A

Project Report on

Ministry Budget

Submitted in partial fulfillment of completion of the course

Advanced Diploma in IT, Networking and Cloud

Submitted by:

Renu

Anju Kumari

Under Guidance of:

(Edunet Mentor) Mrs. Deepika

(Edunet Mentor) Miss. Ankita Shukla

Year 2023

Abstract

The title of our project is '**Ministry Budget**'. This data analysis project is a comprehensive examination of a government ministry's budget, with a specific focus on understanding its allocation and expenditure patterns. Our goal is to leverage data analytics, statistical methods, and visualization tools to gain insights into how financial resources are distributed among various programs, departments, and initiatives within the ministry. By scrutinizing historical budget data and trends, we aim to identify areas for optimization, cost-efficiency improvements, and alignment with the ministry's strategic priorities. This project not only seeks to enhance transparency and accountability in the budgeting process but also to contribute to informed decision-making by providing actionable recommendations for the ministry's resource allocation, ultimately leading to improved public service delivery.

Acknowledgement

We would like to express our sincere gratitude to several individuals and IBM organization for supporting us throughout our diploma study. First, we wish to express our sincere gratitude to our Edunet mentor, Mrs. Mala Mishra Mam & Miss. Ankita Shukla Mam, for her enthusiasm, patience, insightful comments, helpful information, practical advice and unceasing ideas that have helped us tremendously at all times in our study and writing of this project report.

In addition, we are deeply indebted to the Ministry of Skill Development & Entrepreneurship and IBM for granting us the diploma course. Their technical and financial support has enabled us to complete our diploma course studiessuccessfully.

Team Composition and Workload Division

1. Renu – UI (front-end) & Database (backend)

2. Anju Kumari – UI (front-end) & Database (backend)

1. Introduction to Problem

Manual Processes: The existing budget management system relies heavily on manual data entry and approval processes, leading to errors, delays, and a significant administrative burden on staff. This manual approach is prone to inconsistencies and is not scalable as the ministry's financial complexity grows.

Lack of Real-time Insights: Decision-makers within the ministry lack access to real-time financial insights, hindering their ability to respond promptly to budgetary changes, emerging priorities, or financial discrepancies. This lack of agility can impede the ministry's ability to adapt to dynamic circumstances.

Limited Transparency: The absence of a transparent budgetary framework makes it challenging for stakeholders, both internal and external, to understand the allocation of funds and track expenditure. This lack of transparency can erode trust and accountability, essential components of effective financial management.

Inefficient Resource Allocation: Without robust forecasting models and data-driven insights, the ministry faces challenges in optimizing resource allocation. This inefficiency may result in misaligned budget priorities, hindering the ministry's ability to achieve its strategic objectives.

Compliance and Security Concerns: The current system may not fully align with regulatory standards, posing compliance risks. Additionally, the security of financial data may be compromised due to the absence of comprehensive security measures, potentially leading to unauthorized access and data breaches.

2. Proposed Solution

1. **Automated Processes:**

Solution: Implement a modern budget management system that automates key processes such as data entry, approval workflows, and fund allocation. Utilize technologies such as artificial intelligence and machine learning to reduce manual intervention, minimizing errors and streamlining administrative tasks.

2. **Real-time Insights:**

Solution: Develop a dashboard within the budget management system that provides real-time financial insights. Utilize data visualization tools to present key performance indicators, budgetary trends, and expenditure patterns. This empowers decision-makers with timely information to respond proactively to budgetary changes and emerging priorities.

3. **Transparency Enhancement:**

Solution: Introduce features that enhance transparency, such as a publicly accessible budget portal. This portal should provide detailed information on budget allocations, expenditures, and financial performance. Implement user-friendly interfaces and plain-language summaries to make financial information accessible to a wide audience.

4. **Optimized Resource Allocation:**

Solution: Incorporate advanced forecasting models and analytics tools within the budget management system. These tools can leverage historical data and financial trends to assist decision-makers in making informed predictions about resource needs. This ensures that budget priorities align with the ministry's strategic objectives.

5. **Compliance and Security Measures:**

Solution: Conduct a comprehensive review of regulatory requirements and integrate compliance features into the budget management system. Implement robust security measures, including encryption protocols, access controls, and regular security audits, to safeguard financial data and mitigate the risk of unauthorized access or data breaches.

3. Requirements

Technology Stack

Full stack development

Full stack development refers to the development of both front end (client side) and back end (server side) portions of web application. We have used Jupiter Notebook (with python & pandas library Kaggle as the database.

Hardware

- Desktop/Laptop
- Minimum 4GB RAM
- Processor 64-bit
- Hard Drive 250GB
- Internet Connection

Software

Jupiter Notebook

Jupiter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations, and narrative text. It is widely used in data science, machine learning, scientific research, and education. The name "Jupiter" is a combination of three core programming languages it supports: Julia, Python, and R.

