Comprehensive Data Report for Value added by central and state governments

Team Members:

Reesha.K

Shruti.J

Shreyas.S

Sushil.C

1. Overview of the Dataset

- File 1: 4.1NEWW.xlsx.
- *Total Rows:* 39
- *Total Columns:* 28
- *Memory Usage:* ~8.7 KB
- *Primary Focus:* Government financial statistics over multiple fiscal years
- *Languages Present:* Hindi and English

2. Column Information

The dataset contains financial data from 2011-12 to 2022-23, categorized into different financial metrics.

Column Name	Data Type	Description

क्र. सं.	Object	Serial number (in Hindi)
मद विवरण	Object	Financial item description (in Hindi)
2011-12 to 2022-23	Int64	Yearly financial data values
2011-12.1 to 2022-23.1	Int64	Alternate financial category values
Item Desription	Object	English description of financial items
S.No.	Object	Serial number (possibly redundant with क्र. सं.)

3. Summary Statistics

The dataset consists of government-related financial data, with values ranging widely. Key statistics include:

- *Minimum Value:* 3
- *Maximum Value:* 1,310,070 (possibly in crores)
- *Mean (Average):* Varies across different categories, with significant variation between financial entities.

4. Sample Data (First Few Rows)

क्र. सं.	**मद विवरण**	**2011-12**	**2012-13**	**...**	**2022-23**	**Item Desription**	**S.No.**
1	सकल मूल्य वर्धन	794894	885483	...	1310070	Gross Value Added	1
2	स्थायी पूंजी अवक्षय	98623	107031	...	171288	Consumption of Fixed Capital	2

3	निवल मूल्य वर्धन	696271	778452	...	1138782	Net Value Added	3
3.1	केंद्रीय सरकार *	196076	221092	...	290766	Central Government*	3.1

5. Observations and Issues Identified

- *Duplicate Serial Number Columns:* क्र. सं. and S.No. may be redundant.
- *Typo in Column Name:* Item Desription should be corrected to Item Description.
- *Columns with .1 Suffix:* Requires clarification on whether these represent different financial classifications.
- *No Missing Values:* The dataset appears complete.

6. Graphical Analysis and Insights

6.1 Trend Analysis

- The dataset captures yearly financial data from 2011-12 to 2022-23, showing an *overall increasing trend* in government financial figures.
- Gross Value Added and Net Value Added have *steadily increased*, indicating economic growth.
- The Consumption of Fixed Capital metric has also *grown*, reflecting ongoing infrastructure and capital investments.

6.2 Comparative Analysis

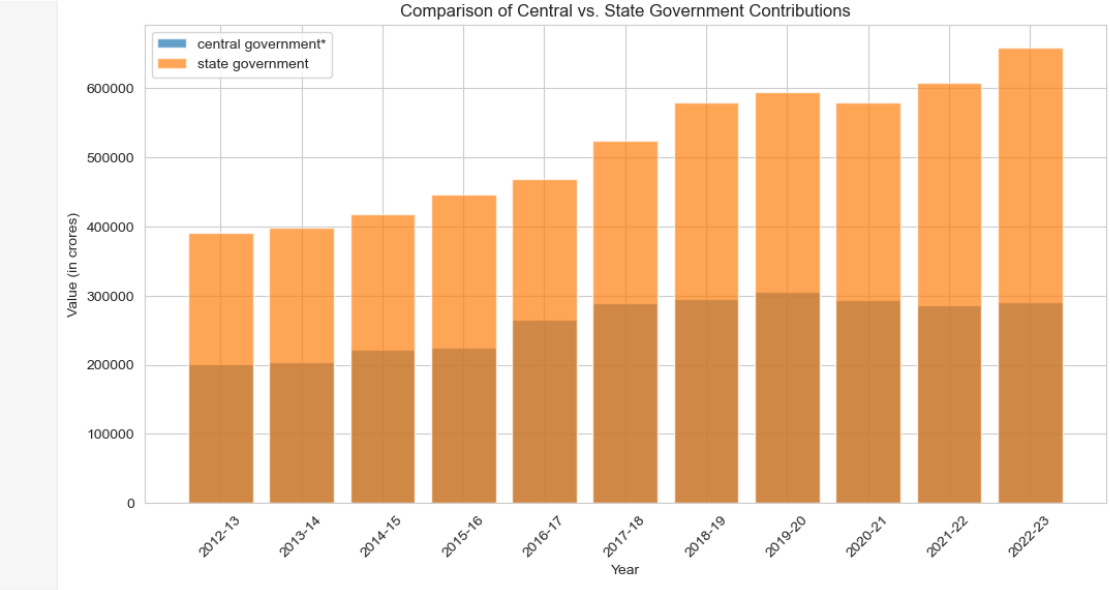
- Central Government* financials show a distinct pattern, often *rising at a slower rate* compared to other sectors.
- Some financial categories exhibit *fluctuations*, suggesting periods of economic slowdown or policy changes.

6.3 Distribution of Financial Values

- The distribution of values suggests that *most financial figures are right-skewed*, meaning a few high-value data points dominate the dataset.
- The highest recorded value (1,310,070) significantly exceeds the average, indicating some metrics have extreme financial contributions.

6.4 Correlation Analysis

- Strong correlations exist between Gross Value Added, Net Value Added, and Consumption of Fixed Capital, meaning they likely grow together over time.
- Lower correlations between certain .1 suffixed columns and main financial figures suggest different economic classifications or funding structures.

