**ASSIGNMENT COVER PAGE**

**BIDA23-043**

**STUDENT ID ;**

**WANGU MAPOSA**

**STUDENT NAMES;**

**BIDA23-043@THUTO.BAC.AC.BW**

**STUDENT EMAIL;**

**2022-2027**

**COHORT;**

**Crimes in L.A 2020-2024**

**STUDENT ID ;** BIDA23-043

**STUDENT NAMES;** WANGU MAPOSA

**STUDENT EMAIL;** BIDA23-043@THUTO.BAC.AC.BW

**COHORT;** 2022-2027

**ASSIGNMENT TITLE ;**WEB AND MULTIMEDIA DEVELOPMENT

**PROGRAMME OF STUDY ;**BUSINESS INTELLIGENCE AND DATA ANALYTICS

**YEAR OF STUDY ;** YEAR 1 SEMESTER 2

**ASSIGNMENT TITLE ;**

**Business intelligence and data analytics**

**PROGRAMME OF STUDY ;**

**Year 2 semester 1**

**YEAR OF STUDY ;**

**L.A CRIME DATA 2020-2024**

This is a report on the crimes in L.A from 2020 to 2024 ,It was taken from the Kaggle website .With this dataset we identified the variables of interest the defects of the data ,as well as using excel to summarize the data and get insights. The analysis of the data will in turn help the department of the police to be able to identify trends in these crimes ,find hotspots and be able to come up with measures that can be used in the future to reduce crime rates in L.A .

**Variables of interest**

The data set provided has several variables some of which are important than the others when it comes to the analysis of the data and generating insights ,the following are the variables that are of interest and will help identify patterns of interest that will help the department to find efficient and effective to try fight crime in L.A

**Date of occurrence –** the date of occurrence is a quantitative datatype that is measured using the ordinal scale .It shows the days these crimes took place .The date of the crime can reveal trends over time, it shows the days the crime rates are at their peak ,this will help get a clear understanding of which days crime rates are expected to be high ,thus on days like Christmas holidays law enforcements and resources will be used the most to try mitigate the crimes in L.A.

**Time of occurrence** -The time of day when crimes happen can help in understanding the pattern of exactly when during the day these crimes occur, we can be able to identify whether these crimes happen the most during the day or at night this will help in knowing when exactly resources and law enforcements should be maximized .The time occurrence variable is a quantitative data type which is measured using the interval scale as it helps us compare between intervals of times of the day where crimes are at their peak

**Area name** -This is a variable of interest as the area in which the crimes are done provide a clear picture of the hotspots of crimes, with the hotspots identified a clear idea of how these crime rates can be tackled as well as knowing which area exactly to focus on allows for efficient allocation of resources to areas affected the most by crimes . The location can also serve as a means of comparing the areas that are affected to try and find a common pattern or characteristic among these areas so that we know which areas are targeted the most based on the characteristic identified .This variable is a categorical datatype that is measured by the nominal scale ,it represents the names of the areas affected by crimes in no specific order.

**DEFECTS OF THE DATASET**

Data defects are the problems that were associated with the dataset, some of them affected the analysis of the data while others did not .The defects associated with this data are ,Missing /null values ,inconsistent data ,invalid data. Below are the descriptions of these defects.

MISSING DATA

The dataset had several blanks in the cells which can cause an inaccurate or biased analysis of the data to clean it I used excel.

**TO CLEAN THE DATA :** I identified the missing values using the filter tool to filter each column for cells with blanks .For columns with numerical values I replaced the blanks using the average of the column so that the skewness of the data is maintained ,I used the formula =AVERAGE to calculate the mean of the column

For columns that included text, or categorical data used the find and replace tool to fill the empty columns with “unknown “ ,which is a place holder to indicate the incomplete information .

INVALID DATA TYPES

The dataset has columns where the datatype is inconsistent with the expected data types ,In our dataset the columns date of occurrence and date reported we expect them to be a date data type however they are saved as text strings ,same applies to the time of occurrence column where the time is saved a text string .

**TO CLEAN THE DATA** : I first selected the columns with the identified defect and used the TEXT TO COLUMN tool to format the text string into the expected datatype ,i then specified the date format as well as time format ,the excel converted the values into the valid and correct data types

EXTRANEOUS DATA /INCONSISTENT DATA FORMATTING

In the dataset we see this defect in the date reported and date of occurrence where the time values which were insignificant were appended to the date values .This defect makes ,analysing challenging as well as filtering data.

