High Level Design (HLD) Crime Data Analysis

Revision Number: 1.0

Last date of revision: 05/10/2022

Nitesh Kumar

# Document Version Control

|  |  |  |  |
| --- | --- | --- | --- |
| **Date Issued** | **Version** | **Description** | **Author** |
| **10th May 2022** | 1.0 | First Version of Complete HLD | Nitesh Kumar |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**Contents**

[Document Version Control 2](#_bookmark0)

[Abstract 3](#_bookmark1)

1. [Introduction 5](#_bookmark2)
   1. [Why this High-Level Design Document? 5](#_bookmark3)
   2. [Scope 5](#_bookmark4)
2. [General Description 6](#_bookmark5)
   1. [Product Perspective & Problem Statement 6](#_bookmark6)

Crime analysis is a law enforcement duty that involves a systematic investigation of

Patterns and trends in crime and disorder. In this Project ,Crime Pattern Information

can help to reduce it and apprehending criminals ………………………………………..6

2.2 [Tools used 6](#_bookmark8)

1. [Design Details 7](#_bookmark9)
   1. [Functional Architecture 7](#_bookmark10)
   2. [Optimization 8](#_bookmark11)
2. [KPIs 9](#_bookmark12)
   1. [KPIs (Key Performance Indicators) 9](#_bookmark13)
3. [Deployment 9](#_bookmark14)

# 

# Abstract

Crime analysis is a systematic approach for identifying the crime. This Data can predict

Region which have high probability for crime occurrences and visualize crime prone area. Using the concept of data mining we can extract previously unknown, useful information from an unstructured data. The extraction of new information is predicted using the existing datasets. Crimes are treacherous and common social problem faced worldwide. Crimes affect the quality of life, economic growth and reputation of nation. With the aim of securing the society from crimes, there is a need for new approaches for improving the crime analytics for protecting their communities. We propose a system which can analysis, detect, and predict various crime probability in given region. This paper explains various types of criminal analysis and crime prediction using several data mining techniques.

# Introduction

## Why this High-Level Design Document?

The purpose of this High-Level Design (HLD) Document is to add the necessary detail to the current project description to represent a suitable model for coding. This document is also intended to help detect contradictions prior to coding, and can be used as a reference manual for how the modules interact at a high level.

The HLD will:

* + - Present all of the design aspects and define them in detail
    - Describe the user interface being implemented
    - Describe the hardware and software interfaces
    - Describe the performance requirements
    - Include design features and the architecture of the project
    - List and describe the non-functional attributes like:
      * Security
      * Reliability
      * Maintainability
      * Portability
      * Reusability
      * Application compatibility
      * Resource utilization
      * Serviceability

## Scope

The HLD documentation presents the structure of the system, such as the database architecture, application architecture (layers), application flow (Navigation), and technology architecture. The HLD uses non-technical to mildly-technical terms which should be understandable to the administrators of the system.

# General Description

## Product Perspective & Problem Statement

Crime analysis is a law enforcement duty that involves a systematic investigation of

Patterns and trends in crime and disorder. In this Project ,Crime Pattern Information

can help to reduce it and apprehending criminals

The objective of the project is to perform data visualization techniques to understand the insight of the data. This project aims apply various Business Intelligence tools such as Tableau or Power BI to get a visual understanding of the data.

## Tools used

Business Intelligence tools and libraries works such as Numpy, Pandas, Excel, R, Tableau, Power BI are used to build the whole framework.





