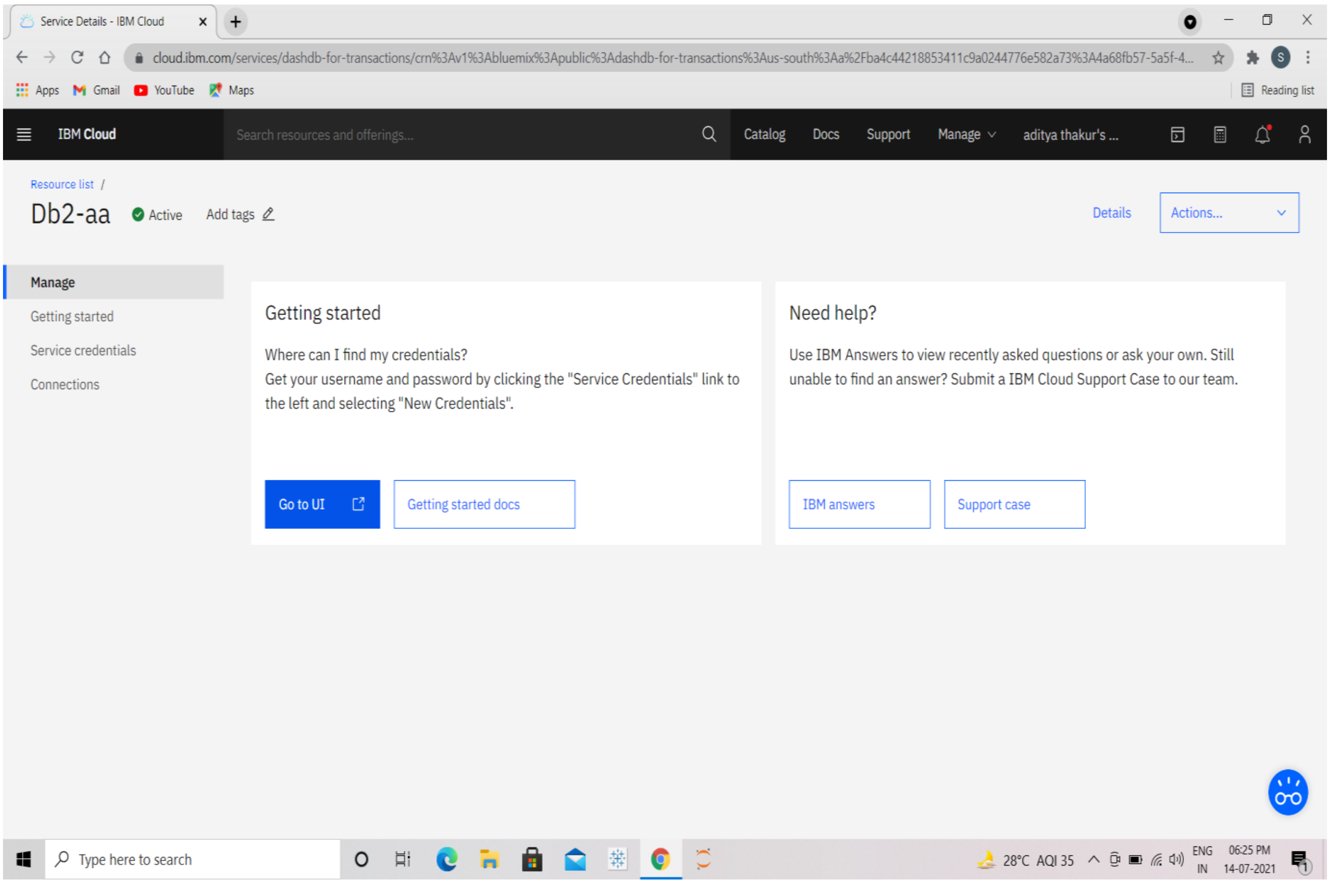
Crime Analysis using SQL with Python

Introduction

In this project I am going to analysis the crime case with help of the sql inside python and IBM cloud as the data server .

Data Preprocessing

Downloaded the 3-csv file of Chicago (crime, public and school) data for the analysis. Next step involves uploading the data in the db2. For the uploading purpose I prefer the IBM cloud as its interface is really easy to handle with the db2 . I used the magic Sql command to perform the analysis.



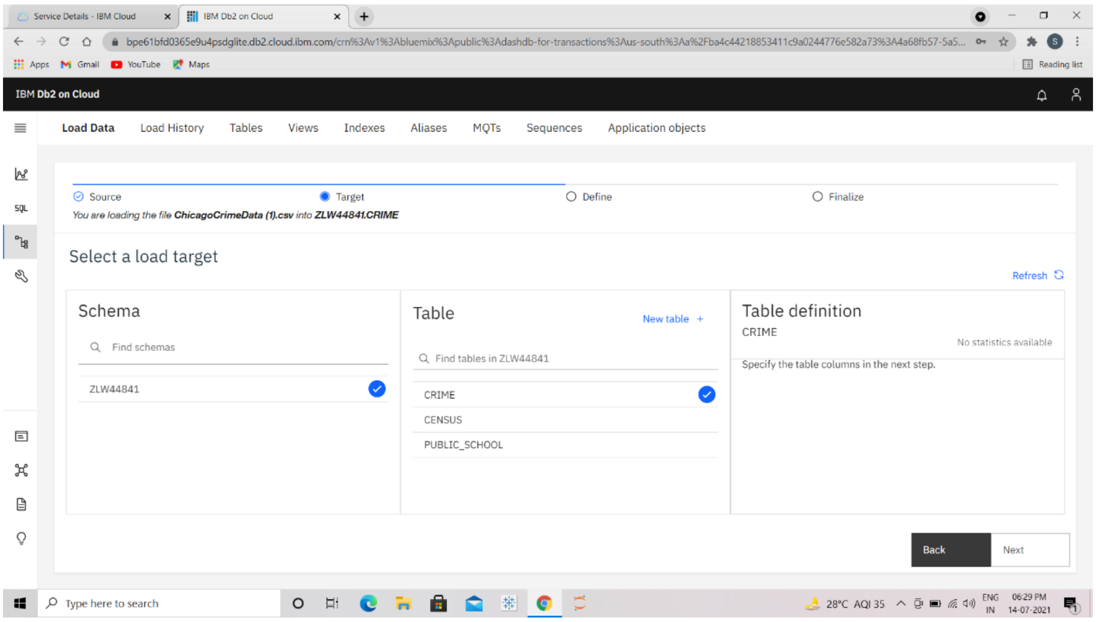
Getting started

Where can I find my credentials?
Get your username and password by clicking the "Service Credentials" link to the left and selecting "New Credentials".

Need help?

Use IBM Answers to view recently asked questions or ask your own. Still unable to find an answer? Submit a IBM Cloud Support Case to our team.

while uploading the data I need to create the schema , table name with its definition along with table then to upload the data .