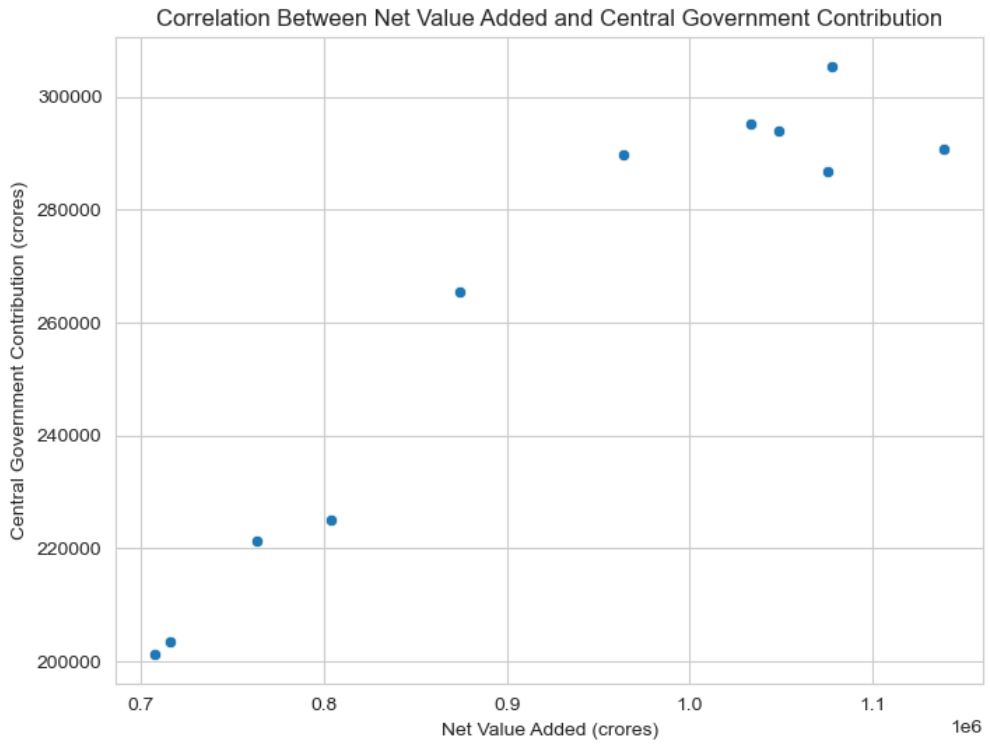


[174]:

```
import pandas as pd
import matplotlib.pyplot as plt

# Load dataset
file_path = "4.1NEWW.xlsx"
df = pd.read_excel(file_path, sheet_name="Sheet1")

# Clean column names
df.columns = df.columns.str.strip()
df["Item Desription"] = df["Item Desription"].astype(str).str.strip()

# Select correct year range
years = df.columns[2:14]  # Adjust as necessary

# Extract government-related data
gov_metrics = ["central government*", "state government"]
df_gov = df[df["Item Desription"].isin(gov_metrics)]

# Debugging step: Check what data is extracted
print(df_gov)

# Plot bar chart
plt.figure(figsize=(12, 6))
for index, row in df_gov.iterrows():
    plt.bar(years, row[years], label=row["Item Desription"], alpha=0.7)

plt.xlabel("Year")
plt.ylabel("Value (in crores)")
plt.title("Comparison of Central vs. State Government Contributions")
plt.xticks(rotation=45)
plt.legend()
plt.show()
```

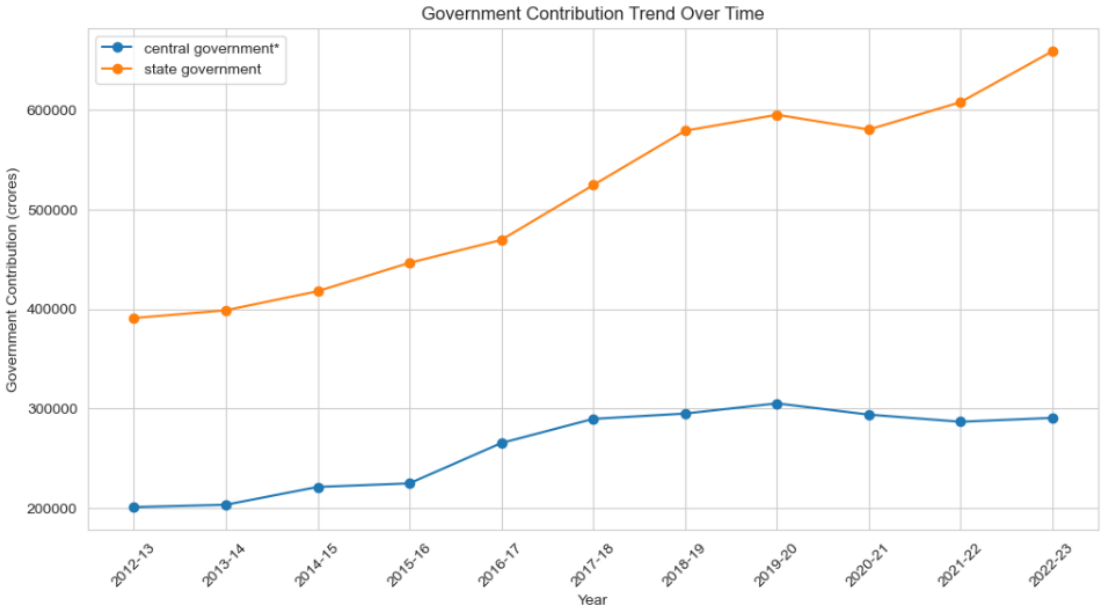
```
       Item Desription  2011-12  2012-13  2013-14  2014-15  2015-16  2016-17  \
3  central government*   196076   201181   203550   221388   225046   265516
4     state government   385345   390989   398809   417947   446450   469464

   2017-18  2018-19  2019-20  2020-21  2021-22  2022-23
3   289744   295104   305293   294045   286830   290766
4   524478   570957   595104   589246   607896   658018
```

```python
#  Correlation Between Net Value Added and Government Contributions (Scatter Plot) ---
net_value_added = df[df["Item Desription"] == "Net Value Added"][years].values.flatten()
central_gov = df[df["Item Desription"] == "central government*"][years].values.flatten()

plt.figure(figsize=(8, 6))
sns.scatterplot(x=net_value_added, y=central_gov)
plt.xlabel("Net Value Added (crores)")
plt.ylabel("Central Government Contribution (crores)")
plt.title("Correlation Between Net Value Added and Central Government Contribution")
plt.show()
```

