Homework Team Orange 2

Agenda: September 3, 2019 - 12pm

1. Analysis
   1. Variables ranking (“From most significant to the least”)
      1. Separately by 4 types
      2. By p-value with an extra column of test statistics, (if too small could use test stats)
   2. Linearity Assumption
      1. Pay attention to variables that have lots of 0 values (Box Tidwell would run only a small subset of data)
         1. Go study the data, is it really 0 or it does not have an account?
            1. “**WHERE**” statement
         2. use **GAM**. (document percentage of 0)
         3. if not converged, go back to Box Tidwell (specify it in report)
   3. Data Considerations (interesting finding)
      1. Within Binary variables (OR): what interesting variables maybe
      2. Within Summary
2. Paired Check in
   1. Sufyan 🡪 Grant 🡪 Evan 🡪 Cathy 🡪 Price 🡪 Sufyan
   2. Correct test
   3. Wald chi-sq p-value
   4. For continuous var, check for 0 / negative, (whether it converge)
3. Writing the report (decide tomorrow)
   1. Executive Summary:
      1. Could write it in the end
      2. briefly mention findings (how many), list some important variables maybe
   2. Results: where we talk about significant variables we found (with tables)
   3. Recommendations?
      1. *(Binary)* variables with high odds ratios
      2. Redundant variables (correlation)
      3. Analytical recommendations
   4. Methodology & Analysis: how we get this result (eg. Box Tidwell)
   5. Conclusion
4. Visualizations (which software)
   1. Explore Tableau
   2. Excel as back up plan
5. Before Next Meeting (9/4/2019 at 4pm):
   1. edit codes and findings
   2. paired check in
   3. interesting findings