题目：

作者：胡逸萍（1710121282）

数据集摘要：

|  |  |
| --- | --- |
| 名称 | Crimes in Boston |
| 特征简介 | ”OFFENSE\_CODE””OFFENSE\_CODE\_GROUP””OFFENSE\_DESCRIPTION”  “DISTRICT””REPORTING\_AREA””SHOOTING””OCCURRED\_ON\_DATE”  ”YEAR””MONTH””DAY\_OF\_WEEK””UCR\_PART””STREET””Lat”  “Long””Location” |
| 记录数 | 319073 |
| 分析目标 | 影响犯罪率的原因 |
| 分析思路及方法 | 1. 读取数据 2. 筛掉不完整数据 3. 数据可视化：按类型画出直方图，计算各类型环比 4. 犯罪的分布：根据经纬度以热力图的形式呈现 |

1. 简介

**来源:** Kaggle

**数据集信息:**

犯罪事件报告由波士顿警察局（BPD）提供，以记录BPD官员应对的事件的初步细节。这是一个数据集，包含来自新的犯罪事件报告系统的记录，其中包括一组简化的字段，集中于捕获事件类型以及事件发生的时间和地点。

记录从2015年6月14日开始，持续到2018年9月3日。

**属性信息:**

每行代表一个犯罪报告，包括：犯罪类型、日期和时间以及地点。

OFFENSE\_CODE:案件代码

OFFENSE\_CODE\_GROUP:案件代码所属分组

OFFENSE\_DESCRIPTION:案件类型

DISTRICT:地区

REPORTING\_AREA:报告区域

SHOOTING:是否有枪击

OCCURRED\_ON\_DATE:日期 YEAR:年 MONTH:月 DAY\_OF\_WEEK:星期

STREET:街道

Lat:经度 Long:维度 Location:坐标

**分析目标:**影响犯罪率的原因

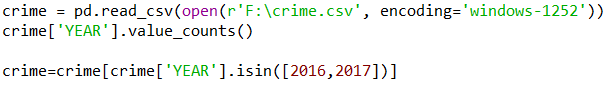
**分析手段和方法**

所需要的库

|  |
| --- |
| import pandas as pd  import numpy as np  import matplotlib.pyplot as plt  import seaborn as sns  import timefrom folium  import Mapfrom folium.plugins  import HeatMap  from folium import Map  from folium.plugins import HeatMap |

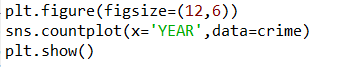
1. 数据处理
2. 读取：利用pandas中的read\_csv()读取crime.csv中的数据作为训练数据(不完整的数据简单去掉)

