

# Today: 1-D Continuous

Sam Ventura

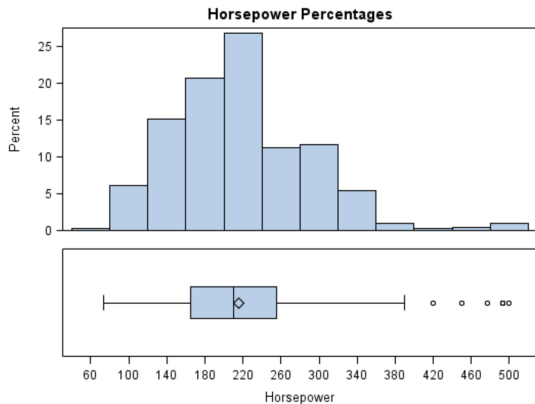
36-315

Today: Boxplots, Histograms, and Density Estimates  
Conditional Distributions for Continuous Variables  
Tartan Data Science Cup – Episode II

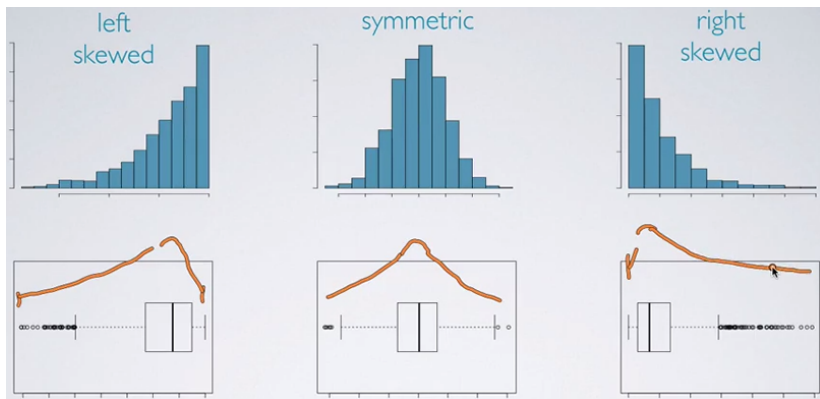
Department of Statistics  
Carnegie Mellon University

September 26, 2016

# Box Plots vs. Histograms: Advantages and Disadvantages



# Determining Skew Visually



# How Can We Improve Boxplots?

Standard Boxplot:

Improved Boxplot?

# Continuous Densities

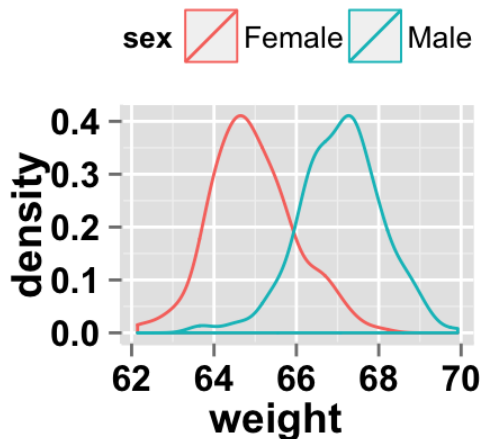
Theoretical:

What if we don't know the underlying distribution?

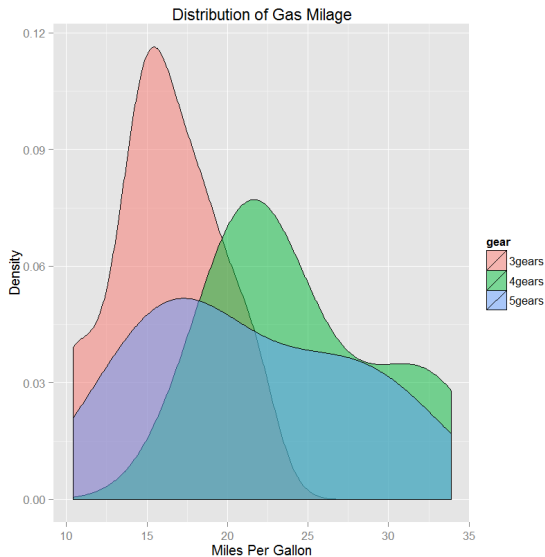
How can we estimate the empirical distribution?

