Measures of Position

Objectives

1. Use z-scores to compare data values
2. Determine and interpret percentiles
3. Determine the five-number summary
4. Create a boxpolot of a dataset

z-scores

z-score

A **z-score** (a.k.a. **standard score**) measures how many standard deviations a data value, x, is from the mean of the data set.

z-scores

z-score

A **z-score** (a.k.a. **standard score**) measures how many standard deviations a data value, x, is from the mean of the data set.

- A positive z-score indicates an above average value.

z-scores

z-score

A **z-score** (a.k.a. **standard score**) measures how many standard deviations a data value, x, is from the mean of the data set.

- A positive z-score indicates an above average value.
- A negative z-score indicates a below average value.

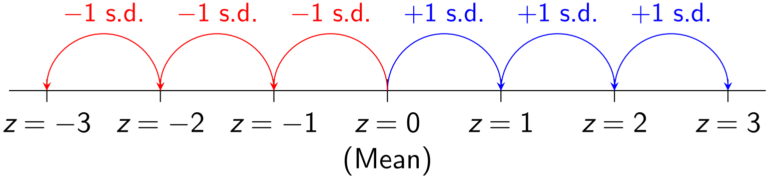
z-scores

z-score

A **z-score** (a.k.a. **standard score**) measures how many standard deviations a data value, x, is from the mean of the data set.

- A positive z-score indicates an above average value.
- A negative z-score indicates a below average value.
- A z-score of 0 indicates an exact average value.

Visual Interpretation of z-Scores



z-Score Formula

$$z = \frac{x - \mu}{\sigma}$$

z-Score Formula

$$z = \frac{x - \mu}{\sigma}$$

z-scores can compare two different data sets that use different scales of measurement.

z-Score Formula

$$z = \frac{x - \mu}{\sigma}$$

z-scores can compare two different data sets that use different scales of measurement.

"Usual" data values have z-scores between -2 and 2.

Example 1

The mean SAT score is 1059 with a standard deviation of 210; meanwhile the mean ACT score is 21 with a standard deviation of 5.4. A student takes both tests and receives a 1350 on the SAT and a 29 on the ACT. On which test did the student score better?

Example 1

The mean SAT score is 1059 with a standard deviation of 210; meanwhile the mean ACT score is 21 with a standard deviation of 5.4. A student takes both tests and receives a 1350 on the SAT and a 29 on the ACT. On which test did the student score better?

$$z_{\text{SAT}} = \frac{1350 - 1059}{210}$$

Example 1

The mean SAT score is 1059 with a standard deviation of 210; meanwhile the mean ACT score is 21 with a standard deviation of 5.4. A student takes both tests and receives a 1350 on the SAT and a 29 on the ACT. On which test did the student score better?

$$z_{\text{SAT}} = \frac{1350 - 1059}{210}$$

$$z_{\text{SAT}} = 1.39$$

Example 1

The mean SAT score is 1059 with a standard deviation of 210; meanwhile the mean ACT score is 21 with a standard deviation of 5.4. A student takes both tests and receives a 1350 on the SAT and a 29 on the ACT. On which test did the student score better?

$$z_{\text{SAT}} = \frac{1350 - 1059}{210} \qquad z_{\text{ACT}} = \frac{29 - 21}{5.4}$$

$$z_{\text{SAT}} = 1.39$$

Example 1

The mean SAT score is 1059 with a standard deviation of 210; meanwhile the mean ACT score is 21 with a standard deviation of 5.4. A student takes both tests and receives a 1350 on the SAT and a 29 on the ACT. On which test did the student score better?

$$z_{\text{SAT}} = \frac{1350 - 1059}{210} \qquad z_{\text{ACT}} = \frac{29 - 21}{5.4}$$

$$z_{\text{SAT}} = 1.39 \qquad z_{\text{ACT}} = 1.48$$

Example 1

The mean SAT score is 1059 with a standard deviation of 210; meanwhile the mean ACT score is 21 with a standard deviation of 5.4. A student takes both tests and receives a 1350 on the SAT and a 29 on the ACT. On which test did the student score better?

$$z_{\text{SAT}} = \frac{1350 - 1059}{210} \qquad z_{\text{ACT}} = \frac{29 - 21}{5.4}$$

$$z_{\text{SAT}} = 1.39 \qquad z_{\text{ACT}} = 1.48$$

The student did relatively better on the ACT.

