Chicago's Data Portal Analysis Report

Using three datasets that are available on the city of Chicago's Data Portal:

1. Socioeconomic Indicators in Chicago
2. Chicago Public Schools
3. Chicago Crime Data

1. Socioeconomic Indicators in Chicago

This dataset contains a selection of six socioeconomic indicators of public health significance and a "hardship index," for each Chicago community area, for the years 2008 – 2012.
A detailed description of this dataset and the original dataset can be obtained from the Chicago Data Portal at:
https://data.cityofchicago.org/Health-Human-Services/Census-Data-Selected-socioeconomic-indicators-in-C/kn9c-c2s2

2. Chicago Public Schools

This dataset shows all school level performance data used to create CPS School Report Cards for the 2011-2012 school year.
A detailed description of this dataset and the original dataset can be obtained from the Chicago Data Portal at:
https://data.cityofchicago.org/Education/Chicago-Public-Schools-Progress-Report-Cards-2011-/9xs2-f89t

3. Chicago Crime Data

This dataset reflects reported incidents of crime (with the exception of murders where data exists for each victim) that occurred in the City of Chicago from 2001 to present, minus the most recent seven days.
A detailed description of this dataset and the original dataset can be obtained from the Chicago Data Portal at:
https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-present/ijzp-q8t2

This dataset is quite large, exceeding 1.5GB in size with over 6.5 million rows. For the purposes of this assignment, I will be using a much smaller subset of this dataset.

Database Used in this Lab: **MySQL learners** database has been used in this Analysis.
I'm creating and inserting data into the below mentioned 3 tables

1.chicago_public_schools
2.chicago_socioeconomic_data
3.chicago_crime

I will be using 3 dump files for this analysis purpose.

1. chicago_public_schools
2. chicago_crime
3. chicago_socioeconomic_data

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

Ask:

Objective: Determine the total number of crimes recorded in the Chicago Crime dataset.

Prepare:

Data Source: Chicago Crime Data
Database Used: My learners database
Table Used: `chicago_crime`

Process:

Query to find the total number of crimes

```
SELECT COUNT(*) AS Total_Number_of_Crime
FROM `chicago_crime`;
```

Analyze:

Result:
Total Number of Crimes: 533

Share:

Reporting Format: PPT
Stakeholders: Law enforcement, city planners.

Act:

Recommendation: Monitor crime trends and allocate resources based on high-crime areas.

Problem 2: Retrieve first 10 rows from the CRIME table.

Ask:

Objective: Retrieve the details of the first 10 recorded crimes.

Prepare:

Data Source: Chicago Crime Data
Database Used: My learners database
Table Used: `chicago_crime`

Process:

Query to retrieve first 10 rows from the CRIME table

```
SELECT * FROM `chicago_crime`
LIMIT 10;
```

Analyze:

Result: Details of the first 10 recorded crimes.

Share:

Reporting Format: PPT
Stakeholders: Law enforcement, analysts.

Act:

Recommendation: Examine specific crime details for initial analysis.

Problem 3: How many crimes involve an arrest?

Ask:

Objective: Determine the number of crimes involving an arrest.

Prepare:

Data Source: Chicago Crime Data

Database Used: My learners database

Table Used: `chicago_crime`

Process:

Query to count crimes involving an arrest

```
SELECT COUNT(*) AS Crimes_involed_an_Arrest
FROM `chicago_crime`
WHERE `ARREST` = 'TRUE';
```

Analyze:

Result: Crimes involving an arrest: 163

Share:

Reporting Format: PPT

Stakeholders: Law enforcement, public safety.

Act:

Recommendation: Focus on strategies to reduce arrests by addressing underlying issues.

Problem 4: Which unique types of crimes have been recorded at GAS STATION locations?

Ask:

Objective: Identify unique types of crimes recorded at GAS STATION locations.

Prepare:

Data Source: Chicago Crime Data

Database Used: My learners database

Table Used: `chicago_crime`

Process:

Query to find unique crime types recorded at GAS STATION locations

```
SELECT DISTINCT(PRIMARY_TYPE) AS Types_of_Crimes
FROM `chicago_crime`
WHERE `LOCATION_DESCRIPTION` = 'GAS STATION';
```

Analyze:

Result:

Types of Crimes at GAS STATION:

THEFT,

NARCOTICS,

ROBBERY,

CRIMINAL TRESPASS.

Share:

Reporting Format: PPT

Stakeholders: Gas station owners, law enforcement.

Act:

Recommendation: Implement security measures at gas stations based on prevalent crime types.

Problem 5: In the CENUS_DATA table list all Community Areas whose names start with the letter ‘B’.

Ask:

Objective: List all community areas starting with the letter 'B'.

Prepare:

Data Source: Chicago Public Schools - Census Data
Database Used: My learners database
Table Used: `chicago_public_schools`

Process:

Query to list all Community Areas starting with 'B'

```
SELECT DISTINCT `COMMUNITY_AREA_NAME`
FROM `chicago_public_schools`
WHERE `COMMUNITY_AREA_NAME` LIKE 'B%';
```

Analyze:

Result:
Community Areas starting with 'B':
BEVERLY,
BELMONT CRAGIN,
BRIGHTON PARK,
BRIDGEPORT,
BURNSIDE.

Share:

Reporting Format: PPT
Stakeholders: Community planners, city officials.

Act:

Recommendation: Explore targeted community development initiatives for areas starting with 'B'.

Problem 6: Which schools in Community Areas 10 to 15 are healthy school certified?

