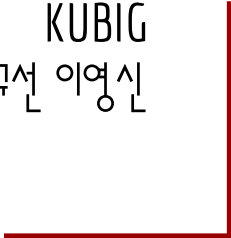




CHICAGO CRIME DATA

KUBIG

박소현 김효익 조송현 조규선 이영신



INDEX

- I . Data description
- II . Visualization(EDA)
- III. Plan for next step

I . Data description

〈Chicago crime data〉

- Incidents of crime that occurred in the City of Chicago from 2012 to 2017
- Data is extracted from the Chicago Police Department's CLEAR (Citizen Law Enforcement Analysis and Reporting) system and includes unverified reports.
- <https://www.kaggle.com/currie32/crimes-in-chicago>

I . Data description

- Variables

Identification variables : ID, Case Number

Time variables : Date, Year, Updated On

Location variables : Block, District, Ward, Community Area, Location Description, X Coordinate, Y Coordinate, Latitude, Longitude, Location

Type variables : IUCR, Primary Type, Description, FBI code, Arrest, Domestic

II. Visualization(EDA)

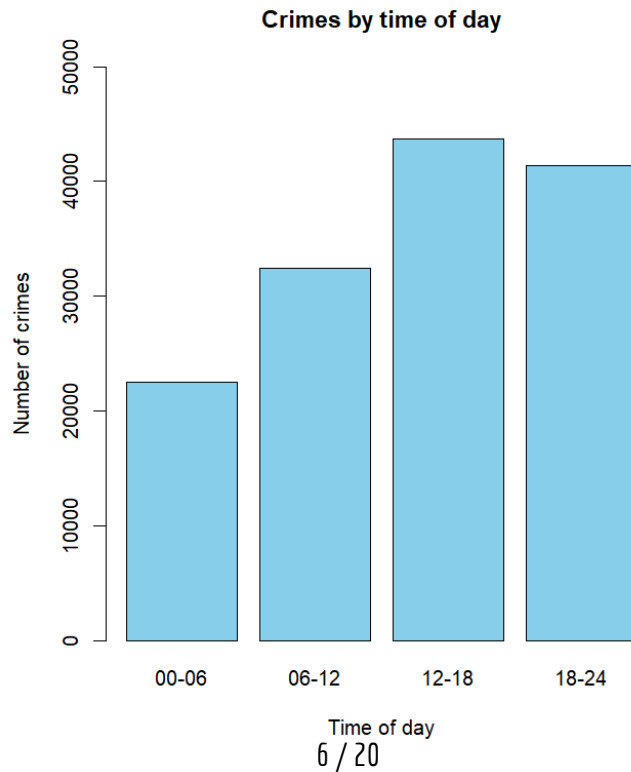
- 시간대 별 범죄 발생 건수

```
#시간대를 00-06,06-12,12-18,18-24로 구분
time.tag <- chron(times=c("00:00:00", "06:00:00", "12:00:00", "18:00:00", "23:59:00"))
crime$time.tag <- cut(crime$time, breaks=time.tag, labels=c("00-06","06-12","12-18",
"18-24"), include.lowest=T)
time <-table(crime$time.tag)
par(mar=c(6,6,4,2))
barplot(time, ylim=c(0,50000), las=1, main="Crimes by time of day", col="skyblue",
xlab="Time of day", ylab="Number of crimes")
```