You are loading the file **ChicagoCrimeData (1).csv** into **ZLW44841.CRIME**

Select a load target

Schema

Find schemas

ZLW44841 ✓

Table

Find tables in ZLW44841

CRIME ✓

CENSUS

PUBLIC_SCHOOL

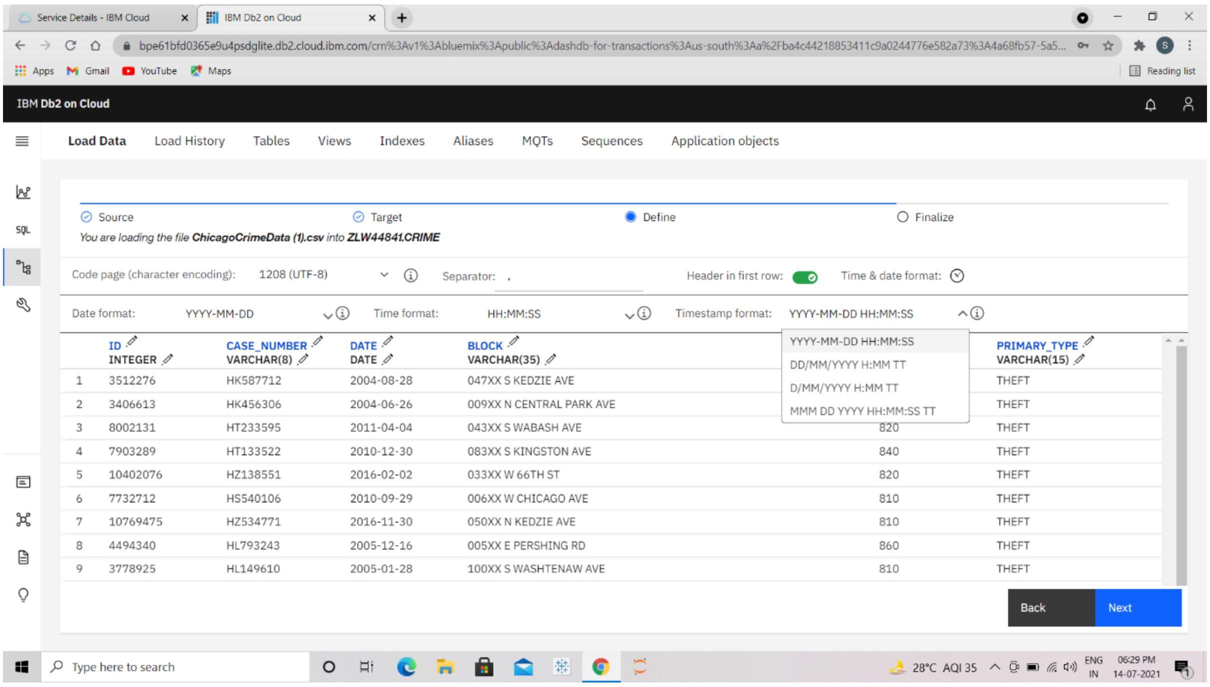
Table definition

CRIME

No statistics available

Specify the table columns in the next step.

After this inside the preview option I have managed to managed to do change the time format , along with including headers, and even changing the data type some time .



You are loading the file **ChicagoCrimeData (1).csv** into **ZLW44841.CRIME**

Code page (character encoding): 1208 (UTF-8)

Separator: ,

Header in first row: ● (on)

Time & date format:

Date format: YYYY-MM-DD

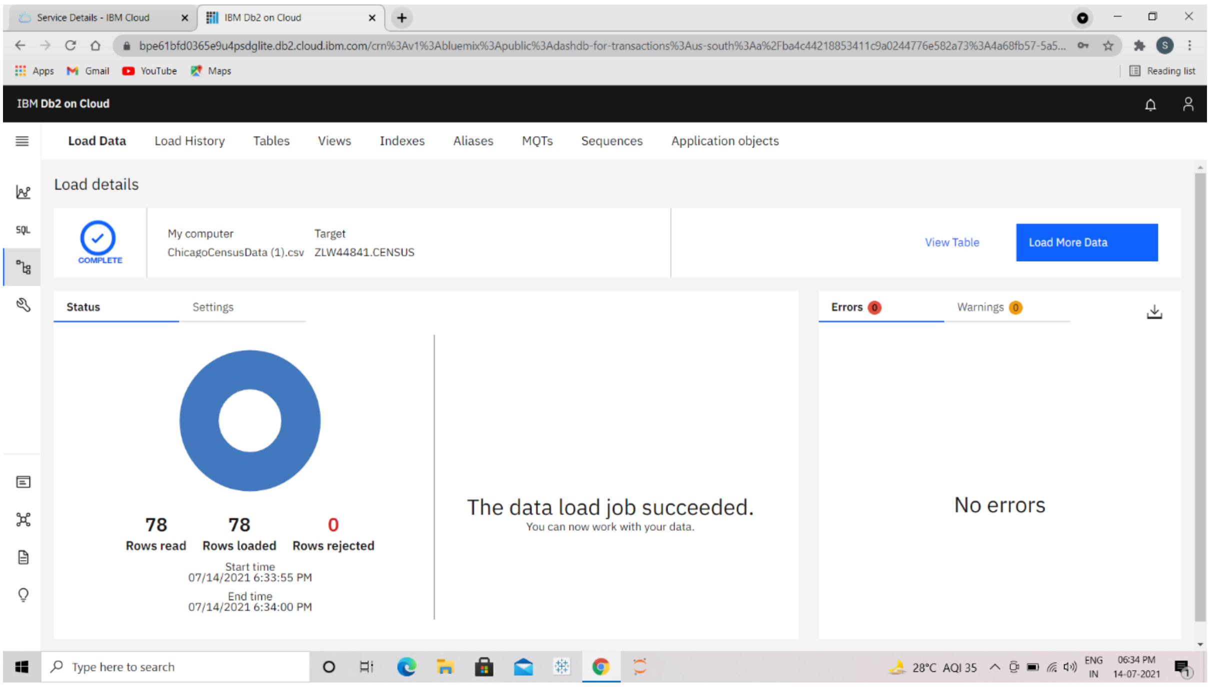
Time format: HH:MM:SS

Timestamp format: YYYY-MM-DD HH:MM:SS

YYYY-MM-DD HH:MM:SS
DD/MM/YYYY H:MM TT
D/MM/YYYY H:MM TT
MMM DD YYYY HH:MM:SS TT

	ID INTEGER	CASE_NUMBER VARCHAR(8)	DATE DATE	BLOCK VARCHAR(35)		PRIMARY_TYPE VARCHAR(15)
1	3512276	HK587712	2004-08-28	047XX S KEDZIE AVE		THEFT
2	3406613	HK456306	2004-06-26	009XX N CENTRAL PARK AVE		THEFT
3	8002131	HT233595	2011-04-04	043XX S WABASH AVE	820	THEFT
4	7903289	HT133522	2010-12-30	083XX S KINGSTON AVE	840	THEFT
5	10402076	HZ138551	2016-02-02	033XX W 66TH ST	820	THEFT
6	7732712	HS540106	2010-09-29	006XX W CHICAGO AVE	810	THEFT
7	10769475	HZ534771	2016-11-30	050XX N KEDZIE AVE	810	THEFT
8	4494340	HL793243	2005-12-16	005XX E PERSHING RD	860	THEFT
9	3778925	HL149610	2005-01-28	100XX S WASHTENAW AVE	810	THEFT

Later on uploaded the data set with no warning or error. To move on the further analysis .



