**Report 1: Temporal Analysis**

Summary

The purpose of this report is to provide deeper insights to when instances of domestic violence (DV) occur. This report considers different timeframes to analyze the frequency of DV over the course of 2021, by month and seasons throughout the year, by the days of the month and seasons, and even by the hour of the day.

Descriptive Analysis

Each figure has a corresponding tabular format attached at the end of the report. Visuals are rendered with Plotly.

Figure 1

DV daily distribution per month


Figure 1 shows the distribution of DV instances per the 12 months. Each data point considered is the number of DV instances that occurred on that given day. Outliers are marked by dots above or below the individual plots. For example, the median daily number of occurrences in January was just over 150, whereas the max daily frequency for January was 303 instances reported. The next highest reported DV instances in January was 203, also an outlier. The January 1 instance is also 67 instances greater than the next most frequent daily total over the entire year (by 2 – 3.94 standard deviations depending on month) suggests that the number for January 1 is more likely due to reporting inconsistencies per precinct rather than by random chance.

Table 1 contains the data that was used for this plot. Below are the descriptive statistics for this data distribution that are used in the visualization in figure 1.

Table

Description automatically generated

Figure 2

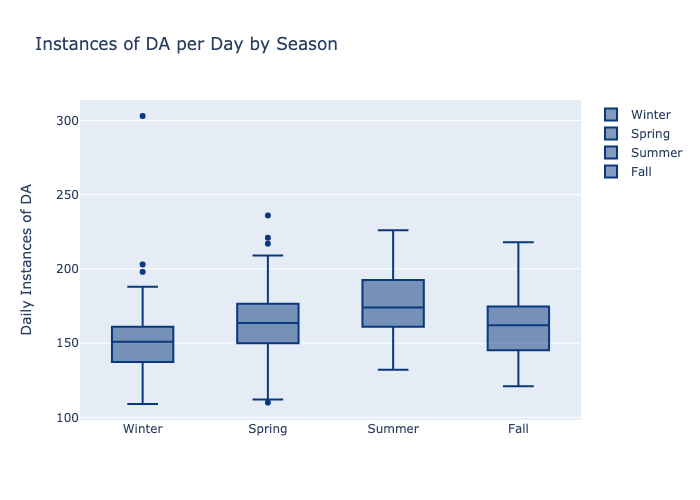
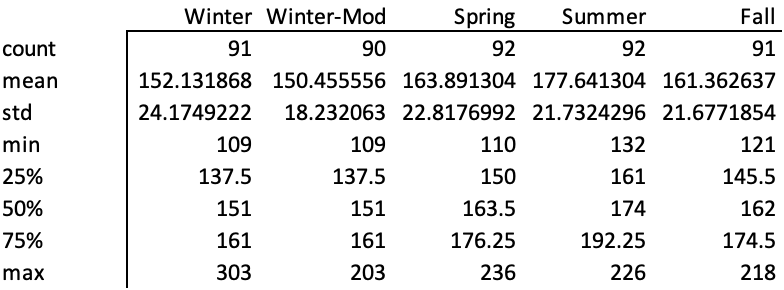


Figure 2 represents the same data as figure 1; however, the data are grouped here by quarterly season rather than by month (i.e. Winter contains months January, February, and December; Spring the next 3 months, and so on). Visually, there is a clear trend that shows instances of DV becoming more frequent during the summer months. The increase in the summer months is statistically significant. The process is detailed below.

Below are the descriptive statistics for this data distribution. Comparisons were drawn to determine significant differences in Season-by-Season occurrences of DV. Immediately, we can see how great an impact the outlier of January 1 creates (Winter-Mod omits January 1).



In order to evaluate the significance of the differences between seasons, the data were assessed for normality in order to ensure appropriate tests were used. Below shows two tests for normality, the Shapiro-Wilcox and the Kolmogorov-Smirnov (KS) tests. It should be noted that for the KS tests, given the lack of prior data or formal model of DV at the time of this report, simulations of normal data were created per season that matched a given season’s mean and standard deviations for comparison. Below shows the results of the tests.

Graphical user interface, text, application

Description automatically generated

Omitting the single extreme outlier in Winter, Summer is the only season that fails one of the tests to assume population normality, that is the distribution of instances of DV is normally distributed per season. Further analyzing Summer to evaluate the effects of randomness 10,000 simulations were run on the KS test for Summer, of which Summer yielded the same KS result in 9,918. For the purpose of this report, the omission of January 1 was maintained, and all seasons were assumed to be reasonably normal.

