Low Level Design

Expenditure Data Analysis

Written By	Swati Khedekar
Document Version	**1.2**
First Revised Date	**19/4/2023**
Last Revised Date	**25/4/2023**
Last Revised Date	**02/05/2023**

Contents:

1. Introduction

1.1 What is Low-Level design document?

The goal of the LDD or Low-level design document (LLDD) is to give the internal logic design of the actual program code for the Expenditure Data Analysis dashboard. LDD 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.

1.2 Scope

Low-level design (LLD) 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 ultimately, performance algorithms. Overall, the data organization may be defined during requirement analysis and then refined during data design work.

4Low Level Design (LLD)

2. Architecture

Power BI Desktop Architecture

1. Get Power BI Desktop

With Power BI Desktop, you can build advanced queries, models, and reports that visualize data. You can also build data models, create reports, and share your work by publishing to the Power BI service. Power BI Desktop is a free download.

2. BI solution architecture in the Centre of Excellence

BI solution architecture can consist of:

- Data sources
- Data ingestion
- Big data / data preparation
- Data warehouse
- BI semantic models
- Reports

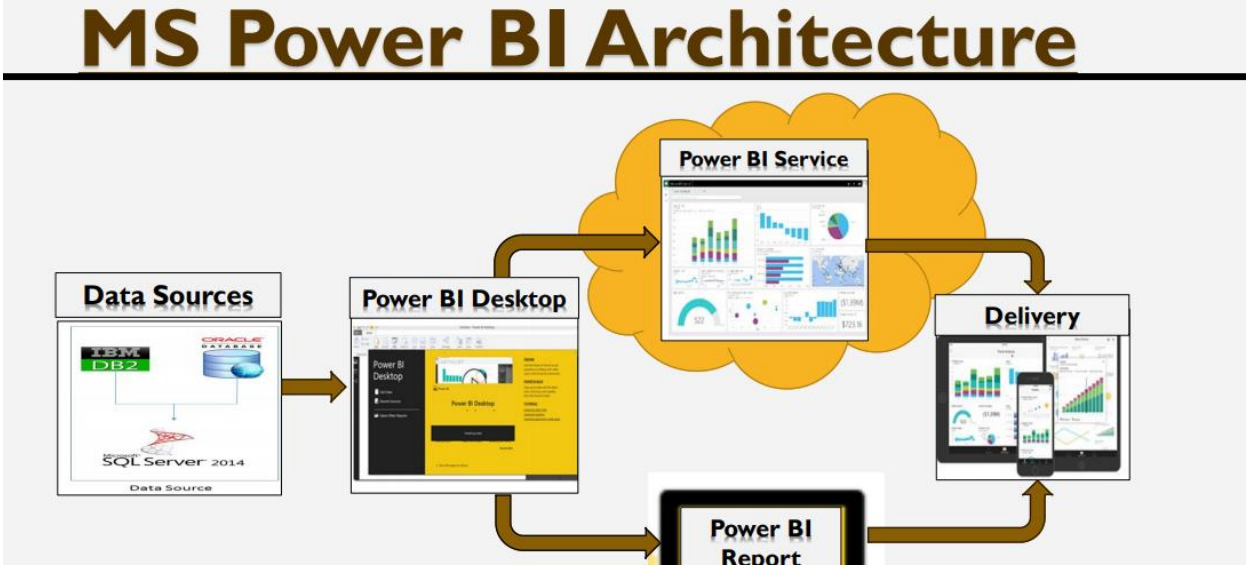


Fig: Power BI Architecture diagram

Microsoft Power BI Desktop is a companion desktop application to Power BI.

With Power BI Desktop, you can:

1. **Get data:**

The Power BI Desktop makes discovering data easy. You can import data from a wide variety of data sources. After you connect to a data source, you can shape the data to match your analysis and reporting needs.

2. **Create relationships and enrich your data model with new measures and data formats:**

When you import two or more tables, oftentimes you'll need to create relationships between those tables. The Power BI Desktop includes the Manage Relationships dialog and the Relationships view, where you can use Autodetect to let the Power BI Desktop find and create any relationships, or you can create them yourself. You can also very easily create your own measures and calculations or customize data formats and categories to enrich your data for additional insights.

3. **Create reports:**

The Power BI Desktop includes the Report View. Select the fields you want, add filters, choose from dozens of visualizations, format your reports with custom colours, gradients and several other options. The Report View gives you the same great report and visualizations tools just like when creating a report on PowerBI.com.

4. **Save your reports:**

With the Power BI Desktop, you can save your work as a Power BI Desktop file. Power BI Desktop files have a .pbix extension.

5. **Upload or Publish your reports:**

You can upload the reports you created and saved in the Desktop to your Power BI site. You can also publish them to Power BI right from Power BI Desktop.

3. Architecture Description

3.1. Data Description: The Dataset contains year wise distribution of all the states of India for the following parameters:

1) Aggregate Expenditure:

Aggregate expenditure is a measure of national income. Aggregate expenditure is defined as the current value of all the finished goods and services in the economy. The Aggregate expenditure is thus the sum of total of all the expenditures undertaken in the economy by the factors during a given time period

2) Capital Expenditure:

Capital expenditure or capital expense is the money an organization or corporate entity spends to buy, maintain, or improve its fixed assets, such as buildings, vehicles, equipment, or land.

3) Gross Fiscal Deficits:

The gross fiscal deficit (GFD) is the excess of total expenditure including loans net of recovery over revenue receipts (including external grants) and non-debt capital receipts. Generally fiscal deficit takes place either due to revenue deficit or a major hike in capital expenditure.

4) Nominal GDP Series:

Nominal GDP is an assessment of economic production in an economy that includes current prices in its calculation. In other words, it doesn't strip out inflation or the pace of rising prices, which can inflate the growth figure.

