Hello everyone My name is Beomseok cho.

And Our team name is ‘Insight.’ It means we should have insight of Data then we finally increase quality of society.

First, I'd like to thank to professor Seonho Kim. I am very happy for 2 weeks. Thank you.

////

Today Our presentation index is ~~

/////

We got the data from the la data city site, and the name of the data is crime data from 2010 to present. Data has Crime data Description. I analysis this data using word count that abdula teached.

//////////////////////////////

It’s the way to Word Count using Python Code. I know everyone know this code. I sorted dictionary and I saved word count dictionary to csv file

And I made Word Cloud using these library. I show to you result.

////

The most common data is theft. Next is Assault. Next is Vehicle. I recommend Be careful your vehicle.

I show you this picture for introducing the list of most careful things. We can visualize which crimes are the most common. And next topic is data reliability 효진 Please

Hello, i'll explain data pre-processing.

The data we collect is two million rows and 28 columns.

The unnecessary columns were deleted. And only 12 columns were used.

These columns are date, crime content, gender, age, latitude and longitude.

And I pre-processed the data , as you can see.

Data pre-processing was done using Python pandas library.

I want to express three kinds of information on one graph.

The x-axis is the date, the y-axis is the number of crime, and each line is the month.

I wanted to make a dynamic graph. This used the ggplot2, gganimate library of R.

Next is Yeom Hee Soo's presentation.

안녕하세요. 저는 인사이트 맴버 염희수입니다.
이제부터 발표할 파트는 데이터분석2 파트입니다.
앞서 전처리과정을 통해 우리는 불필요한 데이터을 없앴습니다..
이후, 나는 지역에따라 그룹으로 묶어,
시간에 따른 범죄빈도를 나타냈다.
보시는바와같이 jupyter를 이용하였고, 그래프를 나타내었다.
그리고 지도를 나타내는 것은"power bi"라는 프로그램을 이용해서 시각화를 하였다.

1. Hi. My name is Heesu Yeom.

Nowthis partisdata analysis2.

Through the pre-processing process, we have deletedunnecessary data.

2. Afterwards, I grouped according to the region.

Then, The frequency of crime over time.

As you can see, it was used a jupyter, it showed a graph.

3. And the map was visualized using a program called "power bi".

저는 데이터에 있는 지역에 대해서 시간 당 범죄수와 그 지역에 많이생기는 범죄유형에대한 그래프를 지도에 나타냈습니다.(링크클릭)
지도에있는 아이콘을 클릭하면 그 지역에대한 그래프를 볼 수 있습니다.
지도를 나타내기위해서 Folium라이브러리를사용했습니다.
그리고 각 지역의 범죄에대한 그래프를 파이썬으로 만든 후 파일로 저장하였습니다.
프레임크기를 파일의 크기에 맞게 설정해주었습니다.
구글맵을 통해서 지역들의 위도와 경도를 찾았습니다.
아이콘을 각 지역의 위도와 경도위치에 맞게 설정하여 지도에 나타냈습니다.
그 지도를 html파일형태로 저장하였습니다.

I showed a graph of the region in the data on the map.
Click on the icon on the map to see a graph of the region.
I used the Folium library to represent the map.
And I made Python a graph of crime in each regionand saved it as a file.
The frame size has been set to fit the picture.
I found the latitude and longitude of the regions through Google Maps.
Icon is set to the latitude and longitude positions of each region and is shown on the map.
I saved the map in html file format.