Executive Summary – Kruskal-Wallis analysis of 911 data

Tony Dunsworth

**Problem and Hypothesis:**

Is there a statistically significant difference in the median call processing times in response to the Department of Emergency and Customer Communications (DECC) of the City of Alexandria’s efforts to preserve the continuity of service to their constituents in response to the SARS-COV2 pandemic?

H0: There is no statistically significant difference in population means meaning the department’s actions had no impact on call processing times.

HA: There is a statistically significant difference in the population means meaning the department’s actions did impact call processing times.

**Summary of the Analytical Process:**

The data from 2019 and 2020 for analysis was extracted from Microsoft SQL Server 2016 databases supporting the city’s Computer Assisted Dispatch systems. After the extraction and preliminary cleaning via a SQL query. The exported csv file was imported into RStudio and converted into a data frame. Four variables were constructed for elapsed time points representing the time from phone pickup to time the call is ready for dispatching, the time from ready for dispatching to the assignment of the first unit, the time from phone pickup to the assignment of the first unit, and the time from phone pickup to phone release. After those measurables variables were created and negative interval outliers were removed, medians were calculated for each of the variables in total. Normality of the data was ascertained through density plots, QQ plots, and Anderson-Darling tests. The data was confirmed to be not normally distributed and was skewed to the right. Based on these finding, non-parametric analytical tests were performed for the remainder of the analysis. The medians were then separated by year and density plots were run against each year to ensure they are similar. Upon that proof, Kruskal-Wallis tests were performed against each of the time points against the independent variable Year. All of the dependent variables showed significant differences in the medians between the two years. Effect sizes were tested to, per Professor Steve Draper of the University of Glasgow, measure the degree to which the certainty the results are not an accident but as an effect of the difference (Draper, 2020). In all the dependent variables the effect size was small. To discern if there is an independent variable which added to the year would yield a more significant effect. Adding in the independent variable WeekNo, using a Scheirer-Ray-Hare test, which is an extension of the Kruskal-Wallis test. In completing the tests, the p-value for each dependent variable was 0, demonstrating a significant difference between the population means.

**Findings:**

After completing the tests and comparing the medians for the weeks between the years for the data sets, there was a significant rise in three of the four dependent variables around the 12-week mark. The three dependent variables which showed the

The final dependent variable, from the time the call is first dispatchable to the time the first unit was assigned to the call, showed more of a marked rise at different points within the year.