5) Own Tax Revenues:

The income generated by states for various activities include revenue receipts like taxes & grants and capital receipts like loans. States which are able to generate more revenue on their own are less dependent on the devolution & central grants.

6) Revenue Deficits:

A revenue deficit occurs when realized net income is less than the projected net income. This happens when the actual amount of revenue and/or the actual number of expenditures do not correspond with budgeted revenue and expenditures.

7) Revenue Expenditure:

Revenue expenditures are short-term expenses used in the current period or typically within one year. Revenue expenditures include the expenses required to meet the ongoing operational costs of running a business, and thus are essentially the same as operating expenses (OPEX).

8) Social Sector Expenditure:

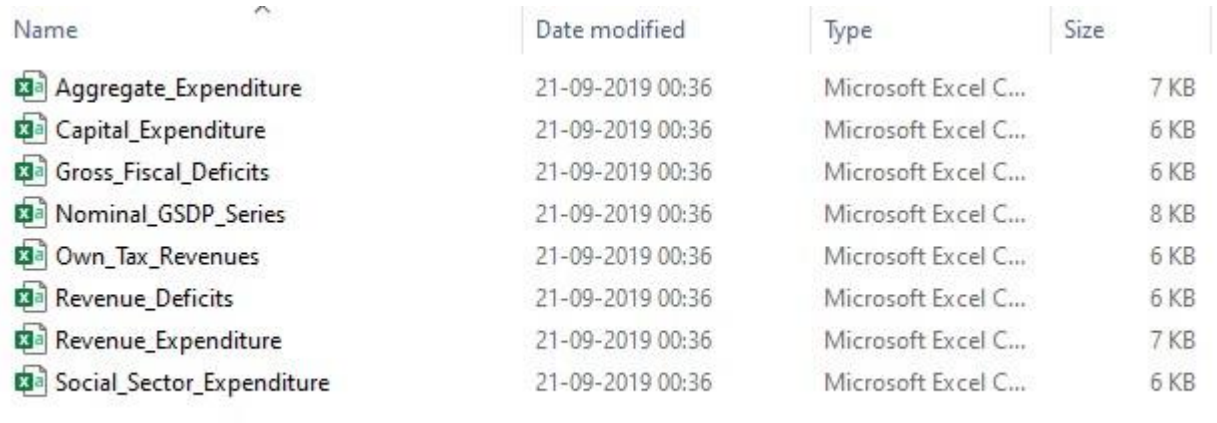
Social sector expenditure has been defined as the total of all expenditures incurred by the central and the state governments on promotional and protective measures.

3.2. Web Scrapping:

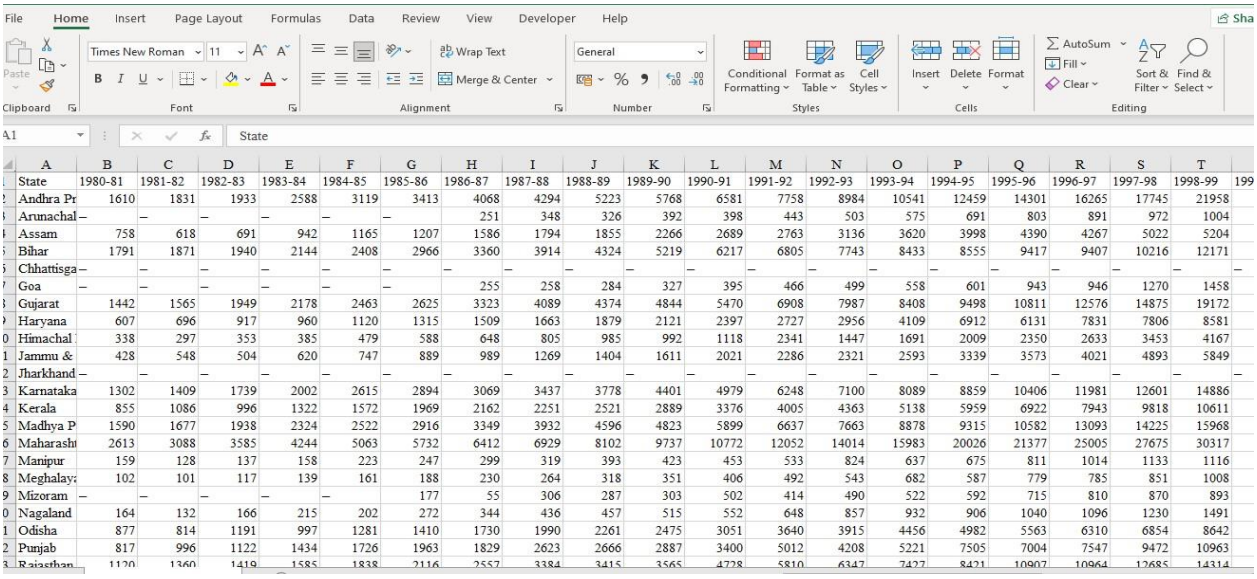
Web scraping is a technique to automatically extract content and data from websites using bots. It is also known as web data extraction or web harvesting. Web scrapping is made simple now days, many tools are used for web scrapping. Some of python libraries used for web scrapping are Beautiful Soup, Scrapy, Selenium, etc.

3.3. Data Preparation:

- In the Preparation Process, we will convert our original datasets with other necessary attributes format. And will merge it with the Scrapped dataset.
- All the 8 datasets are of same format as shown below: Original dataset.