00-06	06-12	12-18	18-24
22475	32463	43706	41356

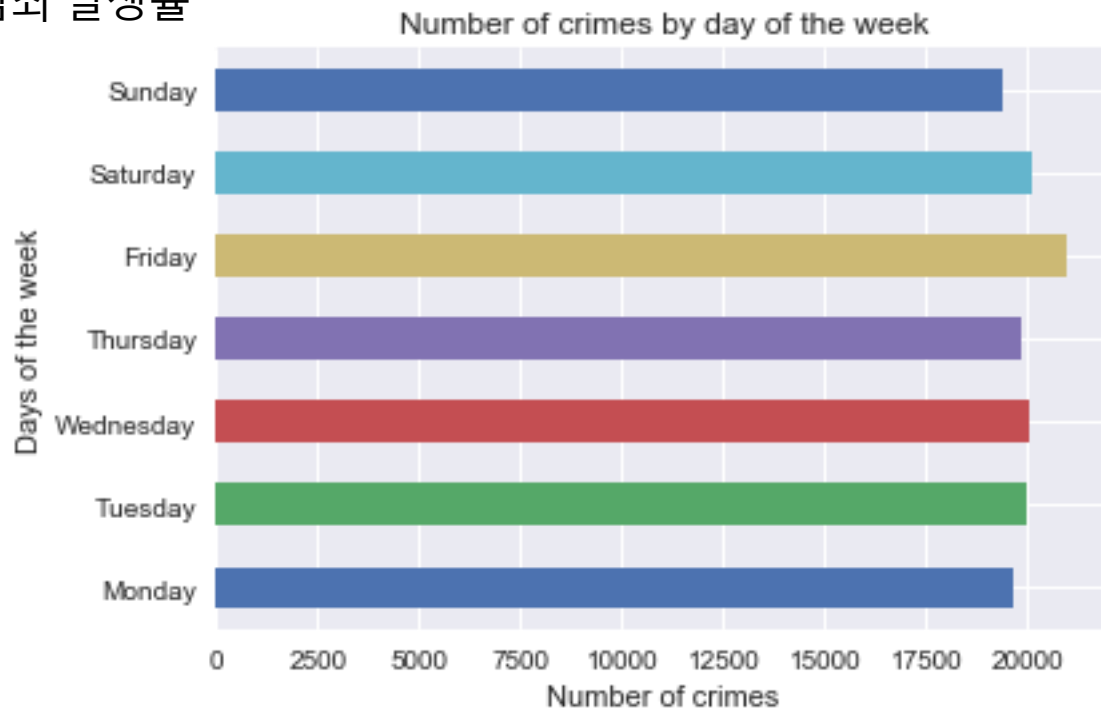
II. Visualization(EDA)

- 시간대 별 범죄 발생률



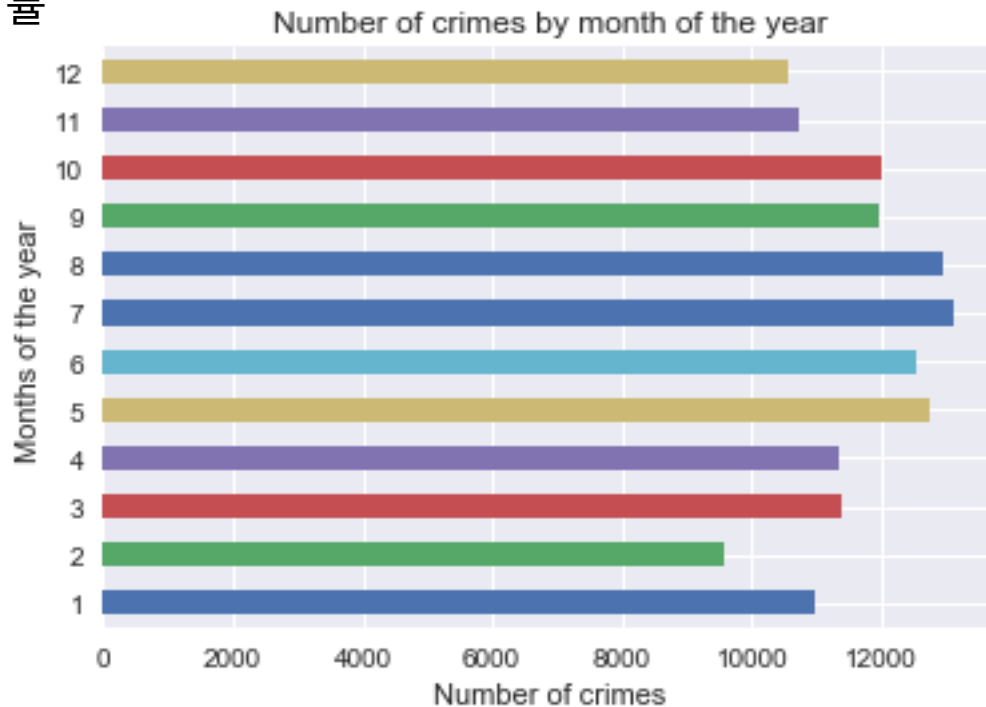
II. Visualization(EDA)