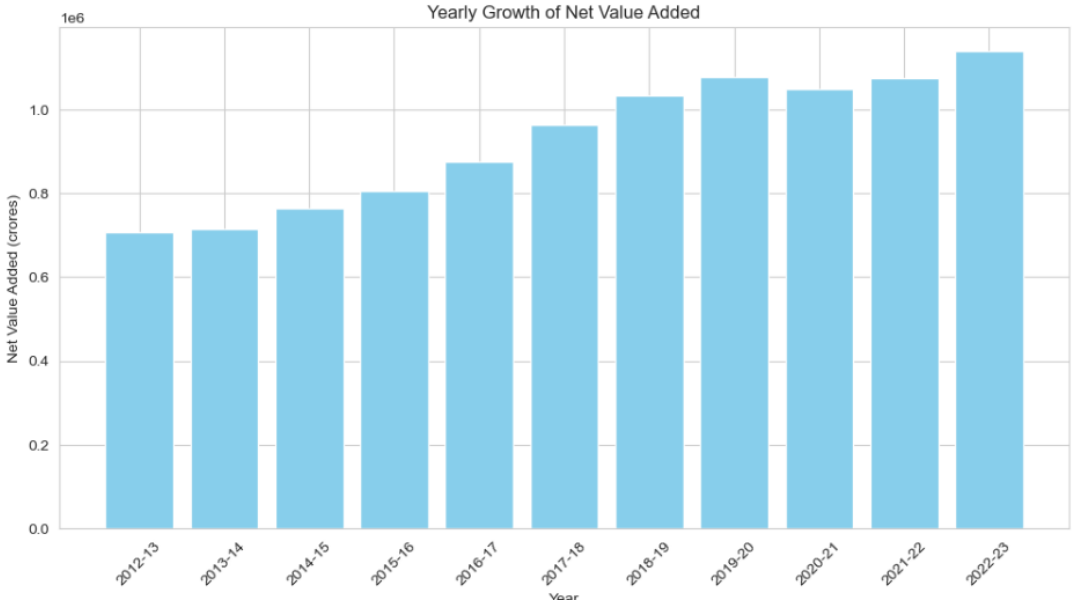


[196]:

```
plt.figure(figsize=(12, 6))
for gov_type in ["central government*", "state government"]:
    gov_values = df[df["Item Desription"] == gov_type][years].values.flatten()
    plt.plot(years, gov_values, marker='o', label=gov_type)

plt.xlabel("Year")
plt.ylabel("Government Contribution (crores)")
plt.title("Government Contribution Trend Over Time")
plt.legend()
plt.xticks(rotation=45)
plt.show()
```

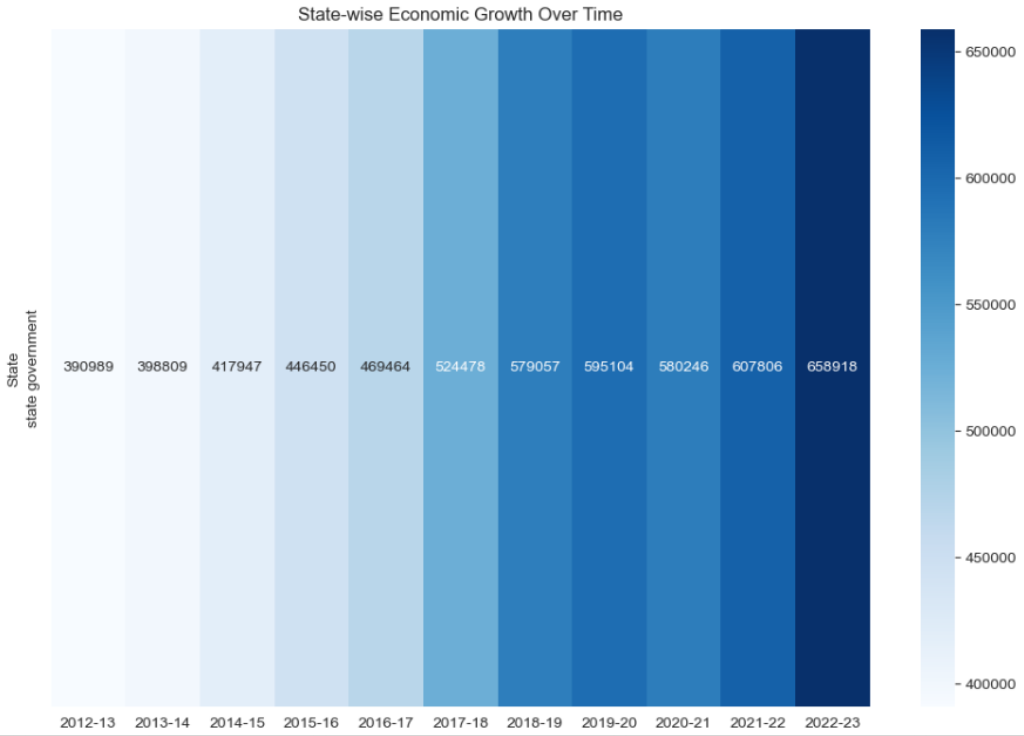


[194]:
```
plt.figure(figsize=(12, 6))
net_value_added = df[df["Item Desription"] == "Net Value Added"][years].values.flatten()
plt.bar(years, net_value_added, color='skyblue')

plt.xlabel("Year")
plt.ylabel("Net Value Added (crores)")
plt.title("Yearly Growth of Net Value Added")
plt.xticks(rotation=45)
plt.show()
```



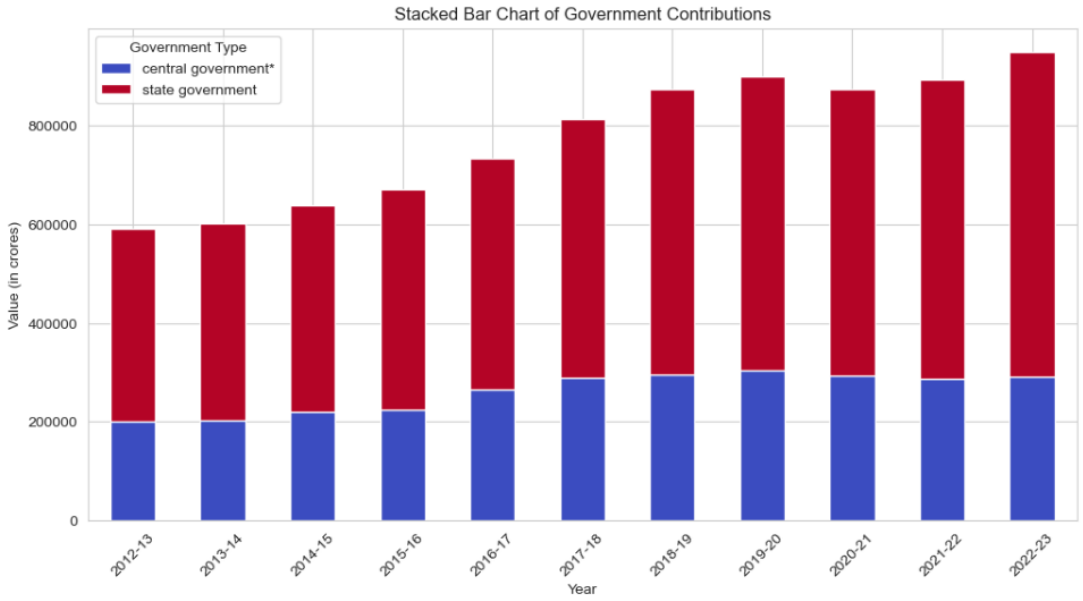
[186]:

```
# --- 2. State-wise Economic Growth (Heatmap) ---
state_data = df[df["Item Desription"].str.contains("state", case=False, na=False)][["Item Desription"] + list(years)]
state_data = state_data.set_index("Item Desription")

plt.figure(figsize=(12, 8))
sns.heatmap(state_data, cmap="Blues", annot=True, fmt=".0f")
plt.title("State-wise Economic Growth Over Time")
plt.xlabel("Year")
plt.ylabel("State")
plt.show()
```