Load details

My computer	Target
ChicagoCensusData (1).csv	ZLW44841.CENSUS

78	78	0
Rows read	Rows loaded	Rows rejected

Start time
07/14/2021 6:33:55 PM

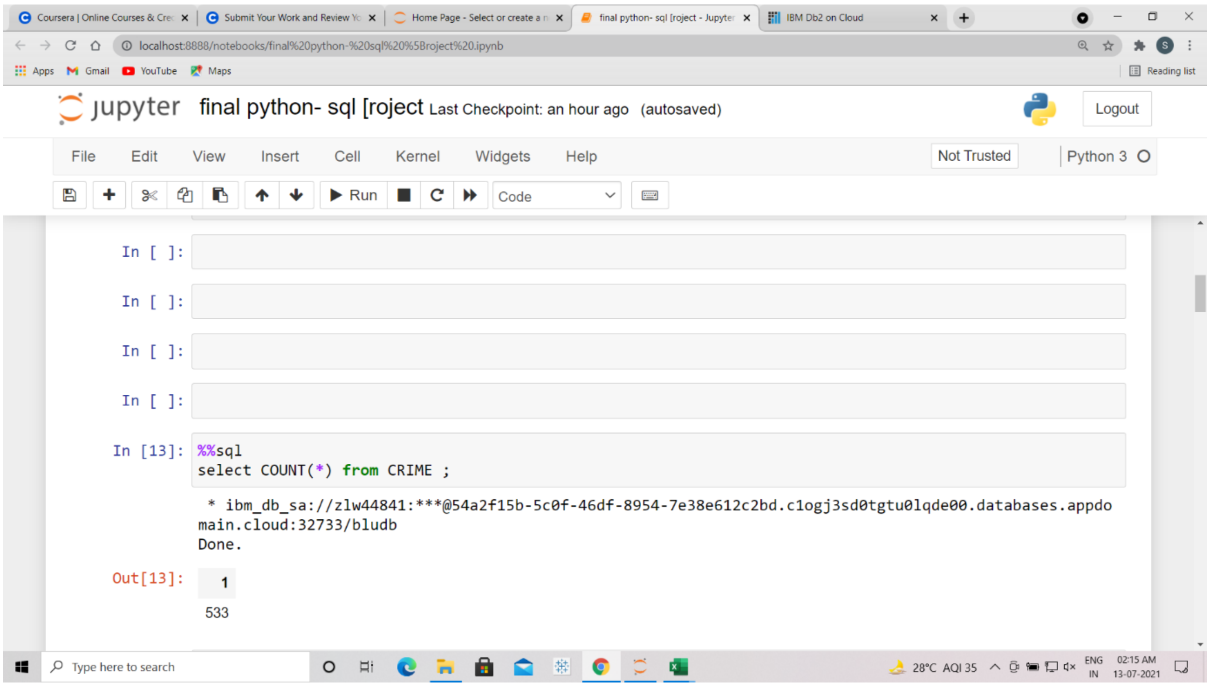
End time
07/14/2021 6:34:00 PM

The data load job succeeded.
You can now work with your data.

No errors

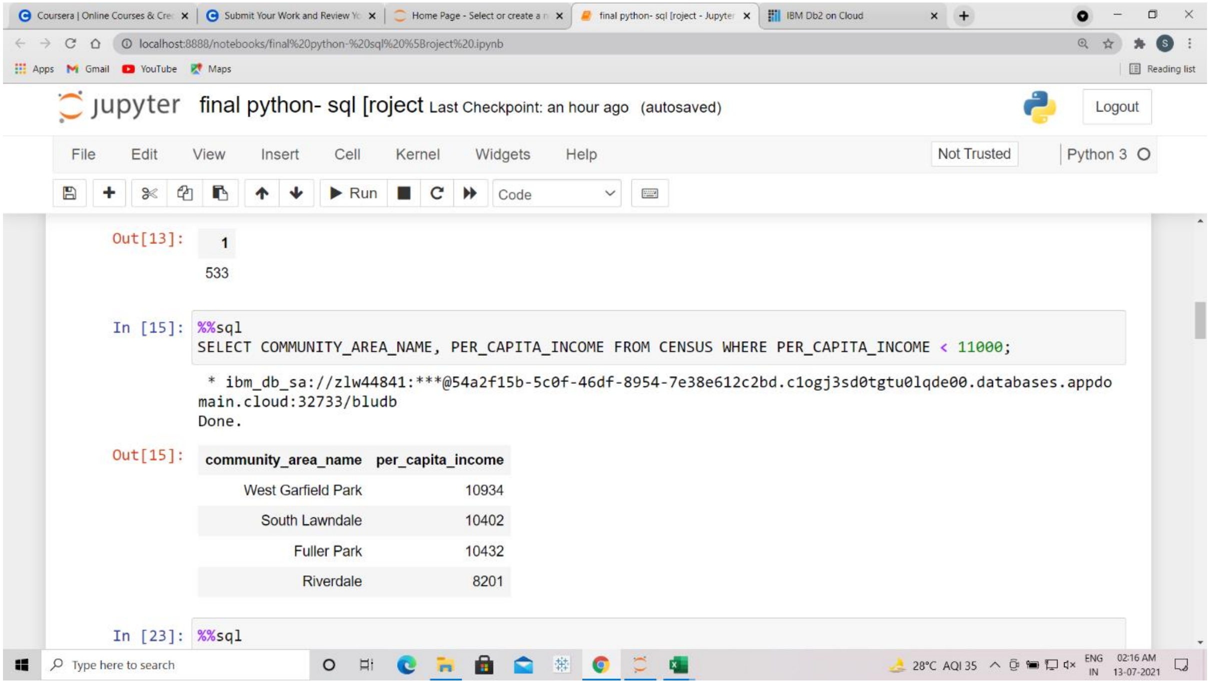
Analysis:

1) Find the total number of crime recorded in the crime table ?


Coursera | Online Courses & Cre
Submit Your Work and Review Yo
Home Page - Select or create a n
final python- sql [roject - Jupyter
IBM Db2 on Cloud
localhost:8888/notebooks/final%20python-%20sql%20%5Broject%20.ipynb
jupyter
final python- sql [roject Last Checkpoint: an hour ago (autosaved)
Logout
File
Edit
View
Insert
Cell
Kernel
Widgets
Help
Not Trusted
Python 3
Run
Code
Type here to search
28°C AQI 35
02:15 AM
13-07-2021

```
In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [13]: %%sql
         select COUNT(*) from CRIME ;

          * ibm_db_sa://zlw44841:***@54a2f15b-5c0f-46df-8954-7e38e612c2bd.c1ogj3sd0tgtu0lqde00.databases.appdo
         main.cloud:32733/bludb
         Done.

Out[13]:   1
         533
```

2) List the community area with per capita income less than 11000?


