

**US Crime Data Exploration and Analysis**

Course-End Project Problem Statement





**Course-End Project:US Crime Data Exploration and Analysis**

### Problem Scenario:An organized quantitative and qualitative investigation is done to find trends in crime and disorder.Information on these patterns helps law enforcement agencies deploy resources more effectively. Crime analysis plays an important role in devising solutions to crime problems and formulating crime prevention strategies.

**Problem Objective:**It is required to delve deeper into data on different types of crimes and figure out the types of crimes which are more frequent and how they are trending over time.

**DataDescription:**

### The file Crime\_us.csv contains the details of crimes, that have occurred in a state of the US in the year 2022.

**Variable Description:**

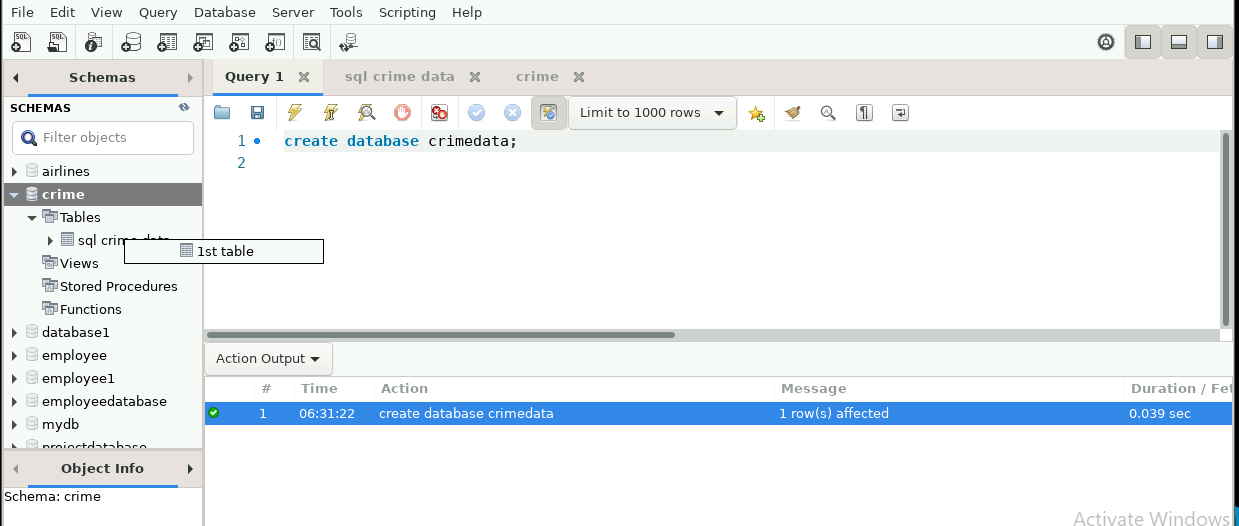
|  |  |
| --- | --- |
| **Variable** | **Description** |
| ID | Shows case id |
| Case Number | Shows case number |
| Date | Explicit the date and time of occurrence of the crime |
| Block | Explains the address where the crime had taken place |
| IUCR | Is a four-digitcode that law enforcement agencies use to classify criminal incidents when taking individual reports |

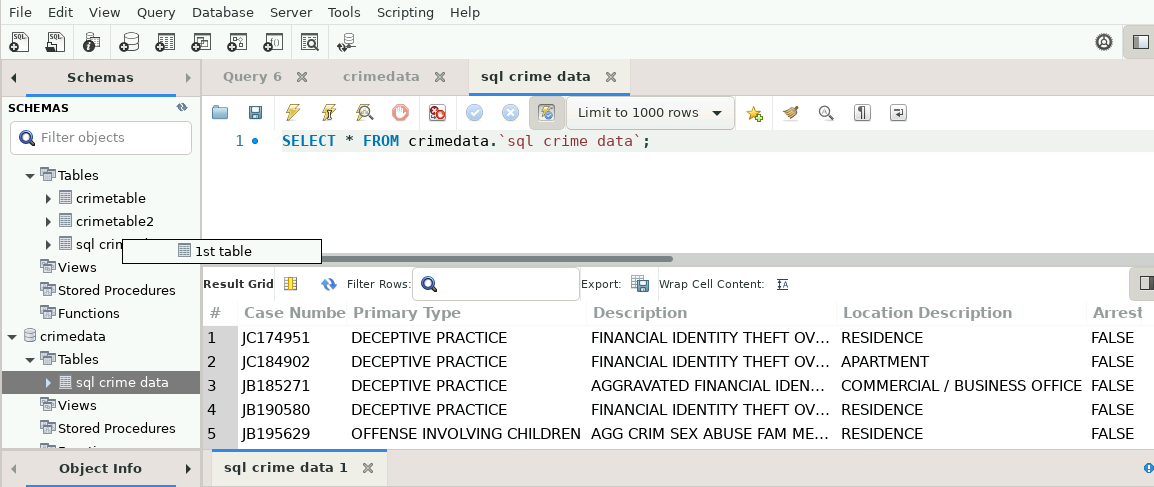
|  |  |
| --- | --- |
| Primary Type | Classify the type of crime |
| Description | Demonstrates the crime event |
| Location Description | Explains the type of location where the crime has taken place |
| Arrest | Shows whether an arrest had been made or not |
| Domestic | Shows whether the crime was of domestic nature |
| Beat | Is the territory that a police officer is assigned to patrol |
| District | Shows the district of the victim |
| Ward | Shows the ward of the victim |
| Community Area | Shows the community of the victim |
| FBI Code | Is the investigation code used to find the criminals |
| X Coordinate | Shows various information about thelocation |
| Y Coordinate | Shows various information about thelocation |
| Year | Shows the year |
| Update On | Signifies the last updating date of the data |
| Latitude | Shows the latitude of the place |
| Longitude | Shows the longitude of the place |
| Location | Shows the location |