*Running One-Way ANOVA for comparing Seasonal means with Tukey HSD post-hoc*

Text

Description automatically generated

In short, all Seasons are pairwise significantly different from one another except for Fall and Spring, which are statistically significantly similar.

Figure 3

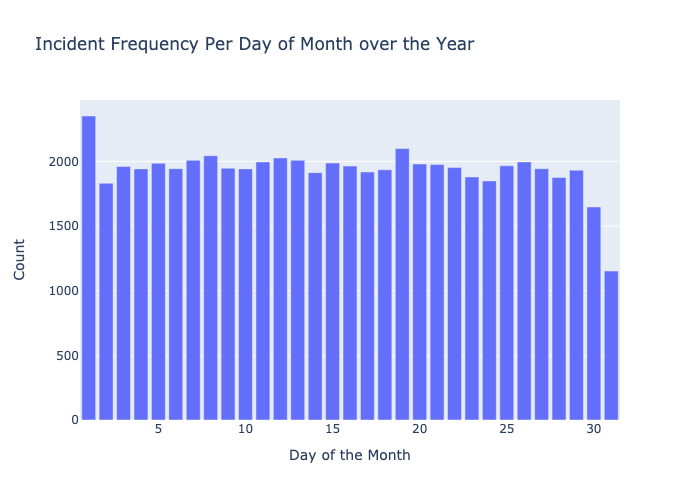


Figure 3 illustrates the frequency of date that DV occurred on. There is a dip on the 31st of the months since roughly half the months do not have a 31st.

Table

Description automatically generated

Figure 8 illustrates this frequency grouped by day of the week rather than by day of the month.

Figure 4

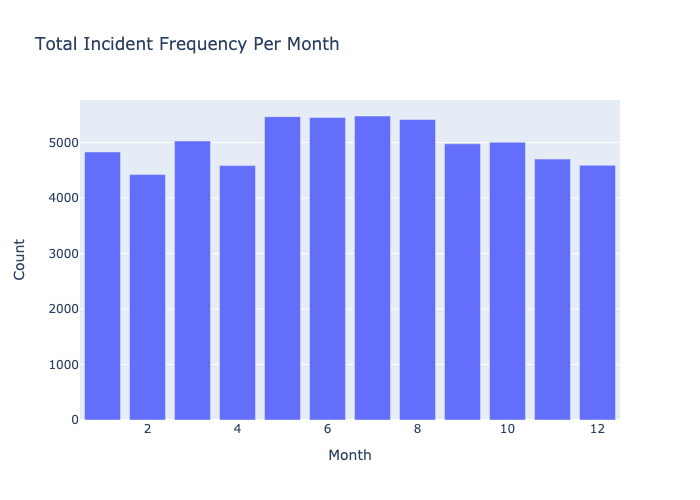
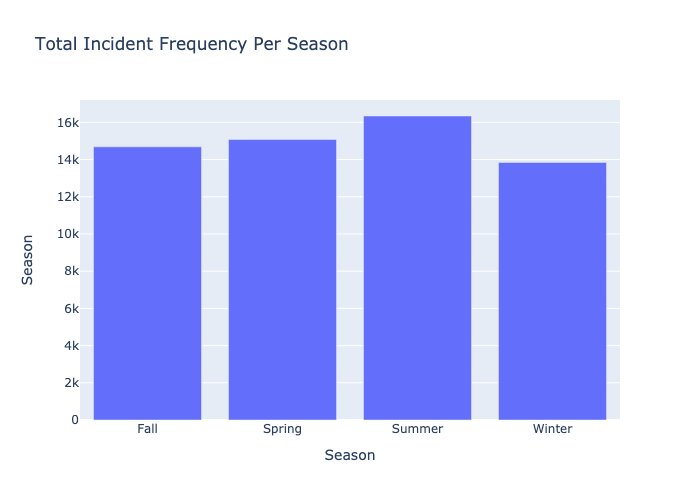


Figure 4 shows raw frequency per month of instances of DV. Below depicts the monthly average over the year. No adjustments were made here based on the length of month.

