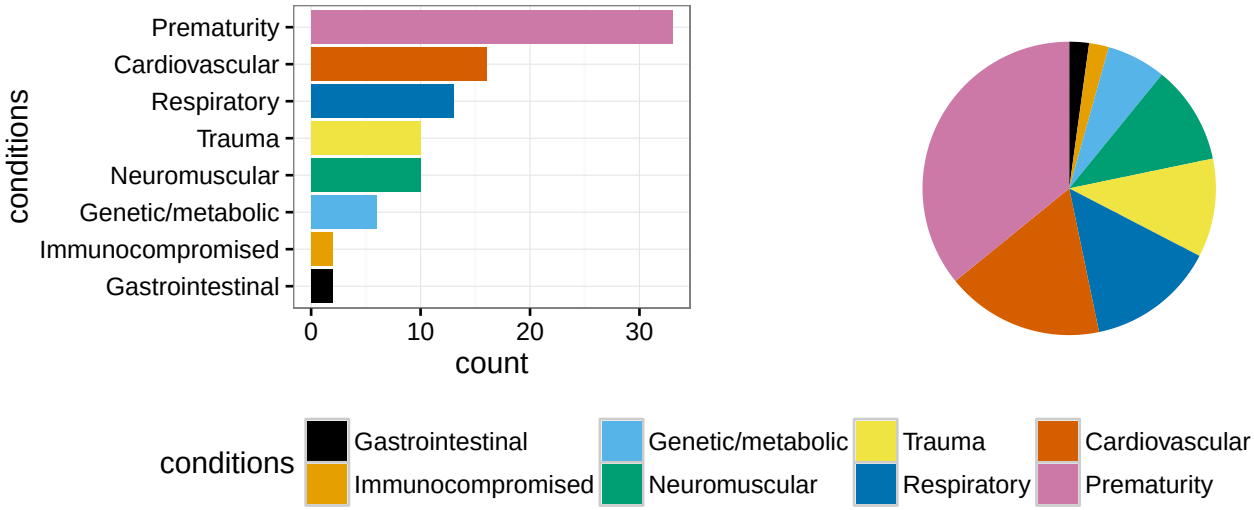
Homework 2

Fall 2016, Math 107, Prof. Adam Loy

Due Tuesday, September 20 by 4:00 p.m.

Problem 1

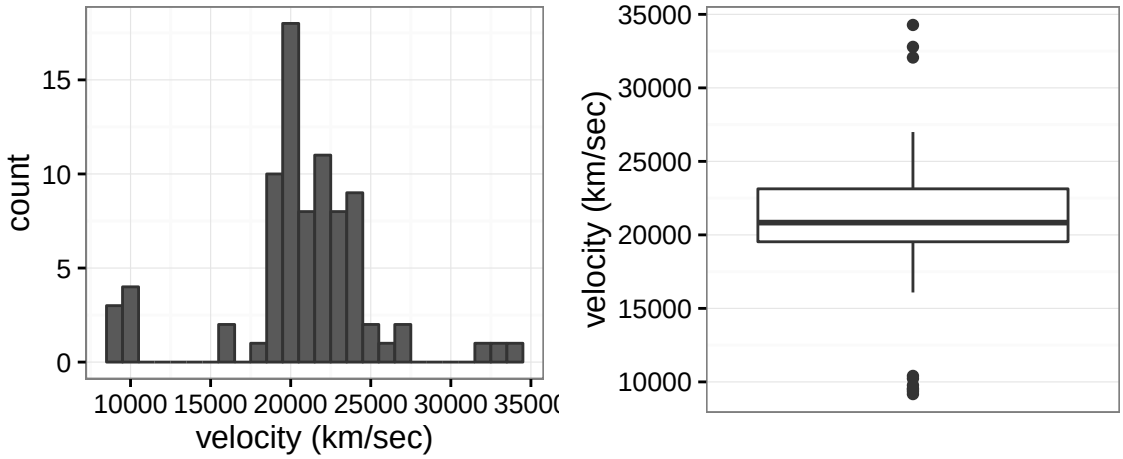
The bar chart and the pie chart below show the distribution of pre-existing medical conditions of children involved in a study on the optimal duration of antibiotic use in treatment of tracheitis, which is an upper respiratory infection.



a. What features are apparent in the bar chart but not in the pie chart?
b. What features are apparent in the pie chart but not in the bar chart?
c. Which graph would you prefer to use for displaying these categorical data?