Name	Date modified	Type	Size
Aggregate_Expenditure	21-09-2019 00:36	Microsoft Excel C...	7 KB
Capital_Expenditure	21-09-2019 00:36	Microsoft Excel C...	6 KB
Gross_Fiscal_Deficits	21-09-2019 00:36	Microsoft Excel C...	6 KB
Nominal_GSDP_Series	21-09-2019 00:36	Microsoft Excel C...	8 KB
Own_Tax_Revenues	21-09-2019 00:36	Microsoft Excel C...	6 KB
Revenue_Deficits	21-09-2019 00:36	Microsoft Excel C...	6 KB
Revenue_Expenditure	21-09-2019 00:36	Microsoft Excel C...	7 KB
Social_Sector_Expenditure	21-09-2019 00:36	Microsoft Excel C...	6 KB



State	1980-81	1981-82	1982-83	1983-84	1984-85	1985-86	1986-87	1987-88	1988-89	1989-90	1990-91	1991-92	1992-93	1993-94	1994-95	1995-96	1996-97	1997-98	1998-99
Andhra Pr	1610	1831	1933	2588	3119	3413	4068	4294	5223	5768	6581	7758	8984	10541	12459	14301	16265	17745	21958
Arunachal	–	–	–	–	–	–	251	348	326	392	398	443	503	575	691	803	891	972	1004
Assam	758	618	691	942	1165	1207	1586	1794	1855	2266	2689	2763	3136	3620	3998	4390	4267	5022	5204
Bihar	1791	1871	1940	2144	2408	2966	3360	3914	4324	5219	6217	6805	7743	8433	8555	9417	9407	10216	12171
Chhattisga	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Goa	–	–	–	–	–	–	255	258	284	327	395	466	499	558	601	943	946	1270	1458
Gujarat	1442	1565	1949	2178	2463	2625	3323	4089	4374	4844	5470	6908	7987	8408	9498	10811	12576	14875	19172
Haryana	607	696	917	960	1120	1315	1509	1663	1879	2121	2397	2727	2956	4109	6912	6131	7831	7806	8581
Himachal	338	297	353	385	479	588	648	805	985	992	1118	2341	1447	1691	2009	2350	2633	3453	4167
Jammu &	428	548	504	620	747	889	989	1269	1404	1611	2021	2286	2321	2593	3339	3573	4021	4893	5849
Jharkhand	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Karnataka	1302	1409	1739	2002	2615	2894	3069	3437	3778	4401	4979	6248	7100	8089	8859	10406	11981	12601	14886
Kerala	855	1086	996	1322	1572	1969	2162	2251	2521	2889	3376	4005	4363	5138	5959	6922	7943	9818	10611
Madhya P	1590	1677	1938	2324	2522	2916	3349	3932	4596	4823	5899	6637	7663	8878	9315	10582	13093	14225	15968
Maharasht	2613	3088	3585	4244	5063	5732	6412	6929	8102	9737	10772	12052	14014	15983	20026	21377	25005	27675	30317
Manipur	159	128	137	158	223	247	299	319	393	423	453	533	824	637	675	811	1014	1133	1116
Meghalaya	102	101	117	139	161	188	230	264	318	351	406	492	543	682	587	779	785	851	1008
Mizoram	–	–	–	–	–	177	55	306	287	303	502	414	490	522	592	715	810	870	893
Nagaland	164	132	166	215	202	272	344	436	457	515	552	648	857	932	906	1040	1096	1230	1491
Odisha	877	814	1191	997	1281	1410	1730	1990	2261	2475	3051	3640	3915	4456	4982	5563	6310	6854	8642
Punjab	817	996	1122	1434	1726	1963	1829	2623	2666	2887	3400	5012	4208	5221	7505	7004	7547	9472	10963
Rajasthan	1120	1360	1419	1585	1838	2116	2557	3384	3415	3565	4728	5810	6347	7427	8471	10907	10964	12685	14314

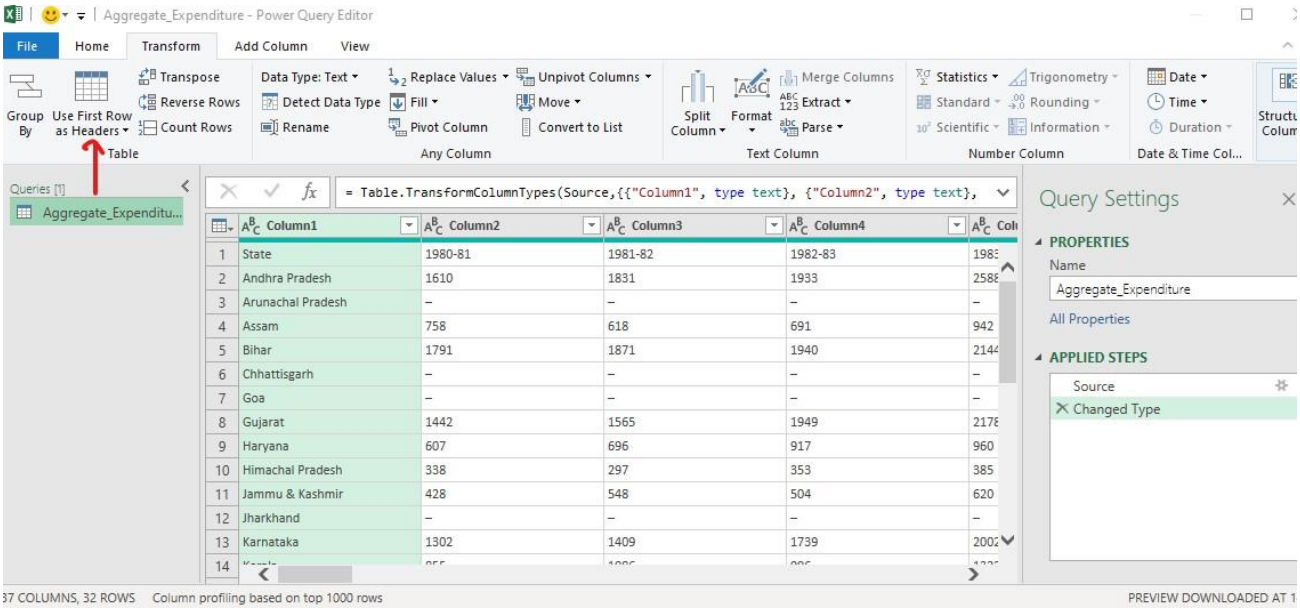
- As you all can notice that format of the data, we have is not good to analyse and visualize. So, we need to reconstruct the structure of the dataset.
- As this is a pivoted data set, we need to unpivot it. That's the only way by which we can make meaningful insights from it. This process is known as converting wide data to long data.

