

edaPlot__Intro

Viraj @DataTasteMaker

April 02, 2016

Contents

Introduction	1
Set up the work place	1
What is the data about	2
Now to EDA	2

Introduction

EDA - Exploratory Data Analysis

The very first step in any and every statistical and analytical problem solving assignment.

We all love to visualize the data, graphs, plots etc. comes in handy and gives us the first understanding report about the data. There are n-number of ways to plot and view the data, there are many easy functions and packages to aid us in this activity.

One such function is [edaPlot\(\)](#)

And the best part of this function is - it gives us *interactive* control on which variables to use, which plots to view.

My personal note on usage of this function is to use graphics window instead of the regular plots section in the R-Studio. Use *windows()* for Windows OS, *X11()* for Linux or MAX OS

Let's start with understanding this function

Set up the work place

As always, do clear the environment and plots before starting, to free up the memory

To ensure same output, set the seed to some number, I use number 7

```
set.seed(7)
```

Load (or install, if not already) the required libraries

```
library(openintro)
library(ISLR)
```

For this I will be using the *Auto* data from the *ISLR* package

```
ipdata <- Auto
```

What is the data about

Auto Data Set

A data frame with 392 observations on the following 9 variables.

- mpg miles per gallon
- cylinders Number of cylinders between 4 and 8
- displacement Engine displacement (cu. inches)
- horsepower Engine horsepower
- weight Vehicle weight (lbs.)
- acceleration Time to accelerate from 0 to 60 mph (sec.)
- year Model year (modulo 100)
- origin Origin of car (1. American, 2. European, 3. Japanese)
- name Vehicle name

Now to EDA

The `edaPlot()` function give a click-interface and eight graphing options. The type of plots, the list of variables are printed out in the console.**

To exit, click “exit” on the plot. *However we will use the graphics Window for this*

```
windows()
X11()
edaPlot(ipdata)
```

The output is like below

