

3. Visualizing data distributions

Lecture learning goals

By the end of the lecture you will be able to:

1. Visualize distributions.
2. Understand how different distribution plots are made and their pros and cons.
3. Select an appropriate distribution plot for the situation.
4. Create density plots to compare a few distributions.
5. Create boxplots and violin plot to compare many distributions.

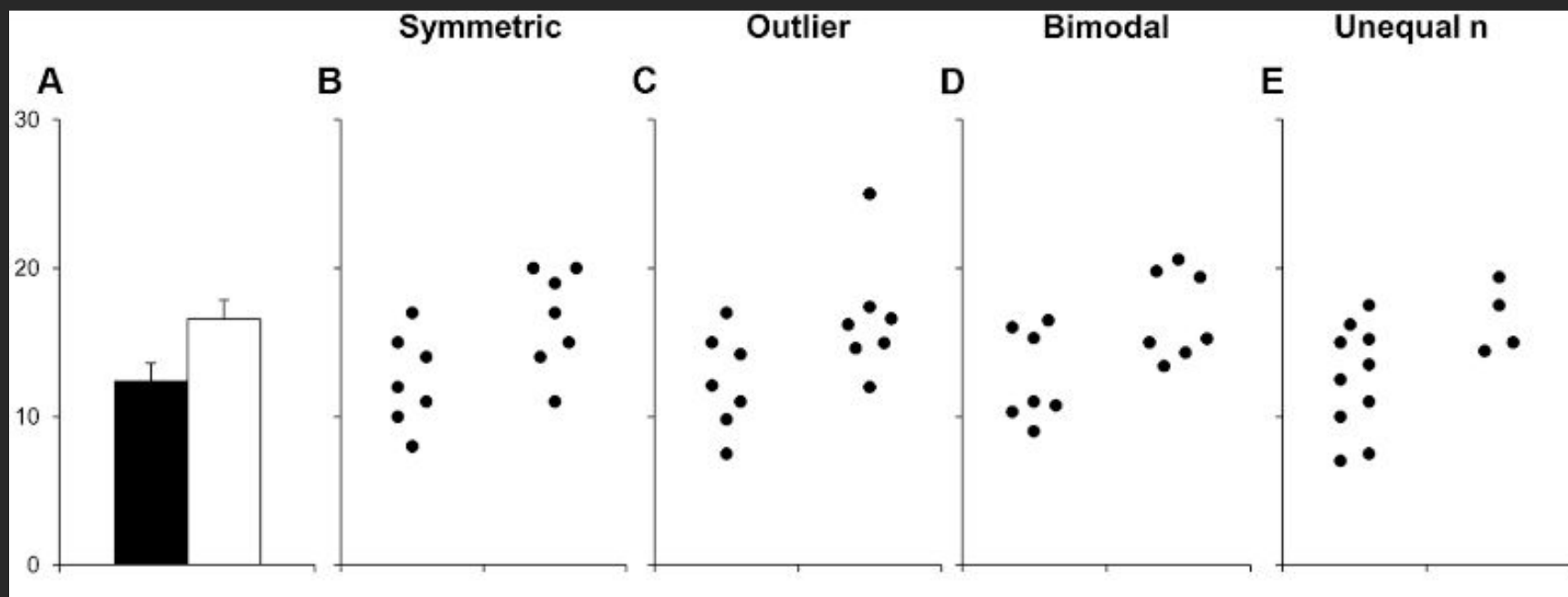
Required activities

Before class:

- Watch [this 30 min video on how distribution plots are created](#) (CWL login required, the screen is intentionally blank for the first minute).

