Analytics Foundations: Lab 12

1. Data were collected in an effort to relate the safety of certain vehicles to different aspects of those vehicles. To obtain this data, submit the following code:

```
safety <- read.csv('https://raw.githubusercontent.com/IAA-Faculty/statistical_foundations/master/safety.csv')
```

This dataset has the following variables:

- **Unsafe**: binary safety designation (1 = below average (unsafe), 0 = average or above average (safe))
- **Type**: type of car (Large, Medium, Small, Sport/Utility, Sports)
- **Region**: manufacturing region (Asia, N America)
- **Weight**: integer value for car weight ranging from 1 to 6
- **Size**: size of car corresponding to **Type** (1 = Small/Sports, 2 = Medium, 3 = Large or Sport/Utility)

Do not create a training and test set and just use the whole dataset for the following analysis.

Perform the following analysis:

a. Which variables are continuous, nominal, ordinal?

b. Examine the association between **Region** and **Unsafe**.
    a. What percentage of cars manufactured in Asia was classified as unsafe?
    b. What percentage of cars classified as safe was manufactured in North America?
    c. What is the appropriate test to use?
        i. State the null and alternative hypothesis for the test.
        ii. At an alpha of 0.05, what is your decision?
    d. Regardless of significance, interpret the odds ratio in the context of the problem.

c. Examine the association between **Size** and **Unsafe**.
    a. What is the appropriate test to use for association?
        i. At an alpha of 0.05, what is your decision?
    b. How strong is the association between these variables?