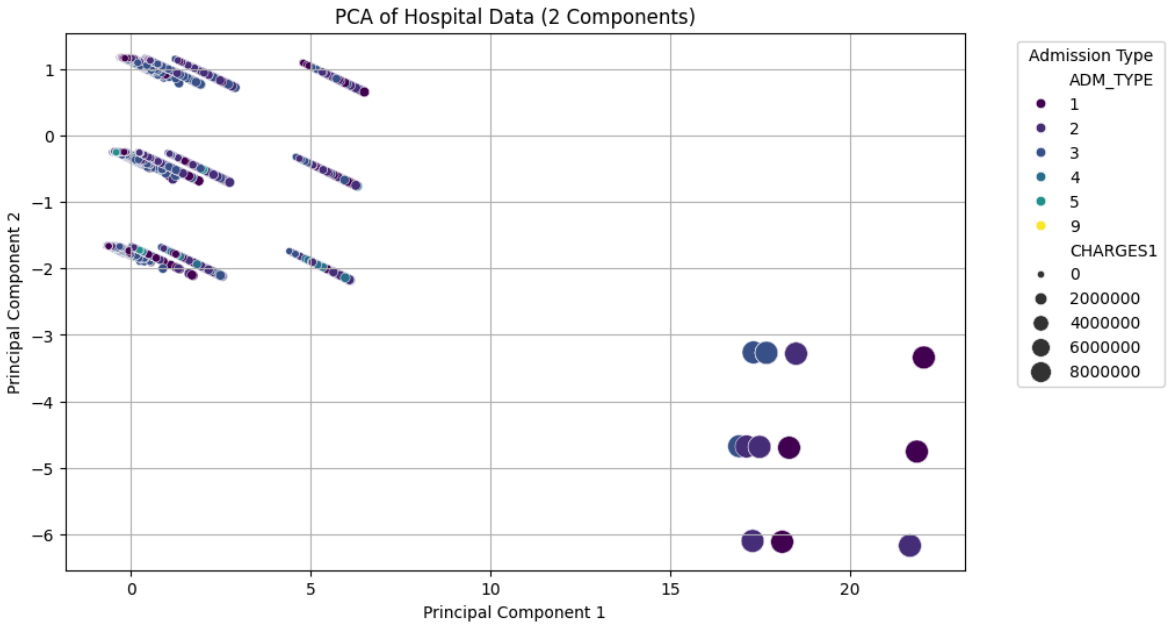
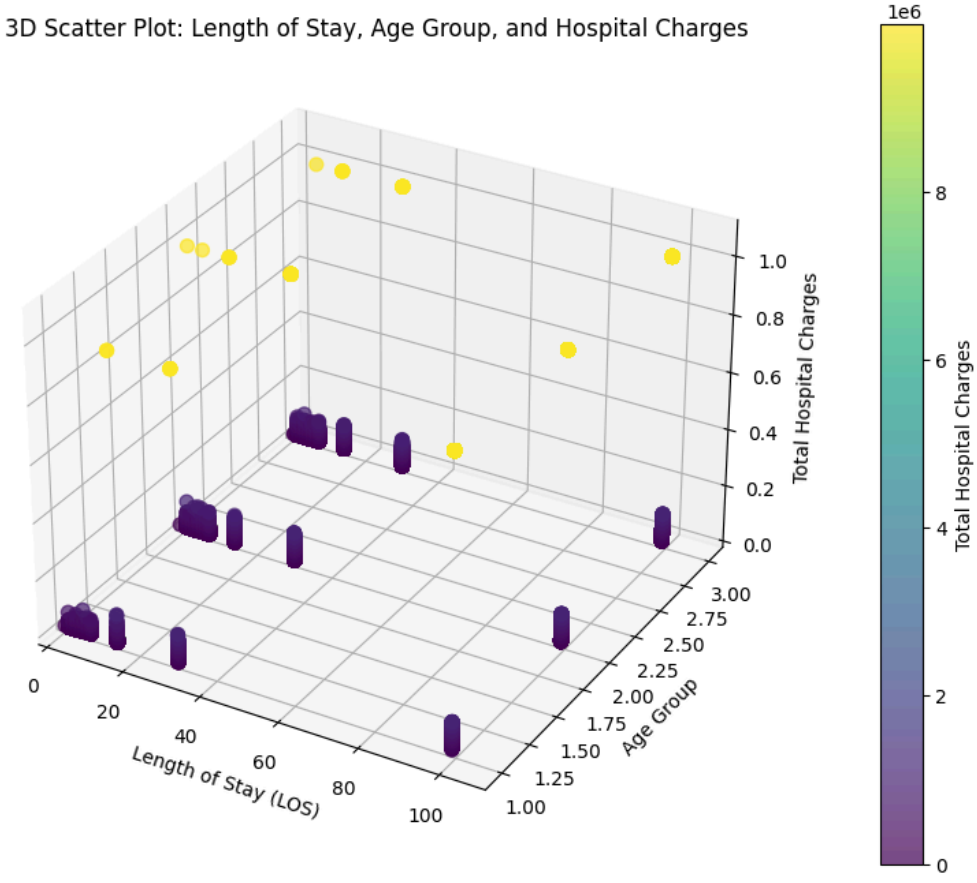