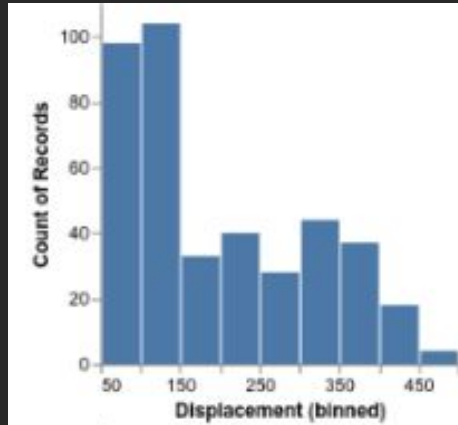
After class:

- Review the lecture notes.
- Chapter 7 and 9 of [Fundamentals of Data Visualization](#).

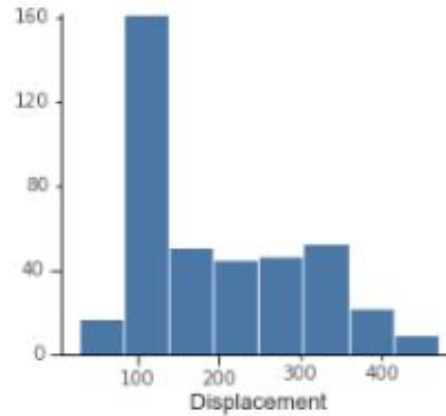


Which histogram do you think best represents the underlying data distribution?

