**Homework Team Orange 2**

**Minutes: September 11, 2019 -12pm**

1. Time Series HW2: (Results finalized)
   1. Interpretations of the first three graphs
      1. Seasonality is not a very important factor, however, it’s still a factor
   2. choosing models:
      1. based on Accuracy stats, Decomposition Graphs, Context
      2. Choose the **Holt Winter Multiplicative (Trend\*Seasonality)**
      3. Why:
         1. Not first one (despite stats): doesn’t include trend
         2. Seasonality: there is a season element, just not big (both in STL plot and context)
         3. Multiplicative: 2017-2018, the variation is getting smaller, also based on Accuracy stats
2. Outline
   1. Executive Summary (Sufyan) – “give context”
      1. Look online (some stats: how’s wake county doing) – context
      2. BLUF: the model we choose is about 79% accurate in predicting
   2. Results (Cathy)
      1. Results
         1. Which model we chose
         2. MAE & MAPE (make sure how to interpret)
      2. Time plot (actual vs. prediction) & detailed interpretation (eg: trend, spikes, how the prediction goes)
   3. Recommendations (Price)
      1. Do more research the cause of irregularities in the data
      2. Refer back to graph on the result section to show why
      3. How we aggregate the data *(depends on future questions we try to answer)*
   4. Methodology & Analysis (Evan)
      1. Dataset (we confirm that the data is normally distributed)
      2. STL decomposition graphs & interpretation (**how** we choose the model)
      3. Do a table for accuracy statistics for all 5 models (Appendix)
   5. Conclusion & Visualizations (Grant)
      1. Next step: random walk/ ARIMA (refine our model) since we do see irregularities in seasonality
   6. **Appendix**: a table for accuracy statistics for all 5 models (Appendix)
3. Before Next Meeting (Thursday 09/12 2:30pm)
   1. Proofread Linear Algebra email
   2. Questions for **report**?
   3. Optional: Visualizations (Grant):
      1. R plot is acceptable (***ggplot*** would be even better)
      2. For the time plot of actual vs. predicted: don’t need Confidence Interval
   4. Logistic Regression HW: (first bullet point)
      1. Use the new training dataset
      2. Check missing values again for each variable
      3. Separation concerns?