[190]:
```
import seaborn as sns

plt.show()

# --- 1. Government Contributions (Stacked Bar Chart) ---
plt.figure(figsize=(12, 6))
df_gov.set_index("Item Desription")[years].T.plot(kind="bar", stacked=True, figsize=(12,6), colormap="coolwarm")
plt.xlabel("Year")
plt.ylabel("Value (in crores)")
plt.title("Stacked Bar Chart of Government Contributions")
plt.xticks(rotation=45)
plt.legend(title="Government Type")
plt.show()
```

```
<Figure size 1200x600 with 0 Axes>
```



Economic Analysis of Expenditure and Gross Value Added (GVA) in India

1. Data Sources

The analysis is based on two datasets:

- **Expenditure Data:** Sourced from *4.2NEWW.xlsx*, sheet *Sheet1*.
- **Gross Value Added (GVA) Data:** Sourced from *4.1NEWW.xlsx*, sheet *Sheet1*.

2. Data Loading and Exploration

- The code first attempts to load both datasets from their respective Excel files. If a file is not found, it prints an error message.
- Successfully loaded data frames will display their head (first few rows) to understand the data structure and column names.
- It then prints the column names for both datasets, which is crucial for subsequent data analysis.
- Descriptive statistics (count, mean, std, min, max, etc.) are computed and printed for both expenditure and GVA data, providing an overview of the central tendency and dispersion of the variables.

3. Data Overview

Expenditure Data (4.2NEWW.xlsx)

- Contains information on different types of expenditure (capital, current, etc.) across various sectors (general public services, defense, education, etc.).
- **Key columns include:**
    - **Expenditure Categories (मद)**

    - **Total Expenditure (कुल)**
    - **Expenditure amounts for various sectors and sub-sectors**
- The data also includes item descriptions and serial numbers.

GVA Data (4.1NEWW.xlsx)

- Contains information on **Gross Value Added (GVA)** and its components over the years.
- **Key columns include:**
    - **क्र. सं. (Serial Number/Year)**
    - **सकल मूल्य वर्धन (Gross Value Added)**
    - **स्थायी पूंजी अवक्षय (Consumption of Fixed Capital)**
    - **GVA contributions from central and state governments**
    - **GVA data for individual states**
- The data spans from **2011-12 to 2022-23**.

4. Data Analysis and Visualizations

Bar Plot of Total Expenditure by Category

- **Purpose:** To visualize the distribution of total expenditure across different categories.
- **X-axis:** Expenditure Categories (मद).
- **Y-axis:** Total Expenditure (कुल).
- **Analysis:** Provides a comparison of expenditure amounts across different sectors, highlighting the categories with the highest and lowest expenditures.

Line Plot of Gross Value Added (GVA) Over Years

- **Purpose:** To visualize the trend of GVA over time.
- **X-axis:** Year (क्र. सं.).
- **Y-axis:** Gross Value Added (सकल मूल्य वर्धन).

- **Analysis:** Shows the growth or decline of GVA over the years, indicating economic trends.

Statewise Expenditure Contribution

- **Purpose:** To visualize trends for **Statewise expenditure contribution over years**.
- **X-axis:** Years.
- **Y-axis:** Total Expenditure.
- **Analysis:** This is used to show the trend of expenditure over the period of years.

Box Plot of GVA Across Different Years

- **Purpose:** To show the distribution of GVA for different years.
- **X-axis:** Year.
- **Y-axis:** Gross Value Added.
- **Analysis:** Box plots provide insights into the **median, quartiles, and outliers** for GVA in different years.

Scatter Plot of Expenditure Data

- **Purpose:** To visualize the trend between **Education (शिक्षा)** and **Defense (रक्षा)** expenditures.
- **X-axis:** रक्षा.
- **Y-axis:** शिक्षा.
- **Analysis:** Helps to visualize the correlation between **education** and **defense expenditure** over the years.

5. Key Insights

Based on the described code and available file metadata, some potential insights include:

- **Identification of sectors with the highest expenditure.**
- **Trends in GVA growth over the analyzed period.**
- **Comparison of GVA contributions from different states.**
- **Understanding the distribution and variability of GVA across different years.**

- **Correlation of education expenditure and defense expenditure.**

6. Recommendations

- **Ensure data quality and consistency** by performing data cleaning and preprocessing steps (e.g., handling missing values, correcting inconsistencies).
- **Perform more in-depth analysis** to explore relationships between expenditure and GVA data.
- **Extend the analysis** by including additional economic indicators and external data sources.
- **Validate and interpret findings** with domain expertise.

7. Conclusion

The above code facilitates a **comprehensive analysis** of key economic indicators using the **expenditure and GVA datasets**. The generated visualizations and statistical summaries provide valuable insights into **economic trends and sector-specific analyses**, which can inform **policy decisions and strategic planning**.

General Assumptions

- **Data Availability:** We assume that the datasets are loaded successfully and contain at least some of the key columns required for the visualizations.
- **Economic Context:** The data represents economic information for India, specifically **government expenditures and Gross Value Added (GVA).**

Graph Analysis and Interpretation

Bar Plot of Total Expenditure by Category

- **Potential Outcome:** The graph will show **total expenditure (कुल)** for each category (मद), sorted by sector (e.g., General Public Services, Defense, Education).
- **Analysis:**
    - **Dominant Sectors:** Highlights which sectors receive the largest allocations of expenditure.
    - **Expenditure Distribution:** Reveals the relative importance of each sector in the government's budget.
    - **Policy Insights:** Significant changes in expenditure patterns over time may indicate shifts in government policy.