- 요일 별 범죄 발생률



II. Visualization(EDA)

- 월 별 범죄 발생률



II. Visualization(EDA)

- 한 달 기준 범죄 발생률 추이



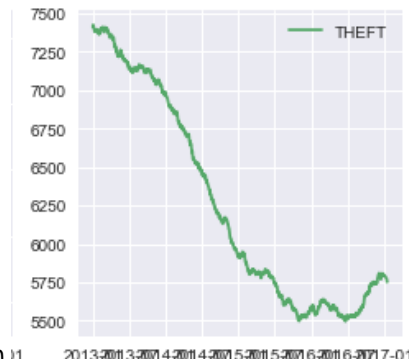
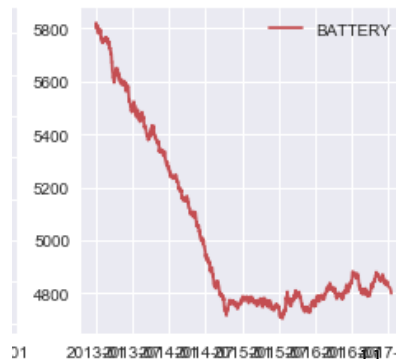
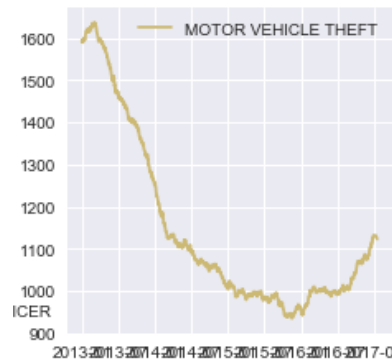
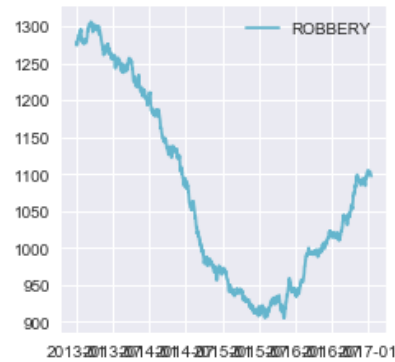
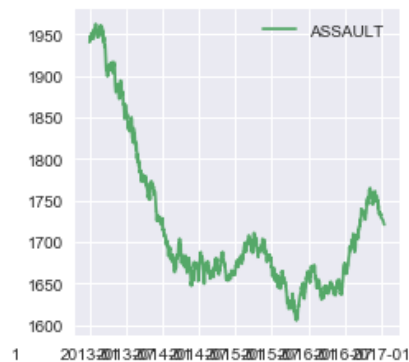
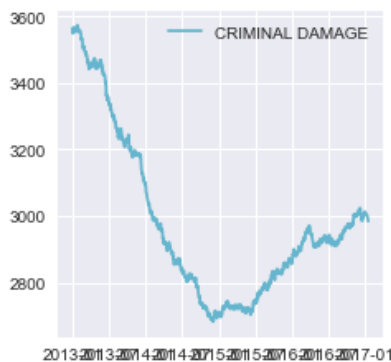
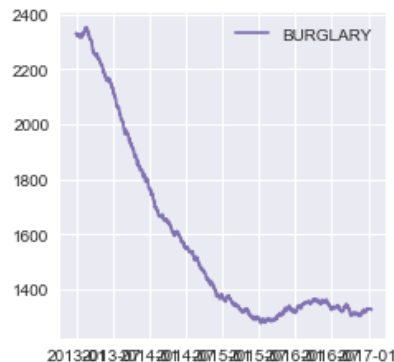
II. Visualization(EDA)

- 하루 기준 범죄 발생률 추이



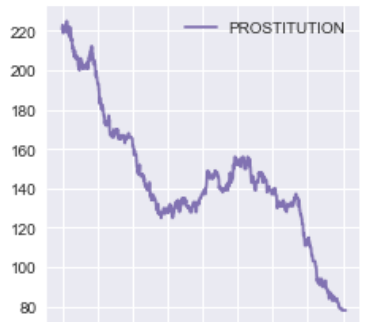
II. Visualization(EDA)

