Fitzwilliam Institute Diploma in Data Analytics - Course Project

Due date – 2nd December 2019 at 11:59pm

Report must be submitted via email to: opsmanager@fitzwilliaminstitute.com

The number and types of motor vehicle crash deaths differ widely among the 50 states and the District of Columbia in the U.S. Fatality rates per capita and per vehicle miles travelled provide a way of examining motor vehicle deaths relative to the population and amount of driving. However, many factors can affect these rates, including types of vehicles driven, single or multiple vehicle collision, whether the passengers were wearing a seatbelt or not (restrained fatally injured occupants versus unrestrained fatally injured occupants), whether the accident occurred in an urban or rural area.

The 2016CarCrashes.csv file contains data from the U.S. Department of Transportation's Fatality Analysis Reporting System (FARS) in 2016.

You work for an insurance company who want you to do an analysis of Fatality rates per capita. You must write maximum a four-page report with your findings. Please provide your report in PDF or DOC. Include your well-commented R code in a separate .R file. No plots are to be included in the report; all graphics used to describe the data must be included in the separate .R file.

You should discuss interesting findings in the R output in your report. The report should have the following headings:

1: Introduction to Data Set

2: Data Analysis

(a) Descriptive data analysis

Describe results from the:

i. Numerical summaries:

A. 5-number summaries

B. correlations

C. variance etc.

ii. Graphical summaries:

A. boxplots

B. histograms

C. scatterplots

D. barplots etc.

(b) Determine the Best Predictive Model

Describe results from:

i. Model Building

ii. t-tests

iii. F-tests , ANOVA

(c) Preform analysis of the best regression model

Describe results from:

i. Coefficients and model fit

ii. Model Diagnostics

iii. Outliers / Influential values

iv. Possible Remedial Measures Section

3. Interpret the results

4. Conclusion and recommendation