Empowering India: Analysing the Evolution of Union Budget Allocations for Sustainable Growth in this topic technology stack

---

Here's a comprehensive and practical Technology Stack you can use for the project:

"Empowering India: Analysing the Evolution of Union Budget Allocations for Sustainable Growth"
This stack supports data collection, storage, processing, analysis, visualization, and reporting.

🔹 1. Data Sources

Primary Sources
Union Budget datasets (year-wise expenditure & revenue statements)
Government Open Data Portal (data.gov.in)
Economic Survey reports (PDF/Excel)
Formats
CSV
Excel (.xlsx)
JSON
PDF

🔹 2. Data Collection & Extraction

Tools
Python
BeautifulSoup / Scrapy (for web scraping)
Tabula / Camelot (PDF table extraction)
Purpose
Automate downloading of budget datasets
Extract tables from PDFs
Convert raw data into structured format

🔹 3. Data Storage

Options
MySQL / PostgreSQL (Relational DB)
MongoDB (NoSQL for semi-structured data)
Local CSV/Parquet files
Why
Store multi-year budget allocations
Enable fast querying and historical comparison

🔹 4. Data Cleaning & Preprocessing

Libraries
Pandas
NumPy
OpenRefine (optional GUI tool)
Tasks
Handle missing values
Normalize column names
Adjust inflation (real vs nominal values)
Standardize sector categories

🔹 5. Data Analysis

Statistical & Analytical Tools
Python (Pandas, SciPy, Statsmodels)
R (dplyr, tidyr)
Techniques

Trend analysis
CAGR calculation
Sector-wise growth comparison
Correlation between allocations & development indicators
🔹 6. Machine Learning (Optional – Advanced)
Libraries
Scikit-learn
XGBoost
TensorFlow / PyTorch
Use Cases
Forecast future budget allocations
Cluster sectors based on spending patterns
Predict impact on sustainability indicators
🔹 7. Data Visualization
Visualization Tools
Tableau
Power BI
Matplotlib & Seaborn (Python)
Plotly (interactive charts)
Dashboards
Year-wise allocation trends
Sector share comparison
Heatmaps for priority sectors
Forecast vs actual spending
🔹 8. Business Intelligence Layer
Tools
Tableau Desktop / Tableau Public
Power BI Service
Purpose
Interactive dashboards
Filters by year, sector, scheme
Drill-down from national → sector → scheme level
🔹 9. Backend (If Web App Required)
Frameworks
Flask / Django (Python)
Node.js (Express)
Functions
Serve processed data
Connect dashboards to database
Handle API requests
🔹 10. Frontend (Optional Web Interface)
Technologies
HTML, CSS, JavaScript
React.js
Bootstrap / Tailwind CSS
Features
Interactive charts
Search & filter
Download reports
🔹 11. Deployment & Hosting
Platforms
AWS / Azure / Google Cloud
Heroku / Render (lightweight apps)
Services
EC2 / App Service
Cloud SQL / RDS
🔹 12. Reporting & Documentation
Tools
Jupyter Notebook

LaTeX / MS Word
PowerPoint
Outputs
Research report
PDF dashboards
Presentation slides
🔹 13. Version Control & Collaboration
Tools
Git
GitHub / GitLab
Benefits
Track changes
Team collaboration
Reproducibility
✅ Sample End-to-End Stack Summary
Collection: Python + BeautifulSoup + Tabula
Storage: PostgreSQL
Processing: Pandas, NumPy
Analysis: Python, Scikit-learn
Visualization: Tableau / Power BI
Deployment: AWS + Flask
Reporting: Jupyter Notebook + PDF
If you want, I can also provide:
✔ Architecture diagram
✔ Tools justification table
✔ Mapping of objectives → tools
✔ Sample dataset schema
Just tell me 👍