BI

Mini Project

Crime analysis in India

Table of Contents

[1. Introduction 3](#_Toc534726633)

[**1.1.** **Setup Checklist for Mini Project** 3](#_Toc534726634)

[**1.2.** **Instructions** 3](#_Toc534726635)

[2. Problem Statement 4](#_Toc534726636)

[**2.1.** **Objective** 4](#_Toc534726637)

[**2.2.** **Technology used:** 4](#_Toc534726638)

[3. Implementation in BI LOT 5](#_Toc534726639)

[**3.1 SUMMARY OF THE FUNCTIONALITY TO BE BUILT:** 5](#_Toc534726640)

[**3.2 Guidelines on the functionality to be built:** 5](#_Toc534726641)

# **Introduction**

This document outlines a mini project for the BI LOT. The project is to analyseCrimes in India. This document contains the work flow of the system and gives guidelines on how to build the functionality gradually in each of the course modules of the BI LOT.

## **Setup Checklist for Mini Project**

Minimum System Requirements

* Intel Pentium 90 or higher (P166 recommended)
* Microsoft Windows 2010 or above
* Memory: 4GB of RAM (4GB or more recommended)
* Internet Explorer 10.0 or higher
* Oracle 11g / SQL Server 2012 or higher
* Informatica 10.2.0 / Data Stage / Abinitio
* Tableau Desktop / SSRS / PowerBI /Qlikview

## **Instructions**

* The code modules in the mini project should follow all the coding standards.
* Create a directory by your name in drive **<drive>**. In this directory, create a subdirectory **MiniProject**. Store your Project here.
* You can refer to your course material.
* You may also look up the help provided in the BI docs and documentation provided in respective tools.

# **Problem Statement**

## **Objective**

* To find the Safest State with respect to crime all over the country (India).
* To find the Most Unsafe State with respect to crime all over the country.
* To find the Most Safe District in each state with respect to crime all over the country.
* To find the Most Unsafe District in each state with respect to crime all over the country.
* The State which has reduced maximum number of Crimes.
* The State which has had an increase in the Crimes.
* Most common crime for each state.

## **Technology used:**

* Oracle 11g / SQL server (Database)
* Informatica power center / SSIS / DataStage / Abinitio(ETL Tool)
* Tableau Desktop / SSRS / Power BI / Qlikview (Reporting Tool)

# **Implementation in BI LOT**

## **3.1 SUMMARY OF THE FUNCTIONALITY TO BE BUILT:**

The participants need to develop the **Crime Analysis**by building the functionality incrementally in each of the course modules of BI LOT using one of the above ETL and reporting tool.

## **3.2 Guidelines on the functionality to be built:**

**Project flow:**

Heterogeneous Sources i.e. operational data (Flat files)

EXTRACT

TRANSFORM

LOAD

Data warehouse (Oracle)

Reports Based On Tool

R

E

P

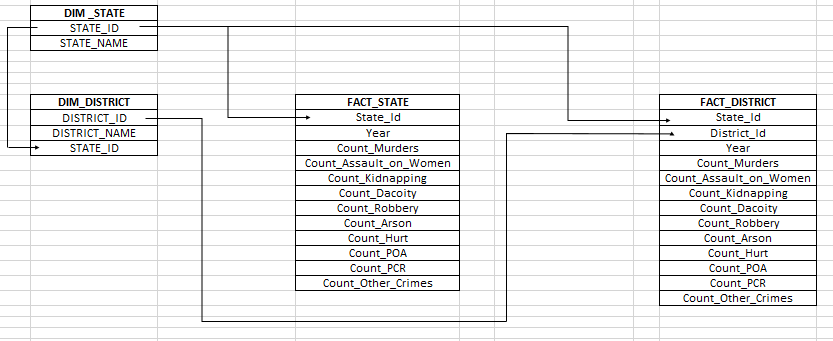
O

R

T

**Schema Design:**

* The project follows the Fact Constellation schema approach.
* A fact constellation has multiple fact tables. It is also known as galaxy schema.
* The Dimension tables are State, District. Fact tables are Fact\_State, Fact\_District.



**NOTE:** The datatype/length for the dimension/fact table attributes can be changed as required.Additional fields can be added,if required.

Dimension tables and their columns:

1. Dim\_State

* State\_Id
* State\_Name

1. Dim\_District

* District\_Id
* District\_Name
* State\_Id

Fact Table and their columns:

1. Fact\_Crime\_District

* State\_Id
* District\_Id
* Year
* Count\_Murders
* Count\_Assault\_on\_Women
* Count\_Kidnapping
* Count\_Dacoity
* Count\_Robbery
* Count\_Arson
* Count\_Hurt
* Count\_POA
* Count\_PCR
* Count\_Other\_Crimes

1. Fact\_Crime\_State

* State\_Id
* Year
* Count\_Murders
* Count\_Assault\_on\_Women
* Count\_Kidnapping
* Count\_Dacoity
* Count\_Robbery
* Count\_Arson
* Count\_Hurt
* Count\_POA
* Count\_PCR
* Count\_Other\_Crimes