Visualizing Progress: Data-Driven Insights and Key Performance Indicators for the Yamuna Canal Rejuvenation Project

Shagun Talwar

Abstract

This report presents a comprehensive data-driven analysis of the Yamuna Canal Rejuvenation Project, utilizing insights visualized through a dedicated Power BI dashboard. The dashboard integrates monitoring data from 260 stations, tracking critical Key Performance Indicators (KPIs) related to water quality, specifically Dissolved Oxygen (DO), Biological Oxygen Demand (BOD), and Total Coliform counts.

The analysis reveals a complex picture of environmental recovery: while the long-term trend in average DO levels shows a positive trajectory, the dashboard concurrently highlights a critical and rapidly increasing challenge posed by rising BOD and Total Coliform concentrations. Geographic analysis identifies major urban centers, including Delhi, Uttar Pradesh, and Uttarakhand, as primary pollution hotspots responsible for contributing to the high maximum BOD values (≈ 116.00 mg/L) and significant bacterial loads (average maximum Total Coliform ≈ 681.37K MPN/100 mL).

The findings underscore the necessity for sustained enforcement of pollution control measures and immediate, targeted interventions in high-risk zones to address organic and pathogenic contamination sources. In conclusion, the successful completion of the rejuvenation project hinges on bridging the gap between improved physical infrastructure and persistent source-point pollution (see References [1]–[3], [8], [17]–[19]).

1 Introduction

Water quality management is a critical component of environmental sustainability, public health protection, and socio-economic development. The chosen domain for this report—environmental monitoring with a focus on water resources—plays a central role in evaluating the ecological condition of rivers, canals, and other freshwater systems. In the context of India, where surface water bodies often face challenges such as pollution, industrial discharge, urban sewage, and agricultural runoff, data-driven assessment becomes essential for informed decision-making (see References [1]–[3], [10], [17]–[19]).

The Yamuna Canal Rejuvenation Project specifically aims to restore the deteriorating water quality of the Yamuna canal system, which is vital for irrigation, local ecosystems, and nearby communities. Monitoring key water-quality parameters such as Dissolved Oxygen (DO), Biological Oxygen Demand (BOD), Total Coliform levels, and pH variations helps determine the canal's ecological health and compliance with national water standards (References [1], [4], [8], [9]). These indicators provide actionable insights into pollution sources, seasonal variations, and overall environmental performance.

In real-world decision-making, reliable water quality analytics support government agencies, environmental researchers, and policy-makers in prioritizing interventions, identifying pollution hotspots, planning resource allocation, and measuring progress over time. By converting raw environmental data into meaningful insights, this report and the accompanying Power BI dashboard contribute to evidence-based strategies for sustainable freshwater management and long-term ecological restoration.

Key Water-Quality Indicators

Parameter	Typical Standard Range	Observation from Yamuna Canal Dashboard
Dissolved Oxygen (DO)	$\geq$ 5 mg/L for Class-B surface water (bathing-quality, CPCB)	Several stations show minimum DO values close to 0 mg/L, indicating severe stress on aquatic life and poor ecological health.
Biological Oxygen Demand (BOD)	$\leq$ 3 mg/L for bathing-class waters	Maximum BOD $\approx$ 116 mg/L and many readings > 20 mg/L indicate heavy organic pollution and high oxygen demand.
Total Coliform	500–5,000 MPN/100 mL (depending on use)	Average maximum values around 681,000 MPN/100 mL reveal extreme microbial contamination and high health risk.
pH	6.5–8.5 (desirable range)	Most stations remain within this range; pH is comparatively less critical than BOD and Coliform in this dataset.

Table 1: Summary of key water-quality indicators and observed status for the Yamuna Canal (based on CPCB and APHA guidelines; see References [1], [4], [8], [9]).

2 Need of the Analysis

Effective environmental management relies on accurate, interpretable data. For the Yamuna Canal system there is a critical gap between the raw water-quality measurements collected by various agencies and actionable insights required for remediation. Although multiple parameters (Dissolved Oxygen, Biological Oxygen Demand, Total Coliform, pH, etc.) are recorded at many monitoring stations, the data often remains fragmented, inconsistent, or under-utilized due to its volume and complexity.

This analysis addresses the following problems:

- **Data gap and fragmentation:** Measurements are collected across different locations and time periods; combining these for consistent analysis requires cleaning and harmonization.
- **Lack of actionable KPIs:** Raw values alone do not indicate compliance or priority—derived indicators and thresholds are needed to support decisions.
- **Difficulty in detecting hotspots and trends:** Without temporal and spatial aggregation, identifying persistent pollution hotspots and seasonal patterns is cumbersome.

Why this data matters: Water quality directly affects public health, agricultural productivity, aquatic biodiversity, and local livelihoods. Timely and accurate interpretation of the data enables authorities to:

- Prioritize remediation and enforcement actions at high-risk sites,
- Allocate resources (manpower, treatment infrastructure) effectively,
- Track the effectiveness of interventions over time, and
- Communicate progress to stakeholders and the public with evidence.

How visualization and DAX help stakeholders: Interactive visualizations (maps, trend-lines, treemaps and gauges) reduce cognitive load by revealing patterns and anomalies at-a-glance. In Power BI, DAX expressions enable calculated KPIs (e.g., `Percent_Records_High_BOD`, moving averages, exceedance counts) and dynamically filterable measures, which:

- Automate aggregation and normalization across stations and years,
- Produce comparable metrics for state-wise and station-wise ranking,
- Allow drill-down analysis (from state $\rightarrow$ district $\rightarrow$ station),
- Support scenario testing and threshold-based alerts.

