Smart Water Quality Management System

Presented by the School of Engineering and Technology, this project report details the development of AquaDash, a web-based platform for comprehensive water quality monitoring. The system integrates data ingestion, cleansing, visualization, and forecasting to support environmental and public health efforts.

Our team, guided by Dr. Shweta Bansal and Dr. Dilraj, has built a modular, interactive dashboard that empowers users to analyze river and groundwater data with ease and precision.

By - Saatvik Tyagi, Sarvesh Chandran , Suyash , Swaraj Chatterjee

Introduction and Project Scope

Challenges in Water Quality Monitoring

Data fragmentation, static reporting, scalability limits, and complex comparative analyses hinder effective water-quality management.

Project Aim

Develop a Smart Water Quality Management System that unifies data ingestion, cleansing, visualization, and forecasting into an interactive web platform.

Key Features

- Multi-dataset upload and cleaning
- Time-series trend visualization
- Robust forecasting models
- Cross-region comparison

	Description
Drop-in data uploads	Accepts up to **5 CSV/XLS(X)** files at once; automatically shows shape & missing-value counts.
Smart column pruning	Slider flags columns that are mostly *NaN* / zeros, letting you hide noisy features before plotting.

System Overview

Interactive Tabs

River Data, Groundwater Data, and Comparative Analysis tabs provide structured workflows from raw data to forecasting.

Local Processing

All data handling and modeling occur client-side using Python libraries, ensuring privacy and responsiveness.

Visualization & Forecasting

Features include multi-year trend plots, five-year forecasts, and side-by-side dataset comparisons.

Architecture & Technology Stack

Frontend

Streamlit framework for UI, interactive widgets, and tabbed navigation.

Data & Processing

Python modules (river.py, groundwater.py, comparative.py) handle data loading, cleaning, and analysis.

Modeling & Visualization

Scikit-learn models for forecasting; Matplotlib for charts embedded in the app.

Data Management

River Data

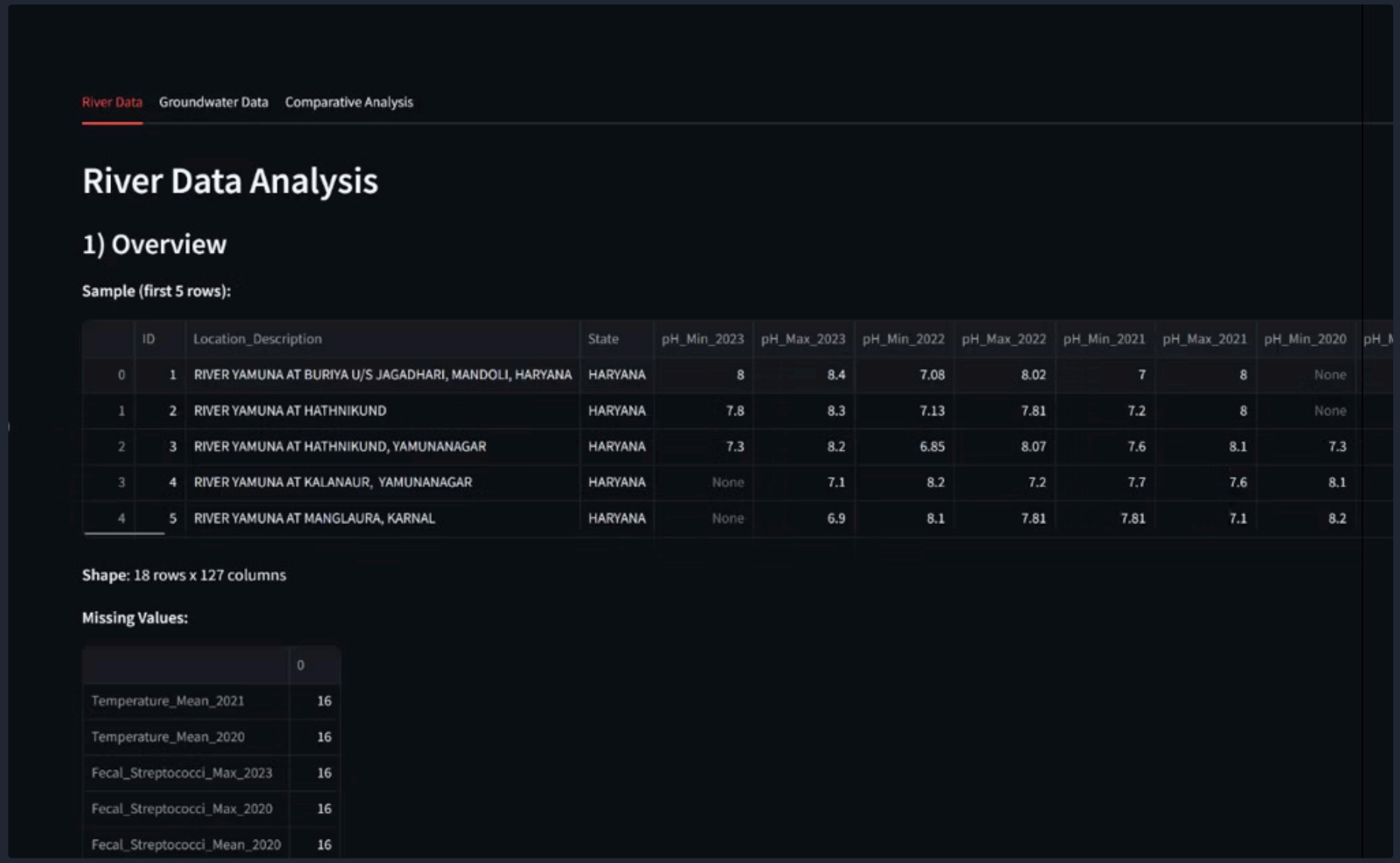
CSV files with min/max yearly parameters; system computes mean values automatically for uniform analysis.