Front End

HTML: Hyper Text Markup Language (HTML) is a markup language for creating a webpage. Webpages are usually viewed in a web browser. They can include writing, links, pictures, and even sound and video. HTML is used to mark and describe each of these kinds of content so the web browser can display them correctly.

CSS: CSS is the language for describing the presentation of Web pages, including colors, layout, and fonts. It allows one to adapt the presentation to different types of devices, such as large screens, small screens, or printers. CSS is independent of HTML and can be used with any XML-based markup language.

4. User Requirements

- Electronic Device: Mobile, Laptop, Desktop or Tablet
- Email Account
- Access to Internet

5. Implementation Details

1. Importing libraries

A library is a collection of functions that can be added to your Python code and called as necessary, just like any other function.

2. Loading dataset

Data loading is the process of copying and loading data or data sets from a source file, folder or application to a database or similar application. It is usually implemented by copying digital data from a source and pasting or loading the data to a data storage or processing utility.

3. Viewing and understanding the Data

Data understanding is the knowledge that you have about the data, the needs that the data will satisfy, its content and location. To be clear, it is much more than current location and a definition of what a data element means in situ within an application or data base.

4. Data Visualization

Data visualization is the representation of data through use of common graphics, such as charts, plots, info graphics, and even animations. These visual displays of information communicate complex data relationships and data-driven insights in a way that is easy to understand.

5. Exploratory Data Analysis

Exploratory data analysis is one of the first steps in the data analytics process. In this post, we explore what EDA is, why it’s important, and a few techniques worth familiarizing yourself with.

6. Future Scope

The future scope of a website or project focused on "Ministry Budget" depends on various factors, including technological advancements, changing organizational needs, and evolving governmental practices. Here are some potential aspects that could shape the future scope of a Ministry Budget website:

Real-time Budget Adjustments: The future could bring real-time budget adjustment capabilities, allowing ministries to adapt quickly to changing circumstances. This could involve automated triggers based on predefined criteria or the integration of scenario modeling tools to assess the impact of potentialchangesonthebudget.

Blockchain for Transparency and Security: The adoption of blockchain technology may enhance the transparency and security of financial transactions within ministries. Implementing blockchain in budget management can provide an immutable ledger, reducing the risk of fraud and ensuring a transparent and auditable budgetary process.

Adoption of Cloud-Based Solutions: Cloud computing offers scalability, flexibility, and accessibility. The future scope may involve the adoption of cloud-based solutions for the Ministry Budget website, enabling seamless collaboration, data storage, and processing across different departments and locations.

Cyber security Measures: With the increasing importance of cyber security, future developments may focus on implementing advanced security measures to protect sensitive financial data. This could include multi-factor authentication, encryption protocols, and regular security audits to ensure the integrity of the budget management system.

International Standards and Interoperability: Aligning with international standards for budget management and improving interoperability with other government systems could be a future focus. This would facilitate data exchange and collaboration between ministries and government agencies on a broader scale.

7. Conclusion

The "Ministry Budget" project represents a significant step towards modernizing and optimizing financial management within the ministry. Through the adoption of innovative technologies and best practices in budgetary control, the project aims to position the ministry for long-term financial sustainability and success. As we embark on this journey, collaboration and commitment from all stakeholders will be essential to achieving the project's objectives and realizing the full potential of improved budget management within the ministry.

Appendix A

Project Code

Importing Libraries

```
Import numpy as np
Import pandas as pd
Import matplotlib.pyplot as plt
%matplotlib inline
Import seaborn as sns
Import plotly.express as px
```

Viewing and Understanding the Data

- data1 = pd.read_csv("dataset.csv")

```
data1
```

- # show top 5 rows

```
data1.head ()
```

- # show lost 5 records

```
data1.tail ()
```

- # general information

```
data1.info ()
```

- # give the number of rows and columns

```
data1.shape
```

- # extract all columns in to list of the dataset

```
list (data1.columns)
```

- # displaye single column type

```
type(data1['Ac_21-22_Capital'])
```

9. # calculate the mean, std, min, max and count of every attributes

```
data1.describe ()
```

Preprocessing the dataset

10. # checking for the missing values dataset

```
data1.isnull ()
```

11. # check for null value

```
data1.isnull ().sum()
```

12. # Check for duplicate values

```
data1.duplicated ().sum()
```

Data Visualization

13.

```
df_filtered_univar_1 = data1.loc[:, 'Sl.No.':'Bu_Es_23-24_Total']
df_filtered_univar_1 = df_filtered_univar_1.select_dtypes([np.int, np. Float])

for i, col in enumerate(df_filtered_univar_1.columns):
    plt. Figure(i)
    sns. Boxplot(x=col, data=df_filtered_univar_1)
    plt. Title("Box Plot of " + col)
```