● 범죄 유형별 발생 추이

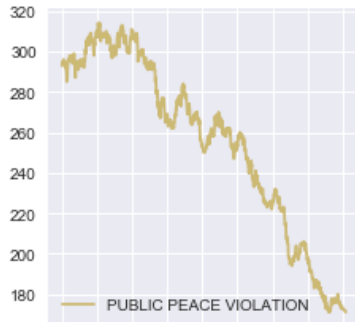


II. Visualization(EDA)

- 범죄 유형별 발생 추이



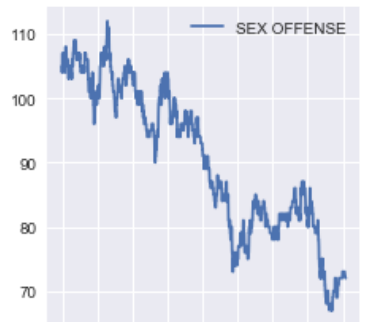
2013 2014 2015 2016 2017-01



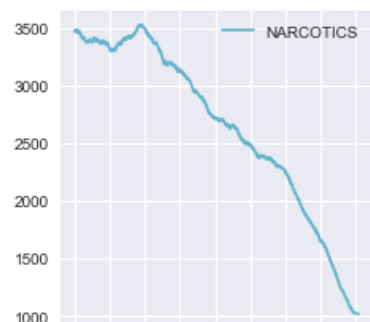
2013 2014 2015 2016 2017-01



2013 2014 2015 2016 2017-01



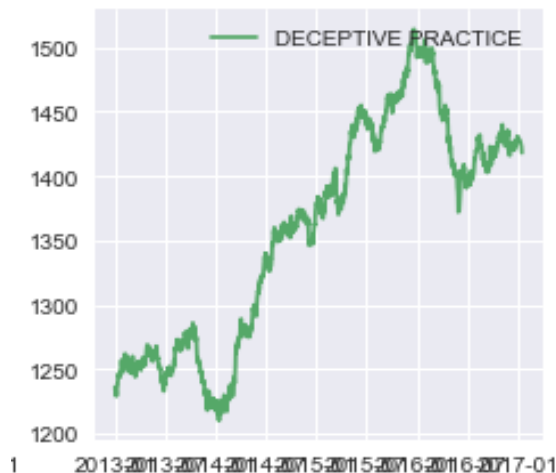
2013 2014 2015 2016 2017-01



2013 2014 2015 2016 2017-01

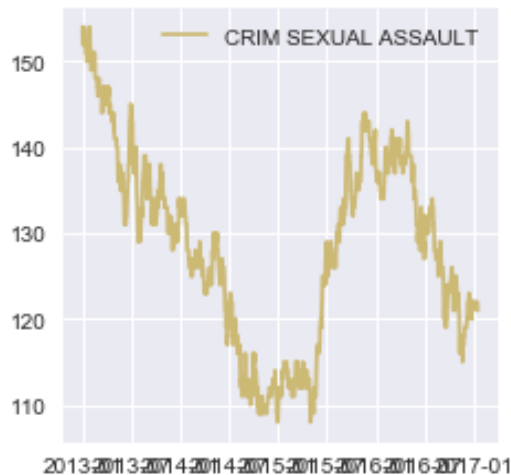
II. Visualization(EDA)

- 범죄 유형별 발생 추이



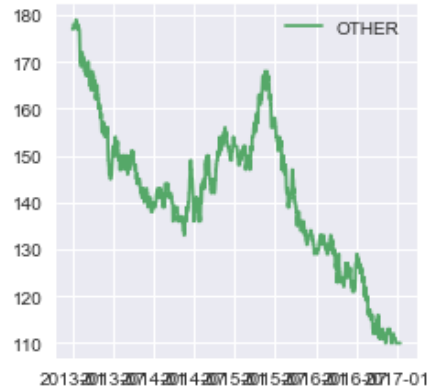
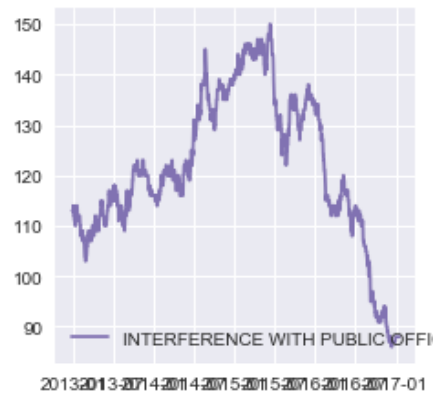
II. Visualization(EDA)