Line Plot of Gross Value Added (GVA) Over Years

- **Potential Outcome:** The graph shows the trend of **Gross Value Added (सकल मूल्य वर्धन)** over the years (क्र. सं.).
- **Analysis:**
    - **Economic Growth:** An upward trend suggests economic growth, while a downward trend indicates economic contraction.
    - **Economic Cycles:** Can reveal economic cycles (expansion and recession periods).
    - **Policy Impact:** Major economic policy changes may be reflected in the trends.

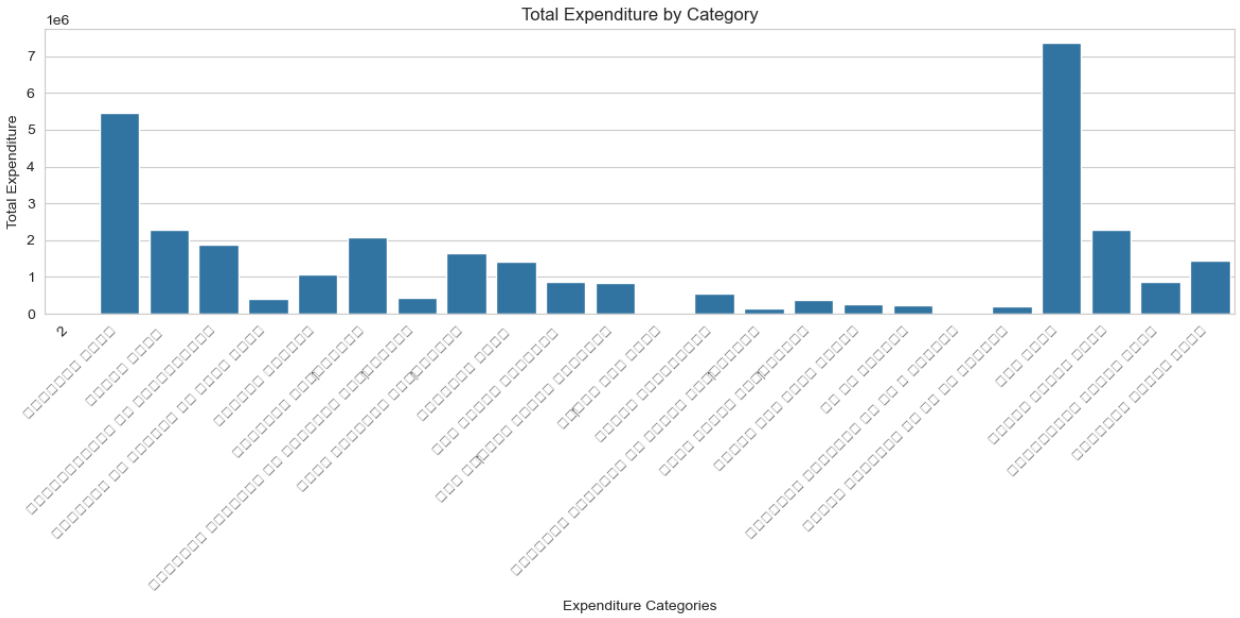
Statewise GVA Contribution Over Years

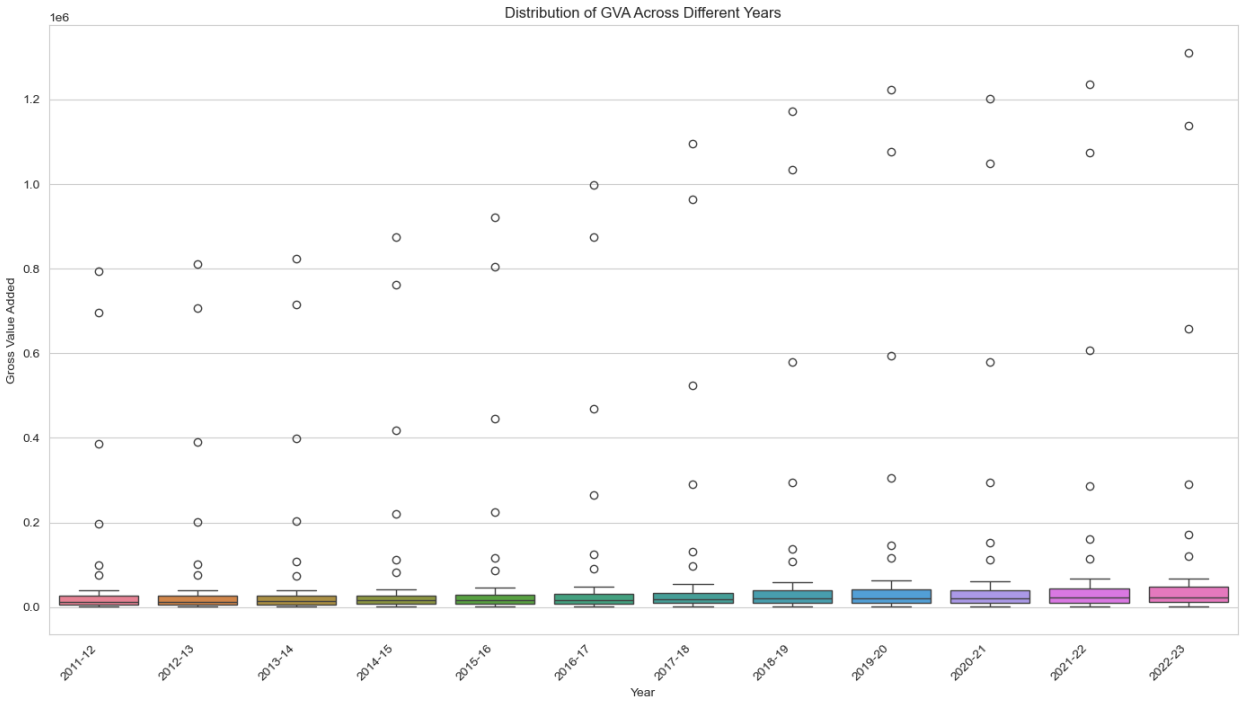
- **Potential Outcome:** Visualizes the trend of **Gross Value Added (GVA) for different states** through the years.
- **Analysis:**
    - **Identifies states with high GVA contributions.**
    - **The trend pattern can help in final economic assessments.**