14.

```
df_filtered_univar_1=data1.loc[:,'Sl.No.':'Bu_Es_23-24_Total']
df_filtered_univar_1=df_filtered_univar_1.select_dtypes([np.int, np.float])

for i, col in enumerate(df_filtered_univar_1.columns):
    plt.figure(i)
    sns.histplot(x=col, data=df_filtered_univar_1,kde=True).set_title("Plot Distribution "+col)
```

15.

```
df_filtered_Bivar=data1.loc[:,'Sl.No.':'Bu_Es_23-24_Total']
df_filtered_Bivar=df_filtered_Bivar.select_dtypes([np.int, np.float])
for i, feature in enumerate(df_filtered_Bivar.columns[1:]):
    plt.figure(i)
    g = sns.histplot(x=df_filtered_Bivar['Sl.No.'], y=feature, data=df_filtered_Bivar).set_title(feature + " vs
Sl.No.")
```

16. # Count plot distribution of "Bu_Es_23-24_Total" where 0 for loan repaid and 1 for client defaulted on loan

```
plt.figure(figsize=(20,15))
sns.countplot(x="Bu_Es_23-24_Total", data=data1, palette="Spectral")
plt. Title("Plot Distribution of Sl.No.")
plt. Show()
```

```
17. # Correlation
data1.corr()
```

Coorelation Matrix

```
18. #ploting the heatmap for correlation
plt.figure(figsize=(18,10))
Corr_data1=data1.corr()
correlation_heatMap = sns.heatmap(Corr_data1, annot=True)
```

Exploratory Data Analysis

```
19. # distplot for Bu_Es_23-24_Total
plt.style.use('fivethirtyeight')
plt.figure(figsize=(13, 7))
sns.distplot(data1['Bu_Es_23-24_Total'], bins=25)
```

Appendix B

Screenshot of Project

- **Loading Dataset**

Out[2]:

	Category	Sl.No.	Ministry/Department	Scheme	Actuals_21-22_Revenue\n	Ac_21-22_Capital	Ac_21-22_Total	Budget_Estimates_22-23_Revenue	Bu_Es_22-23_Capital	Bu_Es_22-23_Total	Revised_Es_22-23_Revenue
0	Demand No. 1	1	Department of Agriculture and Farmers Welfare	Crop Insurance Scheme	13549.24	0.0	13549.24	15500.0	0.0	15500.0	12375.76
1	Demand No. 1	2	Department of Agriculture and Farmers Welfare	Interest Subsidy for Short Term Credit to Farmers	21476.93	0.0	21476.93	0.0	0.0	0.0	0.00
2	Demand No. 1	3	Department of Agriculture and Farmers Welfare	Modified Interest Subvention Scheme (MISS)	0.00	0.0	0.00	19500.0	0.0	19500.0	22000.00
3	Demand No. 1	4	Department of Agriculture and Farmers Welfare	Market Intervention Scheme and Price Support S...	2288.33	0.0	2288.33	1500.0	0.0	1500.0	1500.00
4	Demand No. 1	5	Department of Agriculture and Farmers Welfare	Pradhan Mantri Annadata Aay Sanrakshan Yojna (...	0.00	0.0	0.00	1.0	0.0	1.0	0.00
...	...	...	...	...	...	...	...	...	...	...	...
717	Demand No. 102	718	Ministry of Youth Affairs and Sports	Scheme of Human Resource Development in Sports	0.00	0.0	0.00	0.0	0.0	0.0	0.00
718	Demand No. 102	719	Ministry of Youth Affairs and Sports	National Centre of Sports Science & Research	0.00	0.0	0.00	0.0	0.0	0.0	0.00

- **Viewing and understanding the Data**

Show all head data

	Category	Sl.No.	Ministry/Department	Scheme	Actuals_21-22_Revenue\n	Ac_21-22_Capital	Ac_21-22_Total	Budget_Estimates_22-23_Revenue	Bu_Es_22-23_Capital	Bu_Es_22-23_Total	Revised_Es_22-23_Revenue	Re_I 23_Re
0	Demand No. 1	1	Department of Agriculture and Farmers Welfare	Crop Insurance Scheme	13549.24	0.0	13549.24	15500.0	0.0	15500.0	12375.76	
1	Demand No. 1	2	Department of Agriculture and Farmers Welfare	Interest Subsidy for Short Term Credit to Farmers	21476.93	0.0	21476.93	0.0	0.0	0.0	0.00	
2	Demand No. 1	3	Department of Agriculture and Farmers Welfare	Modified Interest Subvention Scheme (MISS)	0.00	0.0	0.00	19500.0	0.0	19500.0	22000.00	
3	Demand No. 1	4	Department of Agriculture and Farmers Welfare	Market Intervention Scheme and Price Support S...	2288.33	0.0	2288.33	1500.0	0.0	1500.0	1500.00	
4	Demand No. 1	5	Department of Agriculture and Farmers Welfare	Pradhan Mantri Annadata Aay Sanrakshan Yojna (...	0.00	0.0	0.00	1.0	0.0	1.0	0.00	