代码贴图

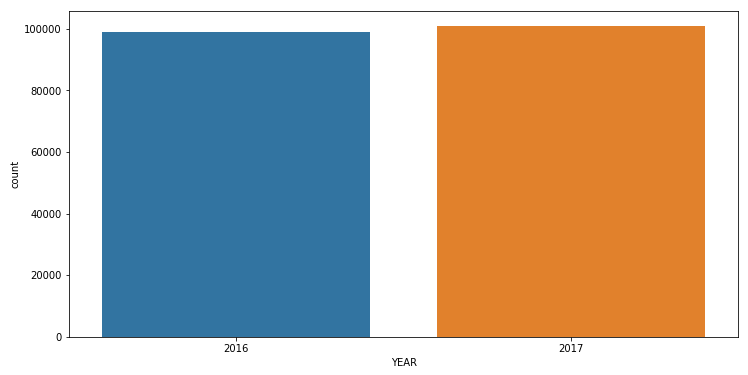


1. 直方图输出2016和2017年的犯罪总数

代码贴图



输出



1. 计算犯罪增长率

代码贴图



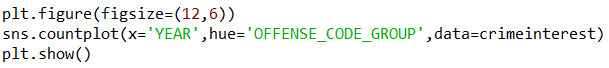
输出



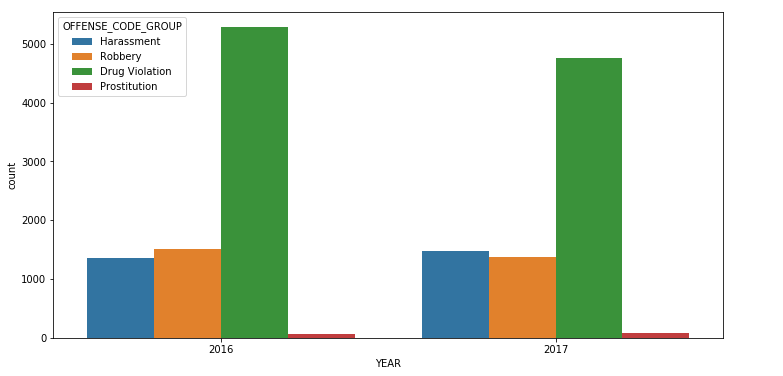
1. 筛选出'Missing Persons Report',’Robbery’,’Drug Violation’,’Prostitution’,’Harassment’五类犯罪类型的案件，以直方图形式输出案件总数

代码贴图





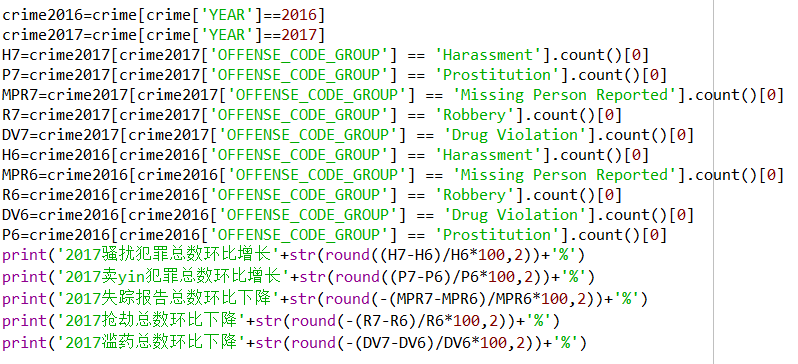
输出



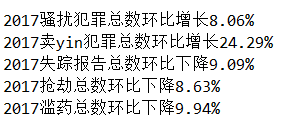
吸毒事件占所有案件的大部分，波士顿贩毒和吸毒属于违法行为，然而波士顿仍毒品猖獗。据美国波士顿环球报报道，位于美国波士顿的马萨诸塞大街则是波士顿著名的戒毒门诊一条街——美沙酮大道。这里既是瘾君子们寻求帮助的避难所，但同时也是波士顿露天毒品交易市场。