**TO CLEAN THE DATA**: To solve this defect I separated the time values that were appended to the dates by using the text to column tool, I used the delimiters option in the tool to separate them ,as a result the time values where now in a separate column which I had to delete as it was not of any use in the analysis of the data.

**DATA SUMMARIZATION**

In this part we are going to be looking at charts that will give us trends that will be of use in the police department ,these charts will provide insights that will help in the decision making of measures that will be put in place to help fight crime.

**TOP FIVE WEAPONS USED**

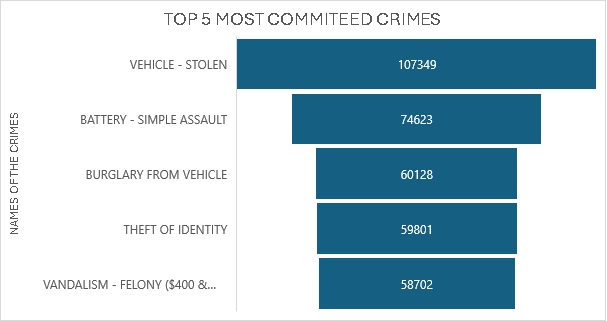
In the pie chart below shows the top 5 weapons used during crimes, it shows that the most used weapon is Strong-arms (hands ,fist, feet or bodily force ) which tells us that the crime that is being committed here is violent therefore this will increase our police department’s readiness as they will know what they are up against from this graph we can also generate

**TOP 5 AREAS OF CRIMES**

The chart below shows the top 5 areas with the highest crime rates recorded, It gives a clear picture of the hotspots of crime in L.A .With the hotspots of the crime identified the department can use it to allocate their resources efficiently around these areas like increasing patrolling groups around these areas ,This chart can also serve a means to track whether or not the strategies implemented have worked or not ,if the crimes recorded in these areas are still high it means that the strategies are not effective hence a different strategy can be used until the numbers significantly reduce.

**TOP 5 CRIMES COMMITED**

**T**he funnel chart below shows the top 5 crimes committed around L.A that were recorded**,** from it we can see the highest committed crime in L.A and the police departments can have the crime context and know which resources can be used so as to avoid wasting the resources .



**DASHBOARD**

****

The screenshot of the interactive dashboard above depicts 4 charts which will be of great use to the department of the police, the first chart shows us the victim sex by the crime ,this chart will show us which gender is most vulnerable to which crimes ,the second chart shows us the crime trend annually ,it will help us track the rates of crimes and check whether or not measures enforced to curb crime rates are effective or not .The chart of the top ten hotspots shows the areas which the police department should focus on ,the last chart shows us the top ten weapon used ,the police are able to know the crime type by just looking at the weapon that is commonly used e.g if guns are frequently used it means that the crime is violent thus the police will know what they are up against and this will help them be well prepared for what they are up against.

**Solutions and recommendations**

**Educate the public and engage it**

For the areas that are affected by crimes , awareness campaigns can be conducted to help raise awareness to the communities so that they are aware of the situation at hand ,frequent community meetings can also be conducted where neighbourhood watch groups can be established so that the neighbours can keep the lookout for each other .Furthermore self defence lessons can be conducted during these campaigns and meetings

**Data resource driven allocation of resources**

From the top 5 areas that are mostly affected by crimes we see these areas which are the 77th street ,central area ,Hollywood ,N-Hollywood and Newton are the hotspots for crimes ,The police department can use this insight to make sure that they allocate resources wisely in tis case the resources would be allocated and increased in these areas ,These resources include the police patrol groups ,crime reduction initiative which include placing cctv in areas such as the mall and property marking

**Implement strict laws**

The data set provided us with the top 5 weapons used during crimes, and these weapons include ,hand guns ,semi automated pistols as well as bodily force and from the weapons used we can tell that these weapons were used to commit violent crimes ,to curb these crime rates the police department can ban the weapons that used during the crimes such as the hand guns ,or alternatively increase the waiting period for the weapon purchases so that a thorough background research of the person purchasing the weapon can be conducted and regular search raids can be conducted by the police to make sure that people don’t go around carrying weapons that have the potential to be used in a crime .

**Regular crime trends reports**

Regular reports on crimes will help the department generate insights from which they can learn which areas are mostly affected the days crimes are at their peak the type of crimes that are mostly committed as well as the weapons used to commit these crimes ,these insights will help the police to make informed decisions such as where resource should be allocated and laws that can be implemented to fight crimes .More over these trends can help track the performance of the measurements put in place to curb the crime rates.