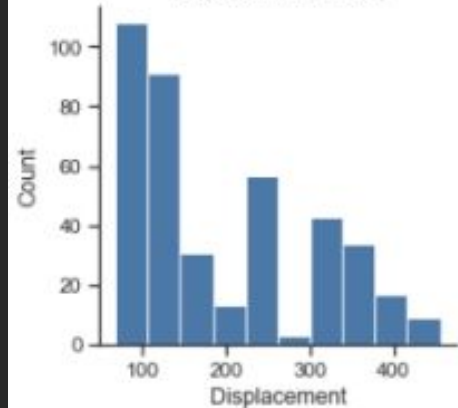
A



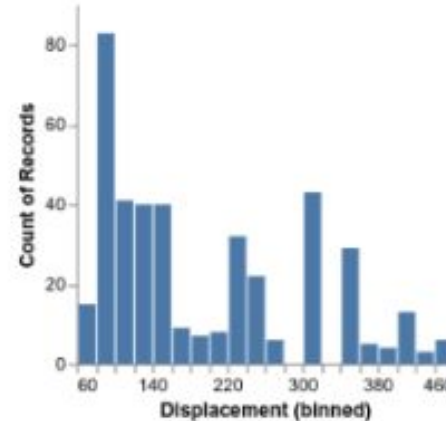
B

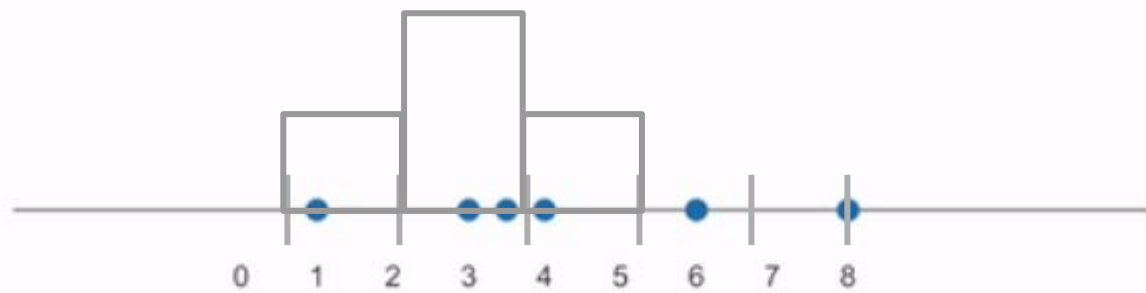


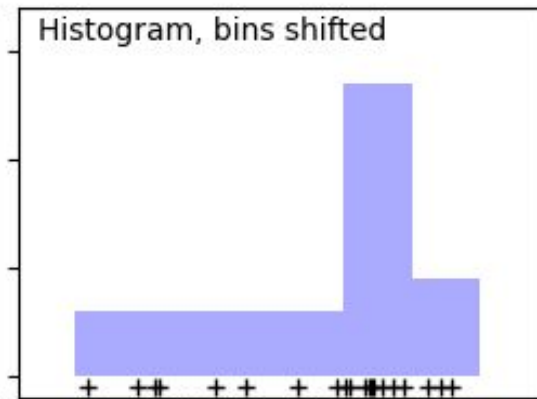
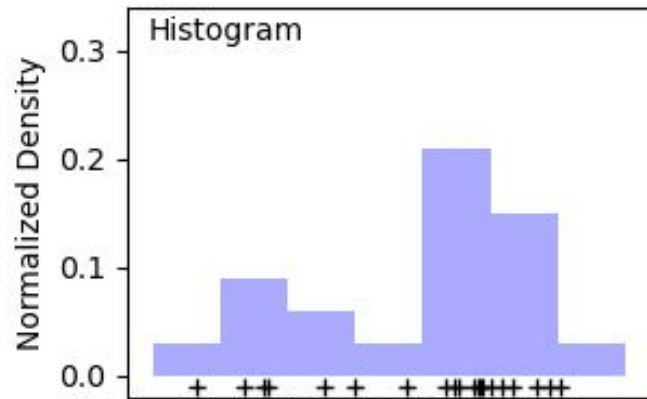
C