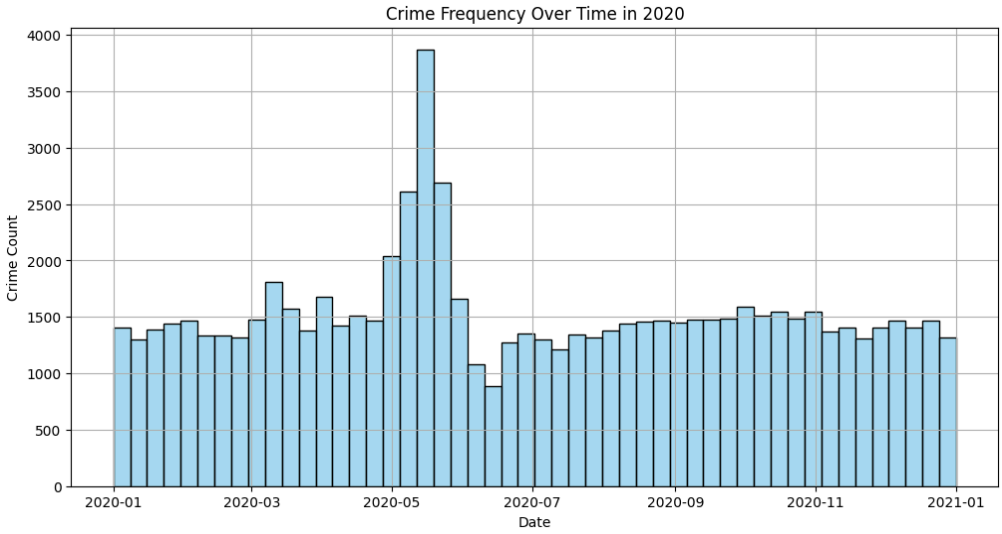


