

Final Project Proposal

Anil Akyildirim

3/18/2020

Introduction

Throughout all the Data Science classes (Data Acquisition and Management, Advanced Programming, Business Analytics and Data Mining and etc. . .), we have been working with variety of different data sets within different domains such as local government issues, education and health. Considering my current Data Analyst role and background in Product Management in the Digital Marketing field, on my final project for Knowledge and Visual Analytics, I would like to focus on Analysis of Digital Marketing Data. Below are some of the possible business problems that I can explore:

** As a business we feel like we are losing a lot of revenue due to technical issues in our web properties? Can we validate this with website data? Analysis of Browser Performance report for a website.

** How well is our website mobile optimized? Analysis of Mobile Performance Report.

** What day of the week and what time of the day our website traffic is highest and lowest? If we know this, we can target launch of certain ecommerce campaigns. Analysis of Hours and Day.

** What referral sites provide the most value for our website? We have partnerships with other non profit organizations, which ones we should focus on? Analysis of Referring Sites.

** Are users coming to our website consuming the content? Page Efficiency Analysis.

** How much it costs to acquire a valuable user or session for our website?

Exploration of browser use against, sessions, visits, traffic source, time spent on site per user based on sessionID and hits can help guide or become a starting point to answer these questions.

About the Data

Data set is part of a Google Analytics Customer Revenue Prediction Competition. Link to the data is here: <https://www.kaggle.com/c/ga-customer-revenue-prediction/data> .

I can also explore creating new features for some of the variables that are not available on the data set, however it may not be possible to drill down some of the dimensions such as Referral Specific sites. The data set is messy and has many missing values. Data set needs to be cleaned and prepared for exploration and visualization.

Purpose

The analysis of the dataset will provide actionable insights to the marketing team. Majority of the creative or marketing agencies use powerpoint presentations. None of the slides can be validated by stakeholder, most of the slides are not interactive or simply screen shot images of not reliable data. The format of powerpoint presentations are also not engaging. Creating an interactive shinny app would help engage and connect with the business stakeholders.

Technology

For this project, I propose to use R with possibly using animate, gganimate, plotly and goolevis packages.