Show all tail data

```
# show lost 5 records
data1.tail()
```

	Category	Sl.No.	Ministry/Department	Scheme	Actuals_21-22_Revenue\n	Ac_21-22_Capital	Ac_21-22_Total	Budget_Estimates_22-23_Revenue	Bu_Es_22-23_Capital	Bu_Es_22-23_Total	Revised_Es_22-23_Revenue
717	Demand No. 102	718	Ministry of Youth Affairs and Sports	Scheme of Human Resource Development in Sports	0.00	0.0	0.00	0.0	0.0	0.0	0.0
718	Demand No. 102	719	Ministry of Youth Affairs and Sports	National Centre of Sports Science & Research	0.00	0.0	0.00	0.0	0.0	0.0	0.0
719	Demand No. 102	720	Ministry of Youth Affairs and Sports	Khelo India	764.28	0.0	764.28	974.0	0.0	974.0	600.0
720	Demand No. 102	721	Ministry of Youth Affairs and Sports	Enhancement of Sports Facility at J& K	15.00	0.0	15.00	50.0	0.0	50.0	5.0
721	Demand No. 102	722	Ministry of Youth Affairs and Sports	Commonwealth Games	59.67	0.0	59.67	30.0	0.0	30.0	1.0

Show general information

```
# general information
data1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 722 entries, 0 to 721
Data columns (total 16 columns):
 #   Column                          Non-Null Count  Dtype
---  ------                          --------------  -----
 0   Category                        722 non-null    object
 1   Sl.No.                          722 non-null    int64
 2   Ministry/Department             722 non-null    object
 3   Scheme                          722 non-null    object
 4   Actuals_21-22_Revenue
          722 non-null    float64
 5   Ac_21-22_Capital                722 non-null    float64
 6   Ac_21-22_Total                  722 non-null    float64
 7   Budget_Estimates_22-23_Revenue  626 non-null    float64
 8   Bu_Es_22-23_Capital             721 non-null    float64
 9   Bu_Es_22-23_Total               721 non-null    float64
 10  Revised_Es_22-23_Revenue        722 non-null    float64
 11  Re_Es_22-23_Revenue             722 non-null    float64
 12  Re_Es_22-23_Total               722 non-null    float64
 13  Bu_Es_23-24_Revenue             722 non-null    float64
 14  Bu_Es_23-24_Capital             722 non-null    float64
 15  Bu_Es_23-24_Total               722 non-null    float64
dtypes: float64(12), int64(1), object(3)
memory usage: 90.4+ KB
```

Show shape, columns, types

```
# give the number of rows and columns
data1.shape
```

```
(722, 16)
```

```
# extract all columns in to list of the dataset
list(data1.columns)
```

```
['Category',
 'Sl.No.',
 'Ministry/Department',
 'Scheme',
 'Actuals_21-22_Revenue\n',
 'Ac_21-22_Capital',
 'Ac_21-22_Total',
 'Budget_Estimates_22-23_Revenue',
 'Bu_Es_22-23_Capital',
 'Bu_Es_22-23_Total',
 'Revised_Es_22-23_Revenue',
 'Re_Es_22-23_Revenue',
 'Re_Es_22-23_Total',
 'Bu_Es_23-24_Revenue',
 'Bu_Es_23-24_Capital',
 'Bu_Es_23-24_Total']
```

```
# displaye single column type
type(data1['Ac_21-22_Capital'])
```

```
pandas.core.series.Series
```

Show describe

```
# calculate the mean, std, min, max and count of every attributes
data1.describe()
```

	Sl.No.	Actuals_21-22_Revenue\n	Ac_21-22_Capital	Ac_21-22_Total	Budget_Estimates_22-23_Revenue	Bu_Es_22-23_Capital	Bu_Es_22-23_Total	Revised_Es_22-23_Revenue	Re_Es_22-23_Revenue	
count	722.000000	722.000000	722.000000	722.000000	626.000000	721.000000	721.000000	722.000000	722.000000	7
mean	361.500000	1027.546357	648.284100	1675.830457	994.103530	775.000610	1638.119626	1153.559751	801.743573	19
std	208.567735	9734.249468	8870.950264	13184.323361	7620.255919	10006.414315	12293.043869	10846.679252	9062.765579	14
min	1.000000	-1732.100000	-194425.540000	-194425.540000	0.000000	-192087.700000	-192087.700000	0.000000	-155539.630000	-155
25%	181.250000	0.000000	0.000000	0.422500	0.000000	0.000000	0.100000	0.000000	0.000000	
50%	361.500000	7.545000	0.000000	40.000000	26.000000	0.000000	50.000000	9.780000	0.000000	
75%	541.750000	125.317500	0.000000	269.387500	198.750000	0.000000	350.000000	124.262500	0.000000	3
max	722.000000	208929.000000	59700.000000	208929.000000	145919.900000	134015.000000	145919.900000	214696.000000	141605.620000	214

```
# checking for the missing values dataset
```

