Homework 2

Insert Name

Math 141, Week 2

Due: 11:59pm, Friday January 28

Instructions

Work through the problems below and submit this document as a knitted .pdf to the Math 141 S22 Wells Lecture gradescope page.

For each problem, put your solution between the bars of red stars.

Acknowledgements

If you work with a classmate, please write a note acknowledging this.

Exercise 1

The following graphic is entitled "A Brief Visual History of MARC Cataloging at the Library of Congress" and was created by Benjamin M. Schmidt so that he could better understand the history of how the Library created their digital card catalogs. The many interesting visual artifacts of the graph also help highlight the significant physical work that was involved in the conversion to digital catalogs.

You can view the image here.. Or if you are working on Reed RStudio server, you can run the following code:

    a. What are the variables displayed in this graphic? For each variable, specify if it is categorical or quantitative.

    b. What **geom** do the variables map to?

    c. What are the **aesthetic**s of the **geom**? For each **aesthetic**, give the variable that sets the values of that **aesthetic**.

    d. Is other context provided? If so, how does it help the viewer better understand the graphic's story?

    e. What does this graph do well?

_____
_____

f. How could this graph be improved?

_____
_____

Exercise 2

The second graphic comes from a 2021 blog post on FiveThirtyEight, titled "Where Climate Change Hits The Hardest". The graph explores the relationship between census region and perception of how climate change impacts communities.

You can view the image here.. Or if you are working on Reed RStudio server, you can run the following code:

a. What are the variables displayed in this graphic? For each variable, specify if it is categorical or quantitative.

_____
_____

b. What **geom** do the variables map to?

_____
_____

c. What are the **aesthetic**s of the **geom**? For each **aesthetic**, give the variable that sets the values of that **aesthetic**.

_____
_____

d. Is other context provided? If so, how does it help the viewer better understand the graphic's story?

_____
_____

e. What does this graph do well?

_____
_____

f. How could this graph be improved?

_____
_____

Exercise 3

As discussed in class, the creator of a graph has many editorial choices that can change the story the graph tells. For example, compare the two graphs from the 2012 nytimes.com article "One Report, Diverging Perspectives". Both tell a story related to unemployment, but the first is from "How a Democrat might see things" and the second from "How a Republican might see things".

You can view the first graph here and the second graph here, or if you are using the RStudio Server, you can run the following code:

a. What story does the first graph tell? What story does the second graph tell?

b. Which graph is told from a Democratic perspective and which from a Republican perspective? Explain your reasoning.

b. Are these graphs neutral? Are they factual? Justify your answer.

Exercise 4

In this problem, we will explore the distribution of finishing times for male and female winners of the New York Marathon between 1970 and 1999. Run the following code to load the `nyc_marathon` data set. *The first time you run the code, the Environment pane will display the words **promise**. Click on the promise text to view the data.*