Low Level Design (LLD)

We will be using only MS Excel with power query for data restructuring and cleaning purpose.

- Also We can upload csv data from Home section, click on Data tab, click on get data ,then from text or csv and upload dataset into Power query editor.
- Power Query Editor window will get popped up.
- Then Click on transform data.

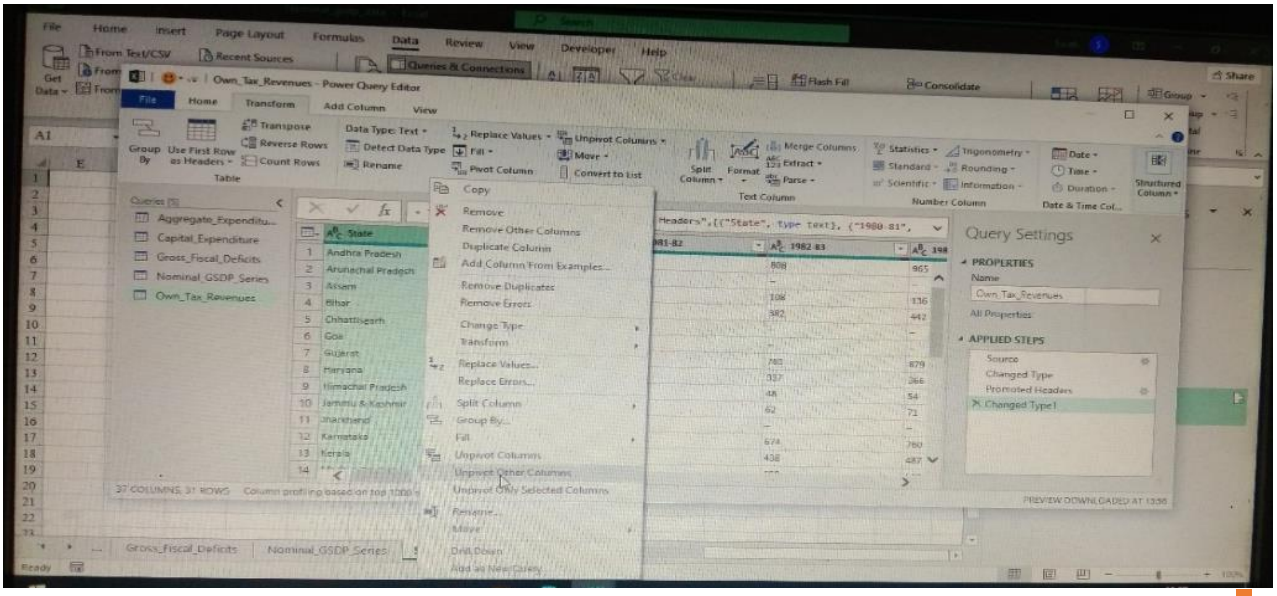
Step 1) Change Headers: In Power Query Editor in transform section Use first row as headers.



	Column1	Column2	Column3	Column4
1	State	1980-81	1981-82	1982-83
2	Andhra Pradesh	1610	1831	1933
3	Arunachal Pradesh	–	–	–
4	Assam	758	618	691
5	Bihar	1791	1871	1940
6	Chhattisgarh	–	–	–
7	Goa	–	–	–
8	Gujarat	1442	1565	1949
9	Haryana	607	696	917
10	Himachal Pradesh	338	297	353
11	Jammu & Kashmir	428	548	504
12	Jharkhand	–	–	–
13	Karnataka	1302	1409	1739

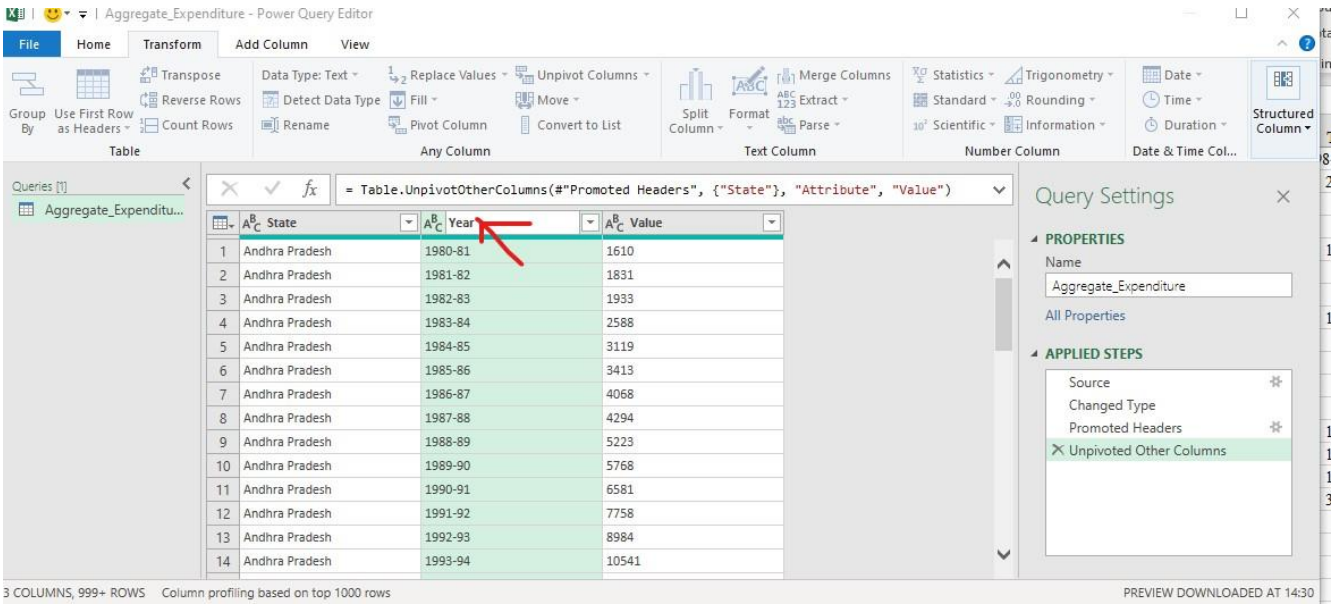
Step 2) Unpivot Dataset:

