



LOW LEVEL DESIGN DOCUMENT

Foreign Direct Investment
Analysis

BY

VIKASH KUMAR MAHAPATRA

Revision Number & Date :- 1.2 30/06/2022

Document Version Control

Date Issued	Version	Description	Author
17 MAY 2022	1.0	Created Document And Add Document Version Control, Content Page, Introduction Page	Vikash
07 JUNE 2022	1.1	Created Architecture Page ,Architecture Description Page & Add page number to contents section	Vikash
30 JUNE 20222	1.2	Final Revision	Vikash

CONTENTS

INTRODUCTION 3-4

- + WHY THIS LOW LEVEL DESIGN DOCUMENT
- + SCOPE
- + PROJECT INTRODUCTION
- + PROBLEM STATEMENT

ARCHITECTURE 5-6

ARCHITECTURE DESCRIPTION 7-10

INTRODUCTION

WHY THIS LOW LEVEL DESIGN DOCUMENT

The purpose of this Low-Level Design (LLD) document is to give the internal logic design of the actual program code for Foreign Direct Investment Analysis .LLD document describes the class diagrams with the methods and relations between classes and program specs. It describes the modules so that the programmer can directly code the program from the document.

SCOPE

The LLD document is a component-level design process that follows a step-by-step refinement process. The process can be used for designing data structures, required software architecture, source code and performance algorithms. Overall, the data organization may be defined during requirement analysis and then refined during data design work.

PROJECT INTRODUCTION

Foreign Direct Investment (FDI) is a type of investment into enterprises located in another country i.e. by either buying a company in the target country or expanding the operation of the existing business in that country. It offers several benefits some them are employment opportunities, introduction to new skill as well as technology, new audiences or can be say as extension to new markets etc. , which reflects on growth of economic income as well as growth of people in any country. Many countries present lucrative schemes to attract foreign investors or can be said as FDI. Different countries have investment ratio i.e. whether it will be complete invested by foreign investors or part investment from foreign investors and part investment from either government of that country or in-house investors. Almost every country have certain set of rules and regulations that is related to how much they invest in any particular sectors as well as no investment on particular sectors as well.

PROBLEM STATEMENT

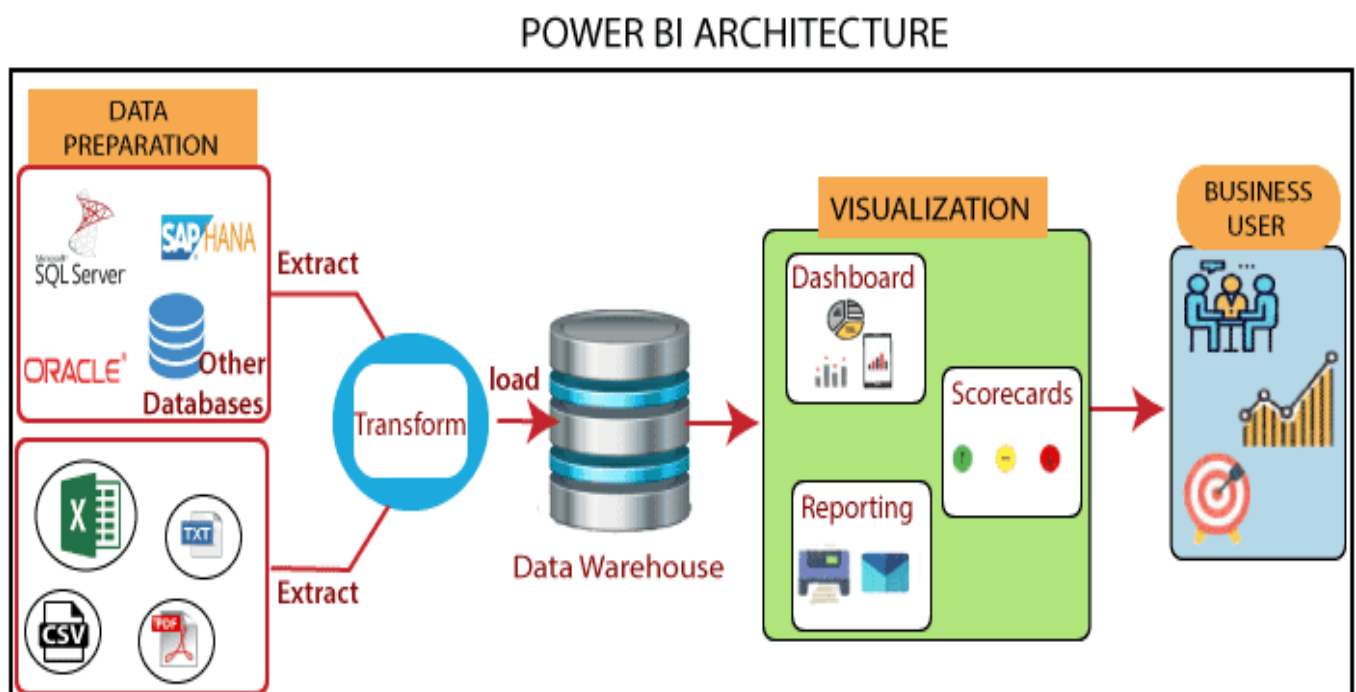
Investment is a game of understanding historic data of investment of investment objects under different events but it is still a game of chances to minimize the risk we apply analytics to find equilibrium investment.