- **Preprocessing the dataset**

Show null values

```
# checking for the missing values dataset
data1.isnull()
```

	Category	Sl.No.	Ministry/Department	Scheme	Actuals_21-22_Revenue\n	Ac_21-22_Capital	Ac_21-22_Total	Budget_Estimates_22-23_Revenue	Bu_Es_22-23_Capital	Bu_Es_22-23_Total	Revised_Es_22-23_Revenue	Re_E 23_Rev
0	False	False	False	False	False	False	False	False	False	False	False	
1	False	False	False	False	False	False	False	False	False	False	False	
2	False	False	False	False	False	False	False	False	False	False	False	
3	False	False	False	False	False	False	False	False	False	False	False	
4	False	False	False	False	False	False	False	False	False	False	False	
...	...	...	...	...	...	...	...	...	...	...	...	
717	False	False	False	False	False	False	False	False	False	False	False	
718	False	False	False	False	False	False	False	False	False	False	False	
719	False	False	False	False	False	False	False	False	False	False	False	
720	False	False	False	False	False	False	False	False	False	False	False	
721	False	False	False	False	False	False	False	False	False	False	False	

722 rows × 16 columns

Now check null values

```
# check for null value
data1.isnull().sum()
```

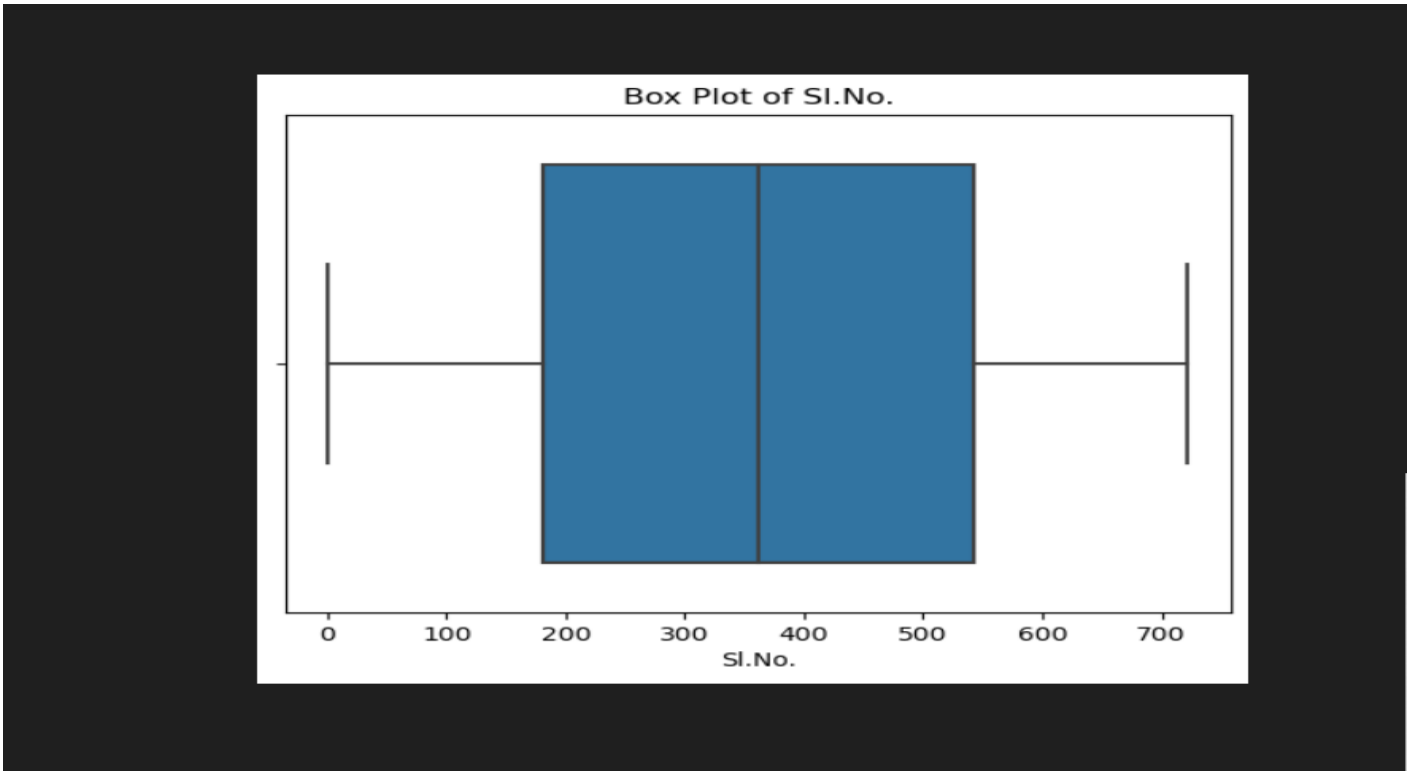
```
Category                           0
Sl.No.                             0
Ministry/Department                0
Scheme                             0
Actuals_21-22_Revenue\n            0
Ac_21-22_Capital                   0
Ac_21-22_Total                     0
Budget_Estimates_22-23_Revenue    96
Bu_Es_22-23_Capital                1
Bu_Es_22-23_Total                  1
Revised_Es_22-23_Revenue           0
Re_Es_22-23_Revenue                0
Re_Es_22-23_Total                  0
Bu_Es_23-24_Revenue                0
Bu_Es_23-24_Capital                0
Bu_Es_23-24_Total                  0
dtype: int64
```