1. 计算各类型犯罪环比增长或下降

代码贴图



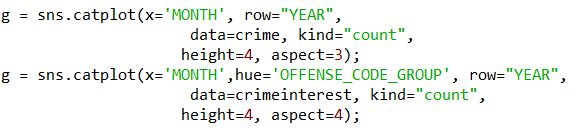
输出



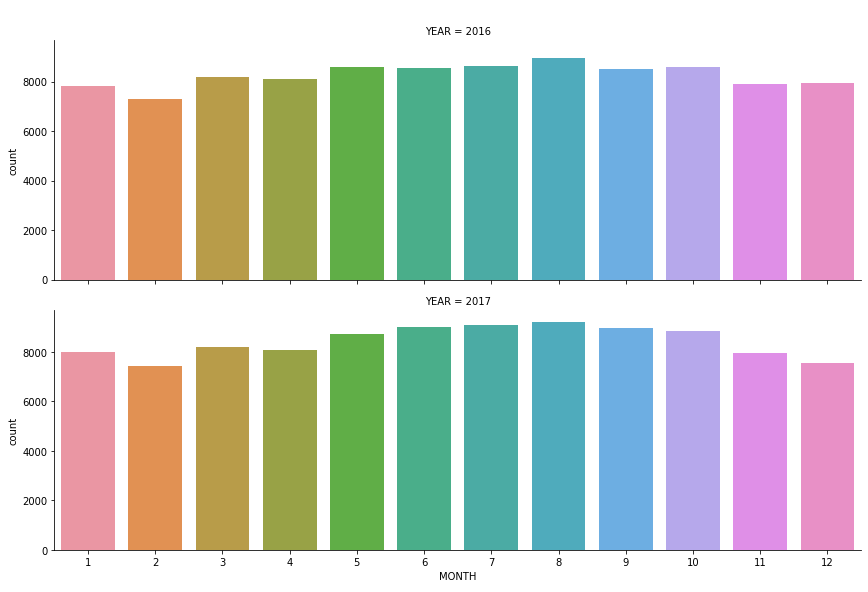
波士顿是马萨诸塞州第一大城市，全美第七大城市，2017年GDP为4029.63亿美元。波斯顿的经济以高新技术产业为主，大量计算机硬件和软件公司聚集于此，金融业和保险业也是波士顿的支柱产业。此外，波士顿还是一个印刷与出版业中心。因为经济情况的好转，贫困导致的抢劫，滥药，人口拐卖等违法事件发生频率有所减少。