- In Transform section Select all year's columns that we have to unpivot and unpivot columns.
- Go to transform section and click unpivot columns. (Or)
- GO to transform section select first row and unpivot others columns.



Step 3) Change attributes Name:

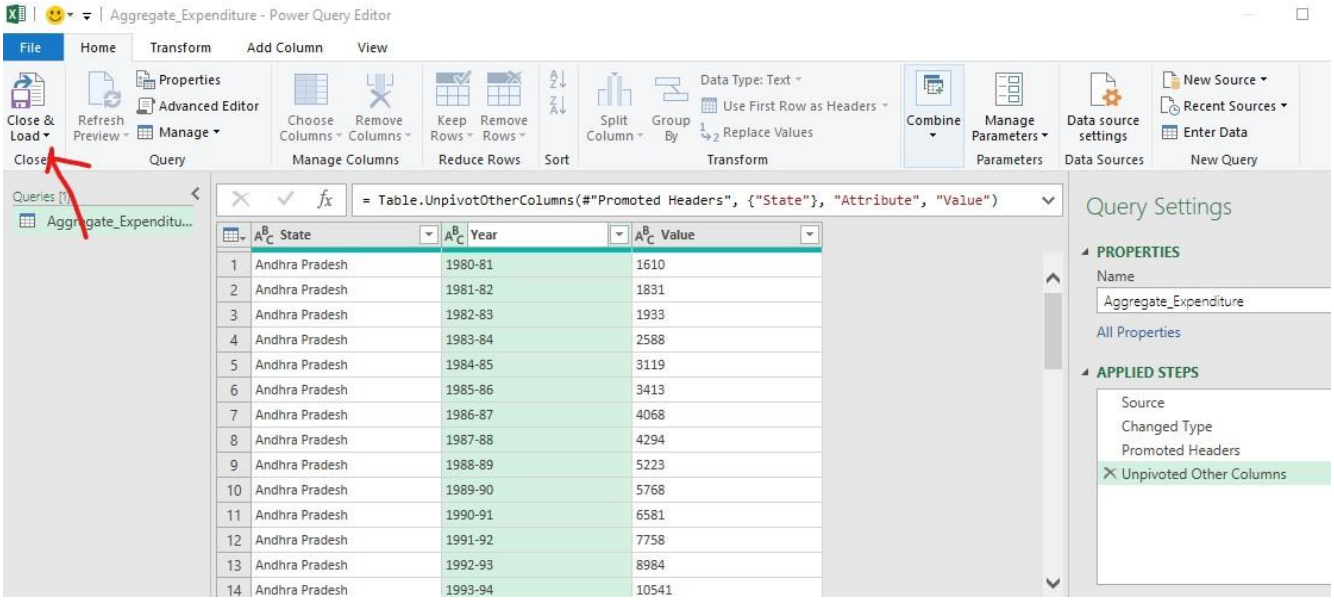
- Now an unpivoted dataset will be shown, change column name of attribute to “Year”.



	State	Year	Value
1	Andhra Pradesh	1980-81	1610
2	Andhra Pradesh	1981-82	1831
3	Andhra Pradesh	1982-83	1933
4	Andhra Pradesh	1983-84	2588
5	Andhra Pradesh	1984-85	3119
6	Andhra Pradesh	1985-86	3413
7	Andhra Pradesh	1986-87	4068
8	Andhra Pradesh	1987-88	4294
9	Andhra Pradesh	1988-89	5223
10	Andhra Pradesh	1989-90	5768
11	Andhra Pradesh	1990-91	6581
12	Andhra Pradesh	1991-92	7758
13	Andhra Pradesh	1992-93	8984
14	Andhra Pradesh	1993-94	10541

Step 3) Close and save:

- Now go to Home section and click close and load.

