

MATH3714 Coursework

Viet Dao

November 24, 2018

Contents

1	Introduction	3
2	Initial Data Analysis	3

1 Introduction

We have been given a data frame $\mathbf{A}_{393 \times 9}$ which is the table of different cars with mpg, cylinders, displacement, horsepower, weight, acceleration, year, origin and name for a given car. Our goal is to be make a model that is capable of predicting mpg from our data given. Now we split up $\mathbf{A}_{393 \times 9}$ into $\mathbf{Y}_{393 \times 1}$ which contains only mpg and $\mathbf{X}_{393 \times 8}$ which contains everything in $\mathbf{A}_{393 \times 9}$ apart from mpg. This sets up our response and explanatory variable.

2 Initial Data Analysis

In this preliminary stage we want to investigate outliers and possible missing data in our dataframe $\mathbf{A}_{393 \times 9}$