To understand FDI in India for the last 17 years from 2000-2017. The dataset contains sector and financial year wise data of FDI in India sector-wise investment analysis, year-wise investment analysis.

Do ETL (Extract-Transform-Load) the dataset and find for me some information from this large data. This is form of data mining. What all information can be achieved by mining this data, would be brainstormed by interns. Find key metrics and factors and show the meaningful relationship between attributes. Do your own research and come up with findings

ARCHITECTURE

Microsoft Power BI architecture consists of four major steps that explain the whole process from data sourcing to creation of reports and dashboard. Various technologies and process work together to get required results with extreme precision. Let us see those steps in pictorial format.



Sourcing data: Power BI extracts data from various servers, Excel sheets, CSV files and databases. The extracted information can be directly imported to Power BI or a live service link is established to receive it. If you directly import the data in Power BI, it will only be compressed up to 1GB post that you can only run live queries on your chunky datasets.

Transforming the data: Before visualizing the data, cleaning and preprocessing it should be done. This means useless or missing values from rows or columns. Following that certain rules will be applied to transform and load datasets into the warehouse.

Report and publish: After cleaning and transforming the data, reports will be created based on requirements. A report is a visualization of the data with different filters and constraints presented in the form of graphs, pie-chart, and other figures.

Creating dashboards: Power BI Dashboards are created by pinning individual elements or pages of live reports. Dashboard should be created after you have published your reports to the BI service. When the reports get saved, the visual maintains the filter setting chosen so that the user can apply filters and slicers.

