

A Perspective of Property Value and Police Proximity to Crime Rates

At the San Diego Blue Line Trolley

Slides by Emily Chen

Last updated on July 2024

Introduction

Why are we doing this project?

The San Diego blue line trolley station opened on November of 2021, and it became well known for its speedy and reliable method of transportation. With the University of California San Diego hosting multiple blue line trolley stops on its campus and La Jolla being known as an affluent neighborhood, we were curious and wanted to investigate the crime rates around these stops. The two factors that we were particularly interested in understanding crime rates around these stops were **property value** and **police proximity**.

Background

Why property value and police proximity?

We believe that using property and police proximity will be an interesting and useful investigation because of the following prior works. One study have found that **crime, housing prices, neighborhood satisfaction and desire to move have a relationship**³. Furthermore, there is evidence that when studies attempt to find causal relationship between arrest rates, they generally find that for crimes, **higher level of police activity implies lower crime rate**⁴. Then, based on these studies, we believe that examining the relationship between crime rates, property values, and police proximities around these stops can advance our understanding of the crime patterns in the area of focus, the San Diego Blue Line trolley station.

Research Question

What are we trying to answer?

Do **La Jolla trolley stops** on the Blue Line experience higher rates of criminal incidents compared to **non-La Jolla trolley stops** within San Diego? If so, do **property value** and **proximity to police stations** correlate with these these patterns?

Data Collection

- 1) [San Diego crimes report](#), 2) [Zillow home values](#) 3) [San Diego police station](#)
- All three data sets were publicly available from credible sources
- Additionally, we used other publicly available files for plotting charts, and we acknowledge the purpose of our usage in our code.

