

1. How many bookings of arrestees were made in 2018?

Total bookings of arrestees: 104277

2. How many bookings of arrestees were made in the area with the most arrests in 2018?

Central: 10951

3. What is the 95% quantile of the age of the arrestee in 2018? Only consider the following charge groups for your analysis: Vehicle Theft, Robbery, Burglary, Receive Stolen Property.

Quantile_95 = 52.0

4. There are differences between the average age of an arrestee for the various charge groups. Are these differences statistically significant? For this question, calculate the Z-score of the average age for each charge group. Report the largest absolute value among the calculated Z-scores. (Only consider data for 2018. Do not consider "Pre-Delinquency" and "Non-Criminal Detention" as these charge groups are reserved for minors. Exclude any arrests where the charge group description is not known.)

The largest absolute value among the calculated Z-scores: 2.128

5. How many arrest incidents occurred within 2 km from the Bradbury Building(https://en.wikipedia.org/wiki/Bradbury_Building) in 2018? [Use (34.050536, -118.247861) for the coordinates of the Bradbury Building. For simplicity, please use the spherical Earth projected to a plane(https://en.wikipedia.org/wiki/Geographical_distance#Spherical_Earth_projected_) equation for calculating distances. Use the radius of the Earth as 6371 km. Note, some arrest records are missing location data and the location is listed as (0, 0). These records should not factor in your calculation.]

Arrest incidents within 2 km of Bradbury Building in 2018: 11461

6. How many arrest incidents were made per kilometer on Pico Boulevard during 2018? For this question, we will need to estimate the length of Pico Boulevard, which mostly stretches from east to west. To estimate the length of Pico Boulevard: Consider all location data which the listed address mentions "Pico". Remove outliers by filtering out locations where either the latitude or longitude is 2 standard deviations beyond the mean of the subset of identified points. To estimate the length, calculate the distance from the most western and eastern coordinate points. As before, use the simplified flat surface equation. Once you have estimated the length of Pico Boulevard, you can proceed to report the number of arrest incidents per kilometer on Pico Boulevard in 2018.

Incidents were made per kilometer on Pico Boulevard: 16.377834818574755

7. Some types of arrest incidents in certain areas occur at a highly disproportionate rate compared to their frequency city-wide. For example, let's say that the rate of larceny arrests (charge group code 6) is 1% in Devonshire (area ID 17). This rate may appear

low but what if larceny arrests constitute 0.1 % city-wide? The ratio between these two probabilities is 10 and we can say that larceny occurs unusually often in Devonshire (Note, these numbers were made up for illustration purposes). Calculate this ratio for all charge group code and area ID pairs. You can view this ratio as the ratio of the conditional probability of an arrest incident of a charge group code given that it occurred in an area ID to the unconditional probability of the arrest incident of a charge group. Report the average of the top 5 of the calculated ratio. Consider all records prior to January 1, 2019. Some arrest incidents don't have a charge group code. These records should not be considered in your analysis. Arrest incidents for charge group code 99 should not be considered in your analysis.

Average of top 5 ratios: 3.515076379865202

8. Felony arrest incidents have been dropping over the years. Using a trend line (linear estimation) for the data from 2010 and 2018 (inclusive), what is the projected number of felony arrests in 2019? Round to the nearest integer. Note, the data set includes arrests for misdemeanor, felonies, etc.

Projected felony arrests in 2019: 31038