Business Analytics
Yelp: "Startup Survival Kit"
—

Nicolás Metallo, Shruti Khandelwal, Karthik Vannela, Karan Sawant, José Collado

INTRODUCTION

Technical report of our presentation Yelp: "Startup Survival Kit". All necessary R files are attached along with a brief explanation in this file.

Data handling

We handled the data from Yelp through different steps:

1. **Data Access**

   The files are readily available here: https://www.yelp.com/dataset_challenge/dataset

2. Data Conversion

We transformed the files from Yelp using the "json_to_csv_converter.py" which can be found here: https://github.com/Yelp/dataset-examples

3. Data Consolidation and Reduction

    a. Yelp_Data_Handling.r (Attached)

Linear Regression

Yelp_Data_LinearRegression.r (Attached)

The model was generated in R and then we cross checked with Rattle. The decision tree was not able to arrive at a solution in rStudio and Rattle was used instead.

Text Analysis

1. Data Preparation

We used a Python script by Ryo Kita (available here) to parse the yelp dataset for easier analysis in R.

2. Data Analysis

Yelp_Data_Text_Analysis.r (Attached)

3. Data Visualization

All charts were plotted again using Tableau, R and Raw (http://raw.densitydesign.org)

4. Web Visualization

The web visualization was done with Shiny (http://shiny.rstudio.com). We start the server with the function RunApp()

    a. ui.R (Attached)
    b. server.R (Attached)