- 범죄 유형별 발생 추이



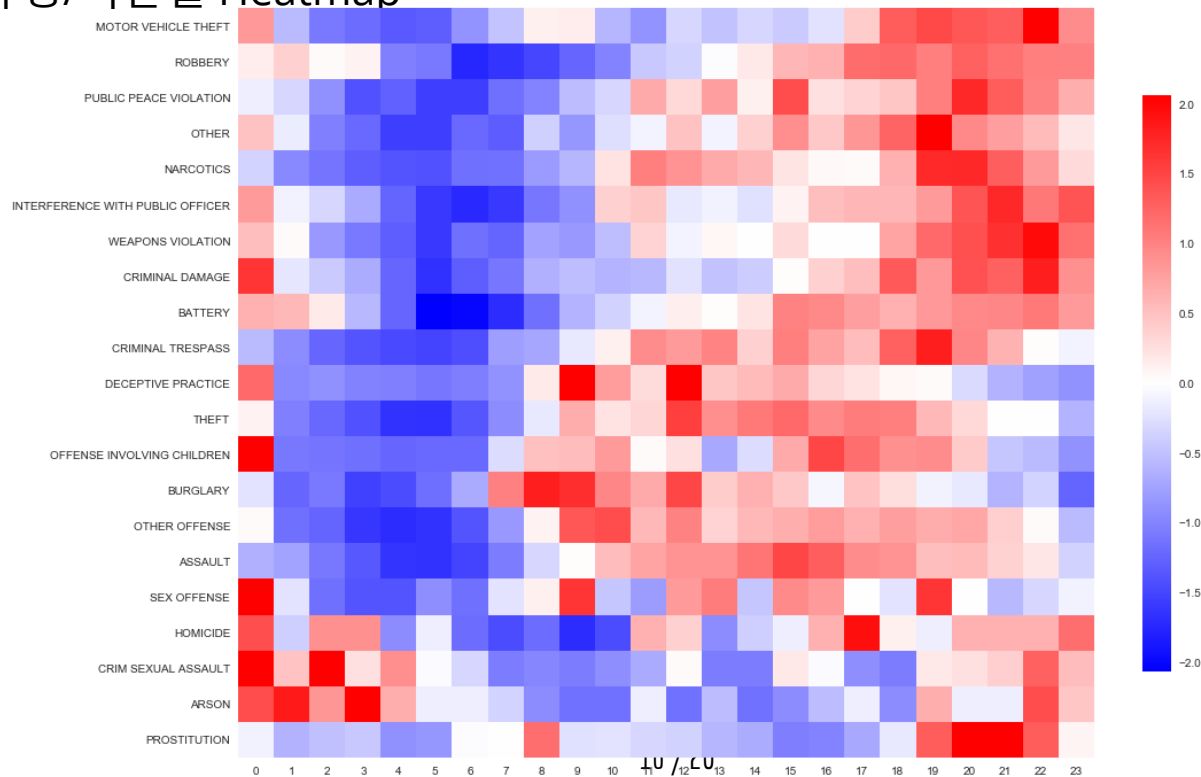
II. Visualization(EDA)

- 범죄 유형별 발생 추이



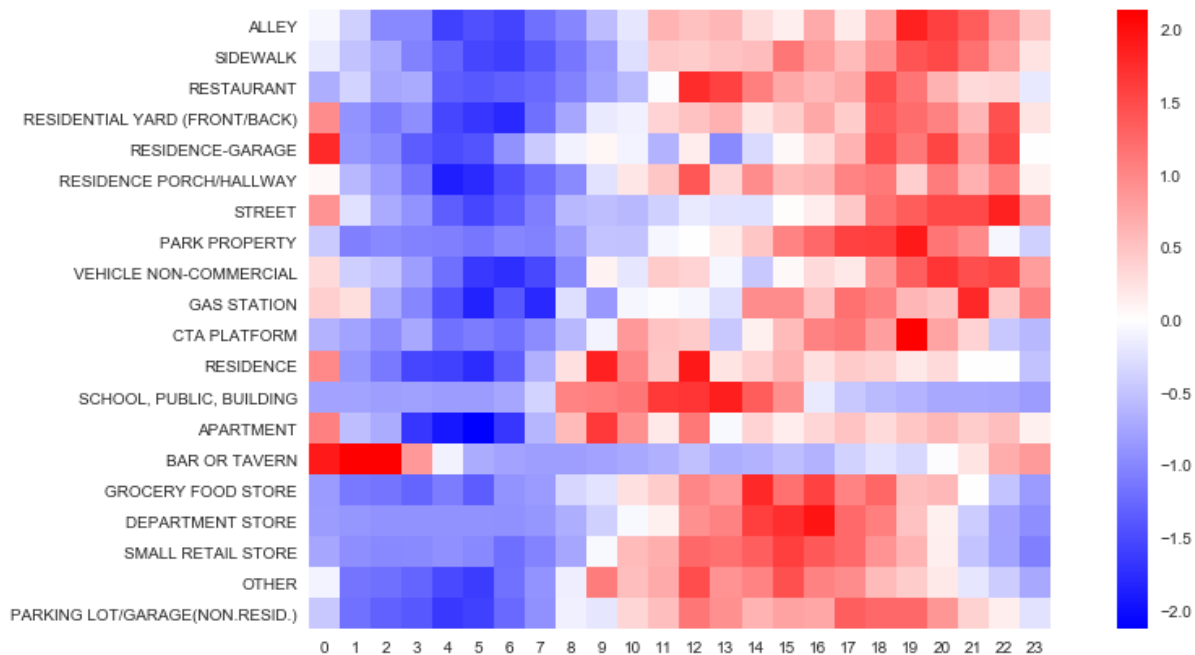
II. Visualization(EDA)