# 1-D Kernel Density Estimation

# Density Estimates and Conditional Distributions

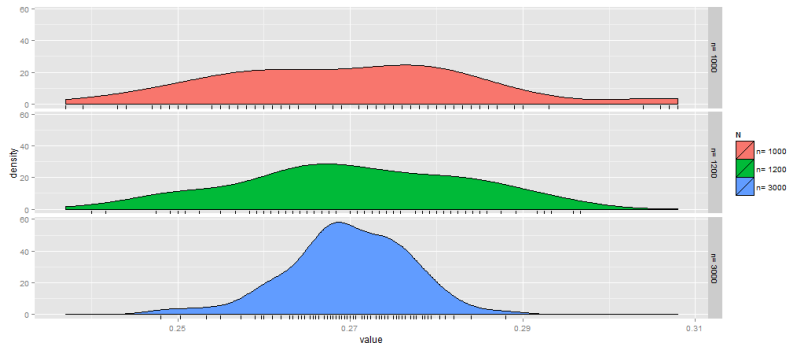


# Density Estimates and Conditional Distributions

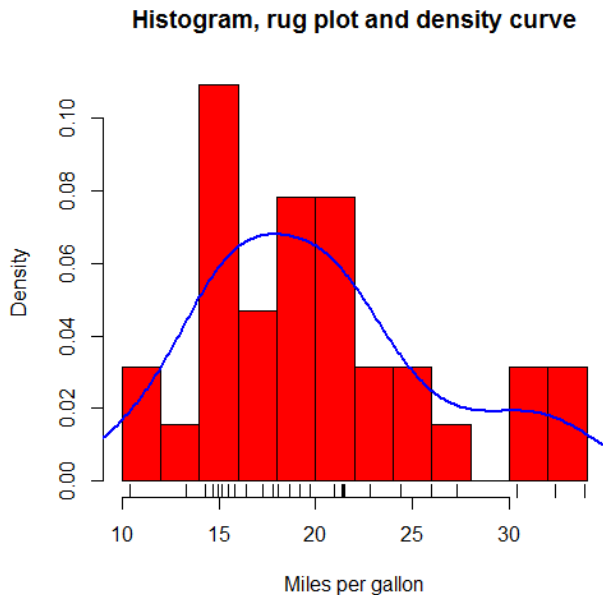




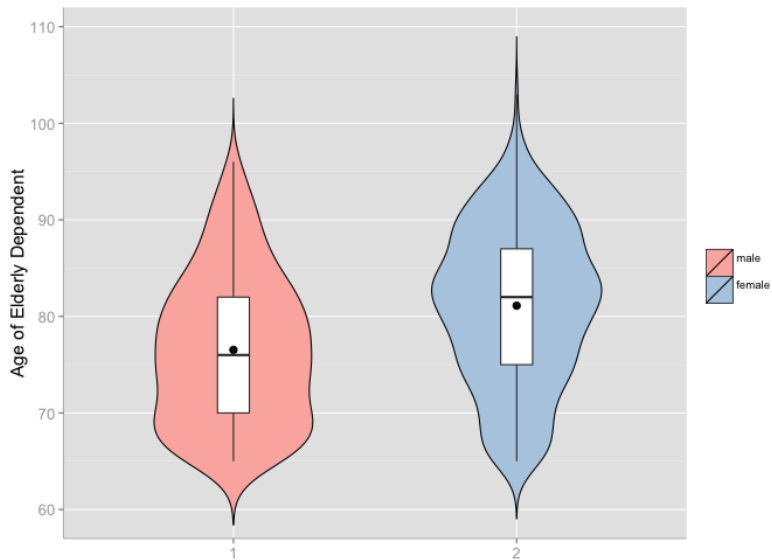
# Density Estimates with “Rugs”



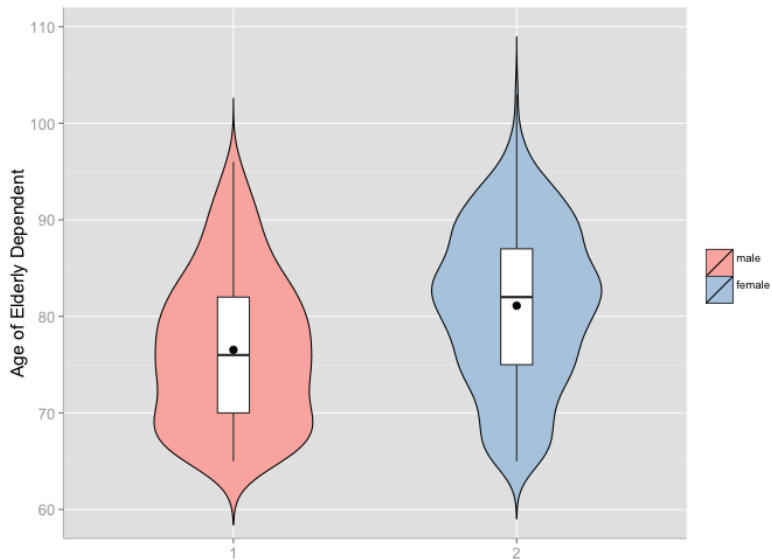
# Histogram with Density Curve and Rug



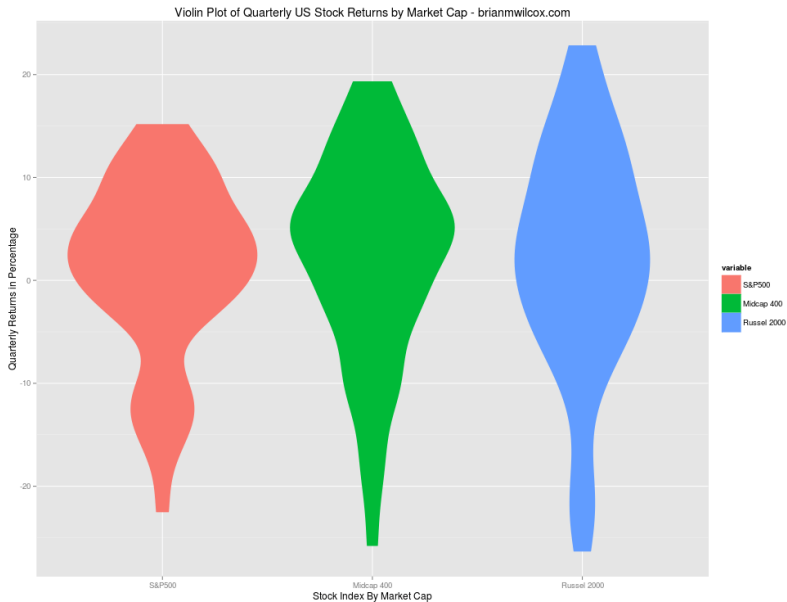
# Violin Plots vs. Boxplots



# Issues with Violin Plots?



# Issues with Violin Plots?



# Improved Violin Plot?