**Steps to Perform:**

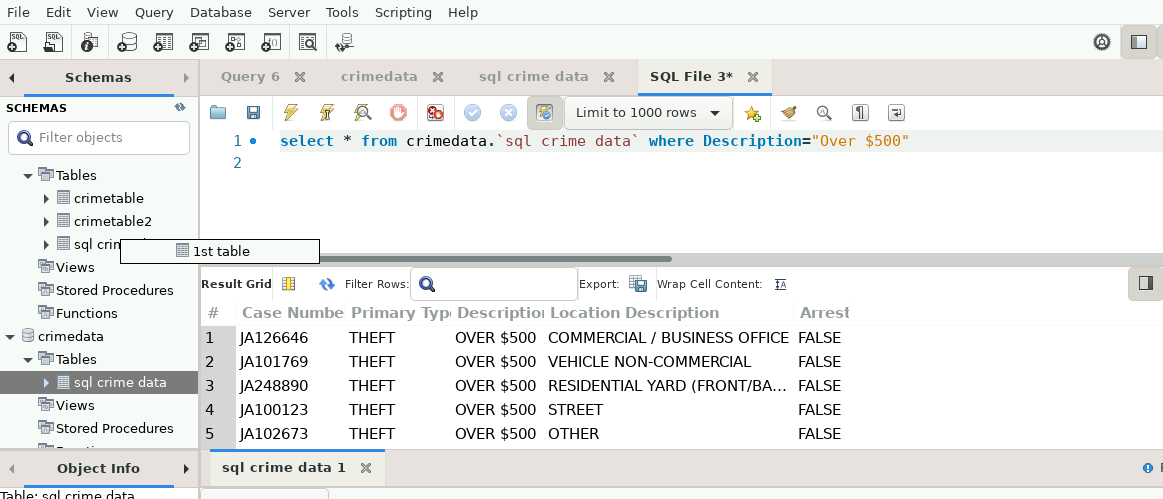
1. Create a SQL database containing data related to the case number, primary crime category, crime description, crime location, and arrest status using the dataset.

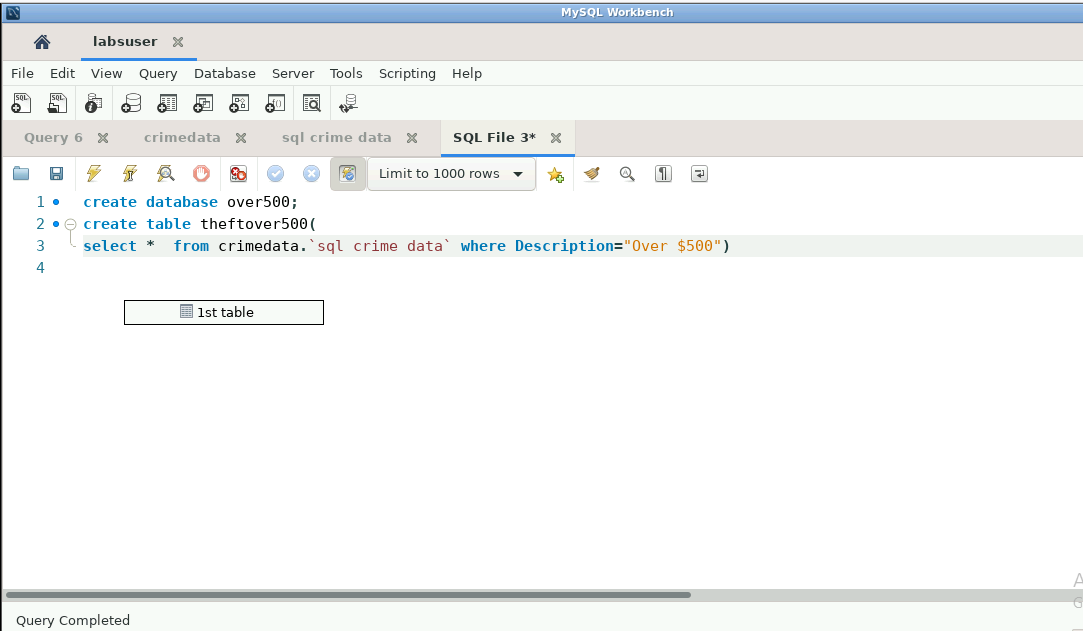
Create database crimedata.