Table

Description automatically generated

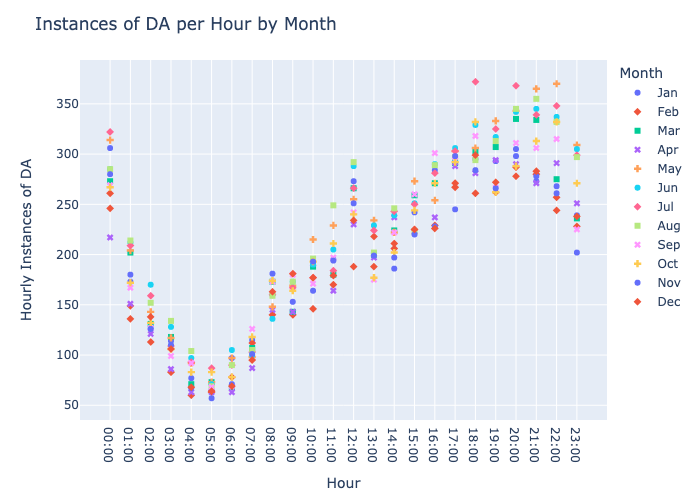
Figure 5



Table

Description automatically generated

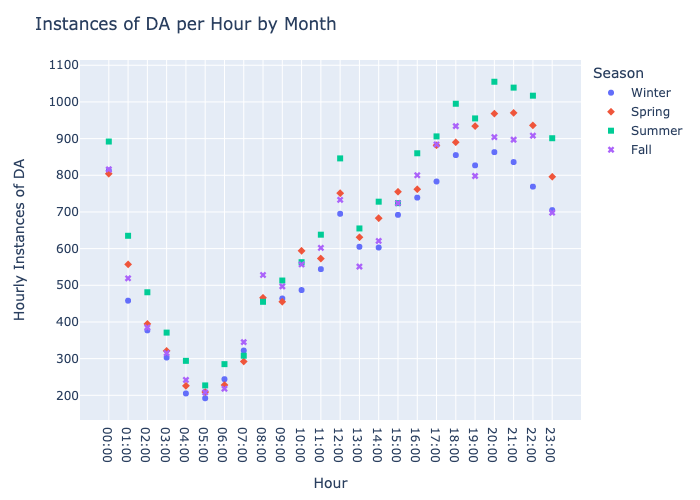
Figure 6



Table

Description automatically generated

Figure 7



Table

Description automatically generated

Figure 8

Chart, bar chart

Description automatically generated

Saturday and Sunday appear to lead the pack as far as most frequent day of the week for DV instances to occur. A chi-squared goodness of fit test returns a statistically small p-value thus rejecting a null hypothesis that the occurrences of DV are roughly the same per day of the week.

Text

Description automatically generated with medium confidence

Figure 9.1

Chart, line chart

Description automatically generated

Figure 9.2

Chart, line chart

Description automatically generated

Figure 9.3

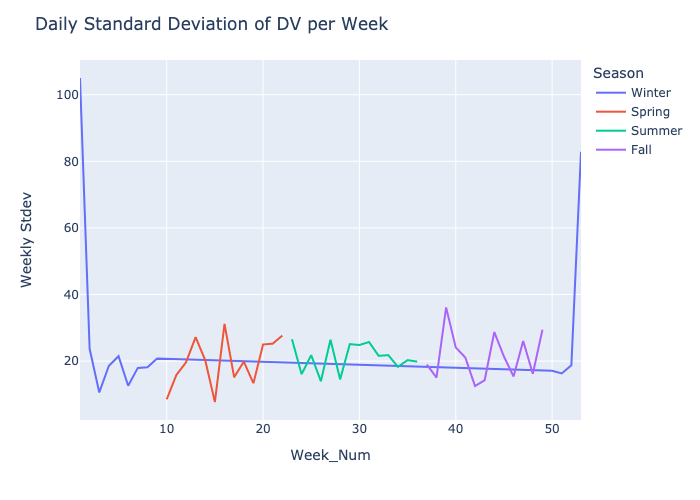


Figure 9.4

Chart, line chart, histogram

Description automatically generated

Figure 9.1 depicts the weekly total frequency of DV instances over the course of the 53 weeks of the year 2020, color-coded by Season. Figures 9.2, 9.3, and 9.4 show the respective weekly metrics over the course of the year. Summary statistics of the weekly analysis are below.

Table

Description automatically generated

Figures 9.3 and 9.4 strongly corroborate the effects of including January 1’s count of 303 DV instances. Below are images 9.1-4 without color coding.

Figure 9.1

Graphical user interface, chart, application, line chart

Description automatically generated

Figure 9.2

Chart, line chart

Description automatically generated

Figure 9.3

Chart, line chart

Description automatically generated

Figure 9.4

Graphical user interface, chart, application, line chart

Description automatically generated