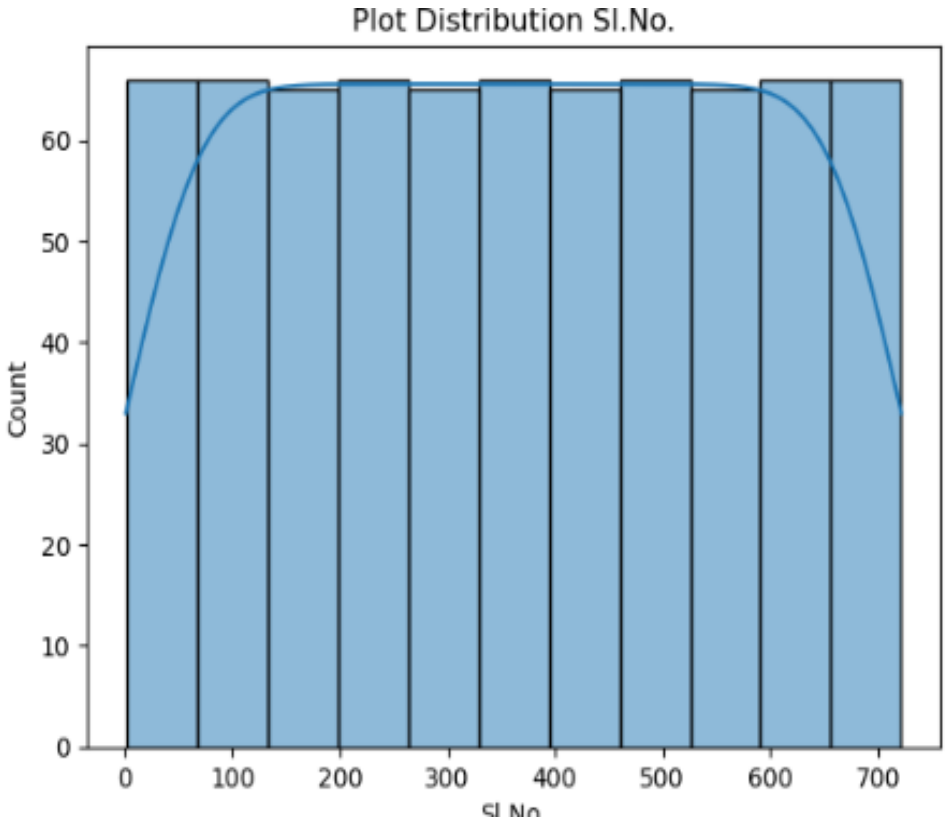
```
# Check for duplicate values
data1.duplicated().sum()
```

