UTAustinX: UT.7.10x Foundations of Data Analysis - Part 1

Center and Spread

SPEAKER: MICHAEL J. MAHOMETA, Ph.D.

When we numerically summarize a variable from our dataset in statistics,

we want to give our reader two main things.

We want to give them a measure of the variable's center,

0:00 / 6:04

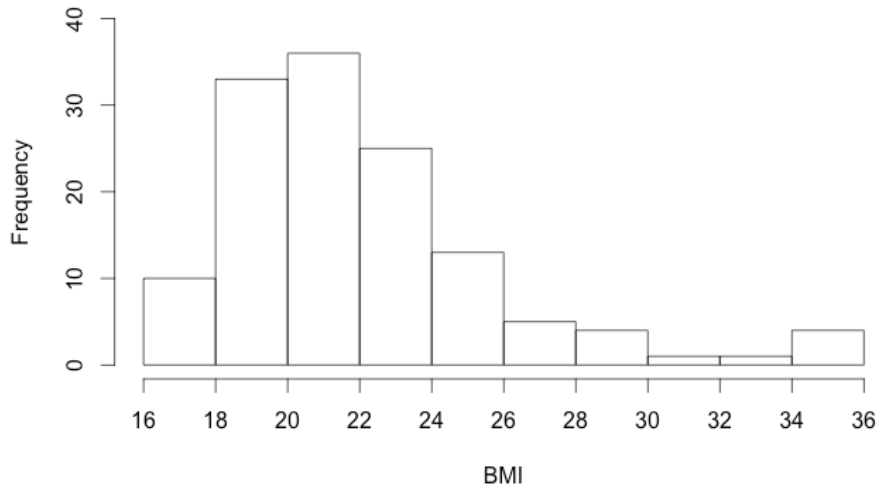
1.0x

Download transcript .srt

Comprehension Check

Below are data from a recent national survey on the BMI (body mass index) of a sample of Americans.

- Week 3: Bivariate Distributions
- Week 4: Bivariate Distributions (Categorical Data)



Identify the missing values in this portion of the frequency table that was used to generate the BMI histogram.

x	f
[16, 18)	(a)
[18, 20)	33
[20, 22)	36
[22, 24)	(b)
[24, 26)	13
[26, 28)	5

(6/6 points)

The value of (a) is:

10 ✔ **Answer:** 10

10

The value of (b) is:

25 ✔ **Answer:** 25

25

The shape of this distribution can best be described as:

- bimodal

- left skewed
- right skewed ✔
- uniform

A BMI of 30 is considered obese. Which of the following is the best description of the BMI of this sample?

- This sample demonstrates that a large percentage of Americans are obese.
- The majority of individuals are not obese, but there are some extreme values in the tail of the distribution. ✔
- The prevalence of obesity cannot be determined from this graph because BMI is an unreliable measure.
- There are no obese individuals in this sample.

When you are comparing two sets of data, and one is strongly skewed and the other is symmetric, which measure of center should you choose for the comparison?

- Mean
- Median ✔
- Mode
- The distributions cannot be compared

Which of the following is an accurate interpretation of the median of a distribution?

- About 50% of the population have the median value.

- The median is the numerical average of the dataset.
- The middle 50% of scores in a distribution constitute the median.
- Half the distribution has a lower score than the median, and half has a higher score than the median. ✔

Twelve coworkers log their hours worked overtime in the past month:

{10, 2, 6, 12, 14, 15, 15, 24, 15, 25, 3, 12}

(12/12 points)

Please attempt to answer the following questions by hand if you can:

What is the mean of the sample? *(Round to 2 decimal places.)*

12.75 ✔ **Answer:** 12.75

12.75

What is the standard deviation of the sample? *(Round to 2 decimal places.)*

7.12 ✔ **Answer:** 7.12

7.12

What is the 5 number summary for the sample?

min =

2 ✔ **Answer:** 2

2

Q1 =

8 ✔ **Answer:** 8

8

Q2 =

13 ✔ **Answer:** 13

13

Q3 =

15 ✔ **Answer:** 15

15

max =

25 ✔ **Answer:** 25

25

The mean number of overtime hours was *(Round to 1 decimal place.)*

12.8 ✔ **Answer:** 12.8

12.8

The median number of overtime hours was

13 ✔ **Answer:** 13

13

The range was

23 ✔ **Answer:** 23

23

The IQR was

7 ✔ **Answer:** 7

7

What is the shape of this distribution of scores?

- (•) approximately normal ✔
- ( ) left skewed

- right skewed
- uniform