🌐 Real-Time Governance System (RTGS) Documentation

For Policymakers | CLI-Based Agentic System | Telangana Open Data Portal

🎯 Clarity First: Project Overview

Item	Description
Mission	Build a terminal-based agentic system that transforms raw public data into clean, trustworthy evidence.
Objective	Create a CLI-first system that ingests open government data, makes it analysis-ready, and surfaces insights for decision-makers.
Scope	Focused on *Telangana District-Level Education Data (2022)* for prototype clarity.
Success Criteria	CLI returns standardized insights, logs are readable, and visuals are generated from cleaned data.

🧩 Step-by-Step Project Approach

1️⃣ Dataset Selection & Scope

Source	Dataset	Scope
Telangana Open Data Portal	District-wise Education Statistics	Year: 2022, District: Hyderabad

Log:

```
Code  ^                                          📋 Copy

[INFO] 2025-09-07 12:00 - Selected dataset: Education_Stats_2022_Hyderabad.c
```

3 Cleaning & Transformation Output

Task	Result
Null values removed	35 rows
Duplicates dropped	12 rows
Categorical encoding	`School_Type` → [Govt, Private, Aided]

2️⃣ Ingestion & Standardization Output

- Loaded CSV using pandas

- Standardized column names: `District`, `School_Type`, `Enrollment`, `Pass_Percentage`
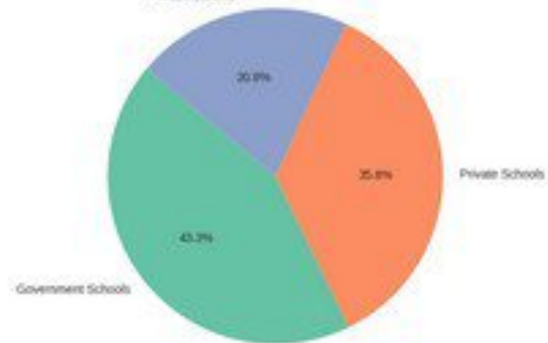
Log:

```
[INFO] 2025-09-07 12:02 - Ingested 1,200 rows
[INFO] 2025-09-07 12:03 - Standardized 4 columns
```

4️⃣ Insights Output

📊 Pie Chart: School Type Distribution in Hyderabad

- **Government Schools**: 520

- **Private Schools**: 430

- **Aided Schools**: 250

School Type Distribution in Hyderabad

Aided Schools

20.9%

Private Schools

35.8%

43.3%

Government Schools

📈 Bar Chart: Average Pass Percentage Across Districts

District	Avg Pass %
Hyderabad	91.2%
Rangareddy	88.5%
Medchal	87.9%
Nalgonda	85.3%
Warangal	84.7%

📈 Visual Summary

- ◆ **Bar Chart: Average Pass Percentage Across Districts**

- **Hyderabad**: 91.2%

- **Rangareddy**: 88.5%

- **Medchal**: 87.9%

- **Nalgonda**: 85.3%

- **Warangal**: 84.7%

Average Pass Percentage Across Districts
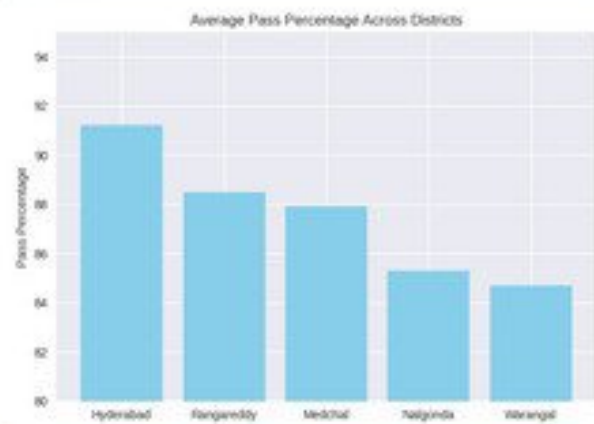
5️⃣ Logs & Documentation Output

Readable Logs Sample:

```
Code  ∧                                              📋 Copy

[INFO] 2025-09-07 12:10 - Analysis complete
[INFO] 2025-09-07 12:11 - Generated insights report: insights_summary.txt
```

Traceability Table:

Stage	File	Description
Raw	`Education_Stats_2022.csv`	Original dataset

6 CLI-Accessible Results

Command	Output
`python run_pipeline.py`	Executes full flow
`python show_insights.py`	Displays summary table
`python generate_plot.py`	Saves pie/bar charts to `/reports`

🔁 Reproducibility Mindset

Component	Status
`run_pipeline.py`	✅
`requirements.txt`	✅
`README.md`	✅
Dockerfile	✅

📂 Findability Map

Folder	Contents
`/data`	Raw and cleaned datasets
`/logs`	Execution logs
`/reports`	Visuals and summaries
`/cli`	CLI scripts

📌 Minimal-Friction Visuals

- Pie chart for school type distribution

- Bar chart for district-wise performance

- Simple tables for metrics