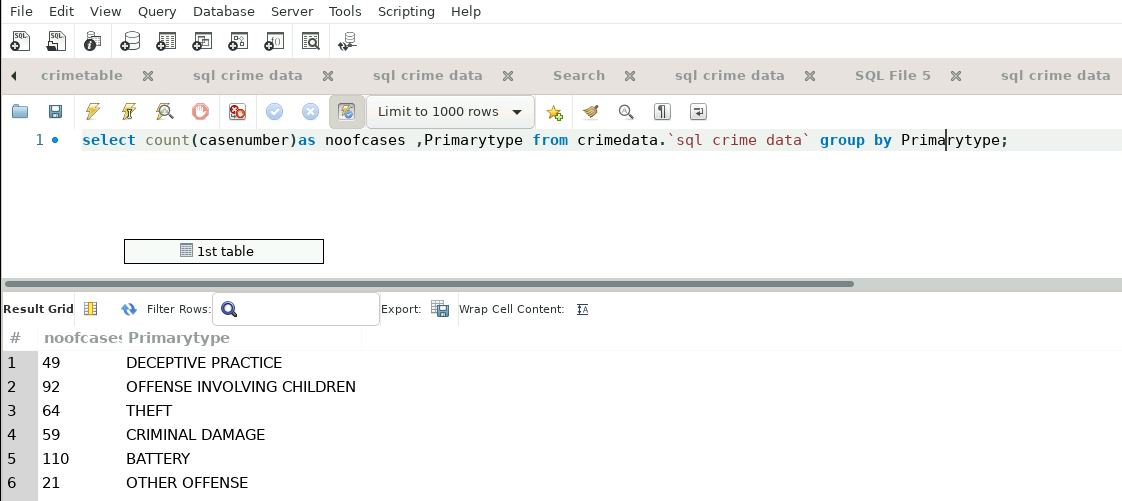


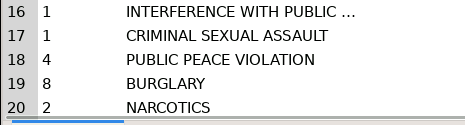
1. Make a database in SQL where theft costs more than $500.