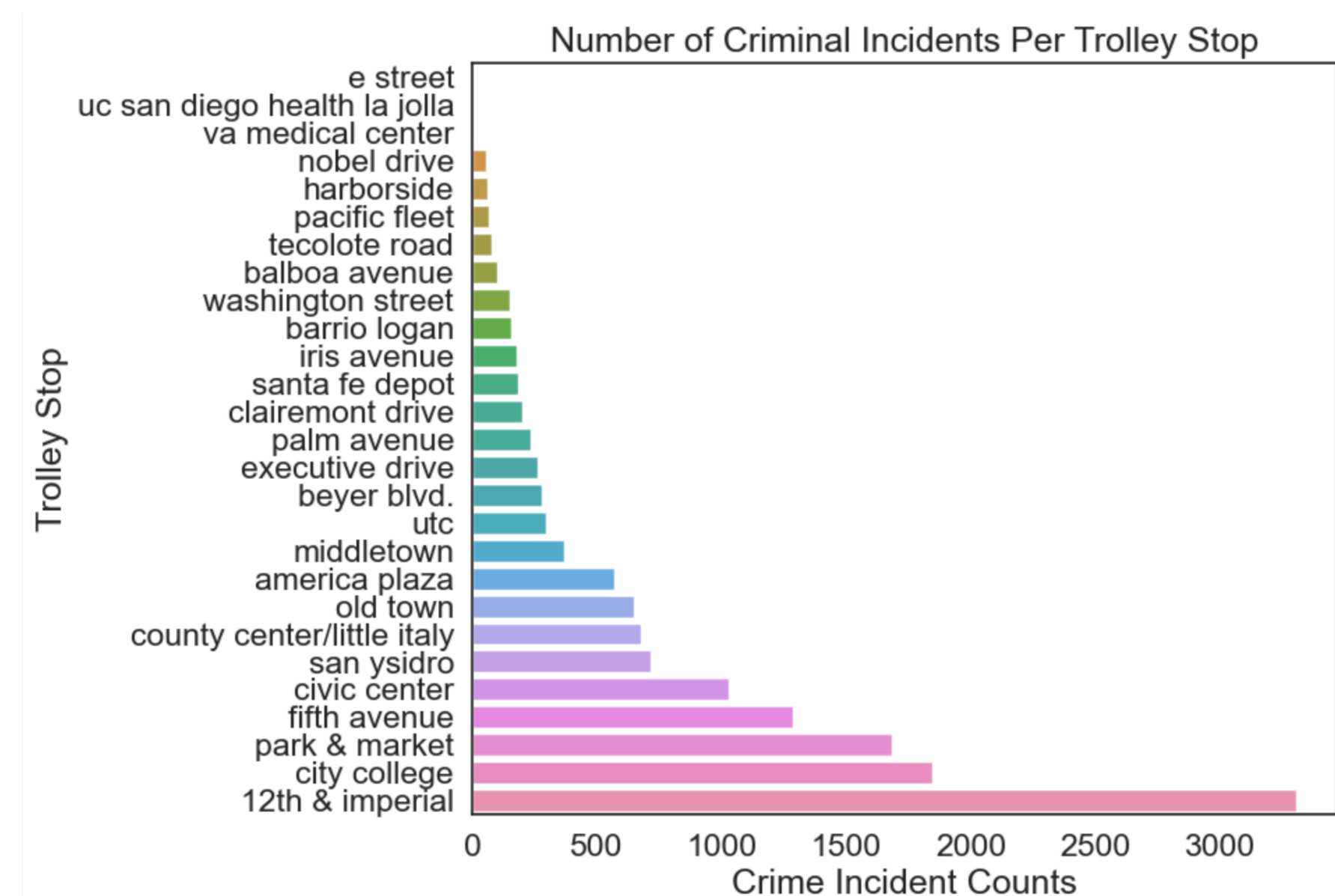
Pitfalls and Obstacles

What were the limitations and how did we address them?

