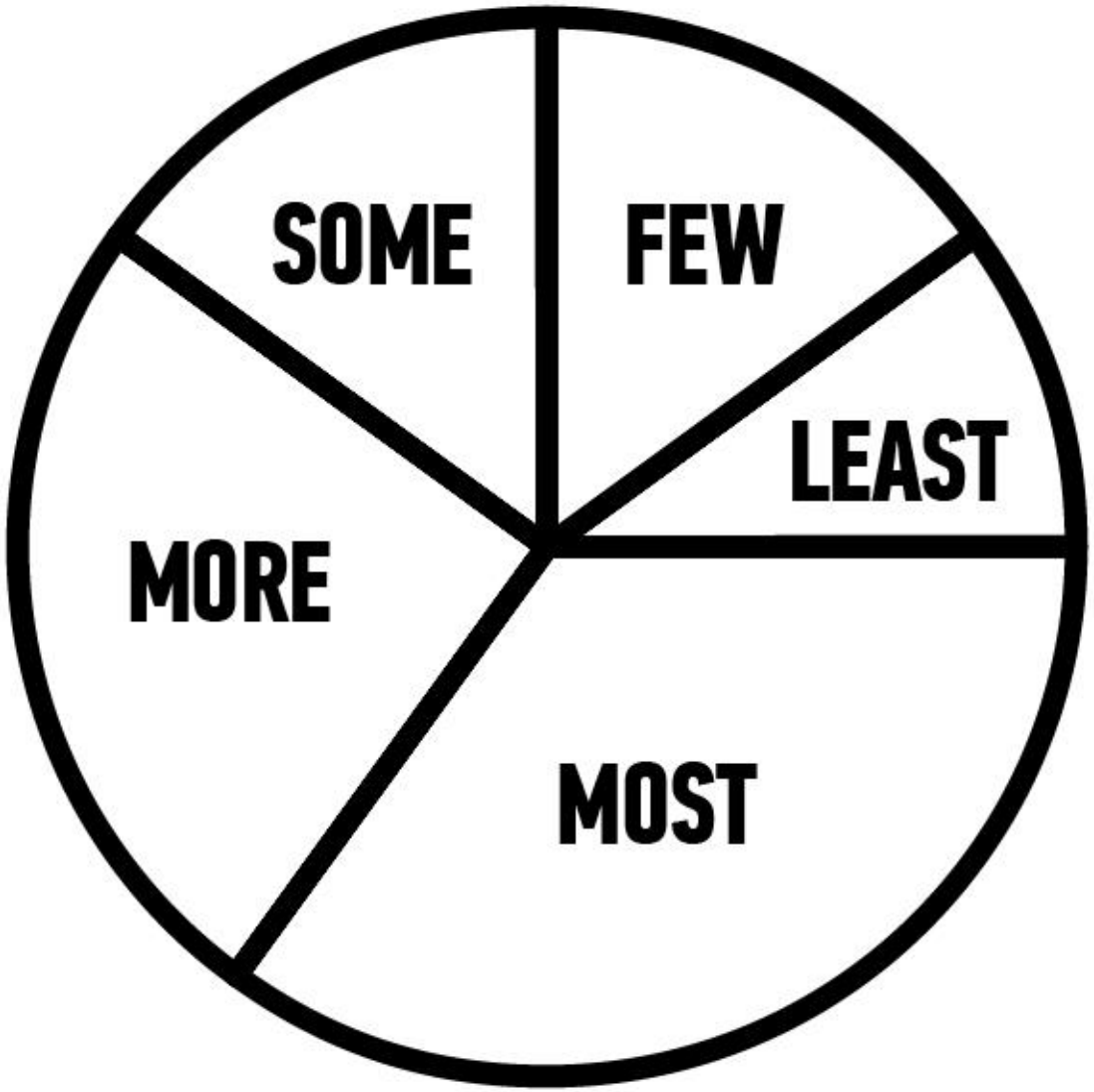


Lecture 7

Range, Quartiles and IQR

Range, Quartiles, and IQR

In this lesson:

1. Definition of a measure of spread
2. Definition of range, quartiles, and interquartile range
3. Finding range, quartiles, and IQR with data lists
4. Five – number summaries and interpretation
5. Finding range, quartiles, and IQR with frequency tables (grouped and ungrouped tables.
6. When to use five – number summaries and IQRs

Topic on summative exam

What you should be able to do:

1. Explain the following terms: measure of spread, range, quartiles, and interquartile range.
2. Calculate the range, quartiles, interquartile range, and five – number summaries with data lists
3. Calculate the range, quartiles, interquartile range, and five – number summaries with ungrouped and grouped frequency tables.
4. Determine when it is appropriate to use these measures of spread in analysis.

Definition: Measure of Spread

Definition | **Measure of Spread**

A **measure of spread** describes **how far away from the measure of center (mean/median)** entries in a data set are.

Why is it important?