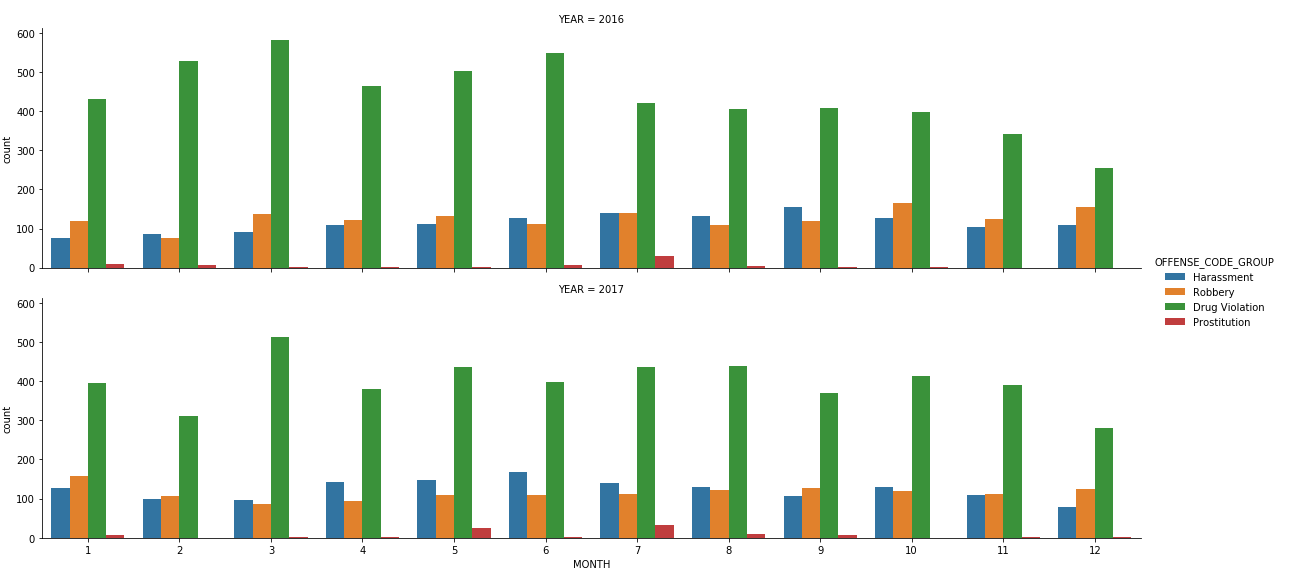
1. 将时间跨度减少至月，以直方图形式输出各类案件总数

代码贴图



输出

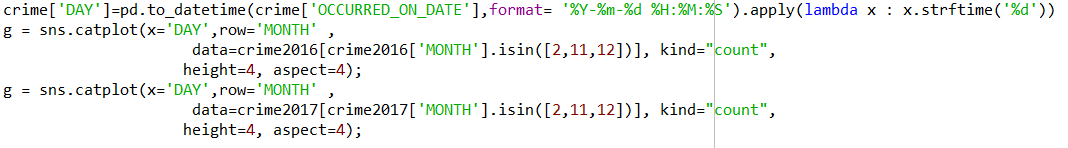




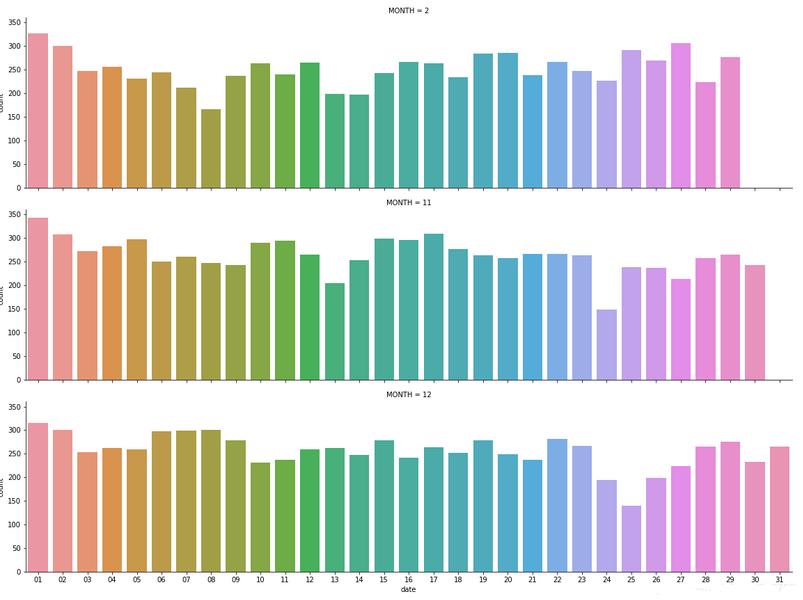
一年之中月犯罪频率相差不大，2月、11月、12月恶行违法事件相对减少，将跨度细分到月查看原因。

1. 将时间跨度减少至月，以直方图形式输出各类案件总数

代码贴图



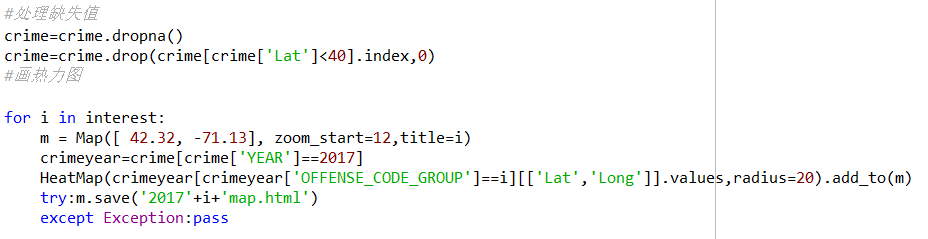
输出



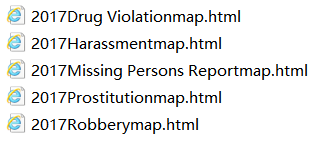
由直方图可直观看到，11.24日感恩节，12.25日圣诞节犯罪数量明显减少，这说明美国重大节日假期期间犯罪率反而下降。