# Design Details

# Process Architecture:-

**Data Data Exploratory**

**Gathering Exploration Data Analysis**

**Data Data Data**

**Cleaning Preprocessing Visualization**

**How BI Works**

**B**

**Presentation**

**Insight**

**Creation**

**Information Integration**

**Organisational Memory**

Online Analytical Processing (OLAP) Tool

-Visualization Tool

-Digital Dashboard

-Score Card

-Text Mining Tool

-Web Mining Tool

-Environmental Scanning

-RFID

-Business Analytical Tools

-Data Mining

-Real Time Decision

-Data Warehouse

-Enterprise resource planning(ERP)

Knowledge Repository

-Content Management System (CMS)

## Optimization

### Your data strategy drives performance

* + - Minimize the number of fields
    - Minimize the number of records
    - Optimize extracts to speed up future queries by materializing calculations, removing columns and the use of accelerated views

### Reduce the marks (data points) in your view

* + - Practice guided analytics. There’s no need to fit everything you plan to show in a single view. Compile related views and connect them with action filters to travel from overview to highly-granular views at the speed of thought.
    - Remove unneeded dimensions from the detail shelf.
    - Explore. Try displaying your data in different types of views.

### Limit your filters by number and type

* + - Reduce the number of filters in use. Excessive filters on a view will create a more complex query, which takes longer to return results. Double-check your filters and remove any that aren’t necessary.
    - Use an include filter. Exclude filters load the entire domain of a dimension, while include filters do not. An include filter runs much faster than an exclude filter, especially for dimensions with many members.
    - [Use a continuous date filter](http://onlinehelp.tableau.com/current/pro/online/mac/en-us/help.htm#filtering_add_dragfields_dates.html). Continuous date filters (relative and range-of-date filters) can take advantage of the indexing properties in your database and are faster than discrete date filters.
    - [Use Boolean or numeric filters](http://www.tableau.com/learn/tutorials/on-demand/logical-calculations). Computers process integers and Booleans (t/f) much faster than strings.
    - Use [parameters](http://onlinehelp.tableau.com/current/pro/online/en-us/help.htm#parameters.html) and [action filters](http://onlinehelp.tableau.com/current/pro/online/en-us/help.htm#actions.html). These reduce the query load (and work across data sources).

### Optimize and materialize your calculations

* + - Perform calculations in the database
    - Reduce the number of nested calculations.
    - Reduce the granularity of LOD or table calculations in the view. The more granular the calculation, the longer it takes.
      * LODs - Look at the number of unique dimension members in the calculation.
      * Table Calculations - the more marks in the view, the longer it will take to calculate.
    - [Where possible, use MIN or MAX instead of AVG](http://onlinehelp.tableau.com/current/pro/online/windows/en-us/help.htm#calculations_aggregation.html). AVG requires more processing than MIN or MAX. Often rows will be duplicated and display the same result with MIN, MAX, or AVG.
    - [Make groups with calculations](http://kb.tableau.com/articles/knowledgebase/creating-groups-using-calculated-fields). Like include filters, calculated groups load only named members of the domain, whereas Tableau’s group function loads the entire domain.
    - [Use Booleans or numeric calculations instead of string calculations](http://onlinehelp.tableau.com/current/pro/online/mac/en-us/help.htm#functions_functions_string.html). Computers can process integers and Booleans (t/f) much faster than strings. Boolean>Int>Float>Date>DateTime>String

# KPIs

Dashboards will be implemented to display and indicate certain KPIs and relevant indicators for the disease.



As and when, the system starts to capture the historical/periodic data for a user, the dashboards will be included to display charts over time with progress on various indicators or factors

## KPIs (Key Performance Indicators)

Key indicators displaying a summary of the Crime and its relationship with different metrics

1. Crime in different Zones of Baton Rouge
2. Count of crimes with respect to Types of Crime
3. Count of Crime with respect to years
4. Count of offence

# Deployment

Prioritizing data and analytics couldn’t come at a better time. Your company, no matter what size, is already collecting data and most likely analyzing just a portion of it to solve business problems, gain competitive advantages, and drive enterprise transformation. With the explosive growth of enterprise data, database technologies, and the high demand for analytical skills, today’s most effective IT organizations have shifted their focus to enabling self-service by deploying and operating PowerBI at scale, as well as organizing, orchestrating, and unifying disparate sources of data for business users and experts alike to author and consume content.