Consider the following scenario between the scores of 2 different classes:

Two different classes in the same grade have the following scores. A score of **50% is needed to pass, 85% is needed to get a 4**. Try to decide which class has **more passing students and which class has more students with a 4.**

Class A:

10, 20, 30, 40, **50, 60, 70, 80, 85, 90, 100**

Mean = 57.72 Median = 60 **7** passing students

3 students with a 4

Class B:

55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65

Mean = 60 Median = 60 11 passing students

0 students with a 4

- If we looked only at the measures of center without looking at the data, we might think that **both classes have equal passing students** [medians are the same] **but class B had more students with a 4** [mean of B is larger].
- The truth, however, is completely different!
  - We were tricked because in **class A, the data values are all far away from each other,** but in **class B, the data values are all close to each other.**

It is important to not just measure what number describes the middle of the data, but to also measure how spread out the data is.

Definition: Range, Quartile, and Interquartile Range

Definition **Range**

The spread between the largest data entry and the smallest data entry.

Range = largest number – smallest number

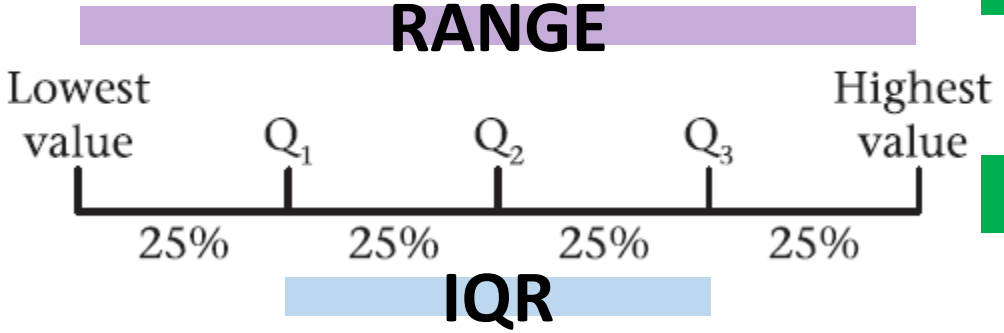
Definition **Quartiles**

The points at which the data set is split into **4 equal parts**.

Quartiles, Q_1, Q_2, Q_3, split the data into four parts.

Q_2 = median!

50% of data < Q_2



Q_1 = lower quartile

25% of data < Q_1

Q_3 = upper quartile

75% of data < Q_3

Definition **Interquartile Range [IQR]**

The data range that contains the **middle 50% of the entries.**

IQR = $Q_3 - Q_1$

Finding Range, Quartiles, and IQR in Data Lists

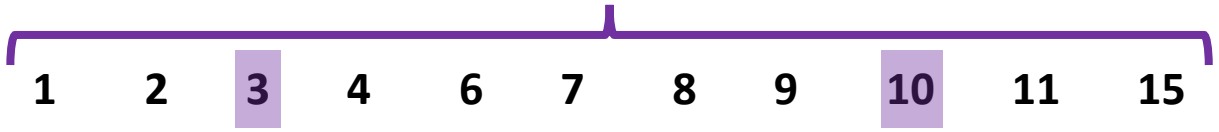
Finding quartiles is similar to finding the median!

Example 1

Find the range and interquartile range of the following data.

7 9 4 6 3 2 8 1 10 15 11

11 numbers total



Range = largest number – smallest number

Range = 15 – 1 = **14**

$$Q_1 = \frac{11}{4} = 2.75 \rightarrow 3$$

Q_1 is third observation → 3

$$Q_3 = \frac{3(11)}{4} = 8.25 \rightarrow 9$$

Q_3 is ninth observation → 10

IQR = $Q_3 - Q_1$

$$IQR = 10 - 3 = 7$$

Range = 14
Q_1 = 3
Q_3 = 10
IQR = 7

Step 1:
Reorganize the data set in order from **smallest to largest**.

Step 2:
To find the range, subtract the **largest number from the smallest.**

Step 3:
To find **Q_1**, divide the number of observations by 4:

- **If $\frac{n}{4}$ is whole**, find the average of that term and the term above.
- **If $\frac{n}{4}$ has a decimal**, then round up and choose that number.

Step 4:
To find **Q_3**, multiply the total number of observations by $\frac{3}{4}$ and repeat analysis from Step 2

Step 5:
To find **IQR**, subtract **Q_3** from **Q_1**

Practice

Find the **Range, Q1, Q3, and IQR** of our classes from the beginning of the lecture

Class A:

10, 20, 30, 40, **50, 60, 70, 80, 85, 90, 100**

Predict the answer before you look!

Range = 90
Q_1 = 30
Q_3 = 85
IQR = 55

Class B:

55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65

Range = 10
Q_1 = 57
Q_3 = 63
IQR = 6

Find the **Range, Q1, Q3, and IQR** of the following number set:

123, 146, 189, 198, 249, 267, 305, 336

Predict the answer before you look!

Range = 213
Q_1 = 167.5
Q_3 = 286
IQR = 118.5

Five Number Summaries and Their Interpretation

A five number summary is a quick and easy way to determine the spread of a data set. With these five numbers, you can determine if a distribution has a big spread, small spread, possible outliers, or none.