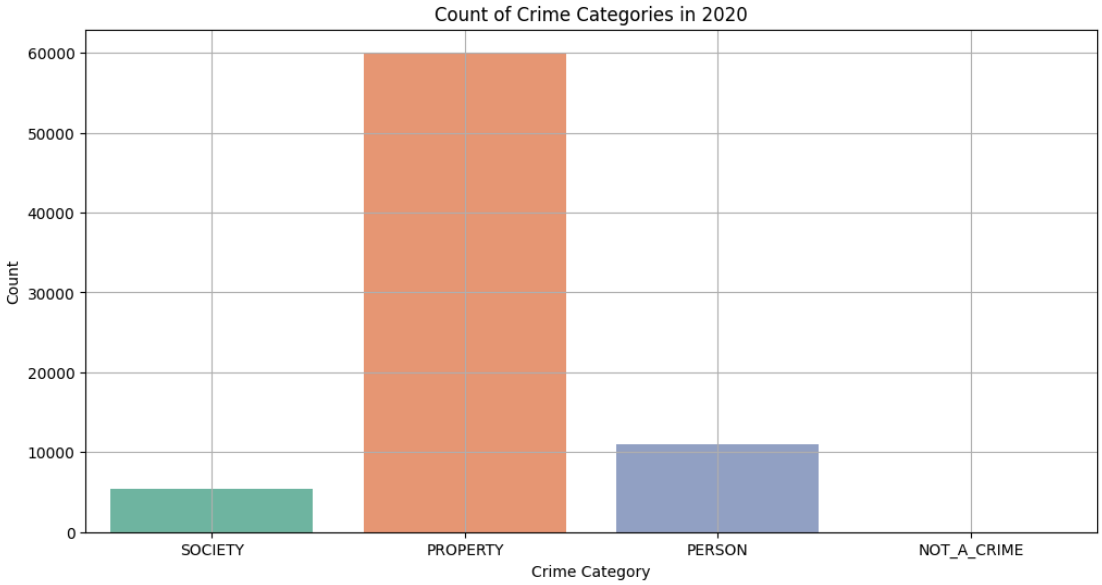


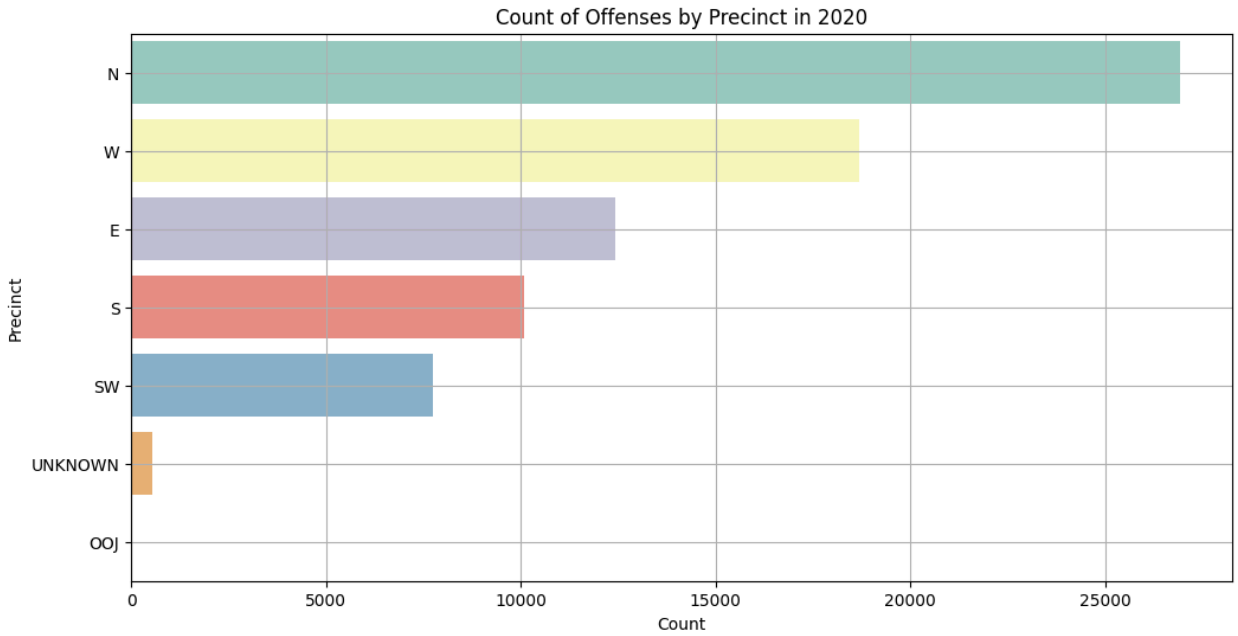
We chose the PCA graph to reduce the high-dimensional hospital data into two principal components, capturing the maximum variance and allowing us to visualize complex patterns.

