

# Sample Quiz 1

*Linda*

*Due: 5:00pm, Jul 19, 2021*

## Contents

<b>1</b>	<b>Overview</b>	<b>2</b>
<b>2</b>	<b>Case study: are Asian cars more efficient?</b>	<b>2</b>
2.1	Question 1: Read data . . . . .	2
2.2	Question 2: EDA . . . . .	3
2.3	Question 3: Comparison . . . . .	3
<b>3</b>	<b>Solution to come</b>	<b>3</b>

# 1 Overview

This is a 30 minute, **individual** quiz. It is an open book exercise. Feel free to copy any R-chunks from the lectures, homework. Ask your TF for help if you get stuck somewhere.

Submit the following two files to the canvas

- quiz1.rmd
- quiz1.html or .pdf (Only 1 compiled one)
- We will keep the submission open until 6:00pm

## Objectives:

We focus on the following aspects

- Read data into R
- Make sense out of the data through EDA
- Solutions/R functions may not be unique
- Be able to use R-Markdown
- Be able to submit your work to the canvas
- No need to beautify your R-markdown .html or .pdf.

# 2 Case study: are Asian cars more efficient?

Asian cars, especially Japanese and Korean cars have been loved by families due to their reputation of being fuel efficiency and relative cheaper in prices. We will use the `car_04_regular.csv` data set (available in the folder) to perform some preliminary analysis.

## 2.1 Question 1: Read data

Read the data first. To make sure you get the data into R we have included the following R-chunk. Notice this chunk works only if you have stored the data `car_04_regular.csv` in a sub-folder called `data`.

```
data1 <- read.csv("data/car_04_regular.csv", header=TRUE)
names(data1)
```

```
## [1] "Make.Model" "Continent" "MPG_City" "MPG_Hwy"
## [5] "Horsepower" "Weight" "Length" "Width"
## [9] "Seating" "Cylinders" "Displacement" "Make"
## [13] "Transmission"
```

- i. How many variables are there in this data? And how many cars are there?
- ii. The variable `Continent` is a categorical variable which labels which region a car is made. How many continent are there?

## 2.2 Question 2: EDA

Fuel efficiency can be measured by `MPG_City`, mileage per gallon in city or `MPG_highway`. Let us look into the `MPG_City`

- i. What are the sample mean and sample standard deviation of `MPG_City` among all cars?
- ii. Show the histogram of `MPG_City` among all cars. Does ‘`MPG_City`’ appear to you a normal variable? Why or why not?

## 2.3 Question 3: Comparison

- i. What is the sample mean of `MPG_City` among all Asian cars?
- ii. What is the sample mean of `MPG_City` among all European cars?
- iii. By comparing the two sample means above, do we have some evidence that Asian cars are more efficient than that of European cars? (The larger `MPG_City` is the more efficient a car gets.) Comment on the limitation of this simple analysis.

End of the sample quiz

## 3 Solution to come