1. 各类犯罪的地理分布

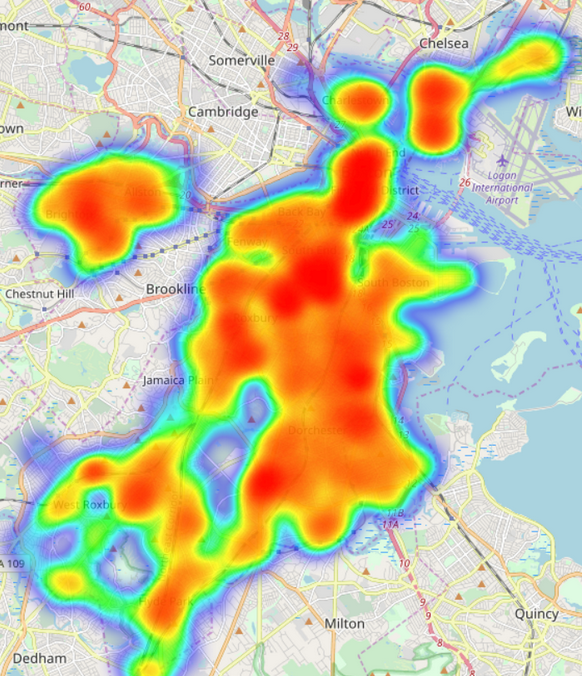
代码贴图



输出



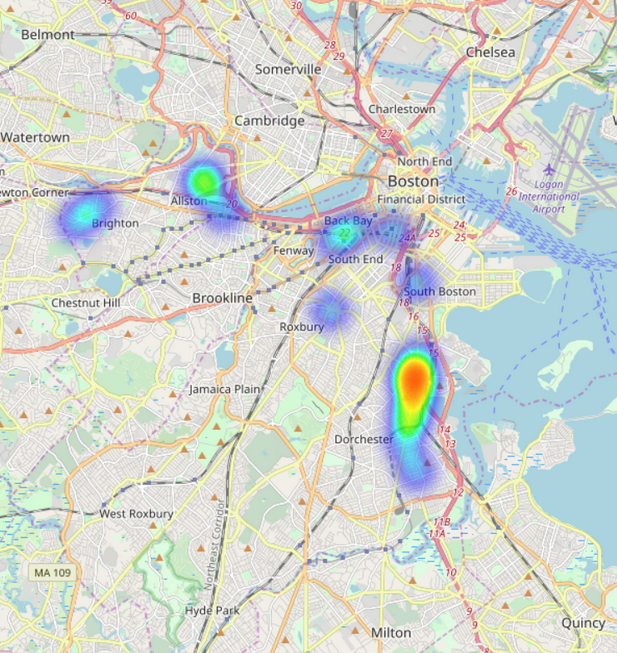
滥药



失踪



卖yin



抢劫



骚扰



1. 分析结果

本次实验报告中通过分析想要得到的结论为波士顿影响犯罪的因素。

1. 作案动机。当城市波士顿的经济呈现较进步的趋势时，抢劫、滥药与人口拐卖等恶性违法事件的总体数量均明显下降，说明经济的发展有可能提供大量工作机会，可以从根本上减少犯罪率。
2. 作案时间。逢美国重大节日如感恩节、独立日、圣诞节等假期，作案数量有反而下降。一方面与放假居民更喜欢在家中休息有关，一方面可能与警察轮休有关。
3. 作案地点。由地图分布可见，滥药问题在波士顿尤为严重，分布最广且频率最高。作为美国马萨诸塞州的首府和最大城市，也是美国东北部的新英格兰地区的最大城市，外来人口众多且集中在城市，基本生存面临问题，故违法案件也是由城市中心向周边扩散。
4. 结论与展望

通过以上的分析可以得到犯罪率与经济发展情况有密切关系。以谋杀率为例，由联合国毒品和犯罪问题办公室（United Nations Office on Drugs and Crime ）(UNODC) 的数据可粗略得到，从数据上，全世界谋杀率最低的国家里面，绝大多数都是发达国家，但也会有少数发展中国家，比如中国，印度尼西亚。而谋杀率最高的几个国家，绝大多数都是发展中国家，也有少数发达国家比如南非。因此，谋杀率与国家经济水平总体成正比。

影响犯罪率的因素有很多，对此数据集的分析不能详尽，不够严谨。对异常数据没有处理，直接筛除，可能导致最终结果有偏差；在将数据可视化之前，应先利用3σ原则识别异常值。