The graph shows how hospital charges and length of stay (PC1) and age groups (PC2) influence patient clustering, revealing distinct groups based on their hospital resource usage and demographics.The 3D scatter plot was selected to explore the interplay between length of stay, patient age, and hospital charges, providing insights into how these factors interact and influence overall expenditures.

Univariate Charts





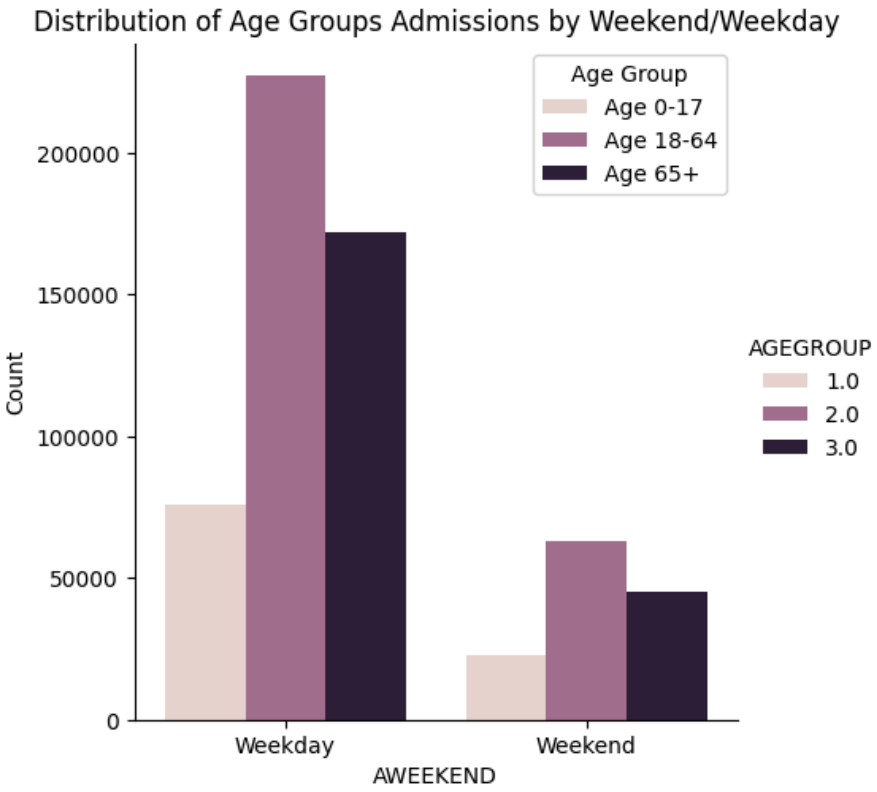


Histogram of Crime Frequency Over Time: This chart was chosen to show the temporal distribution of crime incidents throughout the year, providing insight into patterns and potential spikes in crime activity.

Count Plot of Crime Categories: This visualization was selected to categorize and quantify the different types of crimes, highlighting which categories are most prevalent in the dataset.

Count Plot of Offenses by Precinct: This chart was used to analyze the geographical distribution of crimes, showcasing which precincts experience higher crime rates and helping to identify areas with concentrated crime activity.

