FINAL PROJECT README

This project is completed by Zhaoyun Qin and Zengwenqi Lai.

Introduction

Our project aims to find out an important factor that affects the price of housing. Therefore, we collected the data on shooting incidents as an indicator of violent crime and information on housing sales in New York to find out their relationship.

The goal is to determine the relationship between change in crime rate and change in housing prices in New York City. We hypothesize that a change in the crime rate has a lingering effect, and we would like to verify this hypothesis.

File Explanation

Python Code

In the file PYTHON CODE, it contains 6 coding files.

[Data] Process Sales contains data cleaning codes for sales.

[Data] Process Shooting contains data cleaning codes for shooting.

[Data] Merge Training Data Frame contains codes using df join to merge the two data sets mentioned above. It then becomes the training set.

[Model] Linear Model 1 contains an ordinary OLS model while

[Model] Linear Model on Mean2 contains a OLS with mean value.

[Analysis] Sales contains information regarding the sales data set from first investigation and to produce figures on specific characteristics.

Report

In the file REPORT, it contains final report for this project. Also, inside the file REPORT, there is a file called Figs: it contains tables and images for our project.

Data Source

You can click the title to access

- Speculation Watch List
- NYC Calendar Sales
- DOF Summary of Neighbourhood Sales
- NYPD Shooting Incident Data