

## Introduction to R/RStudio

### Lab Exercise 3: Introduction to Descriptive and Inferential Statistics with R

A researcher is interested in studying the shape of automobiles that are purchased used. She records the shape of the car when it is purchased (`purc`), as well as after 1 (`oneyr`) and 2 years (`twoyr`) post purchase. She also records whether the owner is male or female (`sex`), and the type of automobile (`type`: car, truck, van). The dataset (`cardat.csv`) can be found on my website at [cribbie.info.yorku.ca](http://cribbie.info.yorku.ca).

1. Open the dataset and view the first few lines of the dataset and the names of all variables
2. Calculate:
  - a) mean for the shape of all vehicles when they are purchased
  - b) 20% trimmed mean for the shape of the vehicles after two years, separated by type
  - c) mean shape after one year, separated by sex and type
  - d) histogram of the purchase shape of all vehicles
3. Generate summary statistics for all variables in the dataset
4. Create a plot for the frequency of each type of vehicle
5. Run a two independent samples t-test to determine if the shape of the vehicles after 2 years differs between: a) males and females; and b) cars and trucks. First, check whether the assumptions of normality and variance homogeneity are met or not.
6. Run a paired t-test to see if the shape of the vehicles has changed from purchase to one year post purchase.
7. Run a simple regression to determine if the shape after two years can be predicted from the shape the vehicle was in when it was purchased. Check for any influential cases using Cook's distance, and if any are extreme remove the worst one and rerun the analysis.
8. Run a multiple regression to determine if the shape of the vehicles after two years can be predicted from sex, after controlling for the shape of the vehicles when they were purchased.