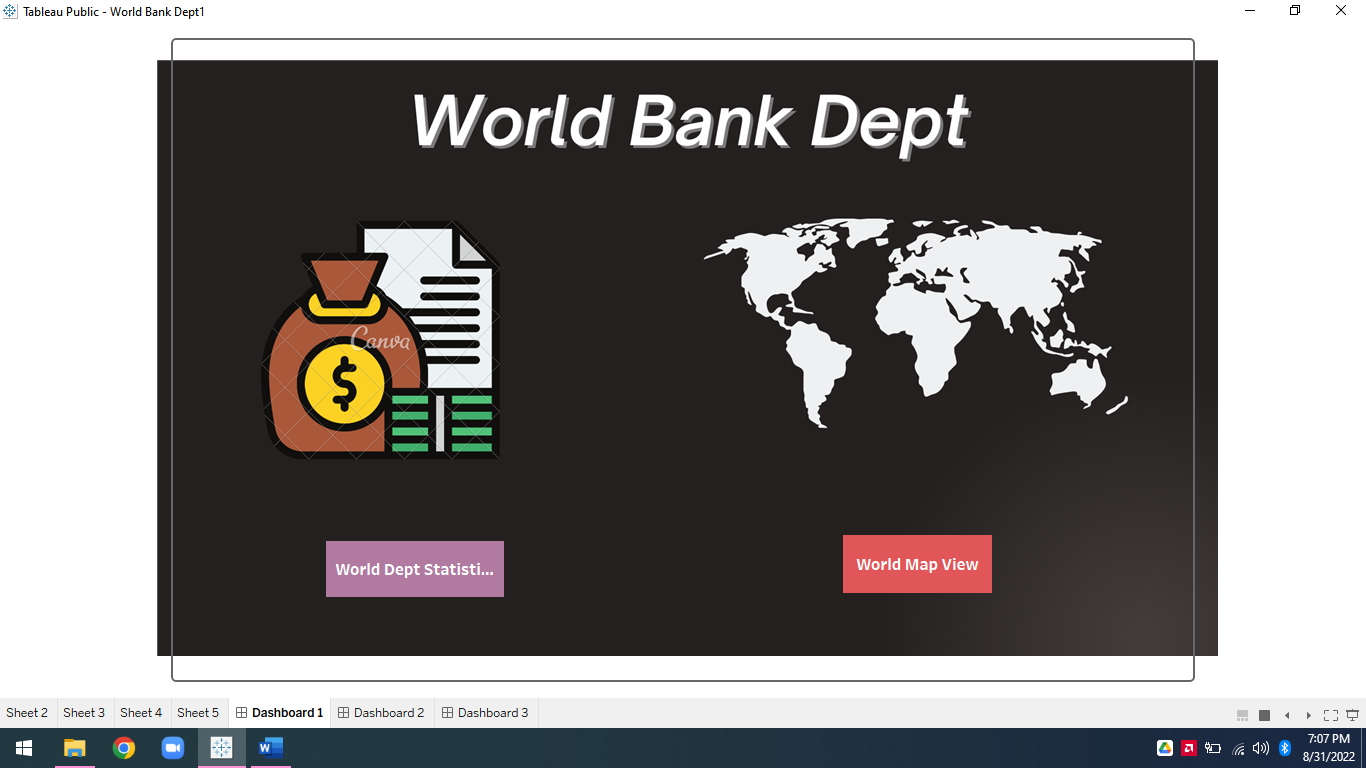
Low Level Design (LLD) Analyze International Debt Statistics



## Revision Number - 1.0

**Last Date of Revision –31/08/2022** Syed Ameer John SK

**Document Control**

|  |  |  |  |
| --- | --- | --- | --- |
| **Date** | **Version** | **Description** | **Author** |
| 31/08/2022 | 1.0 | Introduction, Problem Statement | Syed Ameer John SK |
| 31/08/2022 | 1.0 | Dataset Information, Architecture Description | Syed Ameer John SK |
| 31/08/2022 | 1.0 | Final Revision | Syed Ameer John SK |

## Why this Low-Level Design Document?

The purpose of this document is to present a detailed description of the heart disease prediction analysis technique. It will explain the necessary steps which have to be followed before any analysis can begin. The document would also describe the algorithms and techniques used to predict the presence and absence of the heart disease and present a comparative result for the same. LLD describes the class diagrams with the methods and relations between classes and programs specs. It describes the modules so that the programmer can directly code the program from the document. This document is intended for both the stakeholders and the developers of the system and will be proposed to the higher management for its approval.

