**Unit 1 Project Part 1**

**Background:**

You are asked by the CMO to create a model to determine 12M future spend of customers. They are especially interested in high value customers. This is for the marketing group, and will be used for tailored targeting with offers, promotions, etc.

**Objective:**

Provide linear regression model for ongoing scoring of customer base with regression plotting (facets)

**Deliverable:**

**Initial Deliverable**: Summary of Data Review, Methodology (Universe, Time Frame, Exclusions/Outliers), Summary of Variables Created with explanation and Initial Variable Reduction Results

**Final**: A 10 min presentation outlining your methodology and model overview to the CMO. Model Equation.

**Audience:** Phase One: Director of Advanced Analytics; Phase Two: CMO

Variable Creation Possibilities

1 - **Dummy Variables**

2 - Transformations (log, square, square root, flags) (might use df.apply())

3 - Domain Knowledge

4 - Interactions (var x var) {polynomial\_features - setting for interactions only)

Challenge for Variable Reduction:

1 - P Value (statsmodel)

2 - Reduce Collinearity (Correlation Matrix)

3 - Recursive Feature Elimination (Will also help reduce collinearity)

5 - For Insight - Coefficient Value on Scaled Data

Other Requirements:

Reserve 1 min for reporting on project plan/qc

Evaluation:

1. Did you meet the requirements of the brief?
2. Was the presentation appropriate for the intended audience?
3. Was the presentation professional and polished?