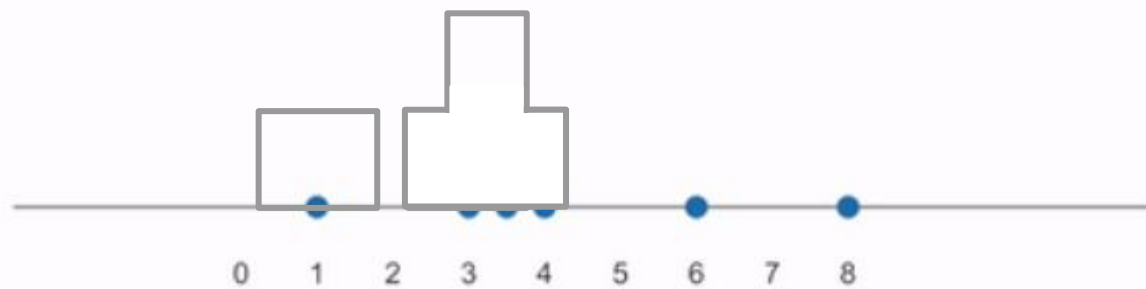


D

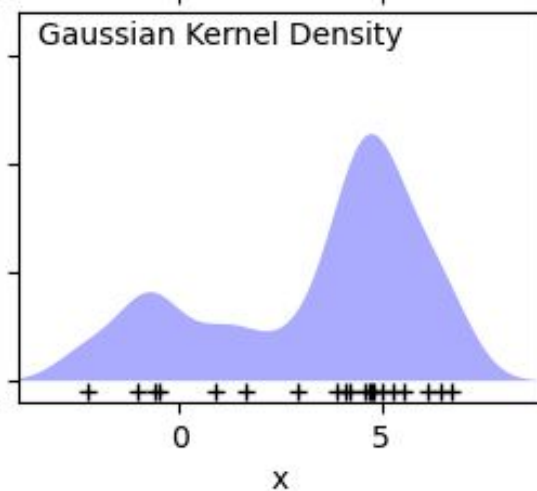
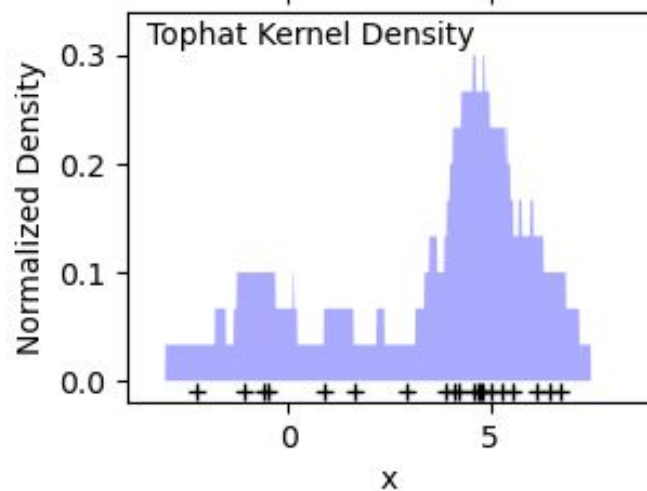
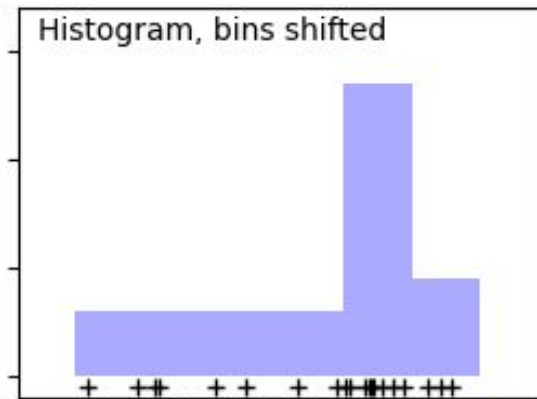
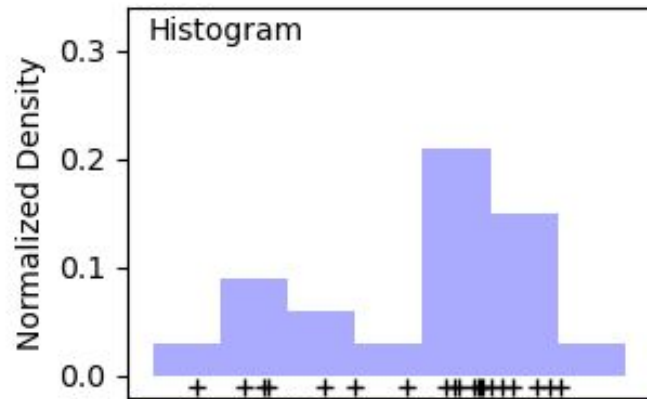