Objectives

1. Use z-scores to compare data values
2. Determine and interpret percentiles
3. Determine the five-number summary
4. Create a boxpolot of a dataset

Percentiles

Percentile

A **percentile** is a measure of location that divide a set of data into 100 groups with about 1% of the values in each group.

Percentiles

Percentile

A **percentile** is a measure of location that divide a set of data into 100 groups with about 1% of the values in each group.

Percentile Score

A **percentile score** is the percent of data values less than a given value. (*Note*: this is <u>not</u> the same as percentage).

Percentiles

Percentile

A **percentile** is a measure of location that divide a set of data into 100 groups with about 1% of the values in each group.

Percentile Score

A **percentile score** is the percent of data values less than a given value. (*Note*: this is <u>not</u> the same as percentage).

Note: There is no universally agreed method to calculate percentiles.

Example 2

Explain the difference between getting 90% on a test and scoring in the 90th percentile on that test.

Example 2

Explain the difference between getting 90% on a test and scoring in the 90th percentile on that test.

Getting 90% means you earned 90% of the total points available for that test.

Example 2

Explain the difference between getting 90% on a test and scoring in the 90th percentile on that test.

Getting 90% means you earned 90% of the total points available for that test.

Scoring in the 90th percentile means that you did better than 90% of all other test takers.

Objectives

1. Use z-scores to compare data values
2. Determine and interpret percentiles
3. Determine the five-number summary
4. Create a boxpolot of a dataset

Quartiles

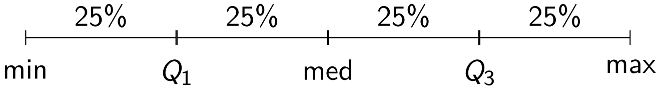
Quartiles

Quartiles are values that divide a data set into 4 groups, with each group holding 25% of the data.

Quartiles

Quartiles

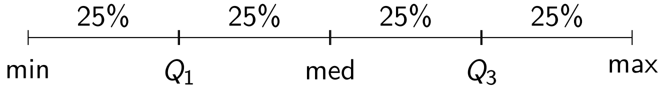
Quartiles are values that divide a data set into 4 groups, with each group holding 25% of the data.



Quartiles

Quartiles

Quartiles are values that divide a data set into 4 groups, with each group holding 25% of the data.



Q_1 is called the first (a.k.a. *lower*) quartile and Q_3 is called the third (a.k.a. *upper*) quartile.

Five-Number Summary

Five-Number Summary

The **five-number summary** are the following values:

Minimum, Q_1, Median, Q_3, Maximum

Example 3

Find the five-number summary of the following dataset:

1	2	2	4	5	7	11	15	15	18	44

Example 3

Find the five-number summary of the following dataset:

1	2	2	4	5	7	11	15	15	18	44

Minimum: 1

Example 3

Find the five-number summary of the following dataset:

1 2 2 4 5 7 11 15 15 18 44

Minimum: 1

Maximum: 44

Example 3

Find the five-number summary of the following dataset:

1 2 2 4 5 7 11 15 15 18 44

Minimum: 1

Median: 7

Maximum: 44

Example 3

Find the five-number summary of the following dataset:

1 2 2 | 4 5 7 11 15 15 18 44

3

Minimum: 1

First quartile: 3

Median: 7

Maximum: 44

Example 3

Find the five-number summary of the following dataset:

1 2 2 | 4 5 7 11 15 | 15 18 44

3 15

Minimum: 1

First quartile: 3

Median: 7

Third quartile: 15

Maximum: 44

Outliers

Outliers

An outlier is an extreme data value in a dataset.

Outliers

Outliers

An outlier is an extreme data value in a dataset.

We can use the five-number summary to detect outliers.

Outliers

Outliers

An outlier is an extreme data value in a dataset.

We can use the five-number summary to detect outliers.

Interquartile Range

The **interquartile range** can be found by subtracting Q_1 from Q_3:

$$\text{IQR} = Q_3 - Q_1$$

Lower and Upper Fences

The **lower fence** is

$$Q_1 - 1.5(\mathrm{IQR})$$

and the **upper fence** is

$$Q_3 + 1.5(\mathrm{IQR})$$

Lower and Upper Fences

The **lower fence** is

$$Q_1 - 1.5(\mathrm{IQR})$$

and the **upper fence** is

$$Q_3 + 1.5(\mathrm{IQR})$$

A data value is an **outlier** if it is less than the lower fence or more than the upper fence.