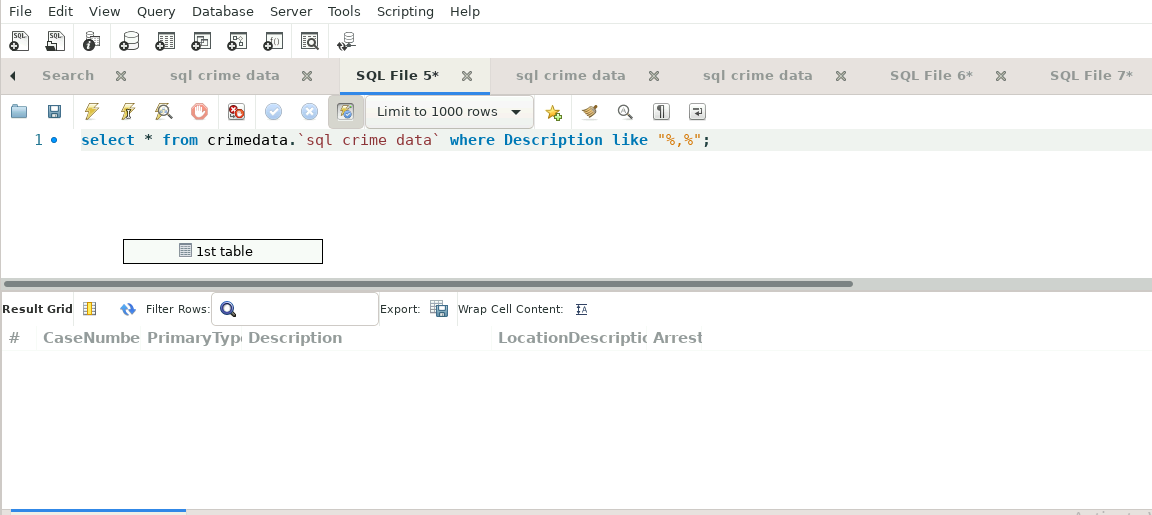
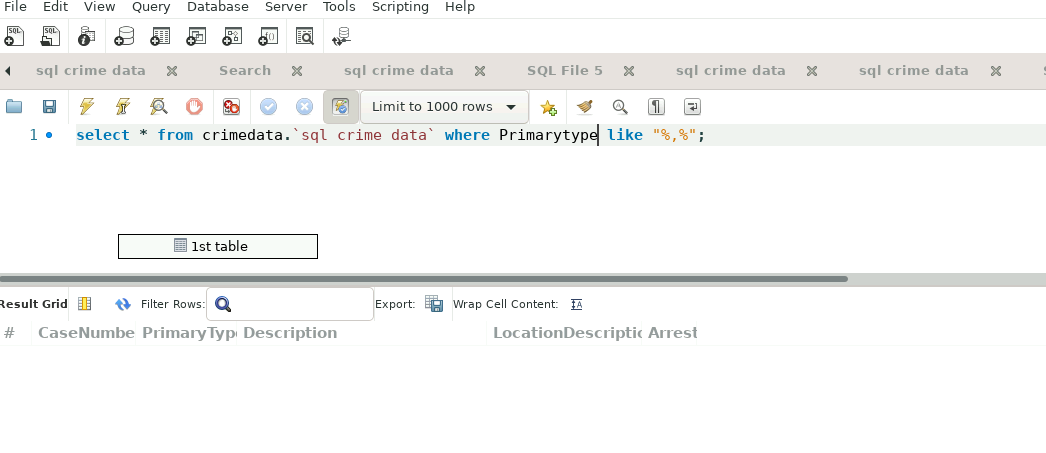
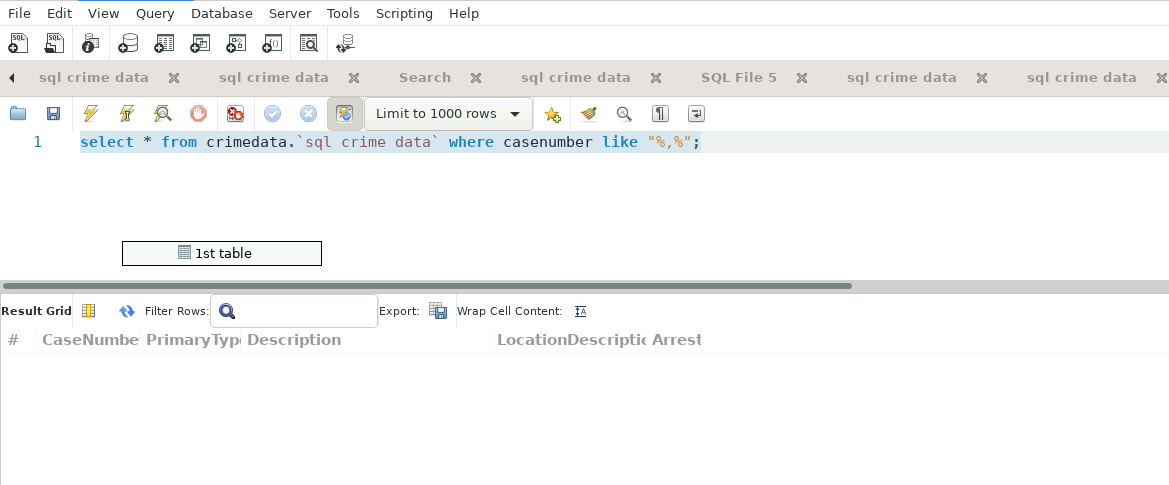
1. Determine the overall number of cases for each major category of crime.