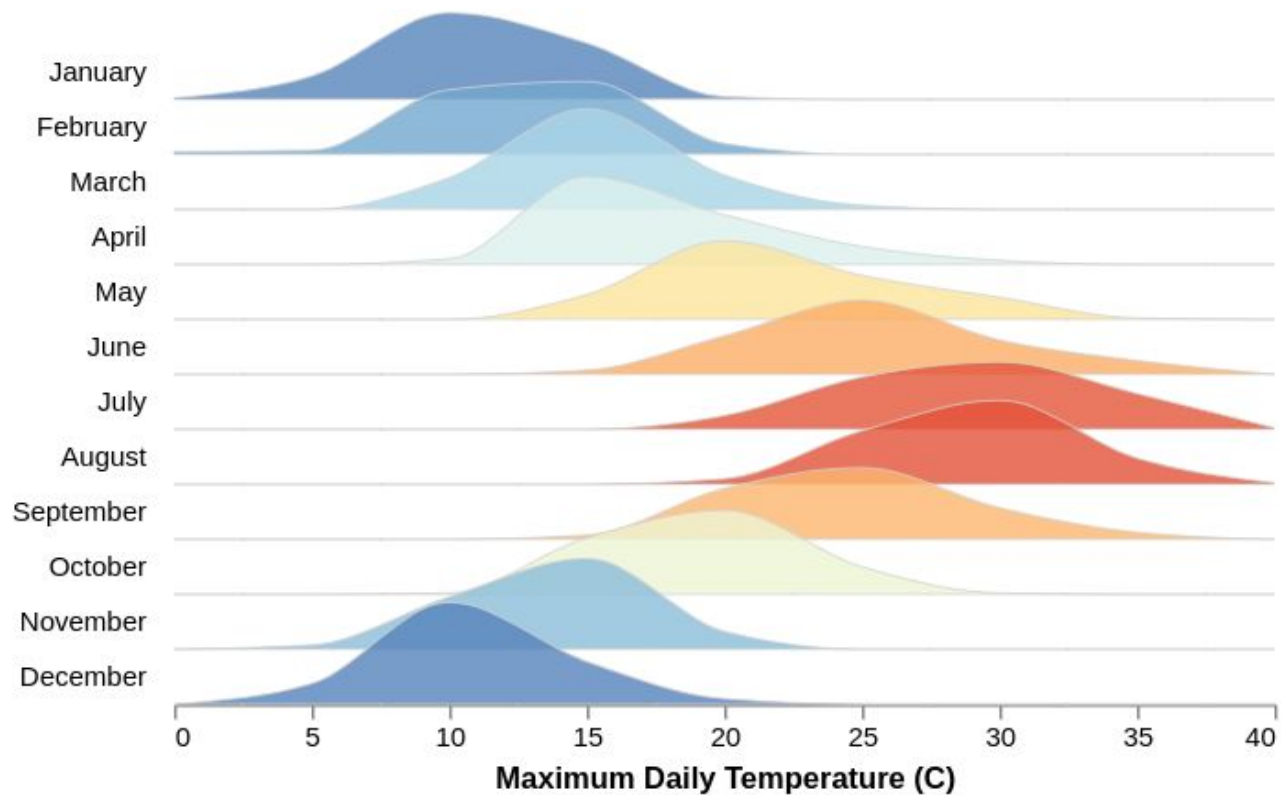


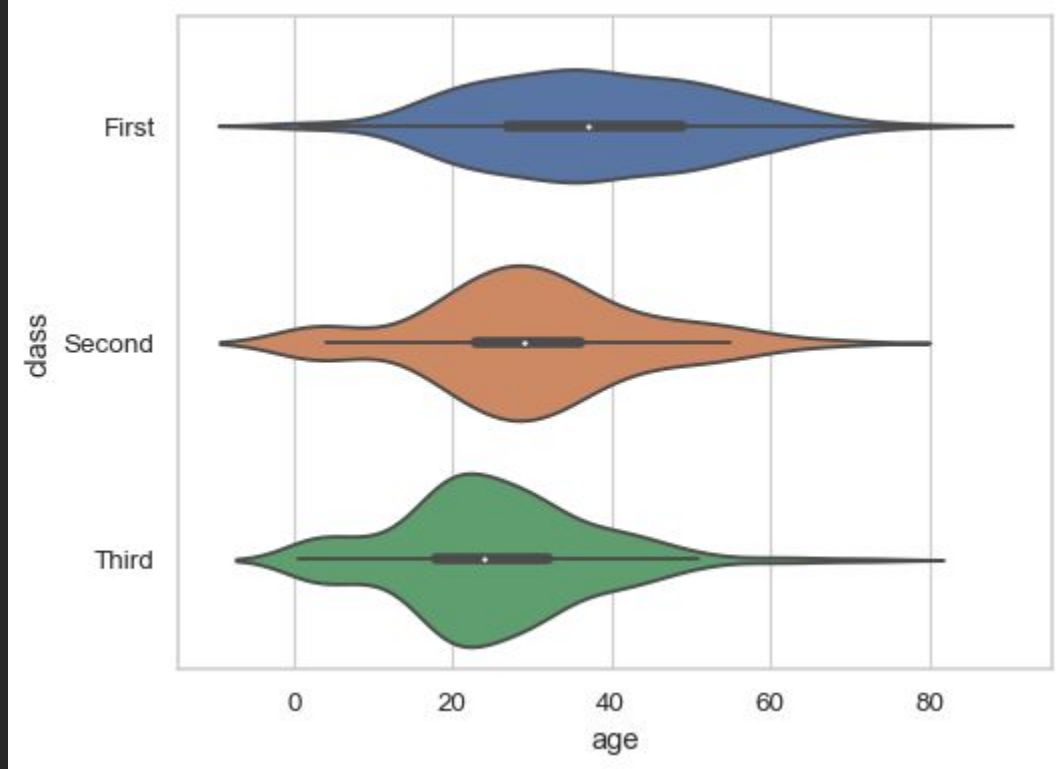


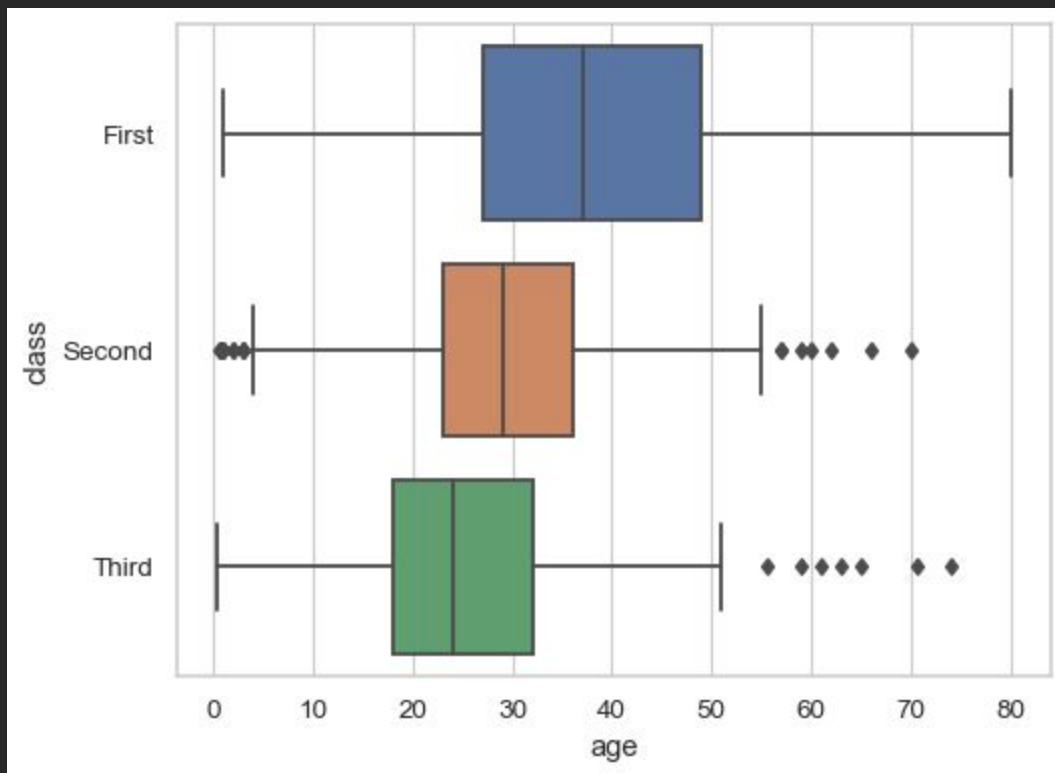