Coursera | Online Courses & Cre
Submit Your Work and Review Yo
Home Page - Select or create a n
final python- sql [roject - Jupyter
IBM Db2 on Cloud
localhost:8888/notebooks/final%20python-%20sql%20%5Broject%20.ipynb
Apps
Gmail
YouTube
Maps
Reading list
jupyter
final python- sql [roject Last Checkpoint: an hour ago (autosaved)
Logout
File
Edit
View
Insert
Cell
Kernel
Widgets
Help
Not Trusted
Python 3
Run
Code
Type here to search
28°C AQI 35
ENG IN
02:16 AM
13-07-2021

Out[13]:

1
533

In [15]:

```
%%sql
SELECT COMMUNITY_AREA_NAME, PER_CAPITA_INCOME FROM CENSUS WHERE PER_CAPITA_INCOME < 11000;
```

```
 * ibm_db_sa://zlw44841:***@54a2f15b-5c0f-46df-8954-7e38e612c2bd.c1ogj3sd0tgtu0lqde00.databases.appdo
main.cloud:32733/bludb
Done.
```

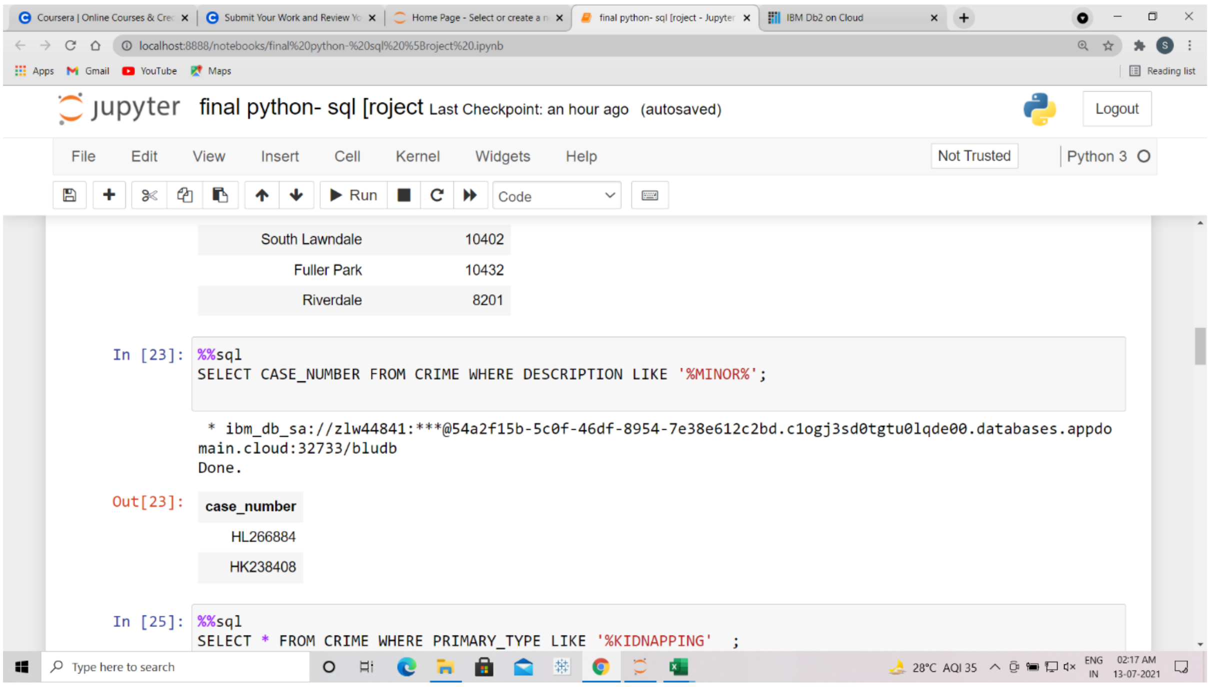
Out[15]:

community_area_name	per_capita_income
West Garfield Park	10934
South Lawndale	10402
Fuller Park	10432
Riverdale	8201

In [23]:

```
%%sql
```

3) List all case numbers for crimes involving minors? (children are not considered minors for the purposes of crime analysis)


Coursera | Online Courses & Cre
Submit Your Work and Review Yo
Home Page - Select or create a n
final python- sql [roject - Jupyter
IBM Db2 on Cloud
localhost:8888/notebooks/final%20python-%20sql%20%5Broject%20.ipynb
Apps
Gmail
YouTube
Maps
Reading list
jupyter
final python- sql [roject Last Checkpoint: an hour ago (autosaved)
Logout
File
Edit
View
Insert
Cell
Kernel
Widgets
Help
Not Trusted
Python 3
Run
Code
Type here to search
28°C AQI 35
ENG IN
02:17 AM
13-07-2021

South Lawndale	10402
Fuller Park	10432
Riverdale	8201

In [23]:

```
%%sql
SELECT CASE_NUMBER FROM CRIME WHERE DESCRIPTION LIKE '%MINOR%';
```

```
 * ibm_db_sa://zlw44841:***@54a2f15b-5c0f-46df-8954-7e38e612c2bd.c1ogj3sd0tgtu0lqde00.databases.appdo
main.cloud:32733/bludb
Done.
```

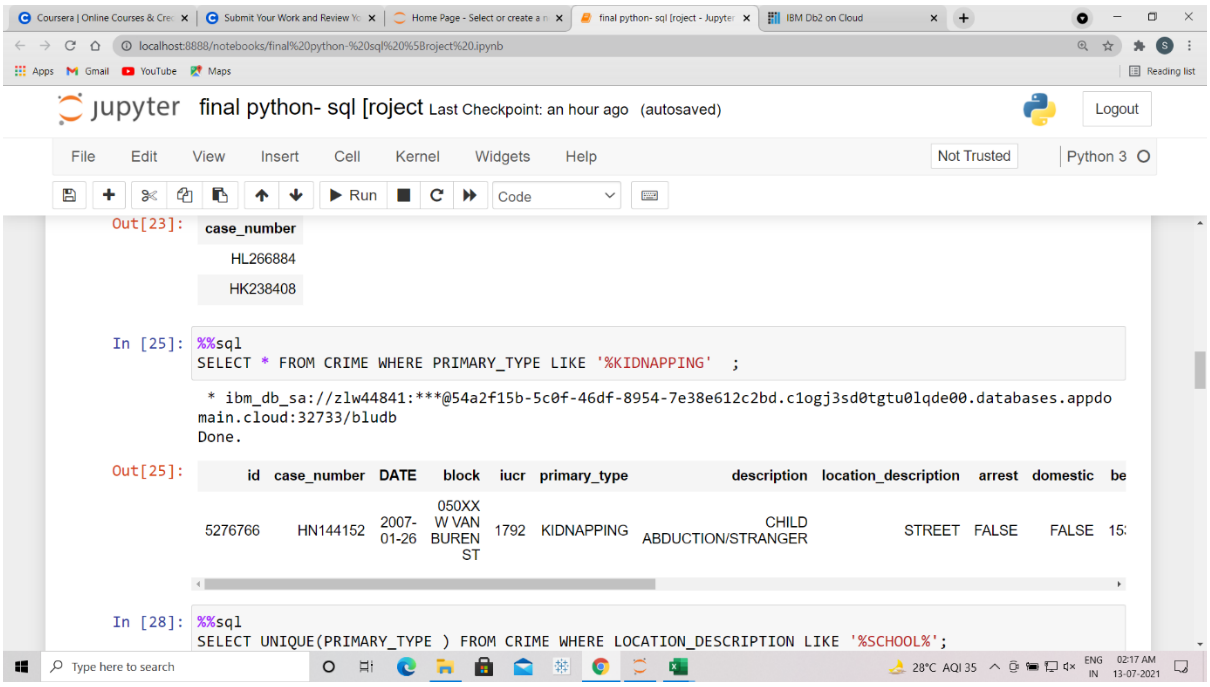
Out[23]:

case_number
HL266884
HK238408

In [25]:

```
%%sql
SELECT * FROM CRIME WHERE PRIMARY_TYPE LIKE '%KIDNAPPING'  ;
```

4) List all kidnapping crimes involving a child?


