**Homework Team Orange 2**

**Agenda: September 5, 2019 --3pm**

1. Ranking Odds Ratios
2. Visualizations: Ideas and suggestions?
   1. Table for significant variables (separated by 4 types)
   2. Appendix table for all variables (separated by 4 types)
   3. Visualization for missing values: Bar chart? Pick top n?
   4. Extra: for interesting findings?
      1. *Comparative horizontal bar chart to show unbalanced distribution of 0/missing data?*
3. Questions and discussions about our **own parts** during writing
   1. Executive Summary**: Cathy**
   2. Results**: Price**
   3. Recommendations: **Evan**
      1. Where to put findings (results or recommendations)
         1. missing values
         2. Small subset of data (the issue with lots of 0’s)
         3. Significant continuous variables that meet linearity assumption
         4. Redundant variables (multicollinearity for continuous, and potential correlation between indicators)
      2. Specific sentences when explaining the importance of this finding.
      3. Recommendations for missing values? Lots of zeros
   4. Methodologies & Analysis: **Grant**
   5. Conclusions**: Sufyan**
4. Review the reports together
   1. Executive Summary**: Cathy**
   2. Results**: Price**
   3. Recommendations: **Evan**
   4. Methodologies & Analysis: **Grant**
   5. Conclusions**: Sufyan**
5. Next Meeting (Friday or Saturday):
   1. discussion about Time Series HW2

Reference: Report Outline

* + Executive Summary: (**Cathy**)
    - *(Could write it in the end)*
    - Business context
    - briefly mention findings (how many variables)
    - list some important variables maybe
  + Results: (**Price**)
    - significant variables we found
    - **table** of Significance (by p-value)
    - **table** of OR (By magnitude), normalize those less than 1
    - Linearity Assumptions for continuous (var name; description; Y/N)
  + *Concerns* & Recommendations (**Evan**)
    - **Bar chart** for Missing values
    - Redundant variables (correlation)
    - *(Binary)* variables with high odds ratios
  + Methodology & Analysis: how we get this result (eg. Box Tidwell) (**Grant**)
    - Data set
    - “where statement” in linearity assumption testing
    - When GAM did not converge, we go back to Box Tidwell (with some data missing)
      * Give 2,3 examples
    - When MH broke, we use EXACT permutation to compute the significance
  + Conclusion **(Sufyan)**
    - *(Could write it in the end)*
    - data consideration?
  + Appendix: Table for all variables