Groundwater Data

Excel workbooks cleaned and standardized; numeric coercion and missing data handling applied.

Automation & Caching

Data loading and transformations are cached to optimize UI responsiveness and avoid redundant computations.



River Data Groundwater Data Comparative Analysis

River Data Analysis

1) Overview

Sample (first 5 rows):

	ID	Location_Description	State	pH_Min_2023	pH_Max_2023	pH_Min_2022	pH_Max_2022	pH_Min_2021	pH_Max_2021	pH_Min_2020	pH_M
0	1	RIVER YAMUNA AT BURIYA U/S JAGADHARI, MANDOLI, HARYANA	HARYANA	8	8.4	7.08	8.02	7	8	None	
1	2	RIVER YAMUNA AT HATHNIKUND	HARYANA	7.8	8.3	7.13	7.81	7.2	8	None	
2	3	RIVER YAMUNA AT HATHNIKUND, YAMUNANAGAR	HARYANA	7.3	8.2	6.85	8.07	7.6	8.1	7.3	
3	4	RIVER YAMUNA AT KALANAUR, YAMUNANAGAR	HARYANA	None	7.1	8.2	7.2	7.7	7.6	8.1	
4	5	RIVER YAMUNA AT MANGLAURA, KARNAL	HARYANA	None	6.9	8.1	7.81	7.81	7.1	8.2	

Shape: 18 rows x 127 columns

Missing Values:

	0
Temperature_Mean_2021	16
Temperature_Mean_2020	16
Fecal_Streptococci_Max_2023	16
Fecal_Streptococci_Max_2020	16
Fecal_Streptococci_Mean_2020	16

River Data Analysis

User Interface & Navigation

Tabbed Navigation

Three main tabs with numbered sub-sections guide users through data overview, analysis, and forecasting.

