

Introduction

Aravind Hebbali

16 April 2018

DSCE Business Analytics Conclave Workshop

19th, 20th & 21st April, 2pm - 5pm

Introduction

Materials for the workshop. Please install all the required softwares and download the data sets and scripts from this repository before the workshop.

Software Prerequisites

R

R can be installed from the R project website. Please install the latest version 3.4.4.

RStudio

RStudio IDE can be installed from the RStudio website. Install the free desktopo version.

R Packages

Below are a list of R packages required for the workshop.

- readr
- readxl
- haven
- dplyr
- ggplot2
- magrittr
- descriptr
- skimr
- inferr
- olsrr
- blorr
- rmarkdown
- knitr
- devtools
- blorr

You can install the packages by running the following code in R:

```
install.packages("package_name")
```

Anaconda

Download and install the Anaconda distribution from their website. We will use Jupyter Notebook for the python session.

R Session

- Quick tour of R & RStudio
- Importing data into R
- Data Wrangling
- Data Visualization
- Exploratory Data Analysis
- Linear Regression
- Logistic Regression

Python Session

- Data Wrangling with Pandas
- Data Visualization
 - matplotlib
 - bokeh
 - plotly
- Modeling with statsmodels

Useful Resources

- Rstudio Webinars
- RStudio Cheat Sheets
- tidyverse
- R Bloggers
- R Weekly