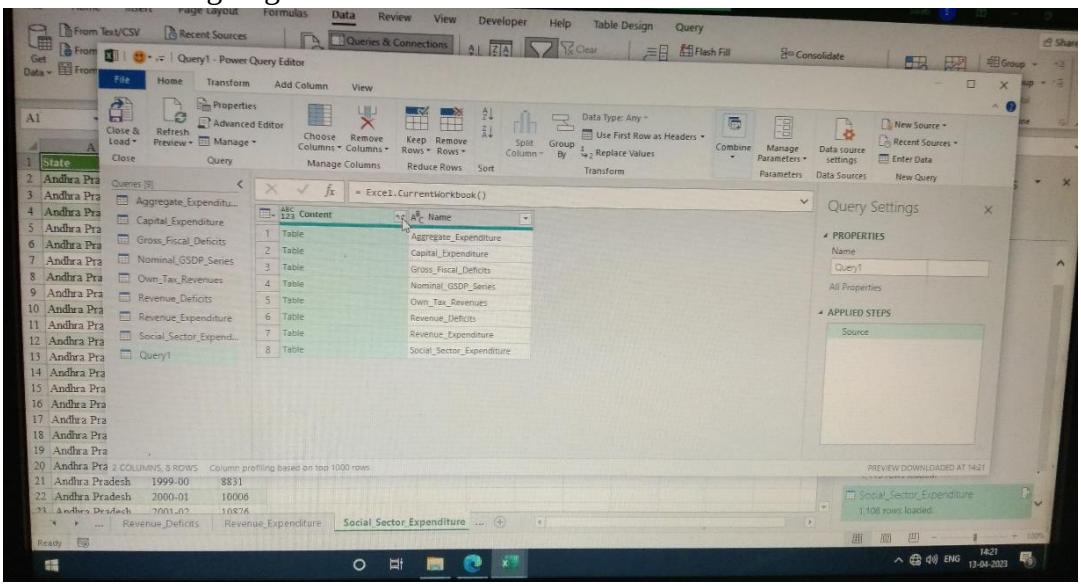


	State	Year	Value
1	Andhra Pradesh	1980-81	1610
2	Andhra Pradesh	1981-82	1831
3	Andhra Pradesh	1982-83	1933
4	Andhra Pradesh	1983-84	2588
5	Andhra Pradesh	1984-85	3119
6	Andhra Pradesh	1985-86	3413
7	Andhra Pradesh	1986-87	4068
8	Andhra Pradesh	1987-88	4294
9	Andhra Pradesh	1988-89	5223
10	Andhra Pradesh	1989-90	5768
11	Andhra Pradesh	1990-91	6581
12	Andhra Pradesh	1991-92	7758
13	Andhra Pradesh	1992-93	8984
14	Andhra Pradesh	1993-94	10541

- Now another table has been created with unpivoted columns.
- In such a way we preparing dataset and now we need to all these steps for all the others datasets we have.
- And save it into new folder as prepared data.

Step 4) Combine dataset:

- We can even combine all the datasets by having a column which specifies the datasets name.
- Open folder prepared data in Ms excels by query editor from Data section, click on get data→ from file→from folder→select folder we want to load→ then click on transform data.
- Then In query editor home section in Formula bar write **"=Excel.CurrentWorkbook()"** for converting large data into one sheet.



Step 5) Final dataset:

- Thus, we have restructured and combined all the datasets into one file save in csv as a name "**Final dataset**".
- In this way we can visualize using filters. And Rename of Expenditure column as "Exp Category", and rest columns are as it is like State, Year and Value.

A1 | Exp Category

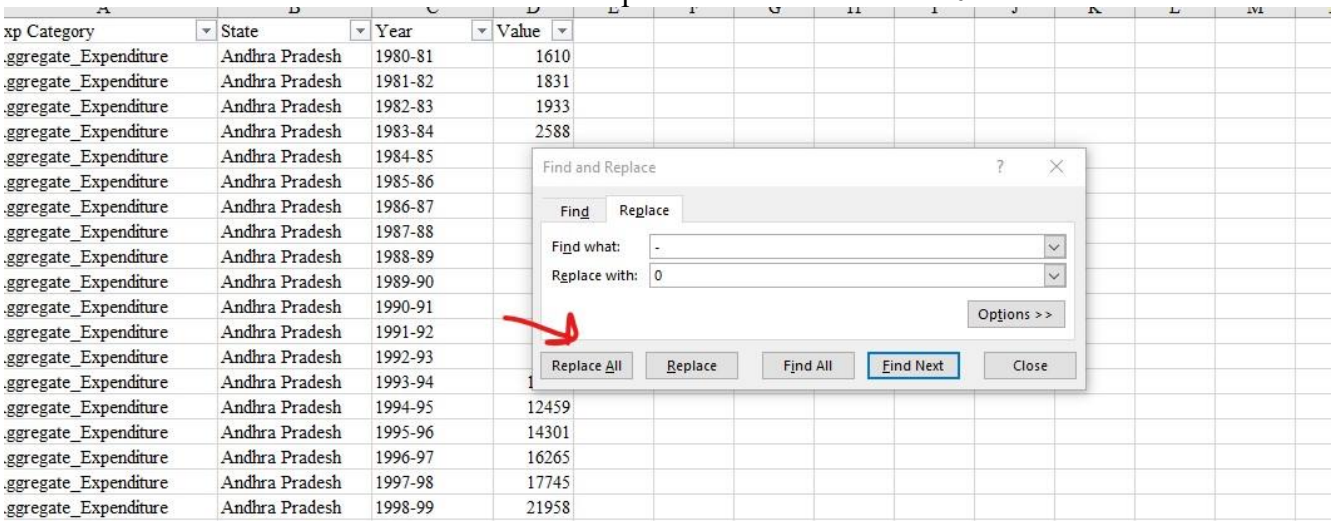
	A	B	C	D	E	F	G
1	Exp Category	State	Year	Value			
2	Aggregate_Expenditure	Andhra Pradesh	1980-81	1610			
3	Aggregate_Expenditure	Andhra Pradesh	1981-82	1831			
4	Aggregate_Expenditure	Andhra Pradesh	1982-83	1933			
5	Aggregate_Expenditure	Andhra Pradesh	1983-84	2588			
6	Aggregate_Expenditure	Andhra Pradesh	1984-85	3119			
7	Aggregate_Expenditure	Andhra Pradesh	1985-86	3413			
8	Aggregate_Expenditure	Andhra Pradesh	1986-87	4068			
9	Aggregate_Expenditure	Andhra Pradesh	1987-88	4294			
10	Aggregate_Expenditure	Andhra Pradesh	1988-89	5223			
11	Aggregate_Expenditure	Andhra Pradesh	1989-90	5768			
12	Aggregate_Expenditure	Andhra Pradesh	1990-91	6581			
13	Aggregate_Expenditure	Andhra Pradesh	1991-92	7758			
14	Aggregate_Expenditure	Andhra Pradesh	1992-93	8984			
15	Aggregate_Expenditure	Andhra Pradesh	1993-94	10541			
16	Aggregate_Expenditure	Andhra Pradesh	1994-95	12459			
17	Aggregate_Expenditure	Andhra Pradesh	1995-96	14301			
18	Aggregate_Expenditure	Andhra Pradesh	1996-97	16265			
19	Aggregate_Expenditure	Andhra Pradesh	1997-98	17745			
20	Aggregate_Expenditure	Andhra Pradesh	1998-99	21958			
21	Aggregate_Expenditure	Andhra Pradesh	1999-00	22767			
22	Aggregate_Expenditure	Andhra Pradesh	2000-01	28119			
23	Aggregate_Expenditure	Andhra Pradesh	2001-02	31074			