Example 4

Calculate the lower and upper fences of the previous example's dataset and use it to find any outliers.

$$1, 2, 2, 4, 5, 7, 11, 15, 15, 18, 44$$

Example 4

Calculate the lower and upper fences of the previous example's dataset and use it to find any outliers.

1, 2, 2, 4, 5, 7, 11, 15, 15, 18, 44

$$\text{Lower fence} = Q_1 - 1.5(IQR)$$

$$\text{Upper fence} = Q_3 + 1.5(IQR)$$

Example 4

Calculate the lower and upper fences of the previous example's dataset and use it to find any outliers.

1, 2, 2, 4, 5, 7, 11, 15, 15, 18, 44

$$\begin{aligned} \text{Lower fence} &= Q_1 - 1.5(IQR) \\ &= 3 - 1.5(15 - 3) \end{aligned}$$

$$\begin{aligned} \text{Upper fence} &= Q_3 + 1.5(IQR) \\ &= 15 + 1.5(12) \end{aligned}$$

Example 4

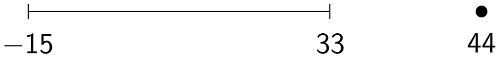
Calculate the lower and upper fences of the previous example's dataset and use it to find any outliers.

1, 2, 2, 4, 5, 7, 11, 15, 15, 18, 44

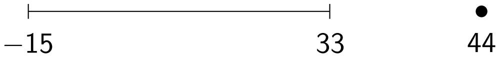
$$\begin{aligned}\text{Lower fence} &= Q_1 - 1.5(IQR)\\ &= 3 - 1.5(15 - 3)\\ &= -15\end{aligned}$$

$$\begin{aligned}\text{Upper fence} &= Q_3 + 1.5(IQR)\\ &= 15 + 1.5(12)\\ &= 33\end{aligned}$$

Example 4



Example 4



44 is an outlier of the dataset

Objectives

1. Use z-scores to compare data values
2. Determine and interpret percentiles
3. Determine the five-number summary
4. Create a boxpolot of a dataset

Boxplots

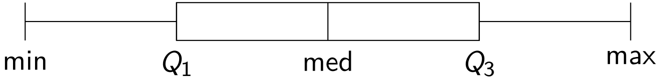
Boxplot

A **boxplot** is a visual display of the five-number summary.

Boxplots

Boxplot

A **boxplot** is a visual display of the five-number summary.



Example 5

Create a boxplot of the dataset

1, 2, 2, 4, 5, 7, 11, 15, 15, 18, 44

Example 5

Create a boxplot of the dataset

1, 2, 2, 4, 5, 7, 11, 15, 15, 18, 44

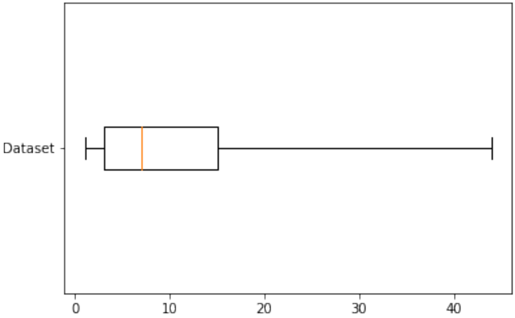
Min $= 1$ $\quad Q_1 = 3$ $\quad$ Med $= 7$ $\quad Q_3 = 15$ $\quad$ Max $= 44$

Example 5

Create a boxplot of the dataset

$$1,\ 2,\ 2,\ 4,\ 5,\ 7,\ 11,\ 15,\ 15,\ 18,\ 44$$

$$\text{Min} = 1 \quad Q_1 = 3 \quad \text{Med} = 7 \quad Q_3 = 15 \quad \text{Max} = 44$$



Modified Boxplot

With a modified boxplot, outliers are shown with symbols such as stars or points.

Modified Boxplot

With a modified boxplot, outliers are shown with symbols such as stars or points.

Whiskers are drawn out to the points that are **not** considered outliers.

Modified Boxplot

1, 2, 2, 4, 5, 7, 11, 15, 15, 18, 44