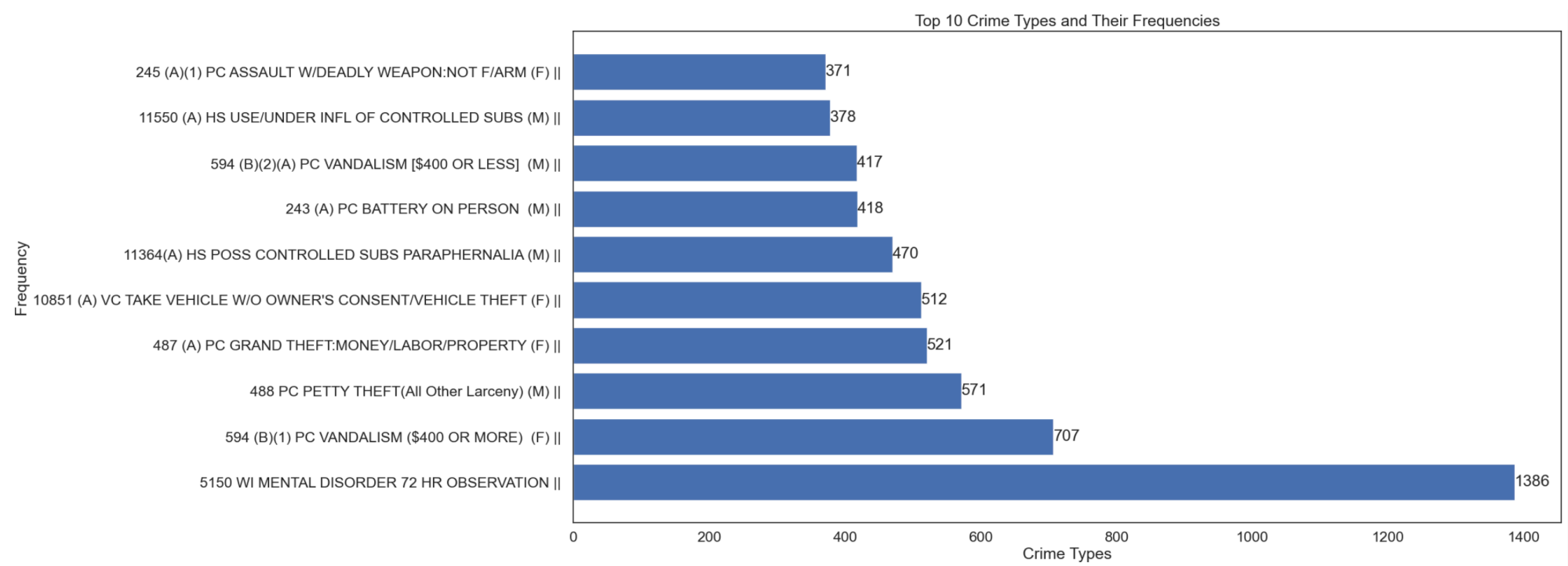
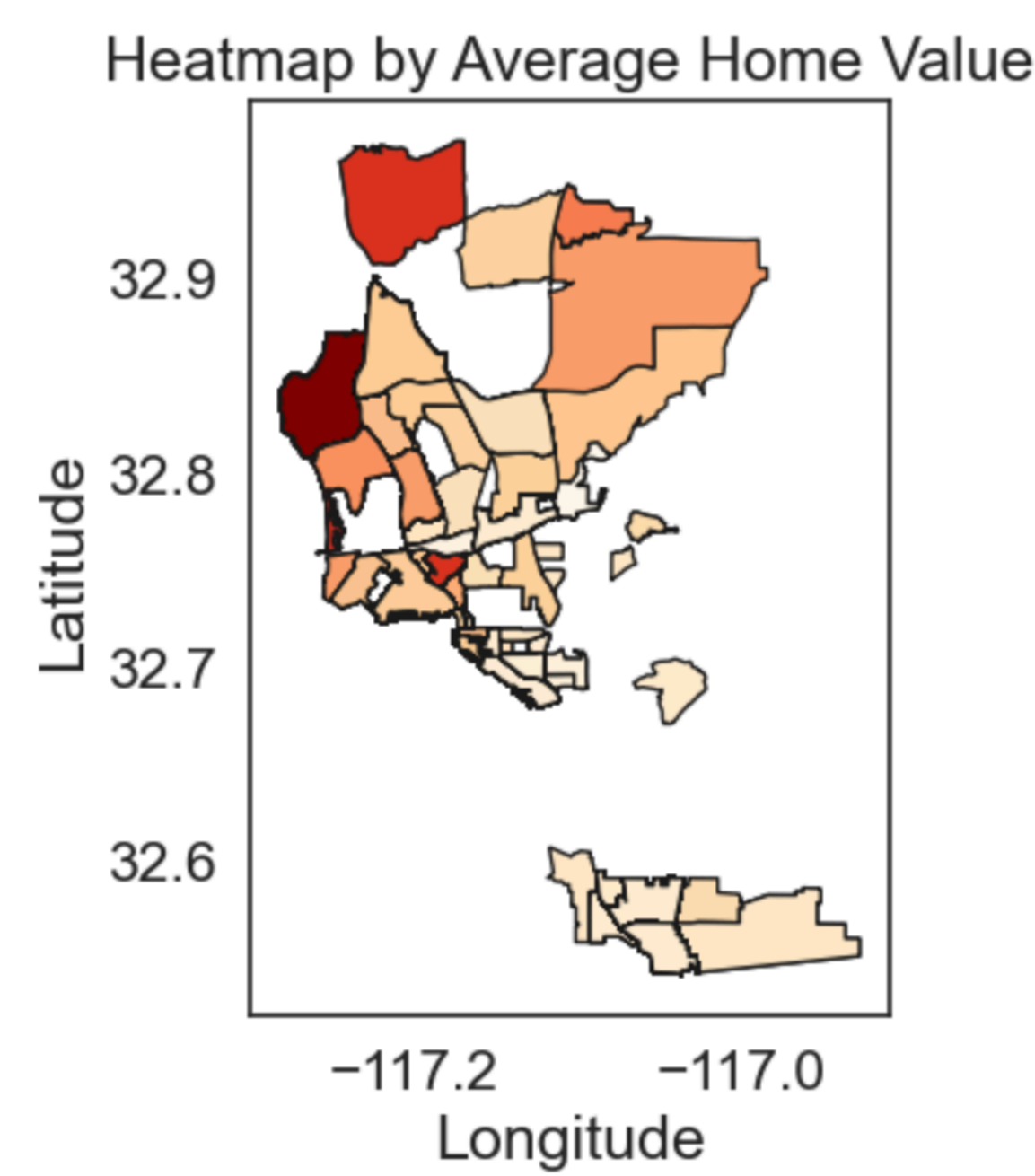
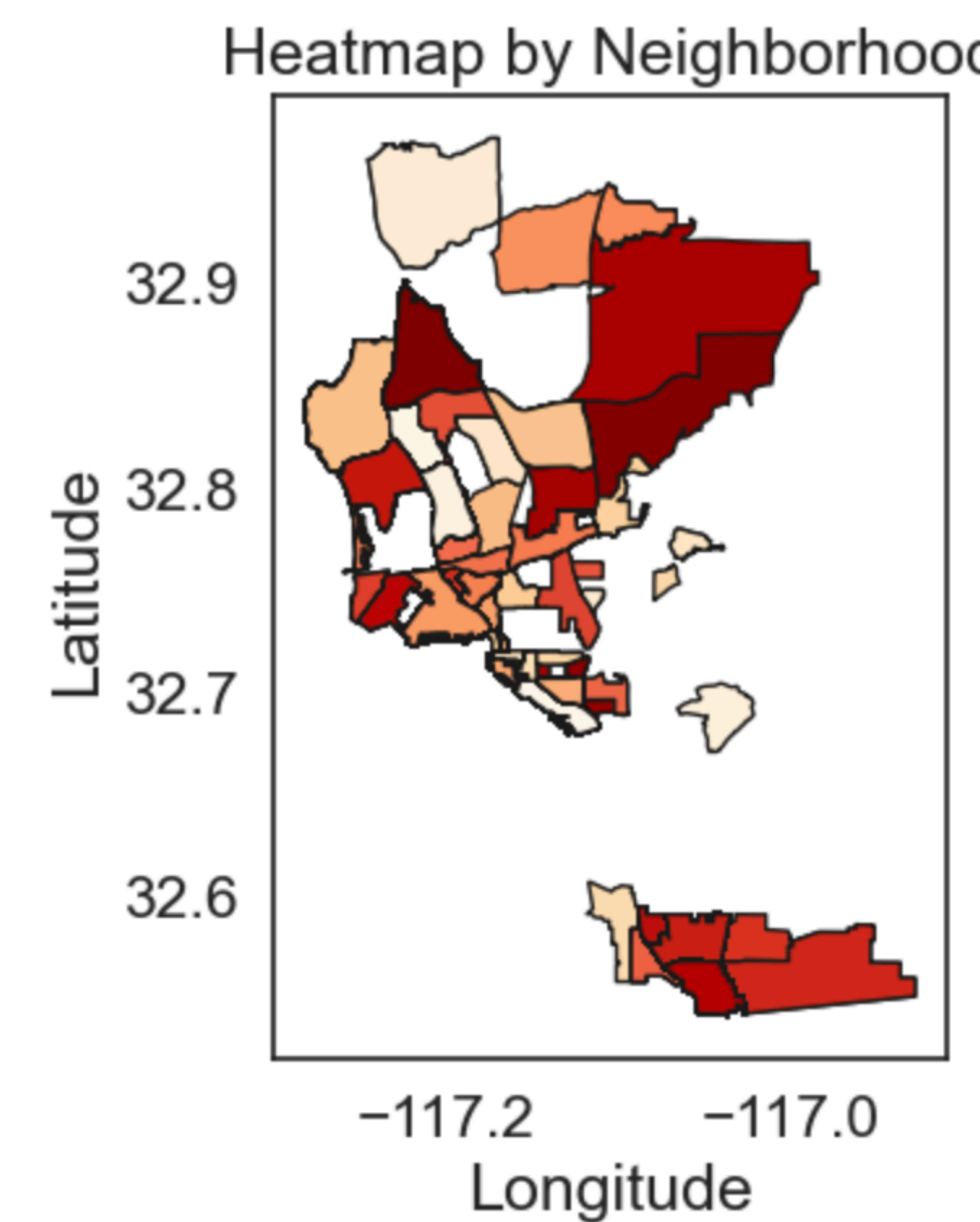
- Zip codes itself did not serve as good proxy for locating crime rates.
- Levels of details about crime locations were scattered among the datasets.
- Zip codes were replaced by San Diego neighborhoods
- Level of details were unified through defining San Diego neighborhoods
- Dataset did not include Chula Vista because it is not part of the San Diego County
- Make the most out of the publicly available data and resume understanding La Jolla versus non-La Jolla regions within the San Diego County instead

