



---

Name: **Iman Mousaei**  
Course: **Data Science**  
Assignment **1**

---

Student Number: **98222099**  
Date: October 24, 2021

## 1 Task 1: Airbnb

### 1.1 What can we learn about different hosts and areas?

Looking at Plot 1-2, we can learn that Manhattan and Brooklyn have most of the hosts.

### 1.2 What can we learn from predictions?

Looking at Plot 1-2 and 2-1, we see that Manhattan and Brooklyn are also the most expensive ones(Although they have most of the hosts).

### 1.3 Which hosts are the busiest and why?

Since this column is not featured in the dataset, I shall estimate with my gut feeling: We can multiply columns that have direct correlation(to the power of correlation's strength) and divide by columns with reverse correlation(to the power of correlation's strength). That results in the following formula:

$$Busyness \sim \frac{roomCount^3 * availability^2 * numberOfReviews}{price^4}$$

So we calculated busyness with that and the result was Plot 3-1 which shows that most of them are equally busy.

### 1.4 Is there any noticeable difference in traffic among different areas and what could be the reason for it?

As we can see in Plot 4-1 the variance of price and roomCount very low so as we can expect, they're almost equally busy.

## **2 Task 2: FBI Crime Reports**

### **2.1 Which crimes are most common?**

According to Plot 1-2: THEFT/LARCENY, DRUGS/ALCOHOL VIOLATIONS, ASSAULT

### **2.2 In which zip codes are crimes more likely to occur?**

According to Plot 2-1: 40214, 40211, 4.203, 40219, 40212, 40215, 40216 So mostly zip codes in format 4021x

### **2.3 Which crimes take the longest to report?**

According to Plot 3-1: Sex Crimes

### **2.4 Is there a trend of some crimes increasing and others decreasing in number over these five years?**

According to Plot 4-1: Only WEAPONS is strictly increasing. Others have their ups and downs.

## References

- [1] [tex.stackexchange.com](https://tex.stackexchange.com)
- [2] [pandas.pydata.org](https://pandas.pydata.org)
- [3] [stackoverflow.com](https://stackoverflow.com)
- [4] [geeksforgeeks.org](https://www.geeksforgeeks.org)
- [5] [overleaf.com](https://www.overleaf.com)
- [6] [matplotlib.org](https://matplotlib.org)
- [7] [w3schools.com](https://www.w3schools.com)