The LLD will be focusing on the below objectives:

* Problem Understanding.
* Data Acquisition.
* Data Pre-Processing and Exploratory Analysis
* Development of models
* Auditing accuracy and retrain if require
* Finalizing the model
* Dashboard report for important activities

**Scope**

The LLD documentation presents the detailed structure of the heart disease prediction system for each of its individual components. The goal of LLD is to give the internal logical design of the actual program code. Low-level design is created based on the high-level design. The LLD documentation contains the complete description of the model used along with the comparisons of the proposed model/library compared with a baseline(existing) model against a set of metrics.

**Project Introduction**

It's not that we humans only take debts to manage our necessities. A country may also

take debt to manage its economy. For example, infrastructure spending is one costly

ingredient required for a country's citizens to lead comfortable lives. The World Bank is

the organization that provides debt to countries.

**Constraints**

Our analysis is done based on International Debt Statistics Dataset. There are many features because, every year dept is separated by one column

1970-2028 Each column for each year

**Risks**

Document specific risks that have been identified or that should be considered.

**Out of Scope**

Delineate specific activities, capabilities, and items that are out of scope for the project.

1.Technical Specification

Dataset

The Dataset is taken from iNeuron’s provided Project Description Document.

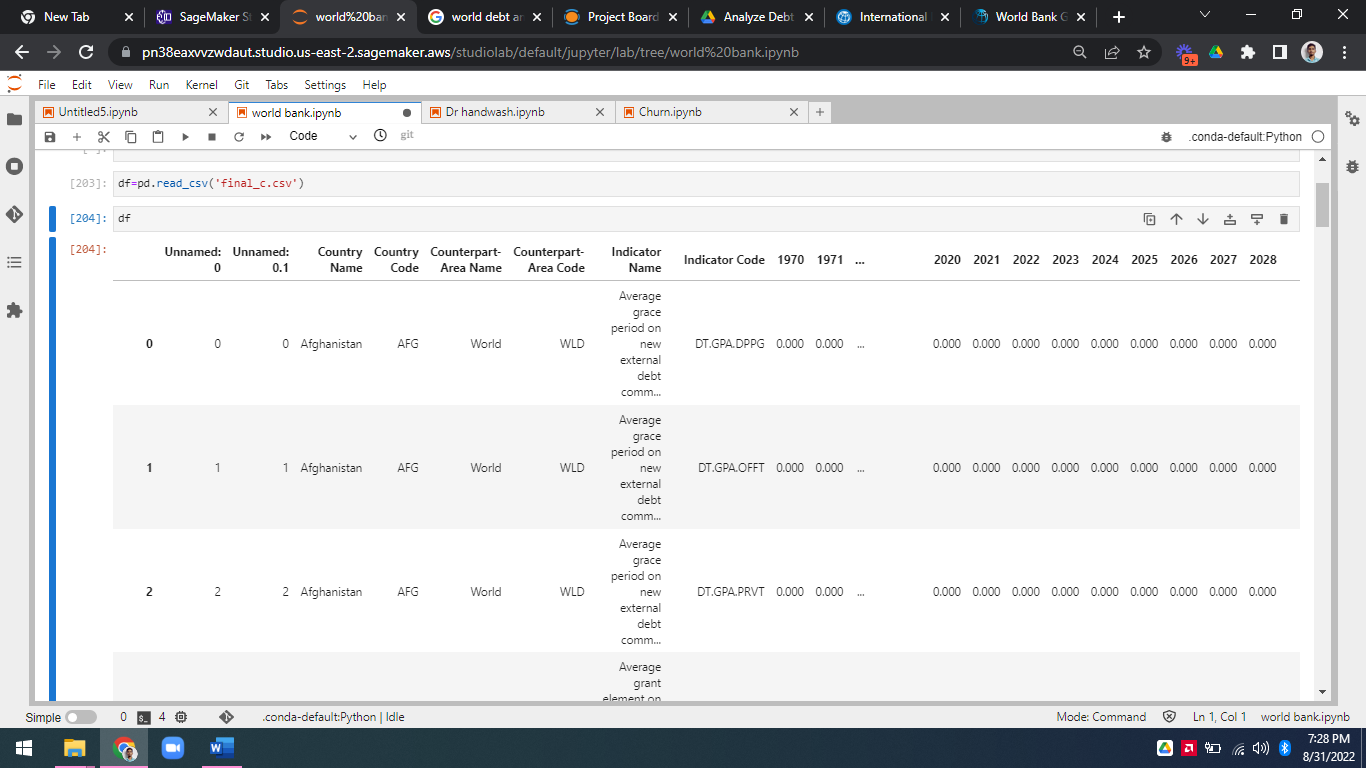


Figure 1: Analyze International Debt Statistics dataset

The dataset consists of 73401 individual data. There are 14 columns in the dataset, which are described below.

* 1. Country Name - Display the Country Name.
  2. Country Code- Display the Country code based on Country Name
  3. **Counterpart-Area Name** - Displays the General as World
  4. **Counterpart-Area Code** - Displays the Code for World

**WLD**- World

* 1. **Indicator Name** - Displays the Reason for Dept
  2. **Indicator Code** – Displays The unique code for Different Indicator
  3. 1970-2028 - Display the dept in years
  4. Dept – Displays the Total dept from 1970 to 2028

# Problem Statement

It's not that we humans only take debts to manage our necessities. A country may also

take debt to manage its economy. For example, infrastructure spending is one costly

ingredient required for a country's citizens to lead comfortable lives. The World Bank is

the organization that provides debt to countries.

In this project, you are going to analyze international debt data collected by The World