Coursera | Online Courses & Cre
Submit Your Work and Review Yo
Home Page - Select or create a n
final python- sql [roject - Jupyter
IBM Db2 on Cloud
localhost:8888/notebooks/final%20python-%20sql%20%5Broject%20.ipynb
Apps
Gmail
YouTube
Maps
Reading list
jupyter
final python- sql [roject Last Checkpoint: an hour ago (autosaved)
Logout
File
Edit
View
Insert
Cell
Kernel
Widgets
Help
Not Trusted
Python 3
Run
Code
Type here to search
28°C AQI 35
ENG
IN
02:17 AM
13-07-2021

Out[23]:

case_number
HL266884
HK238408

```
In [25]: %%sql
SELECT * FROM CRIME WHERE PRIMARY_TYPE LIKE '%KIDNAPPING'  ;
```

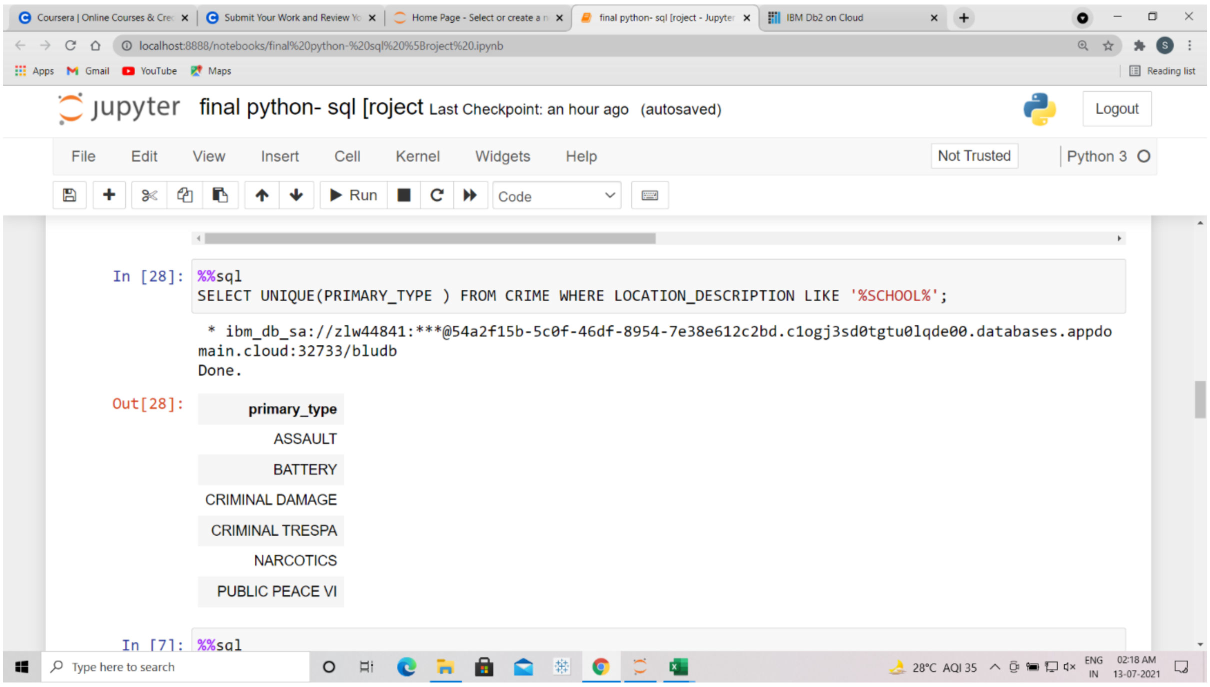
```
 * ibm_db_sa://zlw44841:***@54a2f15b-5c0f-46df-8954-7e38e612c2bd.c1ogj3sd0tgtu0lqde00.databases.appdo
main.cloud:32733/bludb
Done.
```

Out[25]:

id	case_number	DATE	block	iucr	primary_type	description	location_description	arrest	domestic	be
5276766	HN144152	2007-01-26	050XX W VAN BUREN ST	1792	KIDNAPPING	CHILD ABDUCTION/STRANGER	STREET	FALSE	FALSE	15:

```
In [28]: %%sql
SELECT UNIQUE(PRIMARY_TYPE ) FROM CRIME WHERE LOCATION_DESCRIPTION LIKE '%SCHOOL%';
```

5) What kind of crimes were recorded at schools?


Coursera | Online Courses & Cre
Submit Your Work and Review Yo
Home Page - Select or create a n
final python- sql [roject - Jupyter
IBM Db2 on Cloud
localhost:8888/notebooks/final%20python-%20sql%20%5Broject%20.ipynb
Apps
Gmail
YouTube
Maps
Reading list
jupyter
final python- sql [roject Last Checkpoint: an hour ago (autosaved)
Logout
File
Edit
View
Insert
Cell
Kernel
Widgets
Help
Not Trusted
Python 3
Run
Code
Type here to search
28°C AQI 35
ENG
IN
02:18 AM
13-07-2021

```
In [28]: %%sql
SELECT UNIQUE(PRIMARY_TYPE ) FROM CRIME WHERE LOCATION_DESCRIPTION LIKE '%SCHOOL%';
```

