**Senior Project - Meeting 2 Report**

**Relevant Coursework**

My top math, computer science and statistics courses that I have taken are:

|  |  |  |
| --- | --- | --- |
| Calculus III | Database and file structures (SQL) | Probability & Stat. |
| Differential Equations | Data structure and algorithms (C++) | Statistical Computing (SAS, R) |
| Discrete Mathematics | Stat. and Machine Learning (current) |  |
| Linear Algebra |  |  |

**Job Description**

I will be joining Fidelity Investments in June of this year as an Associate Data Engineer where I will be working mostly in the back-end aspect of their applications and developing the workflow/ ETL processes. I hope to eventually climb the ladder and become a Principal Data Scientist at Fidelity or a top leading software/technology or financial company.

**Software Skills & Ranking**

\* 1 - basic knowledge, 3 - advanced familiarity

|  |  |
| --- | --- |
| **Skill** | **Rank** |
| R (RStudio & RGui) | 2 |
| Python | 1 |
| SAS | 2 |
| SQL (SQL Server) | 3 |
| Excel | 2 |
| C++ | 3 |
| Java | 2 |

**Data Summary**

The data was obtained from Zappos.com, an online clothing and shoe store based in Las Vegas, Nevada. Unsurprisingly, Zappos’s parent company is Amazon.com. Zappos captures millions of their users’ data for many purposes, especially for marketing their products and brands. This data obtained from their Analytics team t contains 21,062 records of sample customer transactions. The data include the following fields:

* day: the calendar day
* site: Company site visited by users
* new\_customer: 0 = returning customer; 1 = new customer; null = neither
* platform: the type of device used by a website visitor
* visits: the number of distinct website visits; 1 session may have multiple visits
* distinct\_sessions: the number of distinct website visitors; 1 session may have multiple visits
* orders: the number of website orders
* gross\_sales: the total gross sales for website orders
* bounces: the number of visits that only viewed one page
* add\_to\_cart: the number of visits that added a product to cart
* product\_page\_views: the number of product pages viewed
* search\_page\_views: the number of search pages viewed

Three additional columns/ fields are to be calculated using some of the fields that are already provided above:

* conversion\_rate ='orders'/'visits'
* bounce\_rate ='bounces'/'visits'
* add\_to\_cart\_rate ='add\_to\_cart'/visits'