- 범죄 유형/시간별 Heatmap



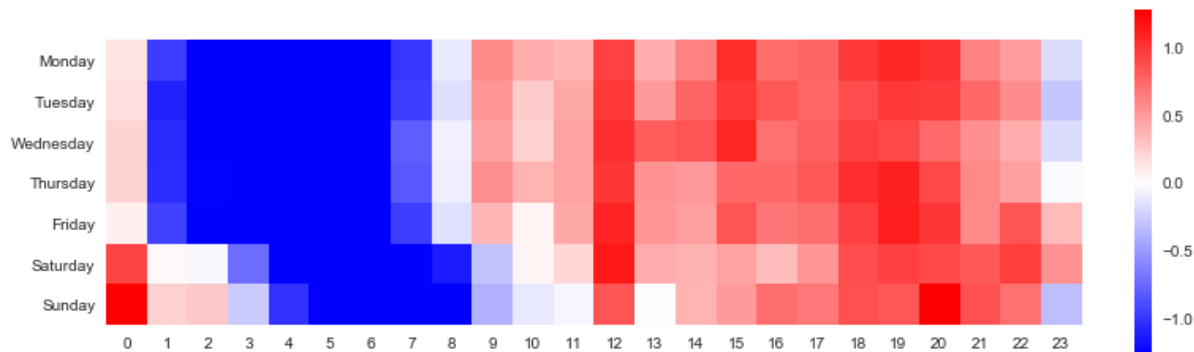
II. Visualization(EDA)

- 범죄 발생 장소/시간별 Heatmap



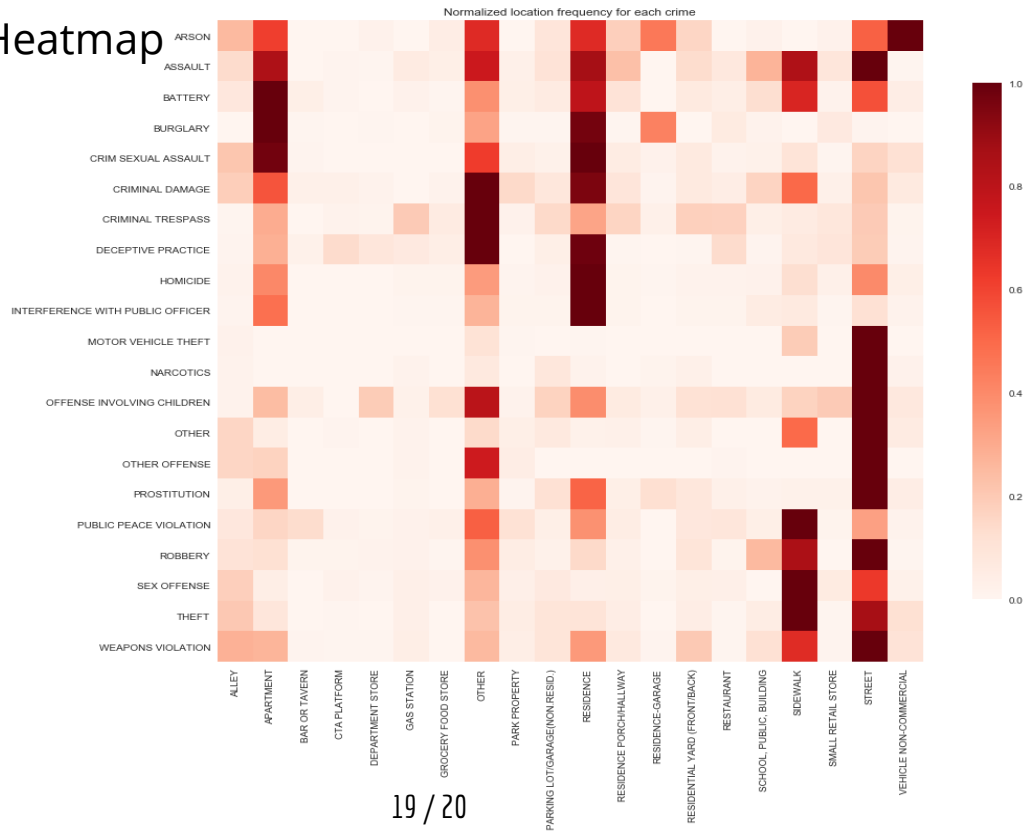
II. Visualization(EDA)