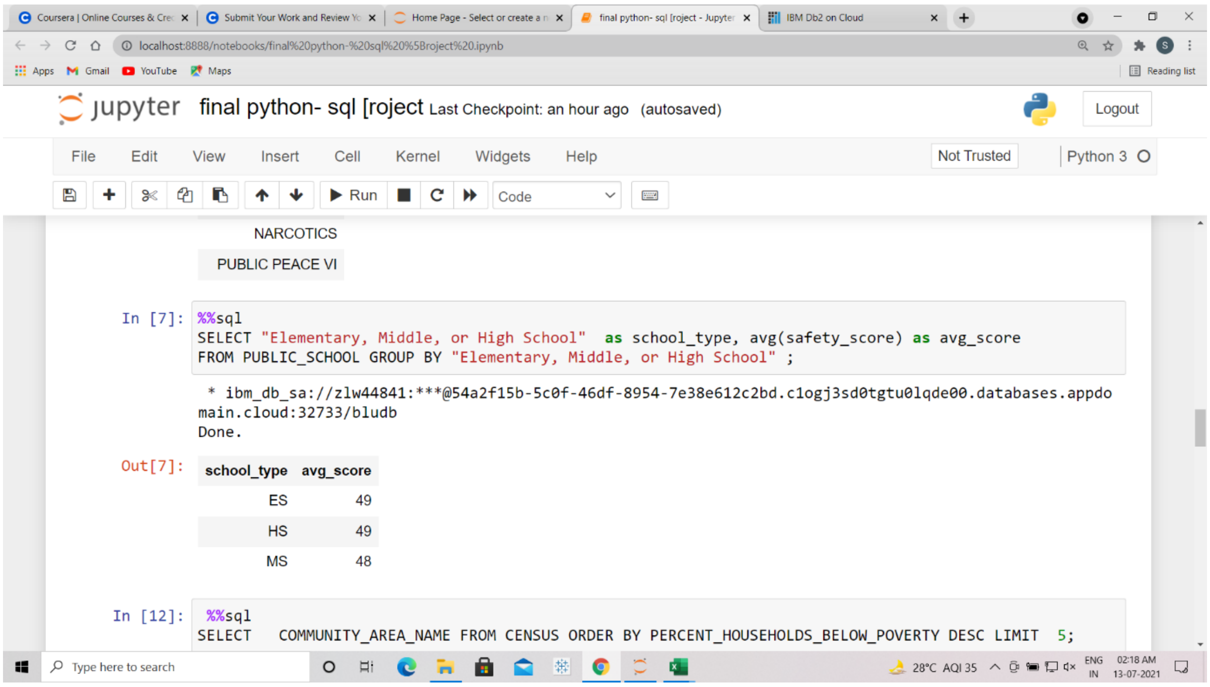
```
 * ibm_db_sa://zlw44841:***@54a2f15b-5c0f-46df-8954-7e38e612c2bd.c1ogj3sd0tgtu0lqde00.databases.appdo
main.cloud:32733/bludb
Done.
```

Out[28]:

primary_type
ASSAULT
BATTERY
CRIMINAL DAMAGE
CRIMINAL TRESPA
NARCOTICS
PUBLIC PEACE VI

```
In [7]: %%sql
```

6) List the average safety score for all types of schools?


Coursera | Online Courses & Cre
Submit Your Work and Review Yo
Home Page - Select or create a n
final python- sql [roject - Jupyter
IBM Db2 on Cloud
localhost:8888/notebooks/final%20python-%20sql%20%5Broject%20.ipynb
Apps
Gmail
YouTube
Maps
Reading list
jupyter
final python- sql [roject Last Checkpoint: an hour ago (autosaved)
Logout
File
Edit
View
Insert
Cell
Kernel
Widgets
Help
Not Trusted
Python 3
Run
Code
Type here to search
28°C AQI 35
ENG IN
02:18 AM
13-07-2021

NARCOTICS

PUBLIC PEACE VI

In [7]:
```
%%sql
SELECT "Elementary, Middle, or High School"  as school_type, avg(safety_score) as avg_score
FROM PUBLIC_SCHOOL GROUP BY "Elementary, Middle, or High School" ;
```

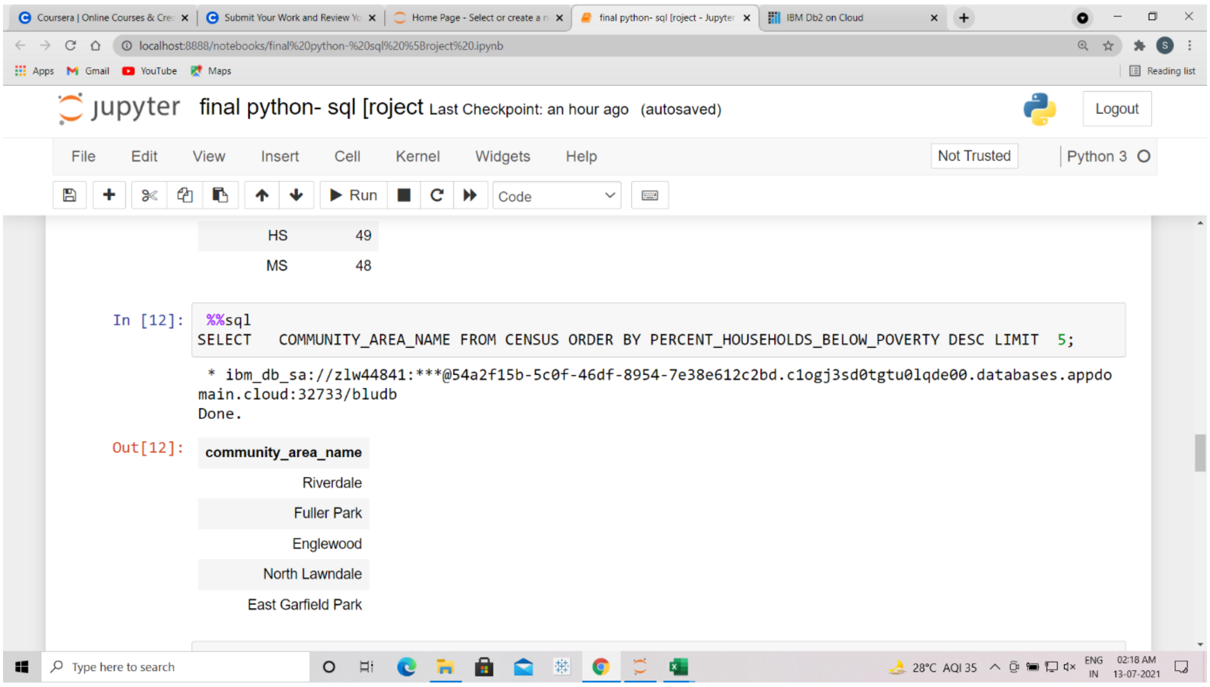
```
 * ibm_db_sa://zlw44841:***@54a2f15b-5c0f-46df-8954-7e38e612c2bd.c1ogj3sd0tgtu0lqde00.databases.appdo
main.cloud:32733/bludb
Done.
```

Out[7]:

school_type	avg_score
ES	49
HS	49
MS	48

In [12]:
```
%%sql
SELECT   COMMUNITY_AREA_NAME FROM CENSUS ORDER BY PERCENT_HOUSEHOLDS_BELOW_POVERTY DESC LIMIT  5;
```

7) List 5 community areas with highest % of households below poverty line?


