Problem 1: Calculations

Consider the vector of data $(1, 5, 6)$. Calculate by hand the mean, median, mode, variance, standard deviation, skew, and kurtosis.

Problem 2: Summary Statistics

A sample of seven grad students reported having eaten the following number of donuts this morning: 1, 1, 6, 2, 1, 6, 6.

(a) Input the data into R.
(b) Obtain the following summary statistics: min, 1st quartile, median, 3rd quartile, max, mean, mode, standard deviation. (Make sure to use type 6 quantiles.)
(c) Obtain the inter-quartile range. (Make sure to use type 6 quantiles.)
(d) Write a function that outputs the information in parts (b) and (c), automatically using type 6 quartiles.
(e) Find the geometric mean.
(f) Make a box plot.

Problem 3: Numerical Data

(a) Go to my website and import `1516PerkinsCDR.xlsx` into R using the `rio` package. It contains data on Perkins college loan defaults.
(b) View the variables within the data set.
(c) Obtain the following summary statistics about the cohort default rate: min, 1st quartile, median, 3rd quartile, max, mean, standard deviation.
(d) Which school(s) have a 100% default rate in this data set?
(e) What was UC-Davis' default rate?
(f) Find the skew and kurtosis of the default rate. Does it seem to be normally distributed?
(g) Plot a histogram and a kernel density plot. Does it look normally distributed?

Problem 4: Categorical Data

(a) Go to my website and import `nytoilets.csv` into R using the `rio` package. It contains data on public toilets in New York City. Yep.
(b) View the variables within the data set.
(c) Show how many toilets are in each borough.
(d) Make a pie chart for borough toilets.
(e) Make a bar plot for borough toilets.