Consider the graphics created by the following code:
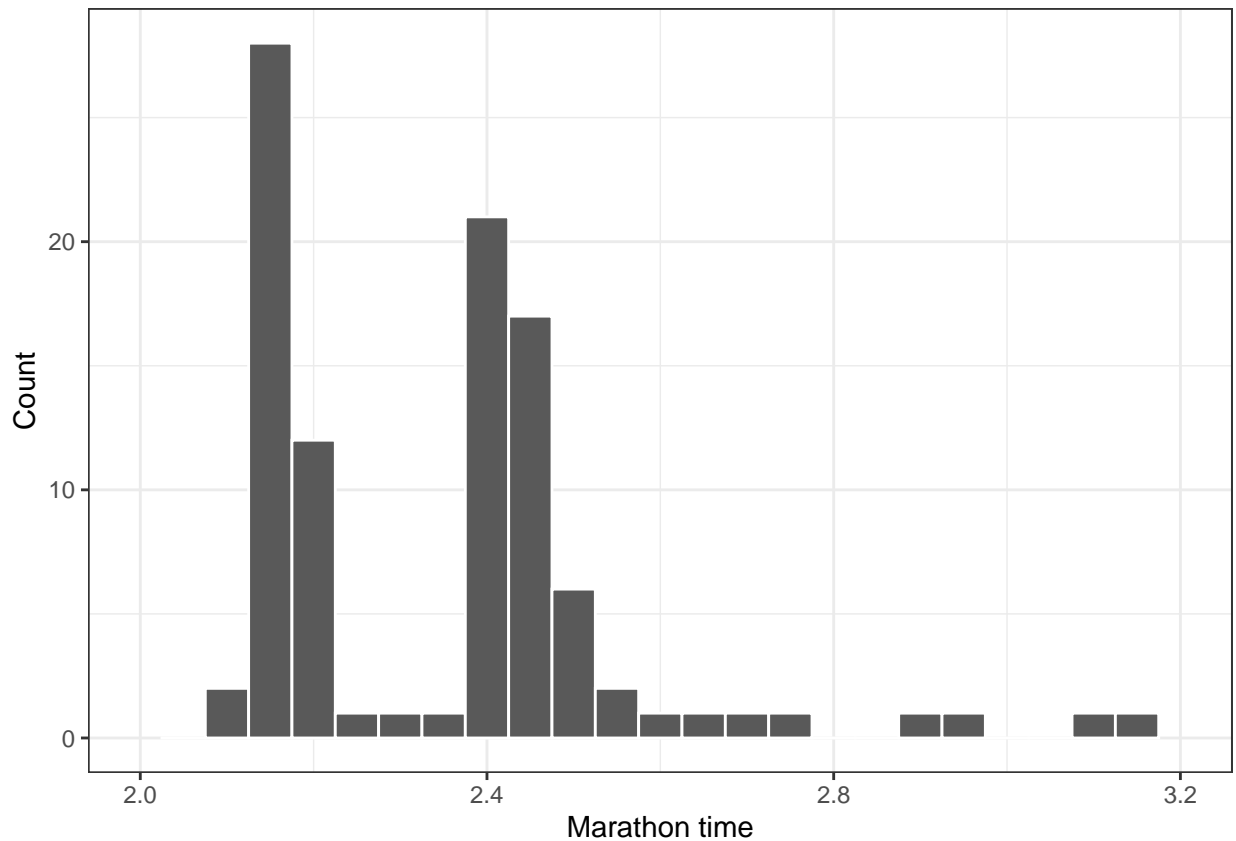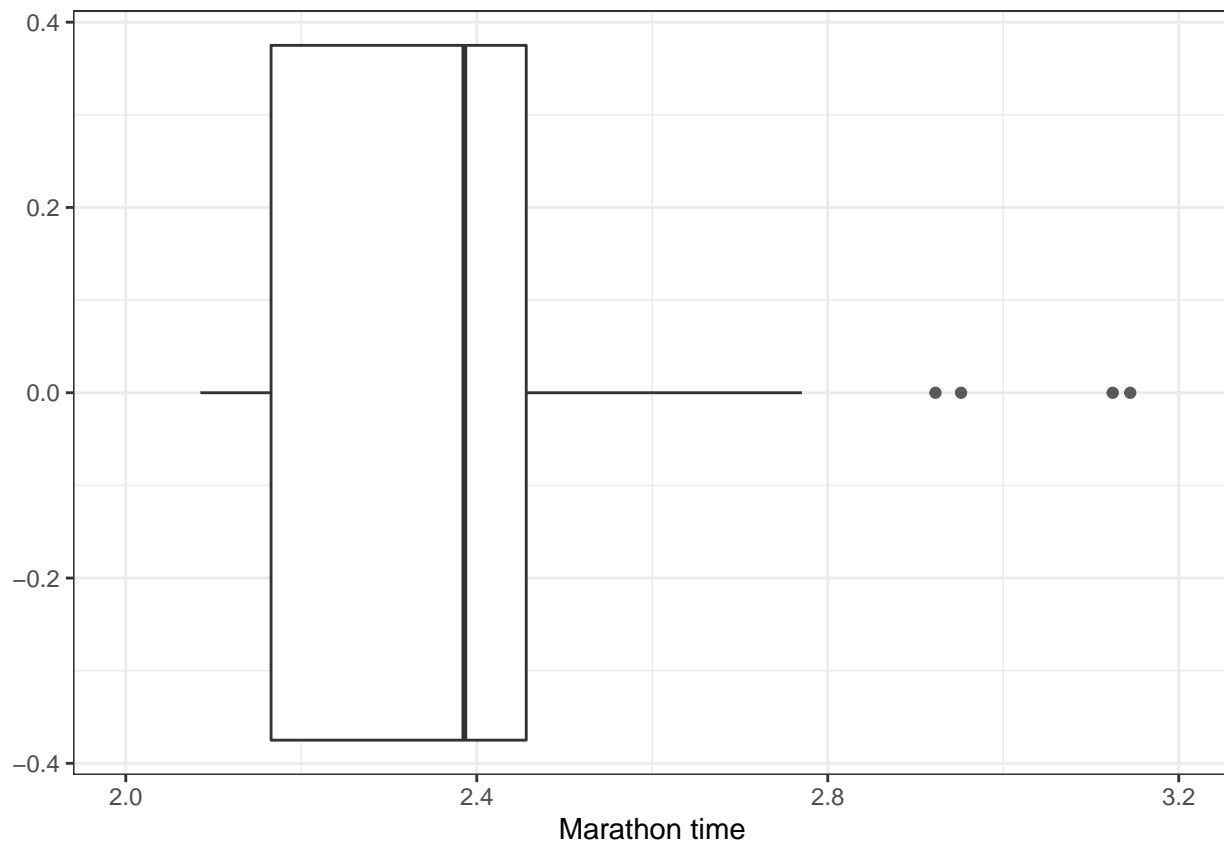
a. What features of the distribution are apparent in the histogram and not the box plot? What features are apparent in the box plot but not in the histogram?
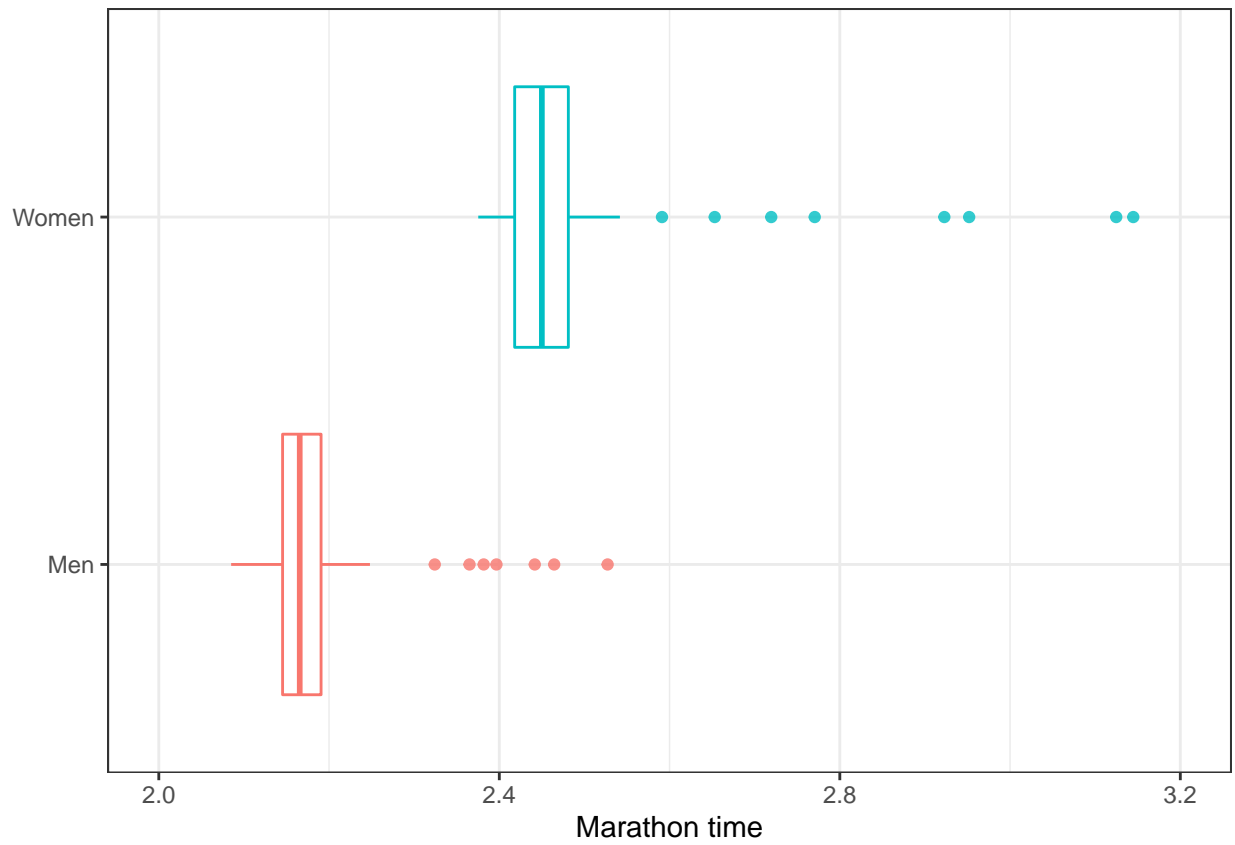
_____

_____

b. What may be the reason for the bimodal distribution? Explain.

_____

_____

c. Compare the distribution of marathon times for men and women based on the box plot shown below.

d. The time series plot shown below is another way to look at these data. Describe what is visible in this plot but not in the others.