expenditure1

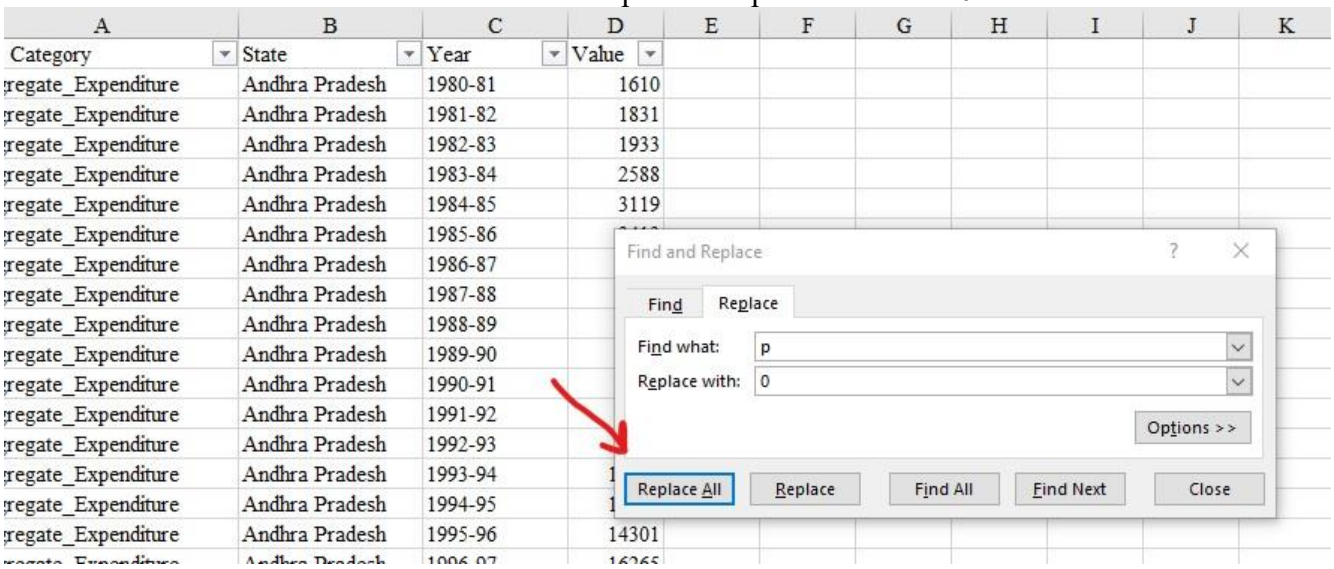
3.4Data Cleaning.

- Filter data, looking for blanks and empty data.
- Set all Title as one place for this go into view section →freeze panes→ select freeze top row.
- The data we have contains blank , “p” or “-” values. We have to fill them with Zeros (0) to make values column a proper numerical column.
- In Excel we can do that by selecting the whole Value column.

1. Click on “Find and select” and then “Find” to replace all “-” values with “0”.



2. Click on “Find and select” and then “Find” to replace all “p” values with “0”.



- Next, In the years column some data points have extra text. To remove them, we can create a new column with formula “=Left (C2,7)” and apply it for whole column, this this creating new column with same name.
- Now we copy and paste with past special as values with again new column and delete existing columns.
- After deleting swap new column Year in C cell. Now over Final dataset is ready. And save it into csv format.
- I am keeping all banks and null values I will cleaning it by using Python programming language.
- Now performing Exploratory Data Analysis with over Final dataset.

3.4 Export Data from Python
Perform Exploratory Data Analysis Using Python:

1. **Import Python Modules**
2. **Load Dataset**
3. **Data Preparation**
4. **EDA: Data Visualization**

- After performing Pre-processing and cleaning dataset.
- Data analyse by using Pandas-profiling.
- After cleaned data, its exported into csv as a expenditure1.csv.
- Now this cleaned dataset uses for creating dashboard in Power BI.

```
In [30]: #save clean profile file

         exp_clean_profile.to_file(output_file="expenditure_after_preprocessing.html")
```

Export report to file: 100% 1/1 [00:00<00:00, 8.93it/s]

```
In [31]: #save clean dataset into csv
         exp_data.to_csv('expenditure1.csv')
```