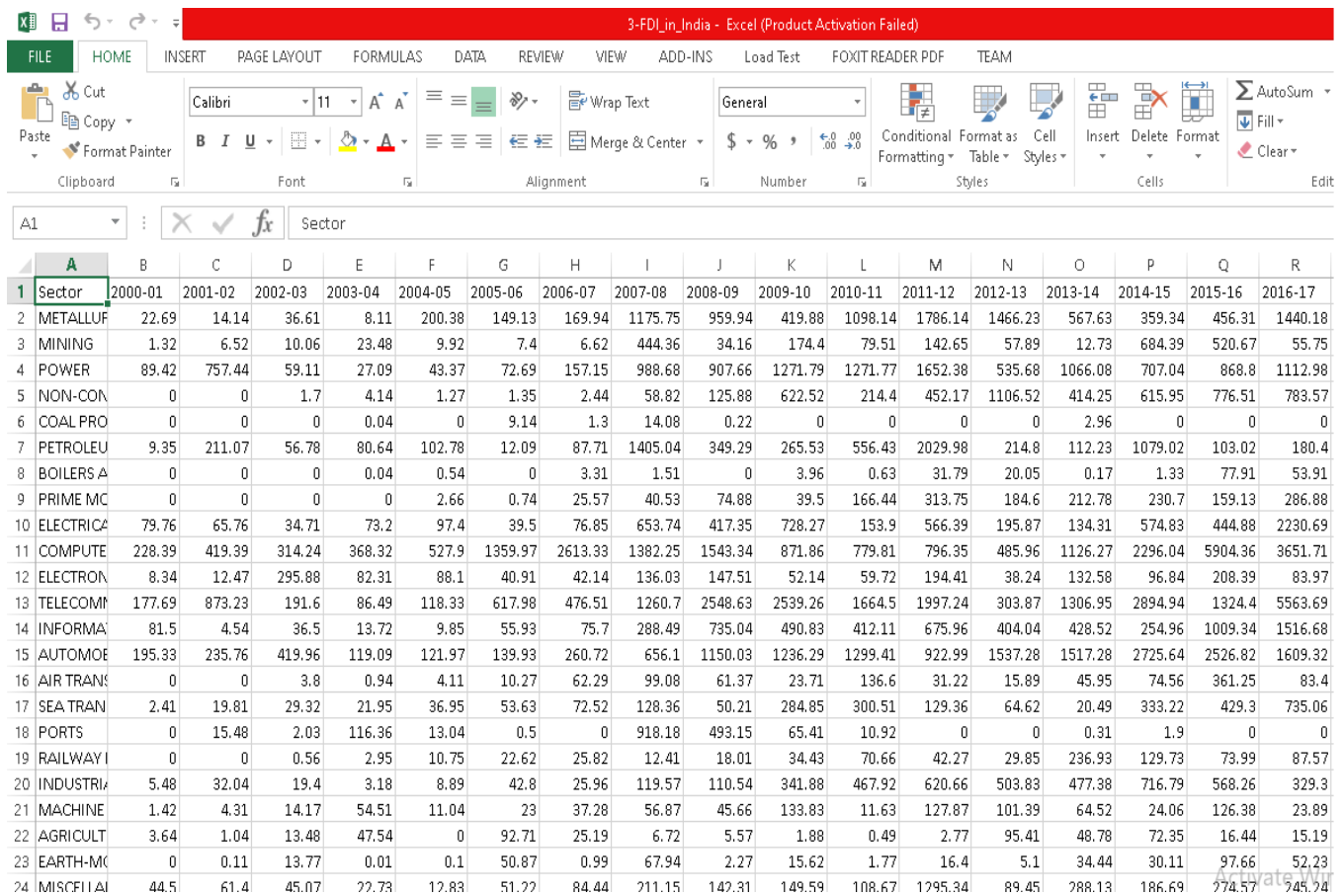
ARCHITECTURE DESCRIPTION

Data Sourcing:

The dataset is in csv (comma separated value) format and Microsoft excel is used to load the data. Dataset is taken from project description document and here goes the drive link <https://drive.google.com/drive/folders/1M5z7z1NmWar7y1eFs67orfjqHL0iSViL?usp=sharing>

Data Description:

The file name FDI_in_India.csv contains different sector names and period interval i.e. from 2000-2017 and contain amount received corresponding years against particular sectors. Sample data below.