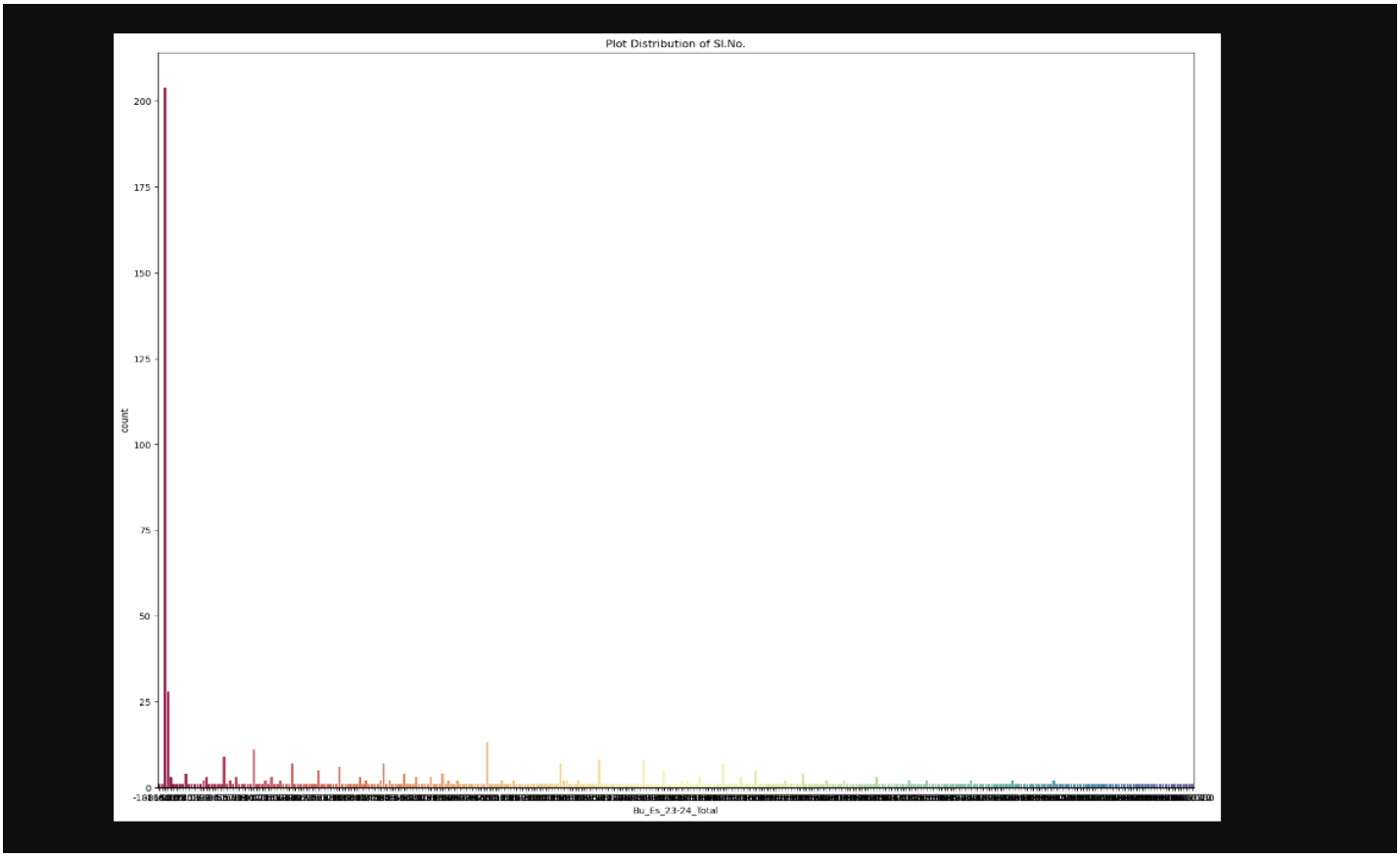
```
0
```

- **Data visualization**

Box plot







Correlation

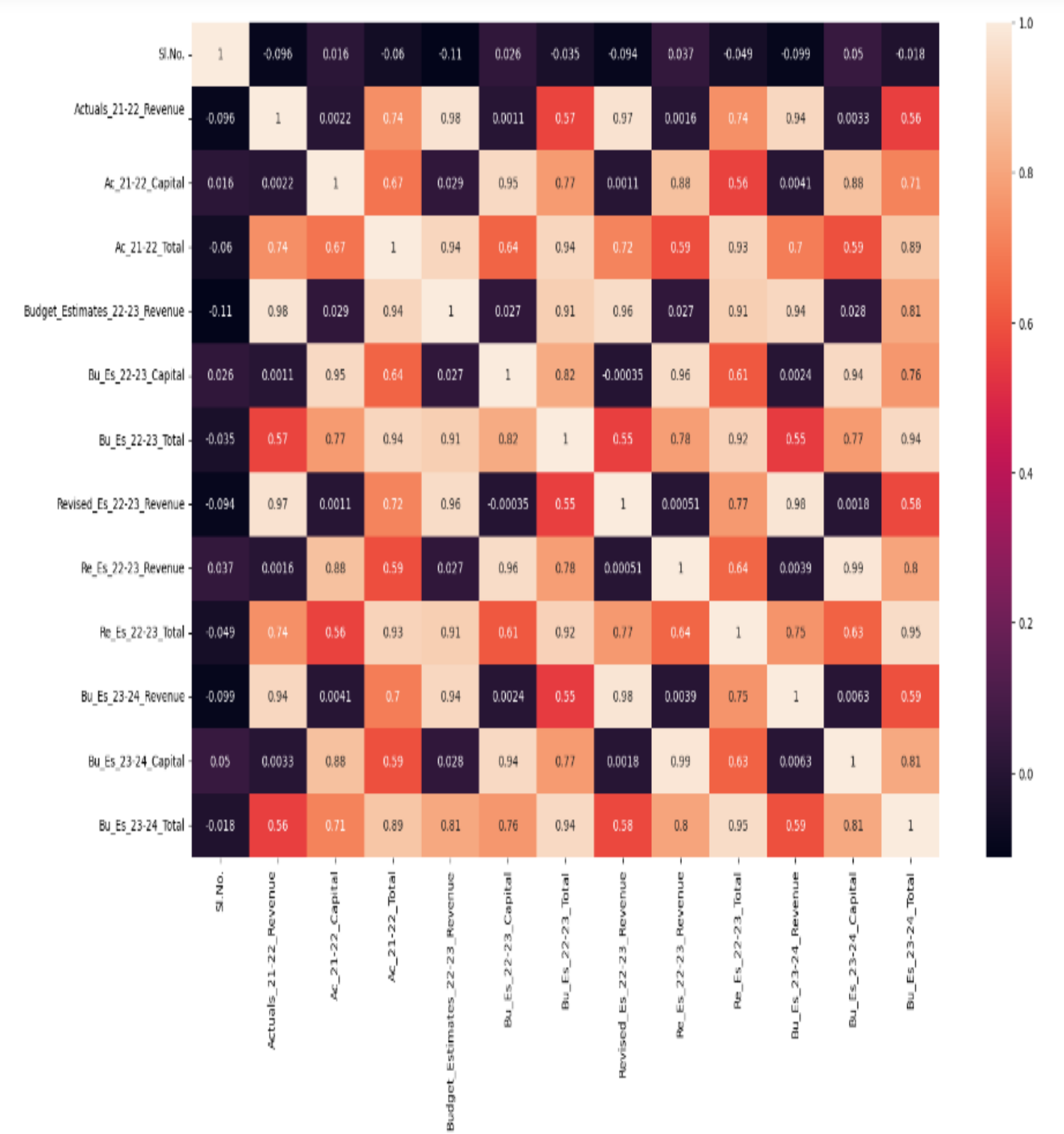
```
# Correlation
data1.corr()
```

```
C:\Users\Renu\AppData\Local\Temp\ipykernel_3472\3085187504.py:2: FutureWarning: The default value of numeric_only in DataFrame.
corr is deprecated. In a future version, it will default to False. Select only valid columns or specify the value of numeric_on
ly to silence this warning.
  data1.corr()
```

	Sl.No.	Actuals_21-22_Revenue\n	Ac_21-22_Capital	Ac_21-22_Total	Budget_Estimates_22-23_Revenue	Bu_Es_22-23_Capital	Bu_Es_22-23_Total	Revised_Es_22-23_Revenue	Re_Es_22-23_Revenue	Re_Es_22-23_Total
Sl.No.	1.000000	-0.095677	0.016268	-0.059694	-0.113175	0.026095	-0.035499	-0.094490	0.037324	-0.048567
Actuals_21-22_Revenue\n	-0.095677	1.000000	0.002183	0.739789	0.977175	0.001096	0.565885	0.969588	0.001590	0.744883
Ac_21-22_Capital	0.016268	0.002183	1.000000	0.674453	0.029146	0.946336	0.773432	0.001123	0.876527	0.562731
Ac_21-22_Total	-0.059694	0.739789	0.674453	1.000000	0.938667	0.637548	0.938206	0.716622	0.590937	0.928590