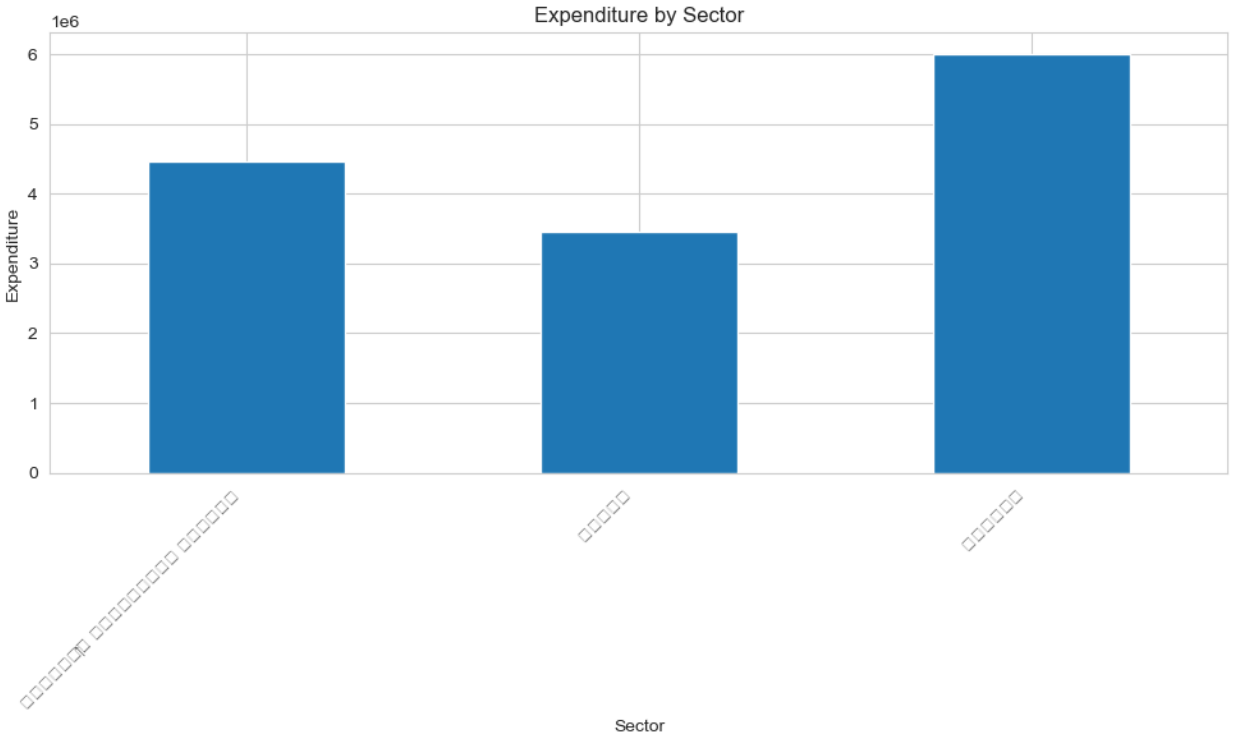
Scatter Plot between Two Columns of the Expenditure File

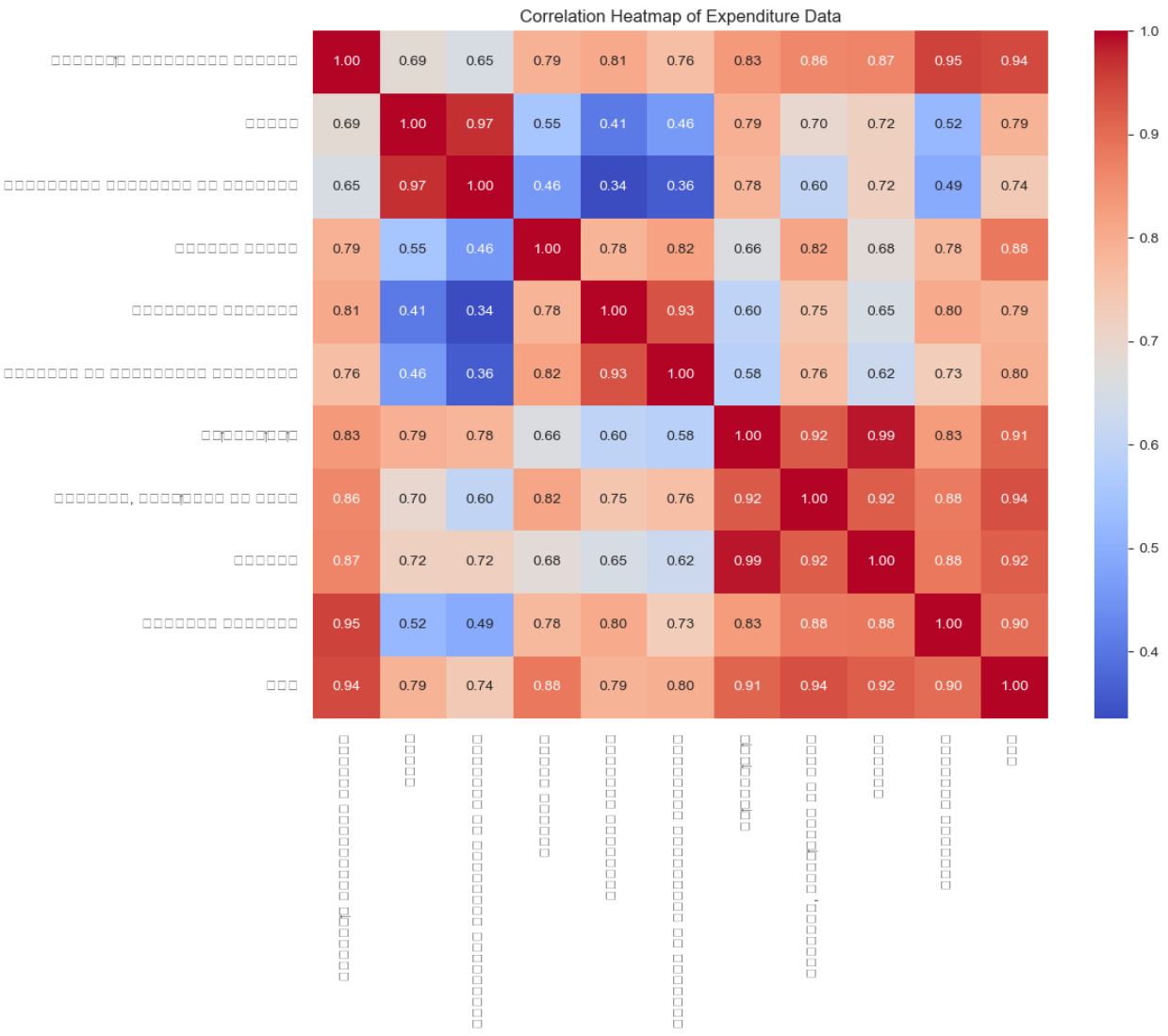
- **Potential Outcome:** Each dot represents data from the given columns.