Bivariate Charts

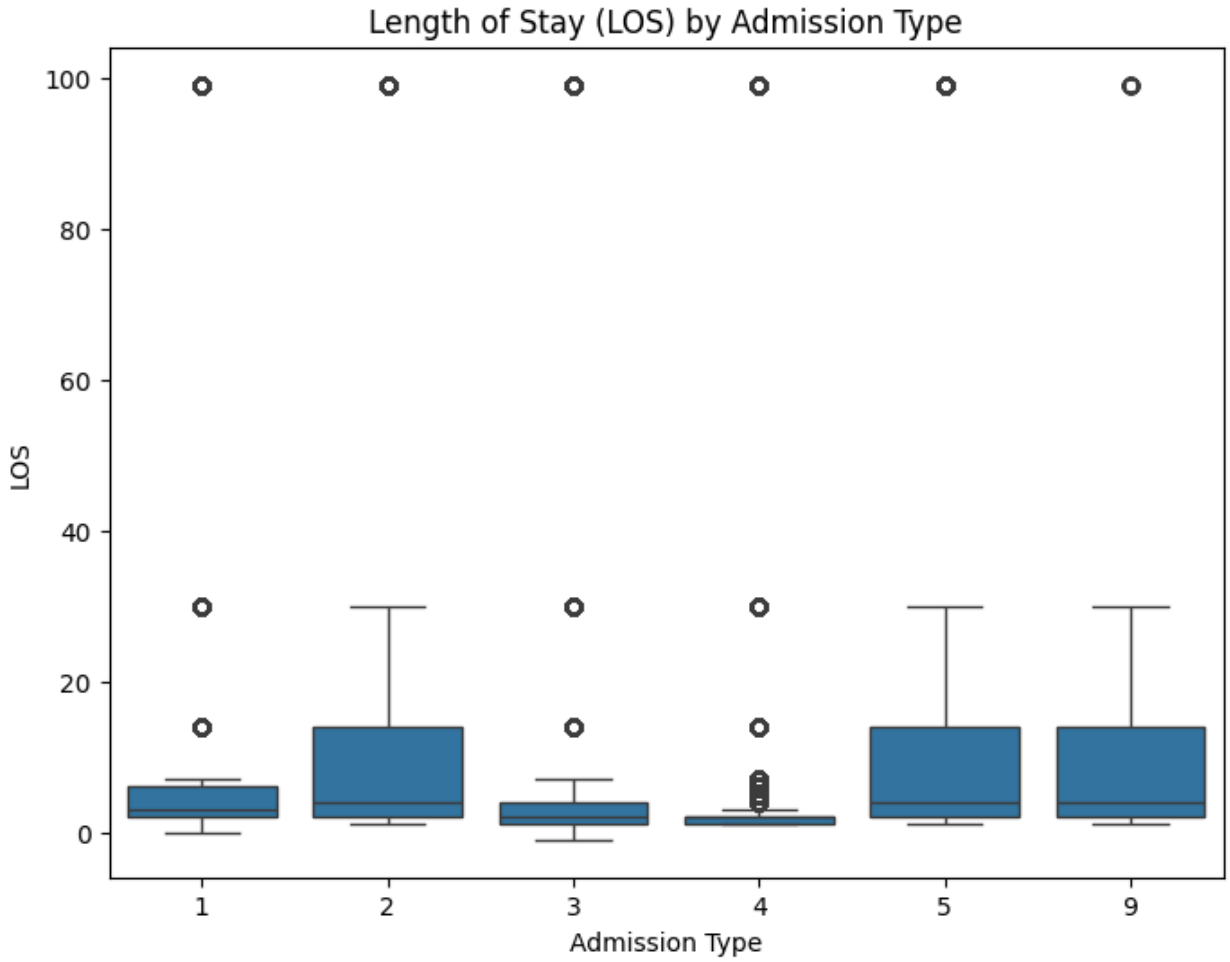


Distribution of age groups admissions by weekend/weekday displays the portion of all admissions across weekdays or the weekend. The distribution looks as expected, with the majority of admissions falling into the largest age group and occuring on weekdays which are the bulk of the week.

