

Day 3 - Statistics

Agenda

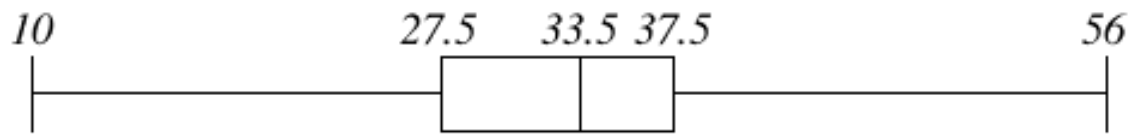
- How long is 30 seconds? Discussion
- How Random is the iPod shuffle?
 - Exploration of randomness

How long is 30 seconds?

- Potential Task:
 - Find the mean, median, and range for:
 - The entire data set
 - Only the men
 - Only the women
 - Create boxplots for all 3 data sets and graph them on the same axis (min, Q1, Median, Q3, max)
 - Median is the center data point (or average of middle two)
 - Q1 is the median of the data below the overall Median
 - Q3 is the median of the data above the overall Median

How long is 30 seconds: Boxplots

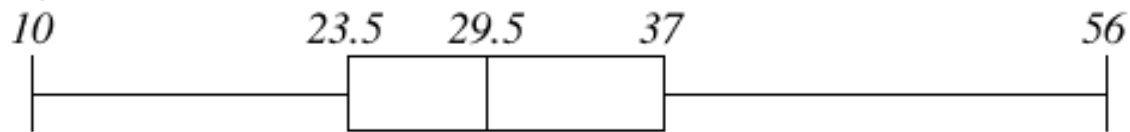
Overall



N = 78

Mean = 32.26

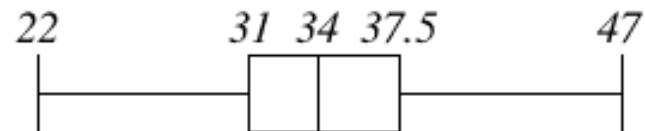
Women Only



N = 50

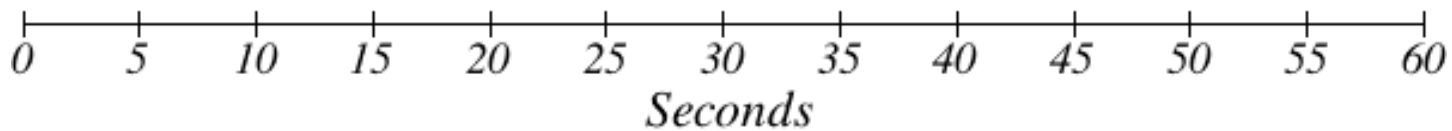
Mean = 30.69

Men Only



N = 28

Mean = 34.32



How long is 30 seconds?

- Discussion questions
 - How do the 3 data sets compare?
 - Does there appear to be a gender difference?
 - Are there any ‘outliers’? What would an outlier look like for each of the data sets?
 - What is the ‘spread’ of each of the data sets? How do they compare?

How long is 30 seconds?

- Discussion Questions about box plots (or box and whisker plots)?
 - Is it possible for a box plot to have no whiskers? If so, what does that mean about the data set?
 - Is it possible for a box plot to have no box? If so, what does that mean about the data set?

How long is 30 seconds?

- Discussion Questions about box plots (or box and whisker plots)?
 - A box plot uses the median as its measure of center (Q1 and Q3 are essentially medians as well). What would you expect to happen if we redefined a boxplot to be drawn with the min, 'lower-mean', mean, 'upper-mean' and max? (lower mean would be calculated by taking the average of only those observations below the overall average). How would this impact the way we interpret our data?
 - After comparing the traditional boxplot with the 'new' boxplot, which do you prefer? Why?

How random is the iPod Shuffle?

- The iPod shuffle feature takes the list of all the songs on an iPod, and rearranges them in a ‘random’ order. Each song will appear in the shuffled playlist only once.
- Many people like to use the shuffle feature, and many wonder:
 - Does the shuffle really play users' songs in random order?
 - Just what makes a playlist of songs “random”?

How random is the iPod Shuffle?

- Group Discussion Questions:
 - What comes to mind when you hear the word 'random'?
 - If the iPod shuffle feature is not producing a random sequence of songs, then what might the sequence of songs look like? What would you expect to see?
 - Do you think you can be 100% certain that a sequence of songs was not randomly generated?

How random is the iPod Shuffle?

- Albert Hoffman, an iPod owner, has written a letter to Apple to complain about the iPod shuffle feature. He writes that every day he takes an hour-long walk and listens to his iPod using the shuffle feature. He believes that the shuffle feature is producing playlists in which some artists are played too often and others are not played enough
- Your group has been asked to respond to Mr. Hoffman's complaint. You will be given several playlists of 20 songs each using the same songs as Mr. Hoffman's library but generating them using a genuine random number generation method.

How random is the iPod Shuffle?

- Group Activity
 - Step 1: Analyze the set of 25 truly random playlists (each with 20 songs) and Develop Rules for what characteristics random playlists appear to have
 - Step 2: Test Rules on the 5 additional truly random playlists, and then revise and adapt rules as necessary
 - Step 3: Evaluate whether or not you believe there is convincing evidence that Albert Hoffman's 3 playlists are not randomly generated
 - Step 4: Summarize your work

How random is the iPod Shuffle?

- Someone could claim that it could be said that Mr. Hoffman's playlists were randomly generated because 'anything is possible'.
- Discuss the difference between the following claims (is one a stronger claim than the other?):
 - Claim 1: the playlists 'may be randomly generated'
 - Claim 2: 'there is not evidence that the playlists are NOT randomly generated'

How random is the iPod Shuffle?

- Common Rules that arise:
 - Artists are usually not repeated more than 3 times in a row.
 - Artists usually do not appear more than 6 times in a playlist.
 - Playlists usually contain at least 7 of the 8 artists.
 - All artists are usually not represented proportionally.

How random is the iPod Shuffle?

- Discussion Questions
 - What made it difficult to come up with a rule to determine whether a sequence of data had been randomly generated? Explain.
 - How might your rules change if there weren't an equal number of songs for each artist? Or a longer set of songs per playlist? (Be specific about how your rules might change.)

How random is the iPod Shuffle?

- Extension Questions
 - What do you now know or believe that you did not know when beginning the activity?
 - What was the process for coming up with rules?
 - Can you prove that something is random?
 - What do you think you were supposed to learn from this activity? What was the purpose?

How random is the iPod Shuffle?

- Follow up Questions
- The Badger 5 is a cash lotto game in Wisconsin. To play, players pick five numbers from 1 to 31. Each number can only be used once. The numbers are listed in numerical order (not necessarily the order in which they were selected). A player wins the Badger 5 Jackpot if all five numbers chosen by that player match the five winning numbers chosen randomly by a computer.
- Here are 4 sets of numbers chosen for the Badger 5
 - 1 2 3 4 5
 - 1 13 25 27 31
 - 10 16 24 25 30
 - 5 10 15 20 25
- Is one of these sets of numbers more likely to win the Badger 5 than any of the others? How about less likely to win?
- Explain why or why not?
- How does the idea of random apply to this situation?

Discussion