- 요일/시간별 범죄 발생 Heatmap



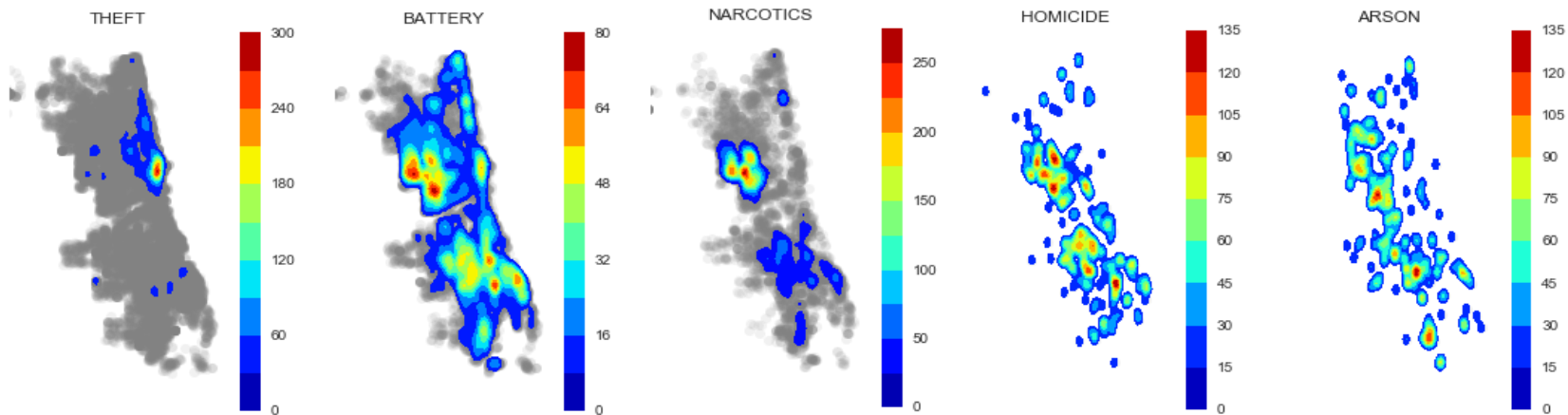
II. Visualization(EDA)

- 범죄 유형/장소 별 Heatmap



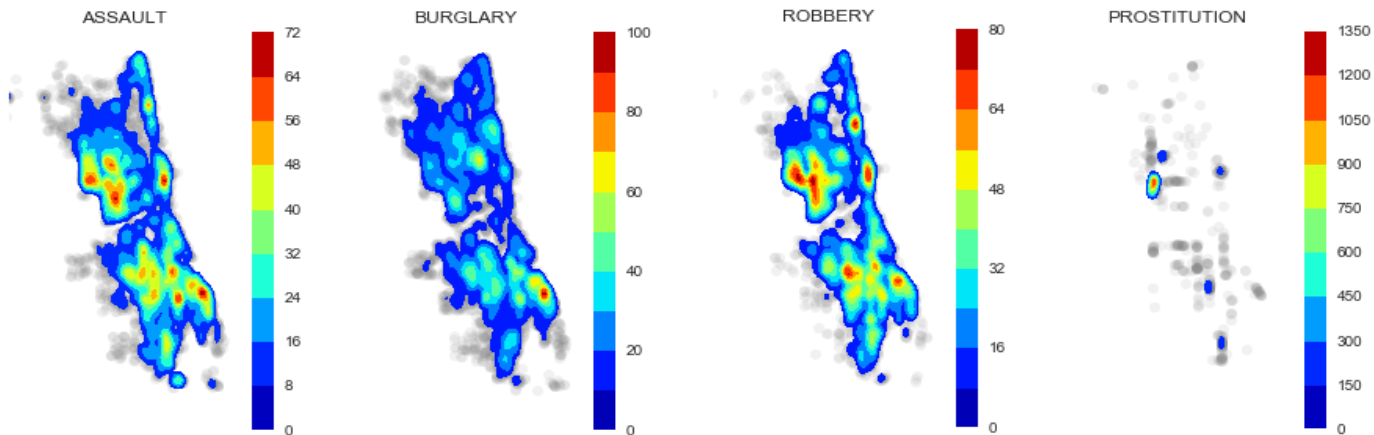
II. Visualization(EDA)

