, 55, 72]
X = numpy.percentile(values, 65)
print(x)
```

Ans. 45.3

Interquartile Range

Interquartile range is a measure of variation, which describes how spread out the data is.

Interquartile Range

Interquartile range is the difference between the first and third quartiles (Q1 and Q3).
The 'middle half' of the data is between the first and third quartile.
The first quartile is the value in the data that separates the bottom 25% of values from the top 75%.
The third quartile is the value in the data that separates the bottom 75% of the values from the top 25%.
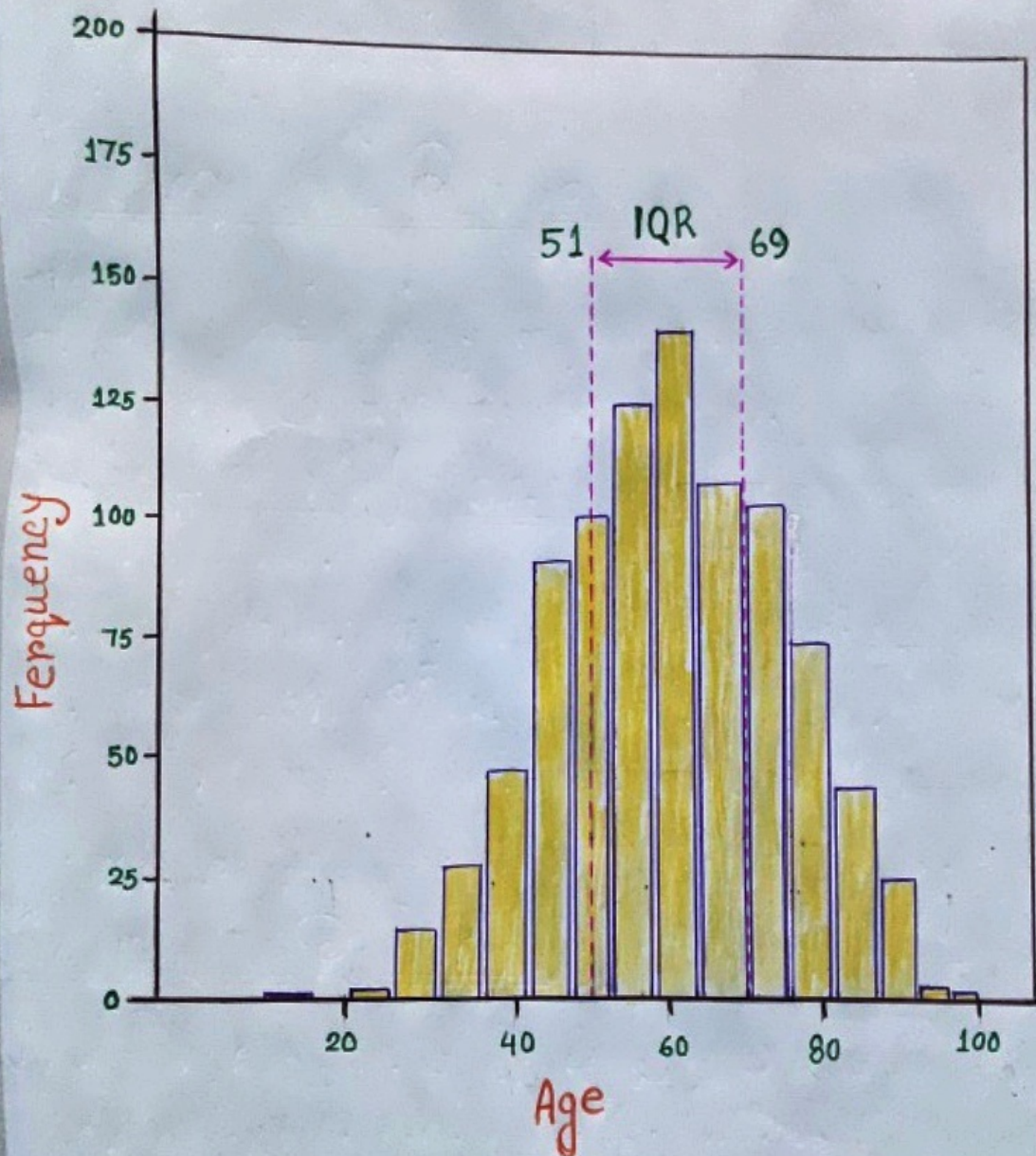
Example in Python

With Python use the SciPy library iqr() method to find the interquartile range of the values 13, 21, 21, 40, 42, 48, 55, 72:

```
from scipy import stats
values = [13,21,21,40,42,48,55,72]
x = stats.iqr(values)
print(x)
```

Ans: 28.75

Here is a histogram of the age of all 934 Bharat Ratna prize winners up to the year 2020, showing the interquartile range (IQR): (It's only an example)

Age of the Bharat Ratna Winners

Here, the middle half of is between 51 and 69 years. The interquartile range for Bharat Ratna prize winners is then 18 years.