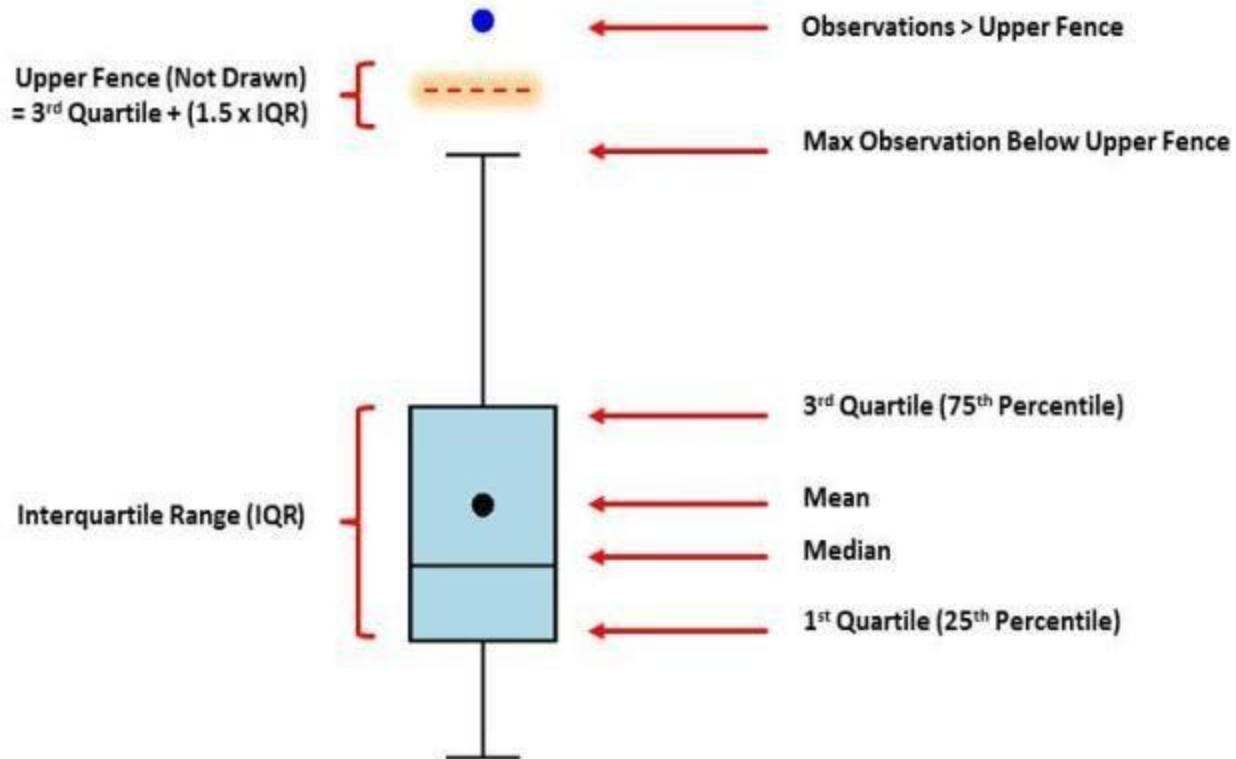


Seattle Weather



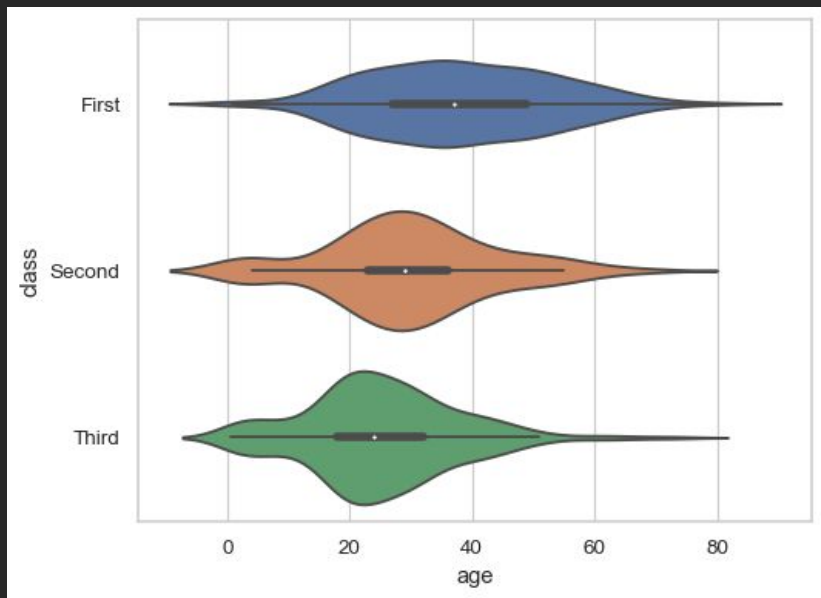






Would you prefer a violin or box plot to compare distributions? Discuss the motivation behind your choice

A



B

