Project Requirements:

Dear Class - the project should have the same sections as any other assignments. So, to start, review the sections (Introduction, Analysis, Results, Conclusions) in the Assignment Doc.

**Specific Assignment Requirements:**

<https://docs.google.com/document/d/166M8Yb-l_T0bWScaNpdCss3ICQH6sGChwVNZrdqm5qI/edit?usp=sharing>

Next, think of the project as a much larger, more robust, and more excellent assignment. This means that each section will be bigger and better - with a lot more vis - and many more subsections. Here is a good outline:

**Introductions:** (4 - 5 paragraphs with references, some history, and some info about the current state of the topic).

**Analysis**

About the Data

Cleaning

<Include measures of before and after>

Preparization

<Things like normalization, feature generation, discretization, transformation, etc)

Visual Exploratory Data Analysis (EDA)

<and stats analysis>

Model 1: Association Rule Mining

The data used (just the first 5 columns)

< Remember - each model will use a different version of the data - so show a bit of the data so the viewer can see what you applied to the model. This is true for all following subsections - but I will not repeat>

Model 2: Clustering

<Use kmeans and hclust>

Model 3: Decision trees & RF

<Have a few at least>

Model 4: Naive Bayes

Model 5: SVM

Model 6: Text Mining

Etc…

**Results**

<You should have a subsection for every model. You should include results, measures, attributes like conf, sup, and lift for association rule mining, like kernels and costs for SVM, etc, etc, . Each model has its own measures and attributes. >

\*\*\* Have a least 2 vis’s for each model

**Conclusions** (4 - 5 paragraphs)

NON\_TECHNICAL

And awesome.

This is the frosting!