Problem 2

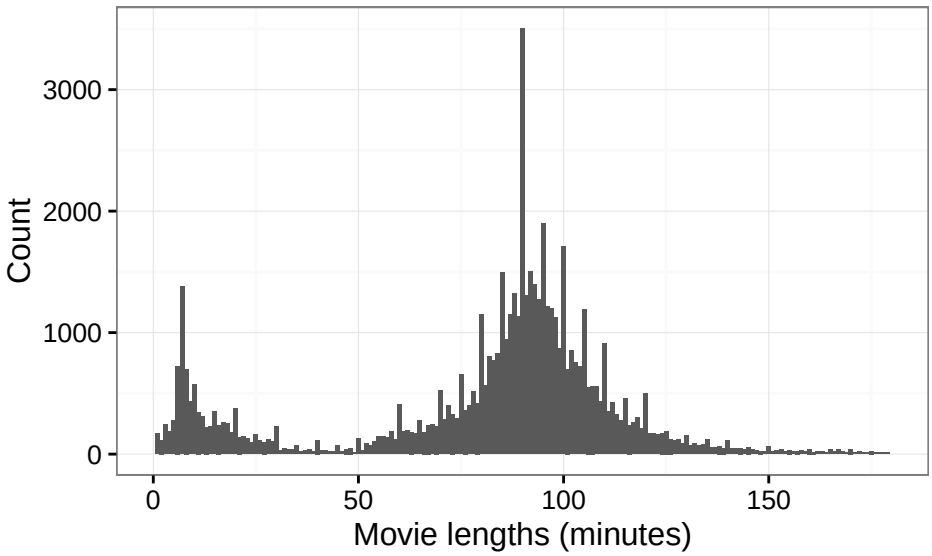
In this problem we will explore a data set containing the velocity (in km/sec) of 82 planets.



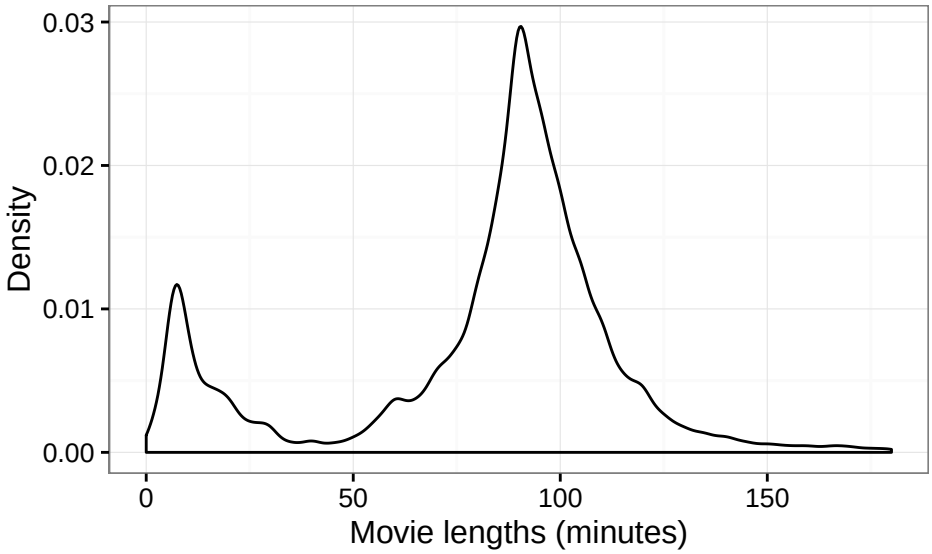
a. Describe the shape of the distribution of velocities. Specifically address the number of modes, symmetry/skew, and any unusual features present.
b. What features are apparent in the histogram but not in the boxplot?
c. Which graph would you prefer to use for displaying these data? Justify your answer.