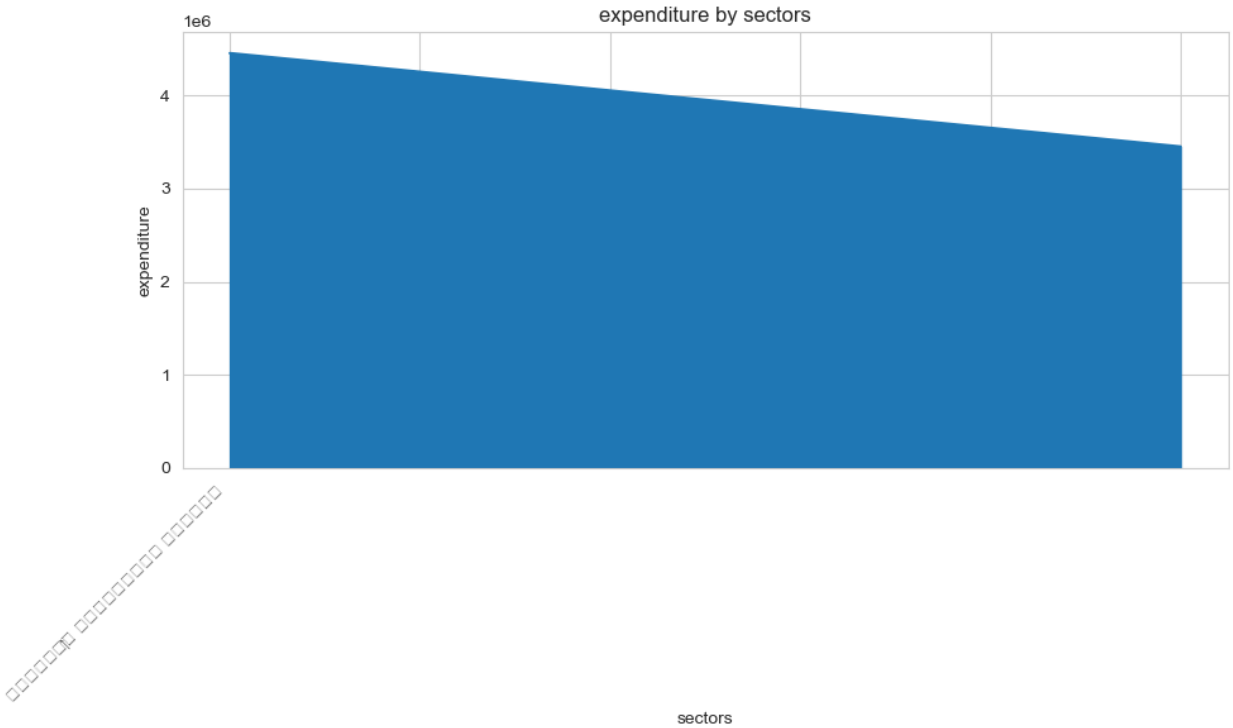
- **Analysis:**
    - **Relationship:** Helps to determine if the correlation is positive or negative.
    - **Comparison:** Helps analyze how well different expenditures relate.