Coursera | Online Courses & Cre
Submit Your Work and Review Yo
Home Page - Select or create a n
final python- sql [roject - Jupyter
IBM Db2 on Cloud
localhost:8888/notebooks/final%20python-%20sql%20%5Broject%20.ipynb
Apps
Gmail
YouTube
Maps
Reading list
jupyter
final python- sql [roject Last Checkpoint: an hour ago (autosaved)
Logout
File
Edit
View
Insert
Cell
Kernel
Widgets
Help
Not Trusted
Python 3
Run
Code
Type here to search
28°C AQI 35
ENG IN
02:18 AM
13-07-2021

HS	49
MS	48

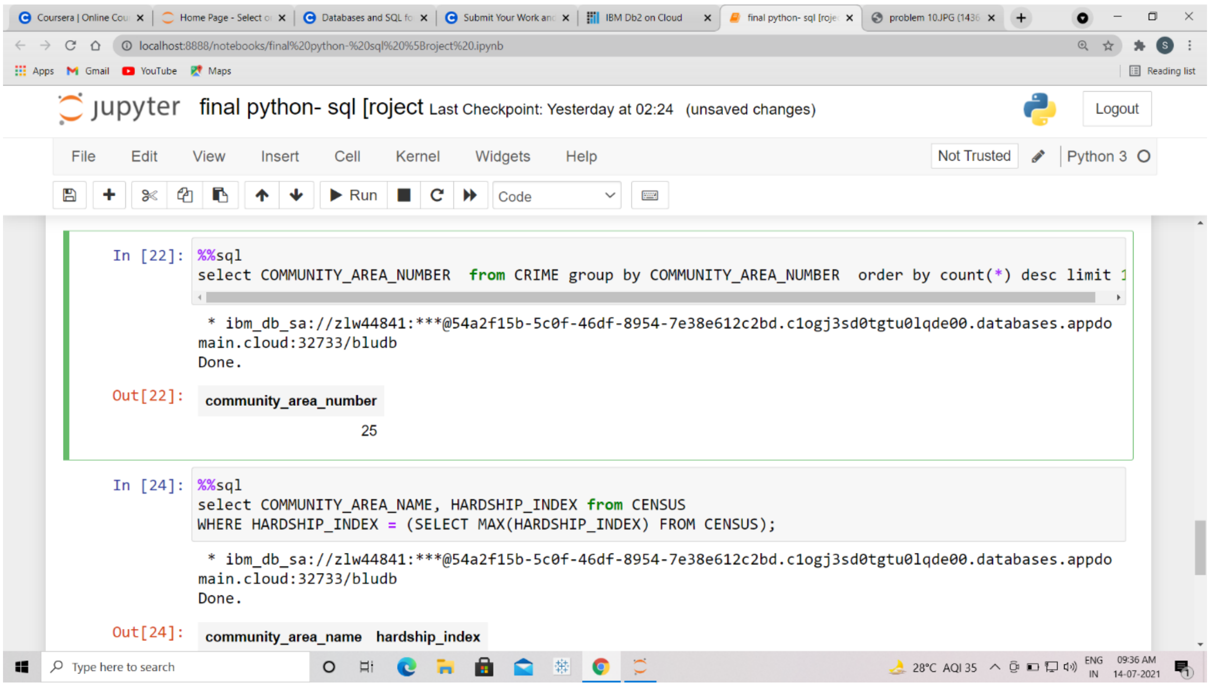
In [12]:
```
%%sql
SELECT   COMMUNITY_AREA_NAME FROM CENSUS ORDER BY PERCENT_HOUSEHOLDS_BELOW_POVERTY DESC LIMIT  5;
```

```
 * ibm_db_sa://zlw44841:***@54a2f15b-5c0f-46df-8954-7e38e612c2bd.c1ogj3sd0tgtu0lqde00.databases.appdo
main.cloud:32733/bludb
Done.
```

Out[12]:

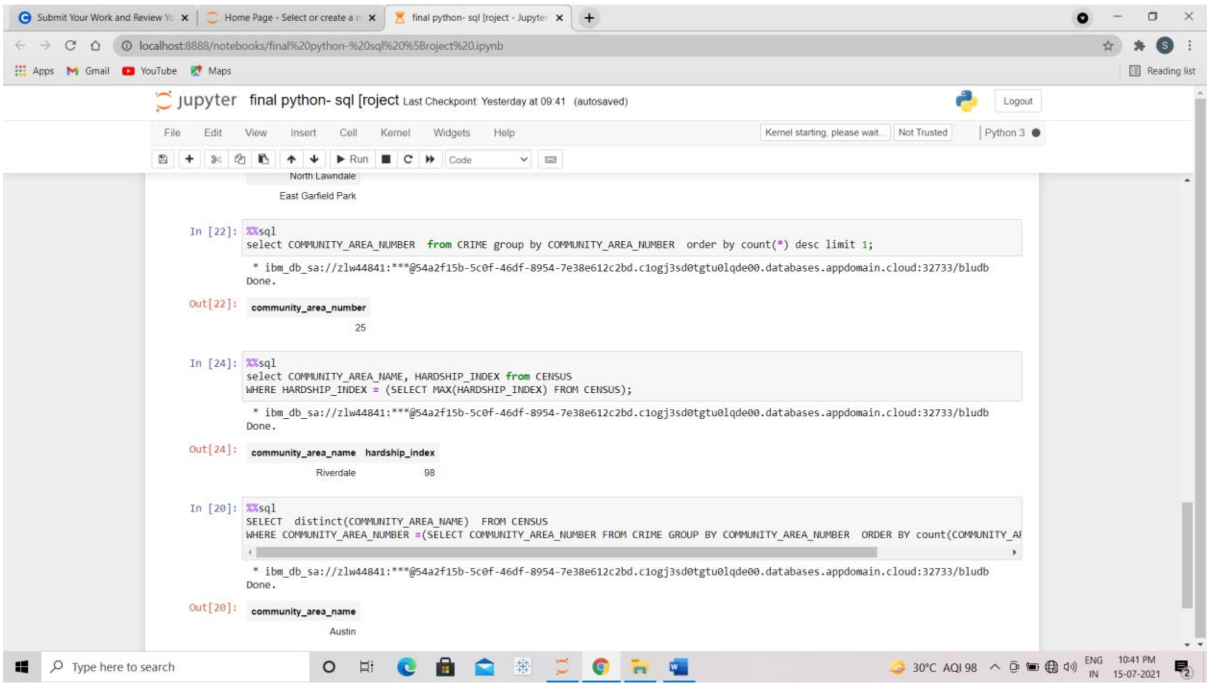
community_area_name
Riverdale
Fuller Park
Englewood
North Lawndale
East Garfield Park

8) Which community area(number) is most crime prone?


jupyter final python- sql [roject Last Checkpoint: Yesterday at 02:24 (unsaved changes)
Logout
File Edit View Insert Cell Kernel Widgets Help
Not Trusted
Python 3
Run
Code
Type here to search
28°C AQI 35
ENG IN
09:36 AM
14-07-2021