Sector	2000-01	2001-02	2002-03	2003-04	2004-05	2005-06	2006-07	2007-08	2008-09	2009-10	2010-11	2011-12	2012-13	2013-14	2014-15	2015-16	2016-17
METALLURGY	22.69	14.14	36.61	8.11	200.38	149.13	169.94	1175.75	959.94	419.88	1098.14	1786.14	1466.23	567.63	359.34	456.31	1440.18
MINING	1.32	6.52	10.06	23.48	9.92	7.4	6.62	444.36	34.16	174.4	79.51	142.65	57.89	12.73	684.39	520.67	55.75
POWER	89.42	757.44	59.11	27.09	43.37	72.69	157.15	988.68	907.66	1271.79	1271.77	1652.38	535.68	1066.08	707.04	868.8	1112.98
NON-CORP	0	0	1.7	4.14	1.27	1.35	2.44	58.82	125.88	622.52	214.4	452.17	1106.52	414.25	615.95	776.51	783.57
COAL PRO	0	0	0	0.04	0	9.14	1.3	14.08	0.22	0	0	0	0	2.96	0	0	0
PETROLEU	9.35	211.07	56.78	80.64	102.78	12.09	87.71	1405.04	349.29	265.53	556.43	2029.98	214.8	112.23	1079.02	103.02	180.4
BOILERS A	0	0	0	0.04	0.54	0	3.31	1.51	0	3.96	0.63	31.79	20.05	0.17	1.33	77.91	53.91
PRIME MC	0	0	0	0	2.66	0.74	25.57	40.53	74.88	39.5	166.44	313.75	184.6	212.78	230.7	159.13	286.88
ELECTRICA	79.76	65.76	34.71	73.2	97.4	39.5	76.85	653.74	417.35	728.27	153.9	566.39	195.87	134.31	574.83	444.88	2230.69
COMPUTE	228.39	419.39	314.24	368.32	527.9	1359.97	2613.33	1382.25	1543.34	871.86	779.81	796.35	485.96	1126.27	2296.04	5904.36	3651.71
ELECTRON	8.34	12.47	295.88	82.31	88.1	40.91	42.14	136.03	147.51	52.14	59.72	194.41	38.24	132.58	96.84	208.39	83.97
TELECOMM	177.69	873.23	191.6	86.49	118.33	617.98	476.51	1260.7	2548.63	2539.26	1664.5	1997.24	303.87	1306.95	2894.94	1324.4	5563.69
INFORMA	81.5	4.54	36.5	13.72	9.85	55.93	75.7	288.49	735.04	490.83	412.11	675.96	404.04	428.52	254.96	1009.34	1516.68
AUTOMOE	195.33	235.76	419.96	119.09	121.97	139.93	260.72	656.1	1150.03	1236.29	1299.41	922.99	1537.28	1517.28	2725.64	2526.82	1609.32
AIR TRAN	0	0	3.8	0.94	4.11	10.27	62.29	99.08	61.37	23.71	136.6	31.22	15.89	45.95	74.56	361.25	83.4
SEA TRAN	2.41	19.81	29.32	21.95	36.95	53.63	72.52	128.36	50.21	284.85	300.51	129.36	64.62	20.49	333.22	429.3	735.06
PORTS	0	15.48	2.03	116.36	13.04	0.5	0	918.18	493.15	65.41	10.92	0	0	0.31	1.9	0	0
RAILWAY	0	0	0.56	2.95	10.75	22.62	25.82	12.41	18.01	34.43	70.66	42.27	29.85	236.93	129.73	73.99	87.57
INDUSTRIA	5.48	32.04	19.4	3.18	8.89	42.8	25.96	119.57	110.54	341.88	467.92	620.66	503.83	477.38	716.79	568.26	329.3
MACHINE	1.42	4.31	14.17	54.51	11.04	23	37.28	56.87	45.66	133.83	11.63	127.87	101.39	64.52	24.06	126.38	23.89
AGRICULT	3.64	1.04	13.48	47.54	0	92.71	25.19	6.72	5.57	1.88	0.49	2.77	95.41	48.78	72.35	16.44	15.19
EARTH-MC	0	0.11	13.77	0.01	0.1	50.87	0.99	67.94	2.27	15.62	1.77	16.4	5.1	34.44	30.11	97.66	52.23
MISCELLA	44.5	61.4	45.07	22.73	12.83	51.22	84.44	711.15	142.31	149.59	108.67	1295.34	89.45	288.13	186.69	274.57	745.74

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 fed to the model to train. This process includes following steps:- Handling Missing/Null values , Handling Skewed Data and 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. a)Remove duplicate or irrelevant observations. b) Filter unwanted outliers. c) 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 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 the easy and self-explanatory report because your model will be used by many stakeholders who are not from a technical background. a)HLD(High Level Document)b)LLD(Low Level Document)c)Architecture Report. d) Wireframe report. e) Detailed Project Report.

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 performing the operations that are need.

Deployment: We create a Power BI Dashboard



