

Assignment 3

```
In [31]: import pymysql
import pandas as panda

conn=pymysql.connect(host='localhost',port=int(3307),user='root',passwd='12345',db='assignment_3')
```

Importing the Data from the Mysql Workbench

```
In [32]: chicagocensusdata = panda.read_sql_query("SELECT * FROM chicagocensusdata",conn)
chicagocrimedata = panda.read_sql_query("SELECT * FROM chicagocrimedata",conn)
chicagopublicschools = panda.read_sql_query("SELECT * FROM chicagopublicschools",conn)
```

Displaying the first few rows of the tables that we are going to use for this assignment

```
In [33]: chicagocensusdata.head()
```

Out[33]:

	COMMUNITY_AREA_NUMBER	COMMUNITY_AREA_NAME	PERCENT_OF_HOUSING_CROWDED	PERCENT_HOUSEHOLDS_BELOW_POVERTY	PERCENT_AGED_1
0	1	Rogers Park	7.7	23.6	
1	2	West Ridge	7.8	17.2	
2	3	Uptown	3.8	24.0	
3	4	Lincoln Square	3.4	10.9	
4	5	North Center	0.3	7.5	

```
In [34]: chicagocrimedata.head()
```

Out[34]:

	ID	CASE_NUMBER	DATE	BLOCK	IUCR	PRIMARY_TYPE	DESCRIPTION	LOCATION_DESCRIPTION	ARREST	DOMESTIC	...	DISTRICT	WARD
0	3512276	HK587712	2004-08-28	047XX S KEDZIE AVE	890	THEFT	FROM BUILDING	SMALL RETAIL STORE	FALSE	FALSE	...	9	14
1	3406613	HK456306	2004-06-26	009XX N CENTRAL PARK AVE	820	THEFT	\$500 AND UNDER	OTHER	FALSE	FALSE	...	11	27
2	8002131	HT233595	2011-04-04	043XX S WABASH AVE	820	THEFT	\$500 AND UNDER	NURSING HOME/RETIREMENT HOME	FALSE	FALSE	...	2	3
3	7903289	HT133522	2010-12-30	083XX S KINGSTON AVE	840	THEFT	FINANCIAL ID THEFT: OVER \$300	RESIDENCE	FALSE	FALSE	...	4	7
4	10402076	HZ138551	2016-02-02	033XX W 66TH ST	820	THEFT	\$500 AND UNDER	ALLEY	FALSE	FALSE	...	8	15

5 rows × 14 columns

In [35]: `chicagopublicschools.head()`

Out[35]:

	School_ID	NAME_OF_SCHOOL	Elementary_Middle_High_School	Street_Address	City	State	ZIP_Code	Phone_Number	
0	610038	Abraham Lincoln Elementary School	ES	615 W Kemper Pl	Chicago	IL	60614	(773) 534-5720	http://schoolreports.cps.edu/SchoolPr
1	610281	Adam Clayton Powell Paideia Community Academy ...	ES	7511 S South Shore Dr	Chicago	IL	60649	(773) 535-6650	http://schoolreports.cps.edu/SchoolPr
2	610185	Adlai E Stevenson Elementary School	ES	8010 S Kostner Ave	Chicago	IL	60652	(773) 535-2280	http://schoolreports.cps.edu/SchoolPr
3	609993	Agustin Lara Elementary Academy	ES	4619 S Wolcott Ave	Chicago	IL	60609	(773) 535-4389	http://schoolreports.cps.edu/SchoolPr
4	610513	Air Force Academy High School	HS	3630 S Wells St	Chicago	IL	60609	(773) 535-1590	http://schoolreports.cps.edu/SchoolPr

5 rows × 78 columns

Problem 1: Find the total number of crimes recorded in the CRIME table

In [36]: `totalnumberofcrimes = panda.read_sql_query("SELECT count(*) FROM chicagocrimedata", conn)
print(totalnumberofcrimes)`

```
count(*)
0      385
```

Problem 2: List community areas with per capita income less than 11000

In [37]: `areas = panda.read_sql_query("SELECT COMMUNITY_AREA_NAME,PER_CAPITA_INCOME FROM chicagocensusdata where PER_CAPITA_INCOME <11000
areas.head()`

