Data Warehouse Automation: Implementation Document

Overview

This document details the steps and implementation required to automate the data collection, transformation, and reporting of macroeconomic indicators. The solution involves extracting data from CSV files, transforming it as per the requirements, and loading it into an Oracle database. The implementation uses C# for the ETL (Extract, Transform, Load) process and PL/SQL for data handling within the Oracle database.

Data Pipeline Design:

Data Sources:
- CorruptionPerceptionIndexDataSet.csv
- WDICSV.csv

Database Schema:
Database: macroeconomic_analysis
Tables: corruption_perception_index and world_development_indicators

ETL Process

- **Extract Data**: Fetch data from CSV files using C#.
- **Transform Data**:transform data extracted and apply the filter. Clean and format data to match the schema
- **Load Staging Tables**: Insert data into Oracle tables in C#.

Reporting: generate the required reports, which can be filtered based on user inputs.
- Connect Power BI to the Oracle database.
- Create a report with filters for countries and indicators.
- Use visualizations to compare the selected indicators for the last two years.