Bank. The dataset contains information about the amount of debt (in USD) owed by

developing countries across several categories. You are going to find the answers to

questions like:

 What is the total amount of debt that is owed by the countries listed in the

dataset?

 Which country owns the maximum amount of debt and what does that amount

look like?

 What is the average amount of debt owed by countries across different debt

Indicators?

# Architecture



Real World

Exploratory Data Analysis (EDA)

Modelling

Deployment

Data Cleaning

Data Pre- Processing

Raw Data Collection

Reporting

* 1. **Architecture Description**
     1. **Raw Data Collection-** The dataset is taken from iNeuron’s which is provided in the introduction document of project.

https://drive.google.com/drive/folders/165Pjmfb9W9PGy0rZjHEA22LW0Lt3Y-Q8

### Data Pre-Processing

Before building any model, it is crucial to perform data pre-processing to feed the correct data to the model to learn and predict. Model performance depends on the quality of data feed to the model to train.

This Process includes-

* + - 1. Handling Null/Missing Values
      2. Handling Skewed Data
      3. Outliers Detection and Removal

### Data Cleaning

Data cleaning is the process of fixing or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete data within a dataset.

* + - 1. Remove duplicate or irrelevant observations
      2. Filter unwanted outliers
      3. Renaming required attributes

### Exploratory Data Analysis (EDA)

Exploratory Data Analysis refers to the critical process of performing initial investigations on data to discover patterns, spot anomalies, test hypothesis

and to check assumptions with the help of summary statistics and graphical representations.

### Reporting

Reporting is a most important and underrated skill of a data analytics field. Because being a Data Analyst you should be good in easy and self explatory report because your model will be used by many stakeholders who are not from technical background.

* + - 1. High Level Design Document (HLD)
      2. Low Level Design Document (LLD)
      3. Architecture
      4. Wireframe
      5. Detailed Project Report
      6. Power Point Presentation

### Modelling

Data Modelling is the process of analyzing the data objects and their relationship to the other objects. It is used to analyze the data requirements that are required for the business processes. The data models are created for the data to be stored in a database. The Data Model's main focus is on what data is needed and how we have to organize data rather than what operations we have to perform.

### Deployment

We created a Tableau dashboard.