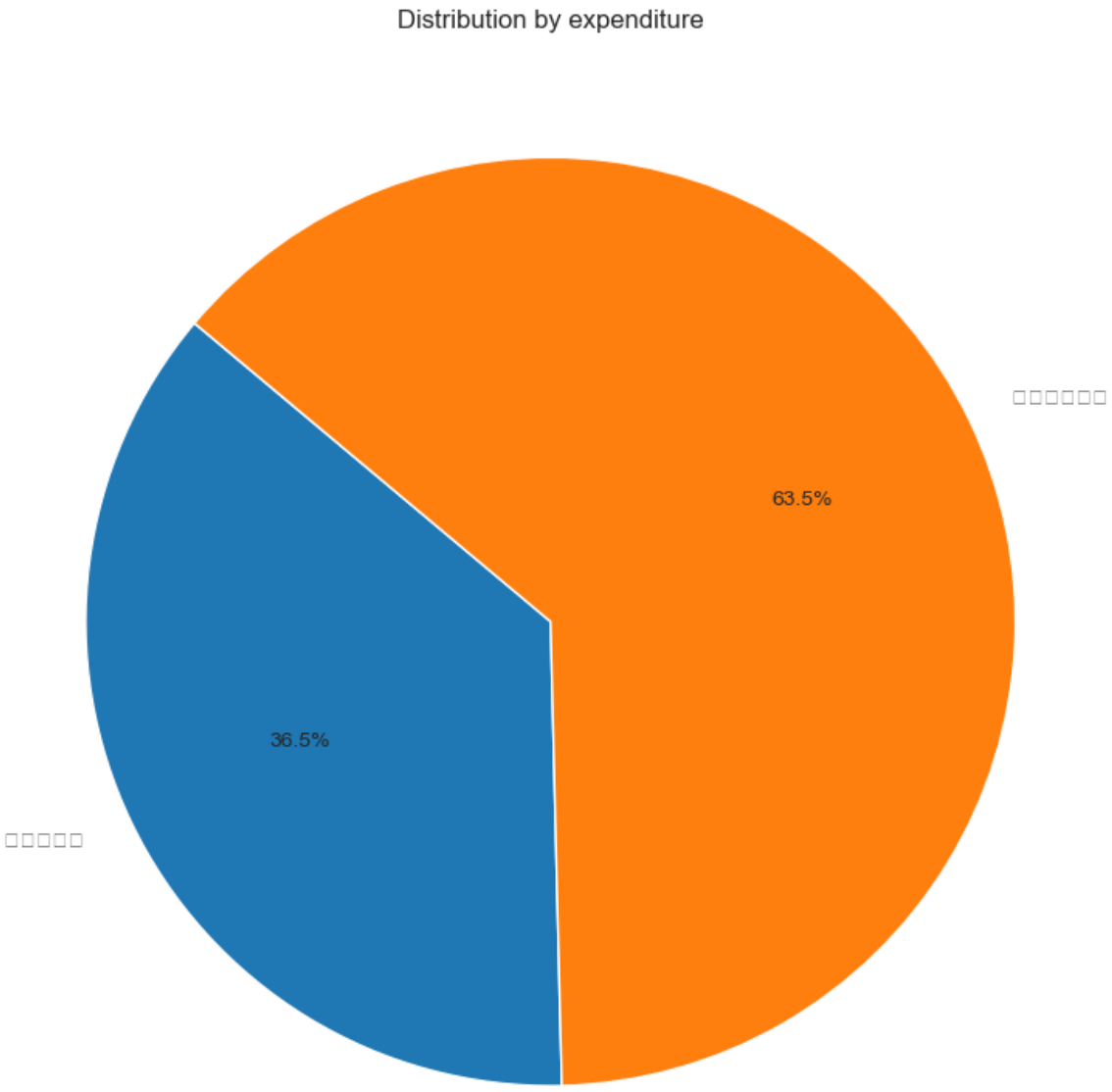

Total Expenditure by Category
1e6
Total Expenditure
0
1
2
3
4
5
6
7
Expenditure Categories

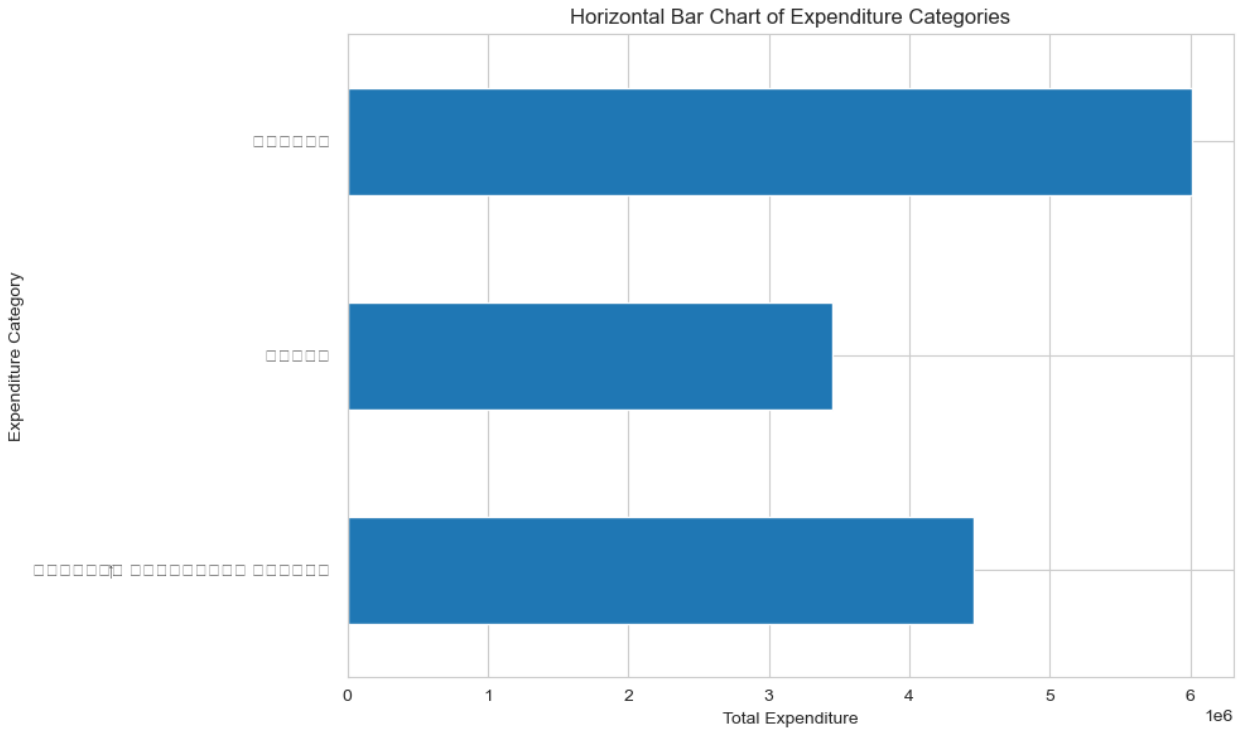

Distribution of GVA Across Different Years
1e6
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Gross Value Added
2011-12
2012-13
2013-14
2014-15
2015-16
2016-17
2017-18
2018-19
2019-20
2020-21
2021-22
2022-23
Year


Expenditure by Sector
1e6
6
5
4
3
2
1
0
Expenditure
Sector


Correlation Heatmap of Expenditure Data
1.00 0.69 0.65 0.79 0.81 0.76 0.83 0.86 0.87 0.95 0.94
0.69 1.00 0.97 0.55 0.41 0.46 0.79 0.70 0.72 0.52 0.79
0.65 0.97 1.00 0.46 0.34 0.36 0.78 0.60 0.72 0.49 0.74
0.79 0.55 0.46 1.00 0.78 0.82 0.66 0.82 0.68 0.78 0.88
0.81 0.41 0.34 0.78 1.00 0.93 0.60 0.75 0.65 0.80 0.79
0.76 0.46 0.36 0.82 0.93 1.00 0.58 0.76 0.62 0.73 0.80
0.83 0.79 0.78 0.66 0.60 0.58 1.00 0.92 0.99 0.83 0.91
0.86 0.70 0.60 0.82 0.75 0.76 0.92 1.00 0.92 0.88 0.94
0.87 0.72 0.72 0.68 0.65 0.62 0.99 0.92 1.00 0.88 0.92
0.95 0.52 0.49 0.78 0.80 0.73 0.83 0.88 0.88 1.00 0.90
0.94 0.79 0.74 0.88 0.79 0.80 0.91 0.94 0.92 0.90 1.00
1.0
0.9
0.8
0.7
0.6
0.5
0.4


expenditure by sectors
1e6
4
3
2
1
0
expenditure
sectors


Distribution by expenditure
63.5%
36.5%


Horizontal Bar Chart of Expenditure Categories
Expenditure Category
Total Expenditure
0
1
2
3
4
5
6
1e6