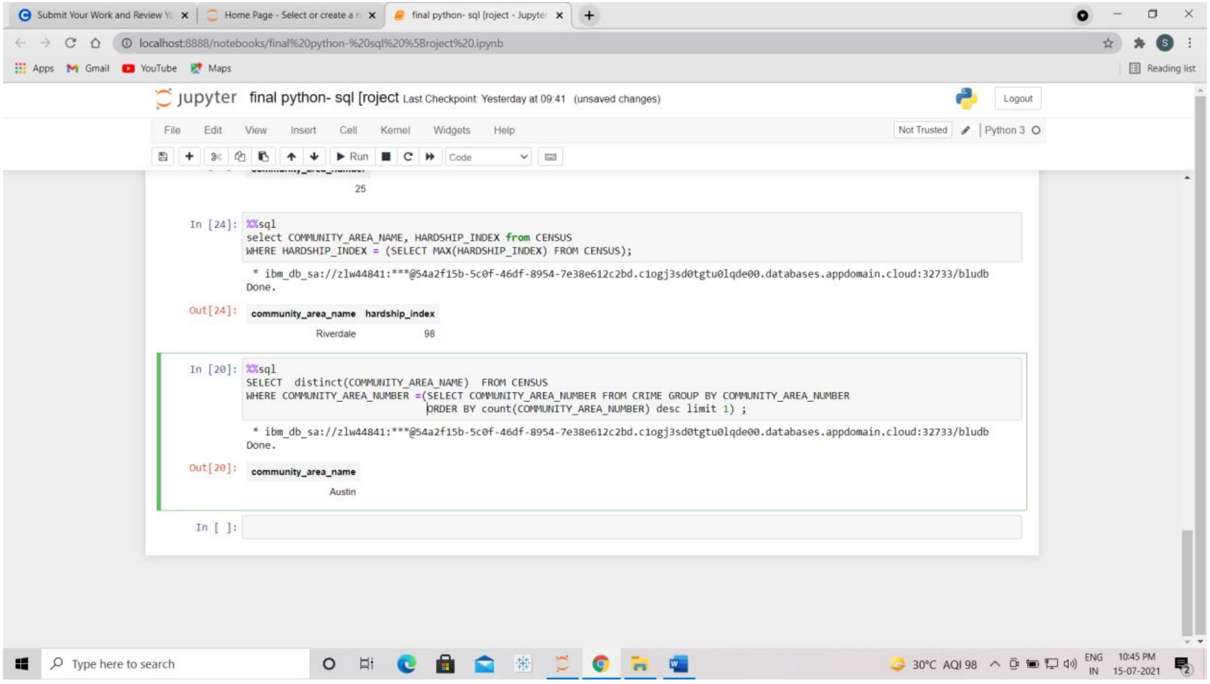
```
In [22]: %%sql
         select COMMUNITY_AREA_NUMBER  from CRIME group by COMMUNITY_AREA_NUMBER  order by count(*) desc limit 1

          * ibm_db_sa://zlw44841:***@54a2f15b-5c0f-46df-8954-7e38e612c2bd.c1ogj3sd0tgtu0lqde00.databases.appdo
         main.cloud:32733/bludb
         Done.

Out[22]: community_area_number
                            25

In [24]: %%sql
         select COMMUNITY_AREA_NAME, HARDSHIP_INDEX from CENSUS
         WHERE HARDSHIP_INDEX = (SELECT MAX(HARDSHIP_INDEX) FROM CENSUS);

          * ibm_db_sa://zlw44841:***@54a2f15b-5c0f-46df-8954-7e38e612c2bd.c1ogj3sd0tgtu0lqde00.databases.appdo
         main.cloud:32733/bludb
         Done.

Out[24]: community_area_name  hardship_index
```

9) Use a sub-query to find the name of the community area with highest hardship index?


jupyter final python- sql [roject Last Checkpoint: Yesterday at 09:41 (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Kernel starting, please wait...
Not Trusted
Python 3
Run
Code
Type here to search
30°C AQI 98
ENG IN
10:41 PM
15-07-2021

```
             North Lawndale
           East Garfield Park

In [22]: %%sql
         select COMMUNITY_AREA_NUMBER  from CRIME group by COMMUNITY_AREA_NUMBER  order by count(*) desc limit 1;

          * ibm_db_sa://zlw44841:***@54a2f15b-5c0f-46df-8954-7e38e612c2bd.c1ogj3sd0tgtu0lqde00.databases.appdomain.cloud:32733/bludb
         Done.

Out[22]: community_area_number
                            25

In [24]: %%sql
         select COMMUNITY_AREA_NAME, HARDSHIP_INDEX from CENSUS
         WHERE HARDSHIP_INDEX = (SELECT MAX(HARDSHIP_INDEX) FROM CENSUS);

          * ibm_db_sa://zlw44841:***@54a2f15b-5c0f-46df-8954-7e38e612c2bd.c1ogj3sd0tgtu0lqde00.databases.appdomain.cloud:32733/bludb
         Done.

Out[24]: community_area_name  hardship_index
                   Riverdale              98

In [20]: %%sql
         SELECT  distinct(COMMUNITY_AREA_NAME)  FROM CENSUS
         WHERE COMMUNITY_AREA_NUMBER =(SELECT COMMUNITY_AREA_NUMBER FROM CRIME GROUP BY COMMUNITY_AREA_NUMBER  ORDER BY count(COMMUNITY_AI

          * ibm_db_sa://zlw44841:***@54a2f15b-5c0f-46df-8954-7e38e612c2bd.c1ogj3sd0tgtu0lqde00.databases.appdomain.cloud:32733/bludb
         Done.

Out[20]: community_area_name
                      Austin
```

10) **Use a sub-query to determine the Community Area Name with most number of crimes?**



Submit Your Work and Review Yo
Home Page - Select or create a n
final python- sql [roject - Jupyter
localhost:8888/notebooks/final%20python-%20sql%20%5Broject%20.ipynb
Apps
Gmail
YouTube
Maps
Reading list
jupyter
final python- sql [roject Last Checkpoint: Yesterday at 09:41 (unsaved changes)
Logout
File
Edit
View
Insert
Cell
Kernel
Widgets
Help
Not Trusted
Python 3
Run
Code
Type here to search
30°C AQI 98
ENG
IN
10:45 PM
15-07-2021

```
25

In [24]: %%sql
         select COMMUNITY_AREA_NAME, HARDSHIP_INDEX from CENSUS
         WHERE HARDSHIP_INDEX = (SELECT MAX(HARDSHIP_INDEX) FROM CENSUS);

          * ibm_db_sa://zlw44841:***@54a2f15b-5c0f-46df-8954-7e38e612c2bd.c1ogj3sd0tgtu0lqde00.databases.appdomain.cloud:32733/bludb
         Done.

Out[24]: community_area_name  hardship_index
                   Riverdale              98

In [20]: %%sql
         SELECT  distinct(COMMUNITY_AREA_NAME)  FROM CENSUS
         WHERE COMMUNITY_AREA_NUMBER =(SELECT COMMUNITY_AREA_NUMBER FROM CRIME GROUP BY COMMUNITY_AREA_NUMBER
                                      ORDER BY count(COMMUNITY_AREA_NUMBER) desc limit 1) ;

          * ibm_db_sa://zlw44841:***@54a2f15b-5c0f-46df-8954-7e38e612c2bd.c1ogj3sd0tgtu0lqde00.databases.appdomain.cloud:32733/bludb
         Done.

Out[20]: community_area_name
                      Austin

In [ ]:
```

Acknowledge:

I would like to thank of YouTube to help me wherever I stuck and IBM too.