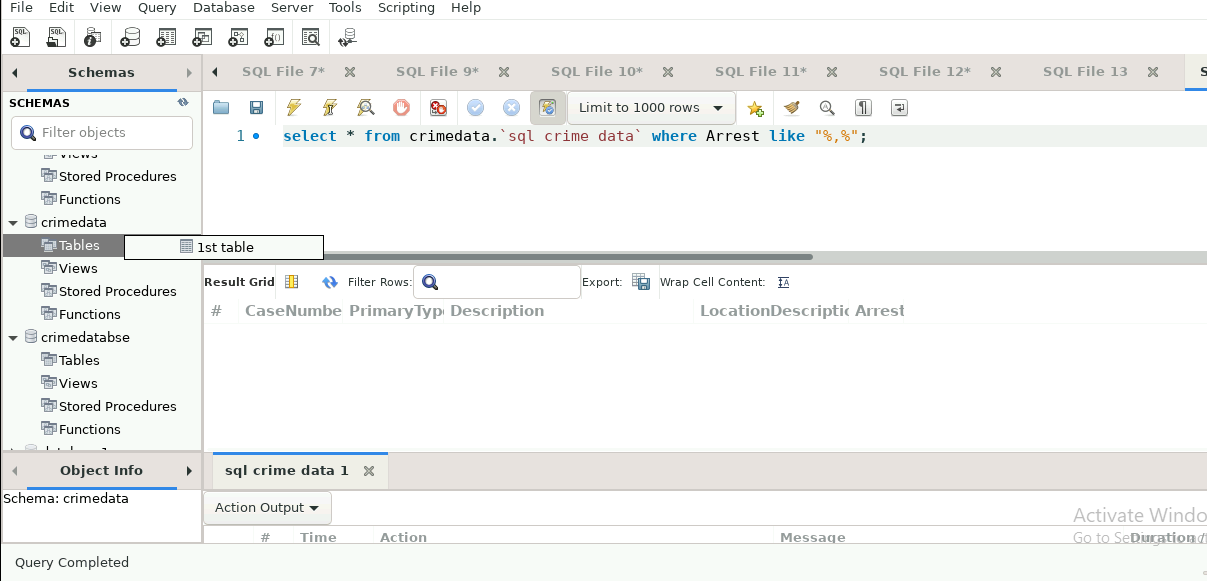


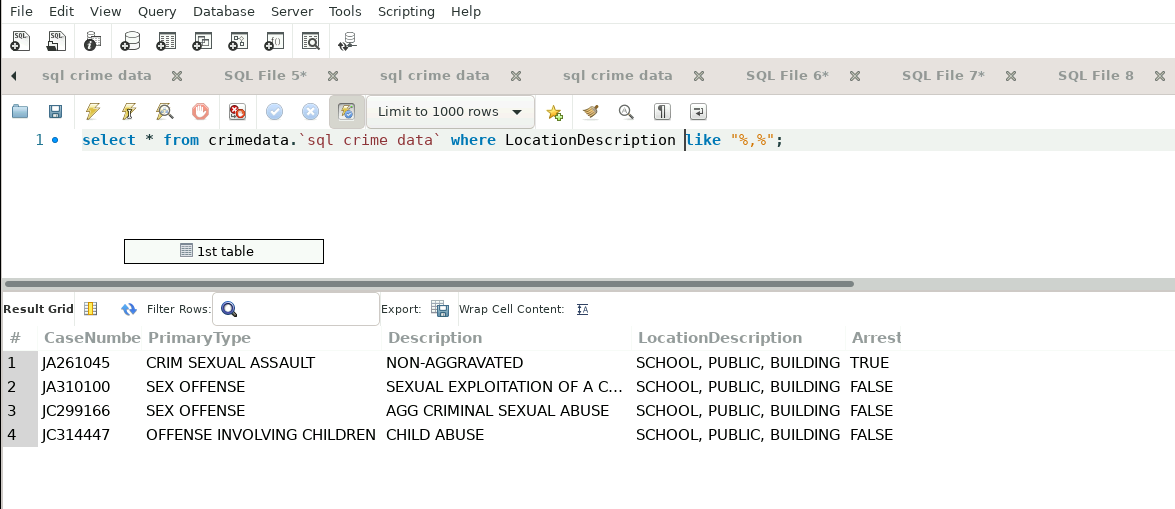


1. Apply 1NF normalization to the dataset provided.

Checking for multiple va;lues if any column have multiple values.







Here we have multiple values in same record so we need to handle it and create seperate columns like location1 ,location2,location3

select casenumber,Primarytype,Description,LocationDescription,

if(LocationDescription like "%,%",(substring(LocationDescription,0,((position("," in LocationDescription))-1))),LocationDescription)as location1,

if(LocationDescription like "%,%",(substring(LocationDescription,((position("," in LocationDescription))+1)),((position("," in LocationDescription))-1)),LocationDescription)as location2,

if(LocationDescription like "%,%",(substring(LocationDescription,((position("," in LocationDescription))+1)),len(LocationDescription)),LocationDescription)as location3,

Arrest

from crimedata.`sql crime data`;