Methodology

Data cleaning, Exploratory Data Analysis, Statistical Inference, Geospatial Analysis



12th & Imperial street seem to have over 3000 crime incidents



$$a = \sin^2 \left(\frac{\Delta \text{lat}}{2} \right) + \cos(\text{lat}_1) \cdot \cos(\text{lat}_2) \cdot \sin^2 \left(\frac{\Delta \text{lon}}{2} \right)$$

$$c = 2 \cdot \text{atan2} \left(\sqrt{a}, \sqrt{1-a} \right)$$

$$d = R \cdot c$$

Haversine Distance Formula

Privacy and Ethics

Data collection: Some potential ethical issues around this could be when we use “neighborhood” to merge crimes and homes value data together, there are potential mis-measurements and inconsistencies around how “near” we’re studying around the trolley stops. For attainability reasons, studying around La Jolla neighborhood which stretches approximately 7 miles (11 km) is a better choice than studying around zip codes 92037 which covers approximately 13 miles (20 km) of land⁵.

Data analysis: A handful of exploratory data analysis were conducted to find graphical relationships between variables and grow the diversity of viewpoints of our investigation. We preserve fairness in our permutation test by refraining to make assumptions about the data. By creating random samples non-La Jolla crime rates and La Jolla crime rates, we could independently calculate each distribution’s statistic, and draw conclusions.

Use and revision: The investigation serves to understand whether or not property value and police proximity were related to crime rates. Thus, we refrained from direct comparison crime rates and neighborhoods because we are aware that this could potentially abet the stigmatization around certain neighborhoods. Moreover, the variable “neighborhood” encompasses many factors that we have not yet understood in this context. Thus, it can give confounding conclusions if misapplied or misinterpreted.

Conclusions

The investigation of the research project attempts to use inferential analysis to determine whether La Jolla property value and crime rates have a statistical significance compared to Non-La Jolla property value and crime rates. From our findings, it seems to be that the two distributions are statistically significantly different.

Acknowledgement

This was originally a final course project that I took at UCSD, and it was done alongside with four other group members. Some notable changes I've done afterward were the addition of this slide deck that serves to communicate a summary of what was accomplished and clarifying ideas in the Jupyter notebook.

References

- (1) BLiggett, R., Loukaitou-Sideris, A., & Iseki, H. (2003). Journeys to Crime: Assessing the Effects of a Light Rail Line on Crime in the Neighborhoods. *Journal of Urban Affairs*, 25(2), 165-184.
- (2) Block, R., & Block, C. (2000). The Bronx and Chicago: Street Robbery in the Environs of Rapid Transit Stations. *Journal of Transportation and Statistics*, 3(3), 29-36.
- (3) Taylor, R. B. (1995). The Impact of Crime on Communities. *The Annals of the American Academy of Political and Social Science*, 539(1), 28-45.
- (4) Weatherburn, D. (2001). What causes crime?
- (5) MapsZipCode. (n.d.). *ZIP code 92037 population*. Zip Code 92037 Population.