Budget_Estimates_22-23_Revenue	-0.113175	0.977175	0.029146	0.938667	1.000000	0.026653	0.914448	0.956297	0.026648	0.913467
Bu_Es_22-23_Capital	0.026095	0.001096	0.946336	0.637548	0.026653	1.000000	0.815904	-0.000351	0.956231	0.612697
Bu_Es_22-23_Total	-0.035499	0.565885	0.773432	0.938206	0.914448	0.815904	1.000000	0.552676	0.781369	0.924889
Revised_Es_22-23_Revenue	-0.094490	0.969588	0.001123	0.716622	0.956297	-0.000351	0.552676	1.000000	0.000514	0.767526
Re_Es_22-23_Revenue	0.037324	0.001590	0.876527	0.590937	0.026648	0.956231	0.781369	0.000514	1.000000	0.641413
Re_Es_22-23_Total	-0.048567	0.744883	0.562731	0.928590	0.913467	0.612697	0.924889	0.767526	0.641413	1.000000
Bu_Es_23-24_Revenue	-0.098586	0.942535	0.004131	0.698672	0.940684	0.002405	0.545924	0.980534	0.003905	0.754765
Bu_Es_23-24_Capital	0.049623	0.003303	0.877931	0.593146	0.028115	0.944517	0.773305	0.001813	0.985776	0.633291
Bu_Es_23-24_Total	-0.017773	0.555188	0.710435	0.887916	0.805134	0.763129	0.943671	0.576261	0.797274	0.953173

- **Coorelation Matrix**

Heat map



- **Exploratory Data Analysis**