Contingency Table for Length of Stay and Admission Type

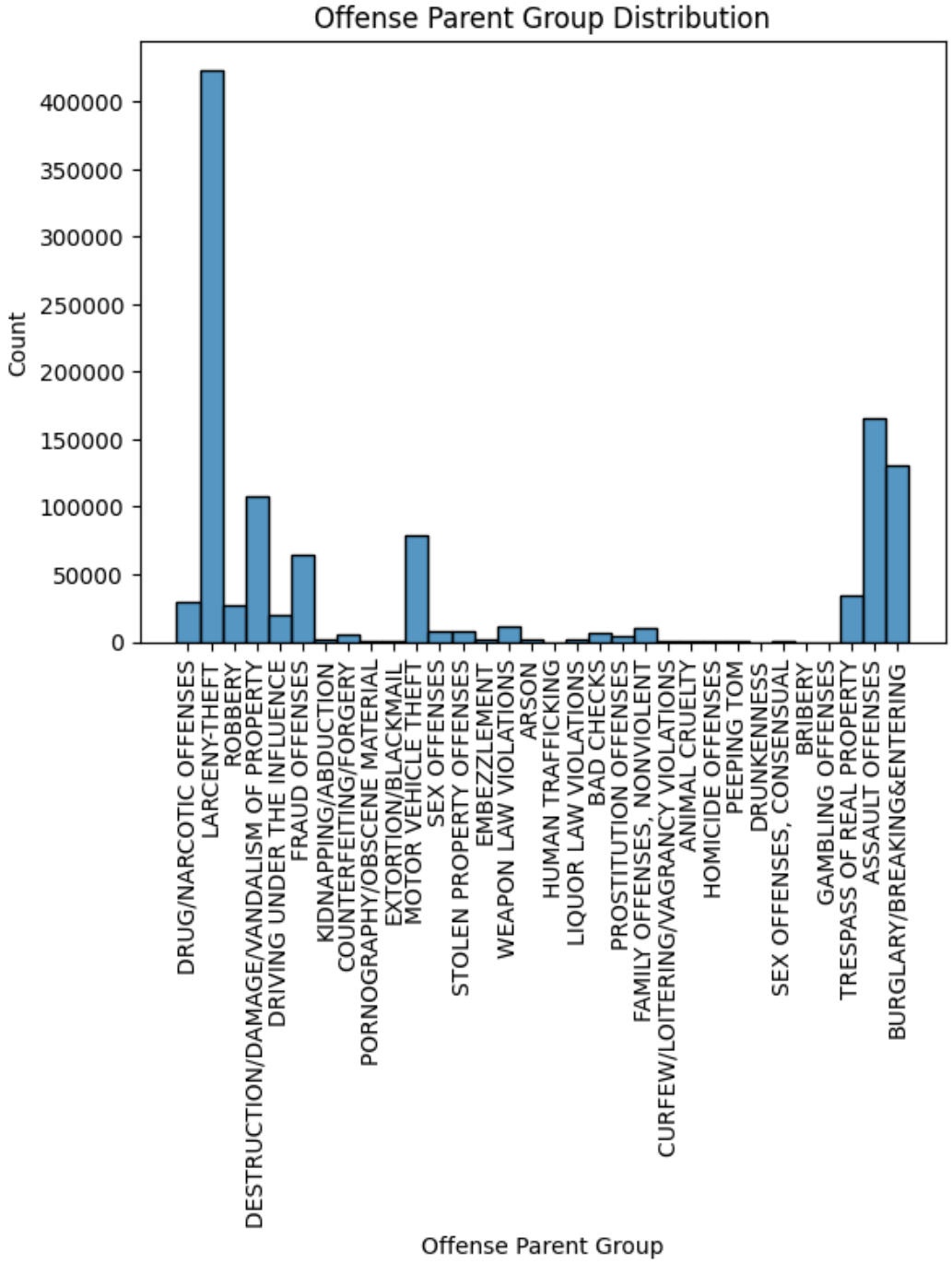
ADM_TYPE	1	2	3	4	5	9
			LOS			
-1.0	0	0	1	0	0	0
0.0	3	0	0	0	0	0
1.0	54735	10667	46425	36199	1154	102
2.0	59521	12837	39245	27248	981	113
3.0	46366	8977	19201	5563	853	105
4.0	33080	6607	10398	1567	751	96
5.0	23529	5255	6139	731	602	56
6.0	16916	4036	3930	387	482	30
7.0	12932	3546	3290	344	339	21
14.0	35577	11821	10134	1414	1152	128
30.0	13128	6581	8589	1395	518	118
99.0	4300	2851	1592	980	210	39

The contingency table of Admission Type and Length of Stay allows us to view the distribution of each categorial variable in relation to eachother, which can help us understand if there is underlying multicolinearity.