Problem 3

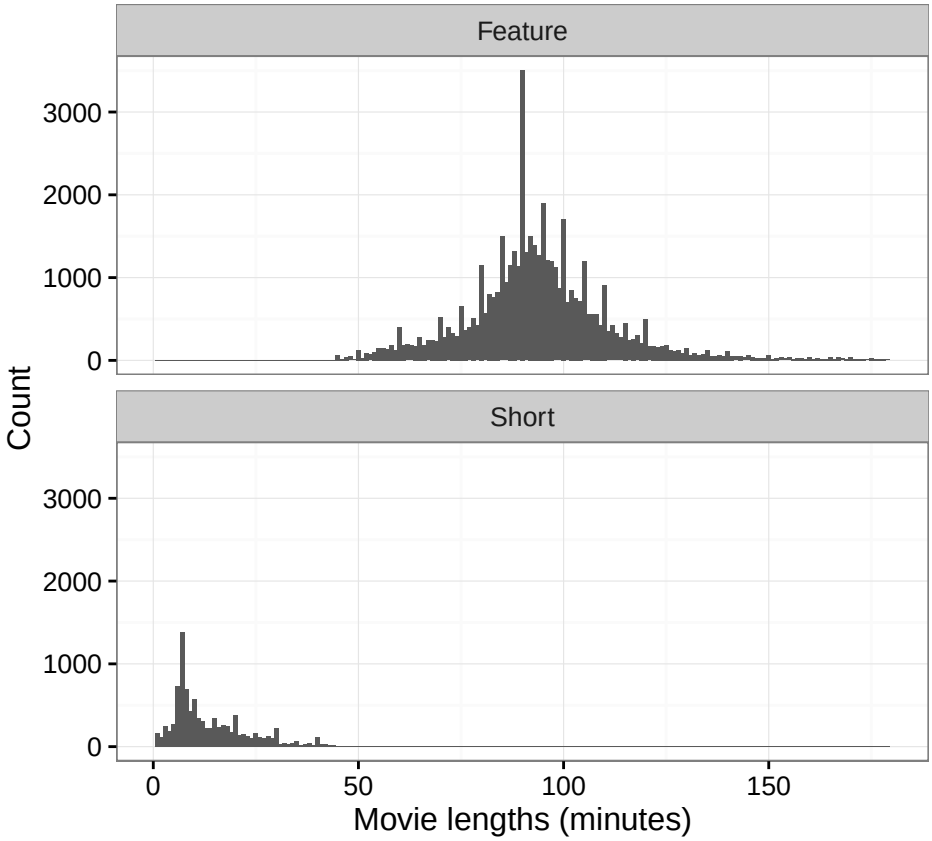
In this problem we will explore the length of 58,788 movies released between 1890 and 2005.



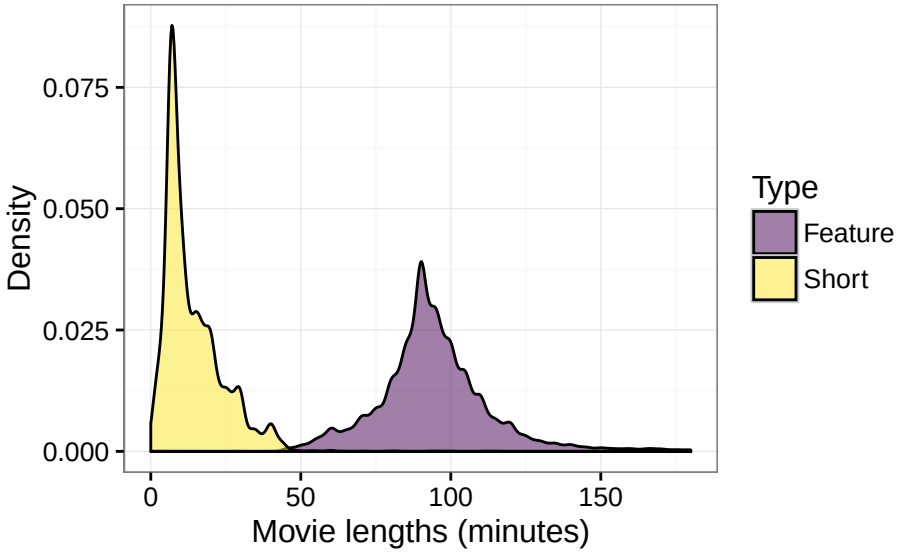
a. Using the histogram, describe the shape of the distribution of velocities. Specifically address the number of modes, symmetry/skew, and any unusual features present.
b. Below is a density plot displaying the same data. What features are easier to see in the density plot? What features are easier to see in the histogram?