- 범죄 유형별 Heatmap



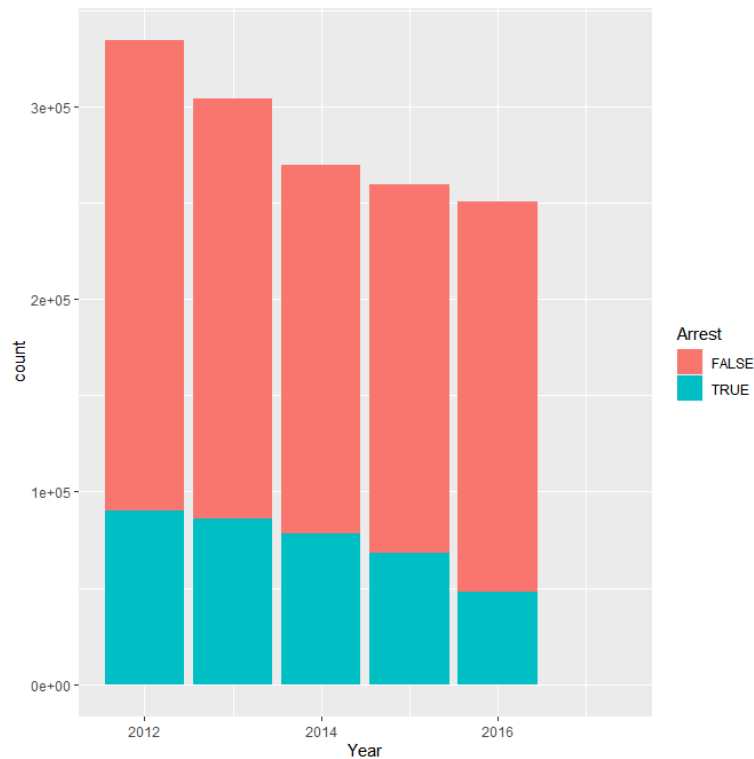
II. Visualization(EDA)

- 범죄 유형별 Heatmap



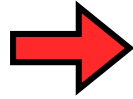
II. Visualization(EDA)

- 연도별 범죄자 체포 성공/실패 건수 비교

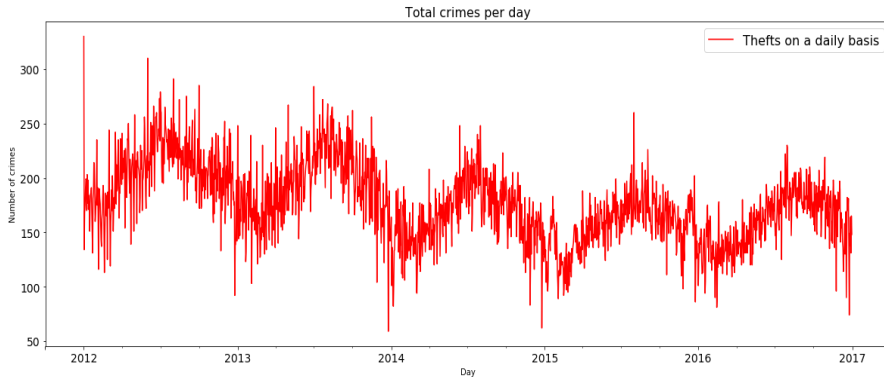


III. Plan for next step

- Multi Linear Regression
- LASSO/ Ridge
- Clustering
- PCA



- Crime possibility index
 1. Location variables: Block, District etc
 2. Crime type: Description, IUCR variables (violent/nonviolent etc)
 3. Time variables: Year, Month, Seasonality etc



Thank you