Ask:

Objective: Identify healthy school certified schools in Community Areas 10 to 15.

Prepare:

Data Source: Chicago Public Schools - Progress Report Cards 2011-2012
Database Used: My learners database
Table Used: `chicago_public_schools`

Process:

Query to find healthy school certified schools in Community Areas 10 to 15

```
SELECT `NAME_OF_SCHOOL`
FROM `chicago_public_schools`
WHERE `COMMUNITY_AREA_NUMBER` BETWEEN 10 AND 15
AND `HEALTHY_SCHOOL_CERTIFIED` = 'YES';
```

Analyze:

Result:
Healthy school certified schools in Community Areas 10 to 15:
Rufus M Hitch Elementary School

Share:

Reporting Format: PPT
Stakeholders: School administrators, health officials.

Act:

Recommendation: Encourage and promote health initiatives in schools.

Problem 7: What is the average school Safety Score?

Ask:

Objective: Determine the average school safety score.

Prepare:

Data Source: Chicago Public Schools - Progress Report Cards 2011-2012

Database Used: My learners database

Table Used: `chicago_public_schools`

Process:

Query to find the average school Safety Score

```
SELECT AVG(`SAFETY_SCORE`) as Average_Safety_Score
FROM `chicago_public_schools`;
```

Analyze:

Result: Average Safety Score: 49.50

Share:

Reporting Format: PPT

Stakeholders: Parents, school administrators.

Act:

Recommendation: Focus on improving safety scores through targeted interventions.

Problem 8: List the top 5 Community Areas by average College Enrollment number of students.

Ask:

Objective: List the top 5 Community Areas by average college enrollment.

Prepare:

Data Source: Chicago Public Schools - Progress Report Cards 2011-2012
Database Used: My learners database
Table Used: `chicago_public_schools`

Process:

Query to list top 5 Community Areas by average college enrollment

```
SELECT `COMMUNITY_AREA_NAME`, AVG(`COLLEGE_ENROLLMENT`) as Avg_Enrollment
FROM `chicago_public_schools`
GROUP BY `COMMUNITY_AREA_NAME`
ORDER BY Avg_Enrollment DESC
LIMIT 5;
```

Analyze:

Result: Top 5 Community Areas by average college enrollment:
ARCHER HEIGHTS, MONTCLARE, BELMONT CRAGIN, WEST ELSDON, NORTH CENTER

Share:

Reporting Format: PPT
Stakeholders: Education administrators, community leaders.

Act:

Recommendation: Support and allocate resources for education in high-enrollment areas.

Problem 9: Use a sub-query to determine which Community Area has the least value for school Safety Score?

Ask:

Objective: Identify the Community Area with the least school safety score.

Prepare:

Data Source: Chicago Public Schools - Progress Report Cards 2011-2012

Database Used: My learners database

Table Used: `chicago_public_schools`

Process:

Query using a sub-query to find the Community Area with the least school safety score

```
SELECT `COMMUNITY_AREA_NAME`, `SAFETY_SCORE`
FROM `chicago_public_schools`
WHERE `SAFETY_SCORE` IN (SELECT MIN(SAFETY_SCORE) FROM `chicago_public_schools`);
```

Analyze:

Result: Community Area with the least safety score:

WASHINGTON PARK

Share:

Reporting Format: PPT

Stakeholders: Education administrators, community leaders.

Act:

Recommendation: Implement safety measures in Washington Park schools.

Problem 10: Without using an explicit JOIN operator] Find the Per Capita Income of the Community Area which has a school Safety Score of 1.

Ask:

Objective: Find the per capita income of the Community Area with a school safety score of 1.

Prepare:

Data Source: Chicago Socioeconomic Data, Chicago Public Schools - Progress Report Cards 2011-2012

Database Used: My learners database

Tables Used: `chicago_socioeconomic_data`, `chicago_public_schools`

Process:

Query without using explicit JOIN to find Per Capita Income of the Community Area with Safety Score of 1

```
SELECT CPS.`COMMUNITY_AREA_NAME`, CSE.`PER_CAPITA_INCOME`
FROM `chicago_socioeconomic_data` CSE, `chicago_public_schools` CPS
WHERE CSE.`COMMUNITY_AREA_NUMBER` = CPS.COMMUNITY_AREA_NUMBER
AND CPS.SAFETY_SCORE = 1;
```

Analyze:

Result: Per Capita Income of Community Area with Safety Score of 1: WASHINGTON PARK, $13,785

Share:

Reporting Format: PPT

Stakeholders: City planners, policymakers.

Act:

Recommendation: Explore socioeconomic development initiatives for Washington Park.

Conclusion

In this analysis of Chicago's crime data, school information, and community statistics, key findings have been derived from a diverse set of problem statements. The dataset, comprising over 6.5 million rows, revealed a total of 533 recorded crimes, with a specific focus on the first 10 entries and the arrest status of these incidents. Crimes at gas station locations included theft, narcotics, robbery, and criminal trespass. Community areas starting with 'B,' such as Beverly and Belmont Cragin, were identified. Additionally, the study highlighted the only healthy school certified in community areas 10 to 15, namely Rufus M Hitch Elementary School. The average safety score across all schools was found to be 49.50. The top 5 community areas by average college enrollment included Archer Heights, Montclare, Belmont Cragin, West Elsdon, and North Center. Washington Park emerged as the community area with the least safety score (1), while its per capita income was determined to be $13,785. These insights provide a holistic view for stakeholders, aiding informed decision-making and community development initiatives.