c. Below is a side-by-side (i.e. faceted) histogram of movie lengths based on whether the film is classified as a "short" film. Have films been classified as a "short" in a consistent way? Explain briefly.

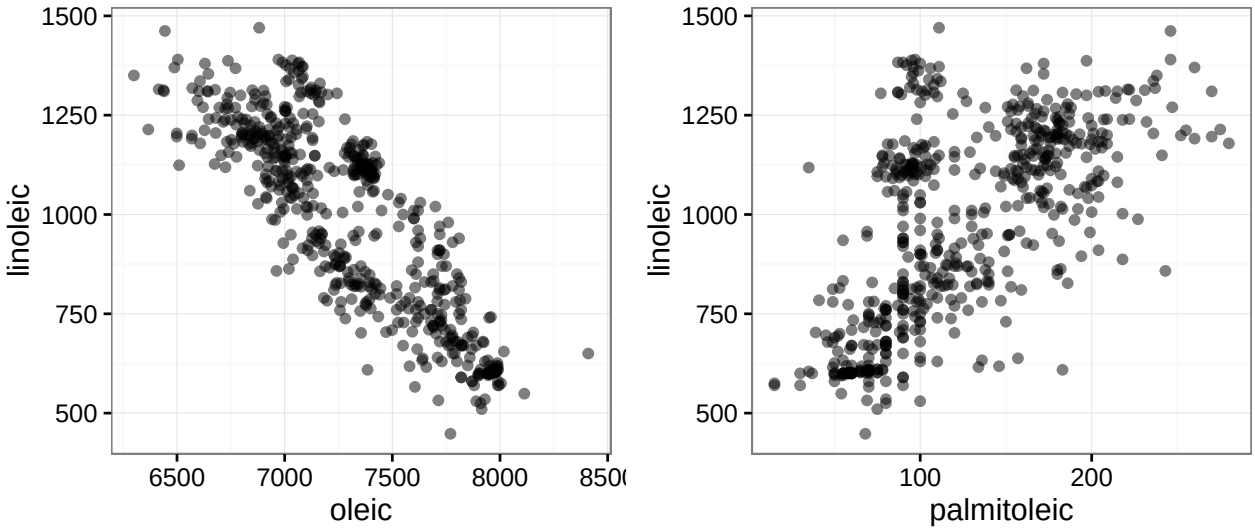


d. The below plot displays the same data, but uses overlaid density plots instead of histograms. What features are easier to see with the density plots? What features are easier to see in the histograms?



Problem 4

Below are two scatterplots displaying the relationship between the percentage composition (% $\times$ 100) of different fatty acids found in the lipid fraction of 572 Italian olive oils. The data is part of a quality control study of olive oils.



a. Describe the relationship between the percentage composition of oleic and linoleic fatty acids.
b. Describe the relationship between the percentage composition of palmitoleic and linoleic fatty acids.
c. It is proposed that the olive oils from different regions have different fatty acid signatures. The below scatterplot uses color and shapes to represent the region of Italy each olive oil is from. What additional information does this provide about the relationship between oleic and linoleic fatty acids?