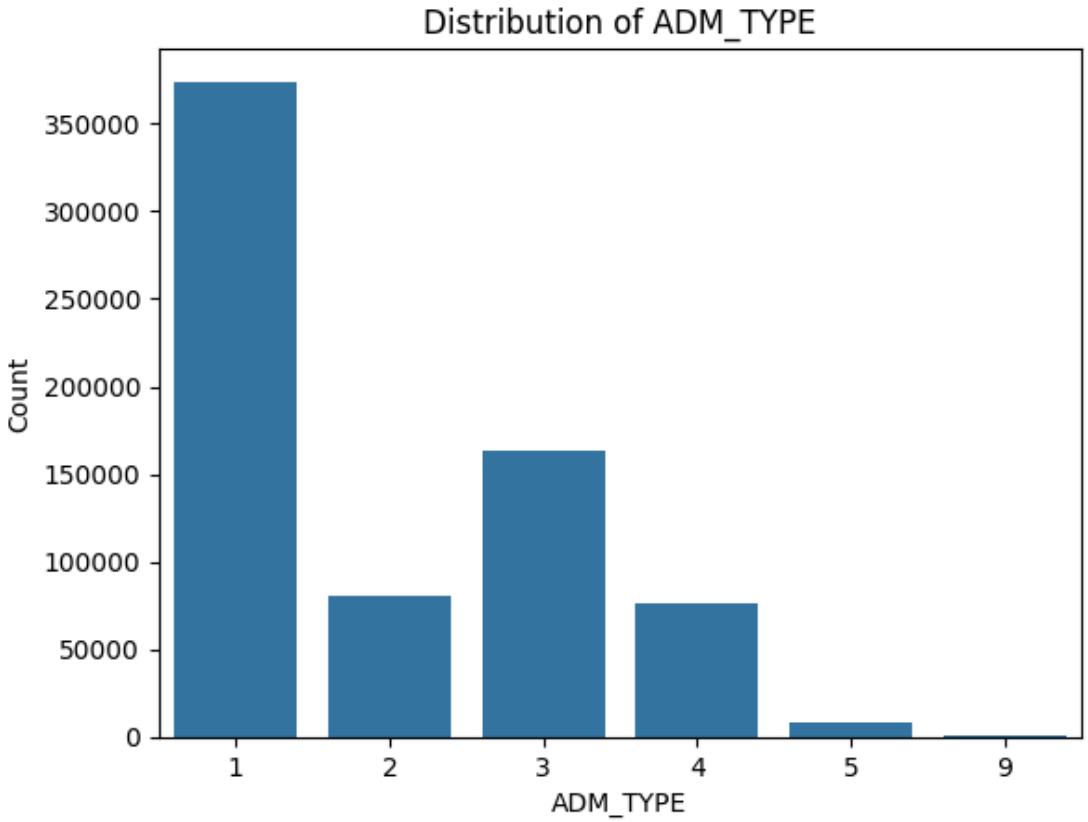


The boxplot of LOS by Admission Type reveals the distribution and variation of hospital stays across different admision categories. This can help identify the effect admission type has on LOS and if further analysis is needed for this relationship.

Graphs that Visualize Issues with Data



This histogram shows the distribution of crime in Seattle by offense parent group. Because larceny/theft has over 2x the amount of offenses as the second largest category, and may not directly correlate with hospital intakes, it points to an issue with the data. We may need to research and consult outside sources to identify crimes that drive hospital admissions. This is important if our goal is predictive modeling so we don't correlate this outlier to hospital admissions.



The histogram shows the distribution of ADM_Type by category. The majority of admissions are type 1. List of categories and numbers of admission type: 1: "Emergency" 2: "Urgent" 3: "Elective" 4: "Newborn" 5: "Trauma" 9: "Information not available"

This points to an issue we may encounter, as elective admissions will likely not be correlated with crime data. We will also have to determine what to do for admissions where information is not available.