Interactive Controls

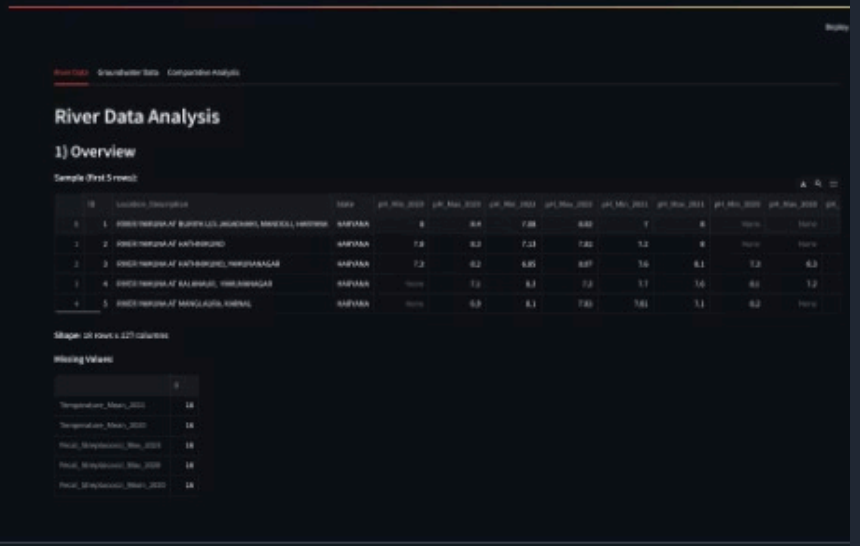
Drop-downs, sliders, and radio buttons allow dynamic parameter selection and real-time updates of tables and plots.

Intuitive Layout

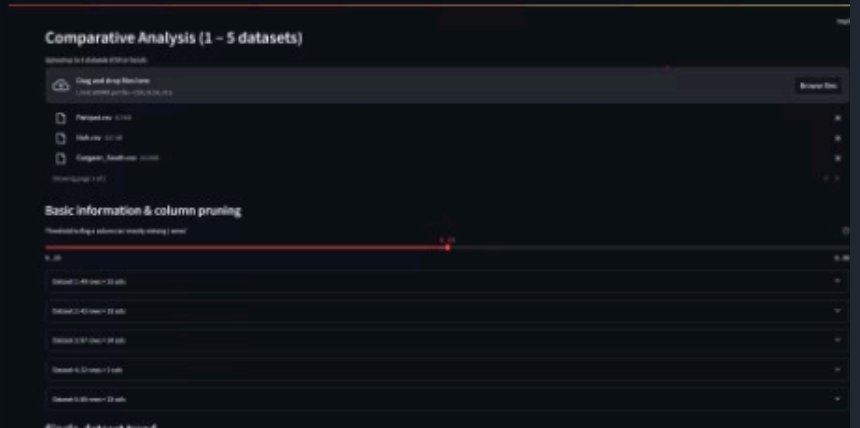
Clear headers, inline widgets, and tooltips enhance usability and support efficient data exploration.

Demo

River overview



Upload & prune



Tab-by-Tab Functionality

1

River Data Tab

Overview, potability classification, yearly parameter analysis, and 5-year forecasting with selectable models.

2

Groundwater Data Tab

Similar workflow as river tab, including data preview, potability checks, parameter trends, and forecasting.

3

Comparative Analysis Tab

Multi-file upload, column pruning, single and multi-dataset trend visualization with line, scatter, or bar charts.

Water Quality Analysis & Statistical Insights

Regional Comparisons

Custom uploads enable side-by-side analysis of chloride, BOD, ammonium, and other parameters across regions.

Key Statistics

- pH near neutral in all regions
- Low BOD indicating minimal organic pollution
- Ammonia levels effectively zero
- Higher chloride in Panipat suggests regional salinity concerns

Conclusions, Implications & Future Work

Conclusions

The system offers a comprehensive, user-friendly platform for water quality monitoring, revealing generally safe conditions with regional variations.

Practical Implications

Enables local authorities to visualize, forecast, and compare water data without specialized software, supporting data-driven decisions.

Future Enhancements

Plans include automated alerts, GIS integration, advanced time-series models, and expanded data sources for broader environmental monitoring.

Conclusions

Region	pH (avg)	BOD (avg mg/L)	NH_3 (avg mg/L)	Chloride (avg mg/L)
Dharuhera	7.51	0.00	0.00	111.31
Gurgaon North	7.23	0.57	0.00	91.67
Gurgaon South	7.13	1.11	0.00	101.89
Panipat	7.79	0.40	0.00	159.82

References

AquaDash README (project documentation)file- **README.md**

main.py – Application entry script defining tabs and loading datafile- **main.py**

river.py – River data module (mean-column creation, potability logic, forecasting routines)file- **river.py**

groundwater.py – Groundwater data module (data prep, potability logic, forecasting)file- **groundwater.py**

comparative.py – Multi-dataset comparison logic (upload, pruning, trend and multi-series plots)file- **comparative.py**