Out[37]:

	COMMUNITY_AREA_NAME	PER_CAPITA_INCOME
0	West Garfield Park	10934
1	South Lawndale	10402
2	Fuller Park	10432
3	Riverdale	8201

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

In [38]: `totalnumberofcrimes = panda.read_sql_query("SELECT DISTINCT CASE_NUMBER, DESCRIPTION FROM chicagocrimedata WHERE DESCRIPTION LIKE
totalnumberofcrimes.head()`

Out[38]:

	CASE_NUMBER	DESCRIPTION
0	HL266884	SELL/GIVE/DEL LIQUOR TO MINOR
1	HK238408	ILLEGAL CONSUMPTION BY MINOR

Problem 4: List all kidnapping crimes involving a child?

In [39]: `kidnapping = panda.read_sql_query("SELECT DISTINCT CASE_NUMBER, PRIMARY_TYPE,DESCRIPTION FROM chicagocrimedata WHERE PRIMARY_TYPE
kidnapping.head()`

Out[39]:

	CASE_NUMBER	PRIMARY_TYPE	DESCRIPTION
0	HN144152	KIDNAPPING	CHILD ABDUCTION/STRANGER

Problem 5: What kinds of crimes were recorded at schools?

```
In [40]: school = panda.read_sql_query("SELECT DISTINCT PRIMARY_TYPE FROM chagocrimedata WHERE LOCATION_DESCRIPTION LIKE '%SCHOOL%'", cc)
school.head()
```

Out[40]:

	PRIMARY_TYPE
0	CRIMINAL DAMAGE
1	NARCOTICS
2	CRIMINAL TRESPASS
3	PUBLIC PEACE VIOLATION

Problem 6: List the average safety score for each type of school

```
In [41]: safety_score = panda.read_sql_query("SELECT Elementary_Middle_High_School, AVG(SAFETY_SCORE) AVERAGE_SAFETY_SCORE FROM chagocpu")
safety_score.head()
```

Out[41]:

	Elementary_Middle_High_School	AVERAGE_SAFETY_SCORE
0	ES	44.8264
1	HS	49.7250
2	MS	53.0000

Problem 7: List 5 community areas with highest % of households below poverty line

```
In [42]: poverty = panda.read_sql_query("SELECT COMMUNITY_AREA_NAME, PERCENT_HOUSEHOLDS_BELOW_POVERTY FROM chagocensusdata ORDER BY PER")
poverty.head()
```

Out[42]:

	COMMUNITY_AREA_NAME	PERCENT_HOUSEHOLDS_BELOW_POVERTY
0	Riverdale	56.5
1	Fuller Park	51.2
2	Englewood	46.6
3	North Lawndale	43.1
4	East Garfield Park	42.4

Problem 8: Which community area is most crime prone?

```
In [43]: crimeprone = panda.read_sql_query("SELECT COMMUNITY_AREA_NUMBER, COUNT(COMMUNITY_AREA_NUMBER) FROM chagocrimedata GROUP BY COM")
crimeprone.head()
```

Out[43]:

	COMMUNITY_AREA_NUMBER	COUNT(COMMUNITY_AREA_NUMBER)
0	0	29

Problem 9: Use a sub-query to find the name of the community area with highest hardship index 📌

```
In [44]: hardshipindex = panda.read_sql_query("SELECT COMMUNITY_AREA_NAME, HARDSHIP_INDEX FROM chagocensusdata WHERE HARDSHIP_INDEX =")
hardshipindex.head()
```

Out[44]:

	COMMUNITY_AREA_NAME	HARDSHIP_INDEX
0	Riverdale	98

Problem 10: Use a sub-query to determine the Community Area Name with the most # of crimes?

```
In [45]: mostcrimes = panda.read_sql_query("SELECT COMMUNITY_AREA_NAME FROM chicagocensusdata WHERE COMMUNITY_AREA_NUMBER = (SELECT COMM
mostcrimes.head()
```

Out[45]:

	COMMUNITY_AREA_NAME
0	CHICAGO

```
In [ ]:
In [ ]:
```