5. Data Visualization:

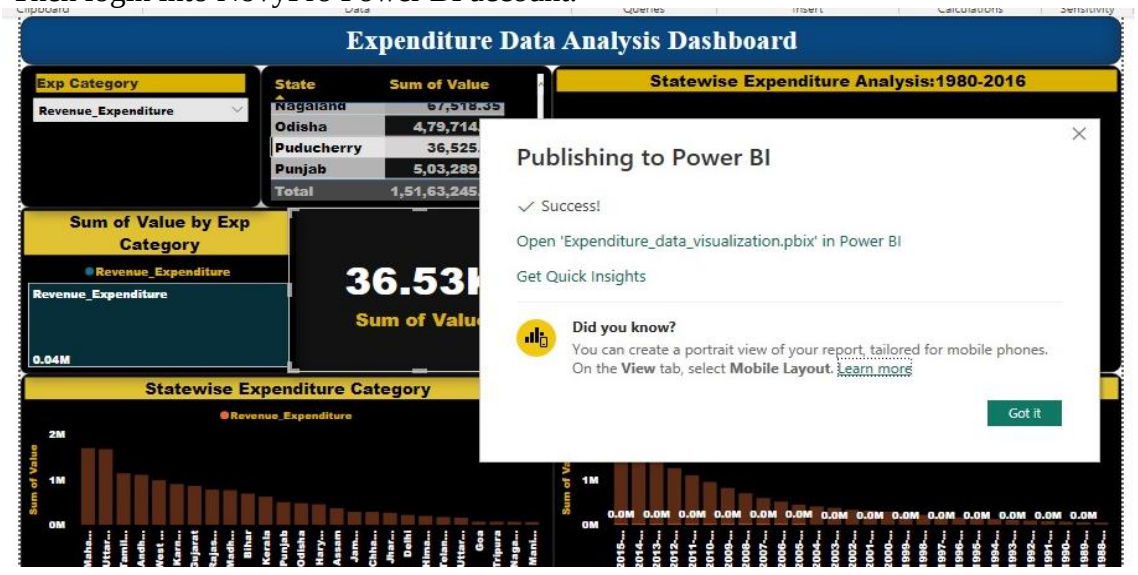
Data visualization is concerned with visually presenting sets of primarily quantitative raw data in a schematic form. The visual formats used in data visualization include tables, charts and graphs.

- In this project we use matplotlib and seaborn python libraries.

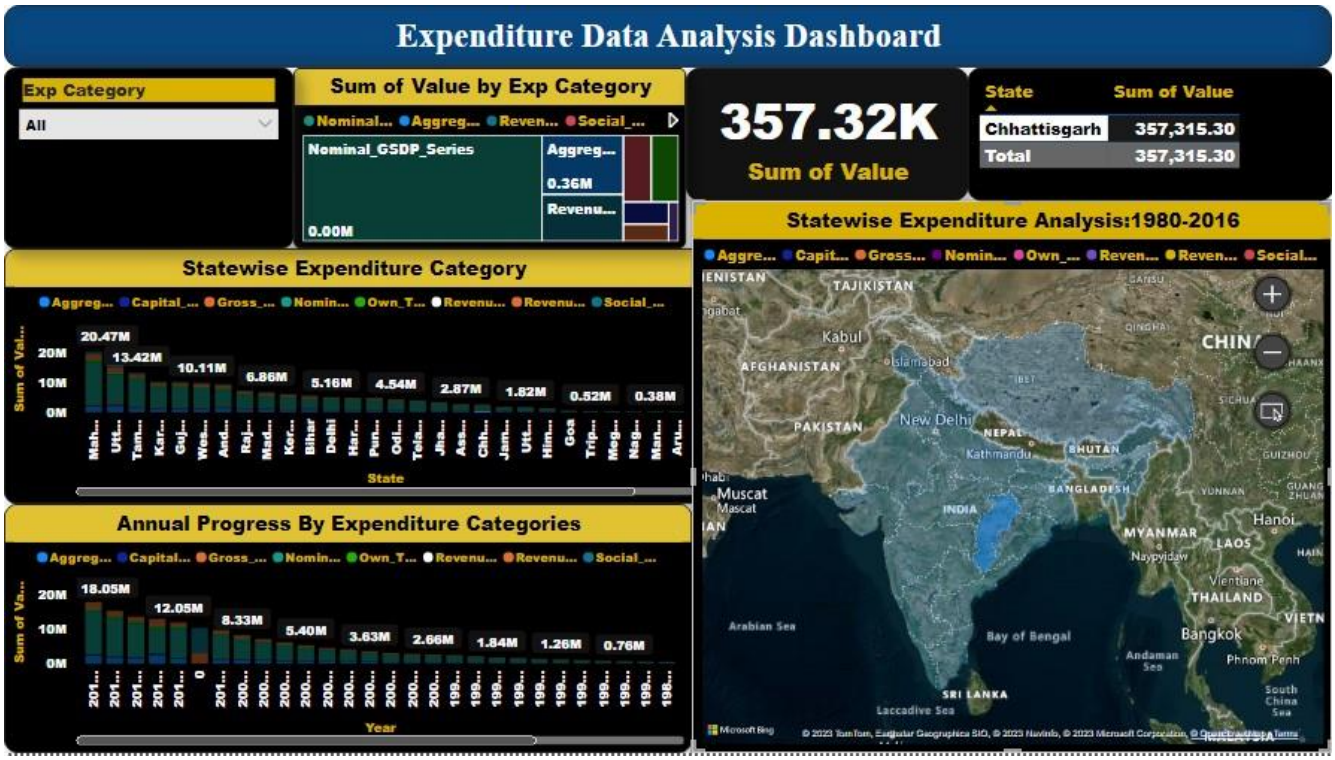
4. Deployment
For Development I used Power BI and NovyPro

Once you've completed your dashboard, follow these steps:

1. Load dataset on Power BI in csv formats and creates visuals for dashboard.
2. After creating all visual, create insightful dashboard.
3. Then Login into Power BI Service by using Microsoft developer account.
4. Then create new project workspace for uploading dashboard and reports into this workspace.
5. Then login into NovyPro Power BI account.



6. Then share dashboard as embedded link as a web into NovyPro portfolio.
7. Here in the below screenshot, we can see that dashboard has been published to NovyPro from Power BI service.



5. Unit Test Cases

TEST CASE DESCRIPTION	EXPECTED RESULTS
Expenditure Category slicer	When click on dropdown of slicer should occur which has various category of the Expenditure.
Select state as a slicer on map	When click on state of map it gives all distribution of expenditure category on tree map, state wise analysis and annual progress and value spending in card
Relation between expenditure category and value	As we select expenditure category corresponding values and state and years will change accordingly.