A five number summary consists of:

- The Minimum
- The First Quartile
- The Median
- The Third Quartile
- The Maximum

Example 1:

- Min: **1**
- Q_1: **2**
- Med: **3**
- Q_3: **4**
- Max: **5**

This distribution has no outliers and a very small spread

These examples are extreme, but they give a general idea of how to interpret a five – number summary.

Example 2:

- Min: **1**
- Q_1: **2**
- Med: **3**
- Q_3: **4**
- Max: **100**

This distribution has a large outlier and a small spread

Example 3:

- Min: **1**
- Q_1: **100**
- Med: **101**
- Q_3: **102**
- Max: **103**

This distribution has a small outlier and a small spread

Example 4:

- Min: **0**
- Q_1: **100**
- Med: **200**
- Q_3: **300**
- Max: **400**

This distribution has no outliers and a large spread

Range, IQR, and Quartiles with Frequency Tables

Finding the quartiles with frequency tables is very similar to finding the median with frequency tables. With ungrouped data, you find the right row and with grouped data you use **interpolation.**

Example 2 **Ungrouped Data:**

Rebecca records the number of CDs in the collections of students in her year.

x	Number of students, f
35	3
36	17
37	29
38	34
39	12

x	Number of students, f	Cumulative frequency
35	3	3
36	17	20
37	29	49
38	34	83
39	12	95

Range = 39 – 35 = 4

$$Q_1 = \frac{95}{4} = 23.75 \approx 24$$

$Q_1 = 37$

$$Q_3 = \frac{(3)95}{4} = 71.25 \approx 72$$

$Q_3 = 38$

$IQR = 38 - 37 = 1$

Range = 4
Q_1 = 37
Q_3 = 38
IQR = 1

Step 1:
Add cumulative frequency column

Step 2:
To find the range, subtract the **largest number from the smallest.**

Step 3:
To find **Q_1**, divide the number of observations by 4:

- **Round the number up if necessary.** Then find the column with that **observation** and choose the appropriate number.

Step 4:
To find **Q_3**, multiply the total number of observations by $\frac{3}{4}$ and repeat analysis from Step 2

Step 5:
To find **IQR**, subtract **Q_3** from **Q_1**

Range, IQR, and Quartiles with Frequency Tables

Finding the quartiles with frequency tables is very similar to finding the median with frequency tables. With ungrouped data, you find the right row and with grouped data you use **interpolation.**

Example 3 **Grouped Data:**

The length of time (to the nearest minute), spent on the internet each evening by a group of students is shown in the table below. Calculate the interquartile range.

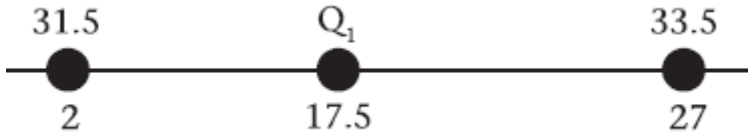
Length of time spent on internet (minutes)	Number of students	Cumulative frequency
30–31	2	2
32–33	25	27
34–36	30	57
37–39	13	70

Range = 39.5 – 29.5 = 10

$$IQR = 36.05 - 32.74 = 3.31$$

Range = 10
Q_1 = 32.74
Q_3 = 36.05
IQR = 3.31

$$Q_1 = \frac{70}{4} = 17.5$$



$$\frac{Q_1 - 31.5}{33.5 - 31.5} = \frac{17.5 - 2}{27 - 2}$$

$$\frac{Q_1 - 31.5}{2} = \frac{15.5}{25}$$

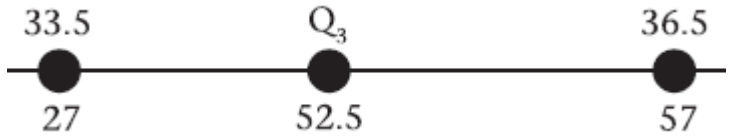
$$Q_1 - 31.5 = \frac{15.5}{25} \times 2$$

$$Q_1 = 31.5 + 1.24$$

$$= 32.74$$

$$Q_1 = 32.74$$

$$Q_3 = \frac{(3)70}{4} = 52.5$$



$$\frac{Q_3 - 33.5}{36.5 - 33.5} = \frac{52.5 - 27}{57 - 27}$$

$$\frac{Q_3 - 33.5}{3} = \frac{25.5}{30}$$

$$= 36.05$$

$$Q_3 = 36.05$$

Step 1:
Add cumulative frequency column

Step 2:
To find the range, subtract the **largest number from the smallest.**

Step 3:
To find **Q_1**, divide the number of observations by 4, then use **Interpolation** to find the answer

Step 4:
To find **Q_3**, multiply the total number of observations by $\frac{3}{4}$ and repeat analysis from Step 2

Step 5:
To find **IQR**, subtract **Q_3** from **Q_1**

When to Use Five – Number Summaries and IQR

Just like how **resistance** is an important point for choosing measures of center, it is also an important point in choosing measures of spread.

Measures of spread are attached to measures of center. Which measure of spread you use depends on which measure of center you choose.

Because the second quartile is also the **median**, and a five number summary uses the median as well, **you use quartiles as a measure of spread when you use the median as a measure of center.**

Quartiles are also the more resistant measure of spread, since they are calculated so similarly to the median.