Collectively, this analysis transforms fragmented water-quality data into actionable insights, enabling evidence-based policy, targeted remediation, and continuous monitoring of the Yamuna Canal rejuvenation efforts.

3 Motivation Behind Choosing the Report / Area

Environmental sustainability and water resource management have become major global priorities, especially in rapidly developing countries such as India. The motivation for selecting the Yamuna Canal Rejuvenation Project stems from the need to understand and address the growing issue of water pollution, which directly impacts public health, agriculture, biodiversity, and long-term ecological balance. As students and researchers, this area connects strongly with academic interests in data analytics, environmental monitoring, and data-driven decision support systems.

This topic is also closely aligned with key United Nations Sustainable Development Goals (SDGs), particularly:

- **SDG 6: Clean Water and Sanitation**
- **SDG 13: Climate Action**
- **SDG 14: Life Below Water**

Analyzing water quality trends helps promote responsible consumption of natural resources, supports climate resilience strategies, and contributes to restoring aquatic ecosystems. By converting raw water-quality measurements into meaningful visual insights, the study strengthens evidence-based decision-making and supports nationwide rejuvenation initiatives (see References [1]–[3], [10], [11]).

At the national level, India is actively implementing large-scale programs such as the *Namami Gange Mission*, the *National Water Quality Monitoring Programme (NWQMP)*, and various state-led canal and river restoration schemes [1–3]. The Yamuna River, being one of the most severely polluted rivers in the country, requires constant scientific monitoring and policy-driven interventions. This dashboard serves as a tool for visualizing the current status of water quality and identifying priority areas requiring immediate attention.

The inspiration for this project was drawn from publicly available datasets provided by the Central Pollution Control Board (CPCB) and state agencies, which offer real-time and historical water quality data (References [1], [4]–[7]). These datasets, along with periodic national water quality reports, highlighted significant variations in pollution levels across different states and monitoring stations, making this a meaningful and impactful area for analysis.

4 Objective of the Analysis

The primary objective of this analysis is to convert raw water-quality data from the Yamuna Canal system into meaningful insights that support environmental monitoring and informed policy-making. The analysis aims to evaluate pollution levels, detect long-term trends, and enable data-driven decision-making through interactive dashboards. The key objectives include:

1. To analyze year-wise and station-wise variations in water-quality indicators such as Dissolved Oxygen (DO), Biological Oxygen Demand (BOD), pH, and Total Coliform levels.
2. To identify patterns, anomalies, and pollution hotspots where water quality consistently deviates from national standards, enabling targeted environmental intervention.
3. To derive actionable insights through Power BI dashboards, including visualizations, DAX-based KPIs, and interactive filters for enhanced interpretation and comparison.
4. To recommend potential environmental improvements and monitoring strategies based on the observed trends, patterns, and performance gaps across different states and monitoring stations.

5 Report Details

The dataset used for this analysis was obtained from official PDF reports published by the Central Pollution Control Board (CPCB). The CPCB provides annual and real-time water-quality information for rivers and canal systems across India. The PDF reports contain detailed tables with recorded values for key parameters such as Dissolved Oxygen (DO), Biological Oxygen Demand (BOD), pH, and Total Coliform levels across multiple monitoring stations (References [1], [4], [5]).

The source PDF reports were downloaded from the CPCB Water Quality Data portal:

- **CPCB Water Quality Data (PDF Reports):**
  https://cpcb.nic.in/nwmp-data/

After downloading the PDFs, relevant tables were extracted using the Power BI PDF connector. Power BI automatically detected tabular structures within the documents, enabling the selection of required tables such as:

- Station-wise water-quality measurements,
- State-wise monitoring summaries,
- Year-wise DO, BOD, pH, and Total Coliform readings,
- Reference permissible limits.

These tables were cleaned, standardized, and merged in Power Query to ensure consistency and analytical readiness. The extracted data formed the foundation for creating interactive dashboards, DAX-based KPIs, and trend visualizations (References [13]–[16]).

Table - 4: Water Quality of river Yamuna for the year 2023

Station Code	Monitoring Location	State Name	Temperature (°C)		Dissolved Oxygen (mg/L)		pH		Conductivity (µmho/cm)		BOD (mg/L)		Nitrate N (mg/L)		Fecal Coliform (MPN/100ml)		Total Coliform (MPN/100ml)		Fecal Streptococci (MPN/100ml)	
			Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
Primary Water Quality Criteria notified under The E (P) Rules, 1986					> 5.0 mg/L		6.5 - 8.5				< 3.0 mg/L				< 2500 MPN/ 100 Ml				< 500 MPN/ 100 Ml	
1492	RIVER YAMUNA AT YAMUNOTRI	UTTARAKHAND	4.0	4.0	11.2	11.2	7.8	7.8	31	31	1.0	1.0	0.32	0.32	2	2	2	2	2	2
1493	RIVER YAMUNA AT HANUMANCHATTI	UTTARAKHAND	10.0	10.0	10.8	10.8	7.6	7.6	56	56	1.0	1.0	0.33	0.33	2	2	2	2	2	2
1494	RIVER YAMUNA AT U/S OF LAKHWAR DAM	UTTARAKHAND	9.0	21.0	9.4	10.8	7.7	8.4	105	214	1.0	1.0	0.32	0.60	2	2	2	2	2	2
30025	RIVER YAMUNA AT DAK PATHER (UTTARAKHAND)	UTTARAKHAND	17.0	21.0	9.4	10.8	8.2	8.4	128	286	1.0	9.0	0.32	0.47	2	4000	13	17000	2	230
1490	RIVER YAMUNA AT U/S DAKPATHER, DEHRADUN	UTTARAKHAND	10.0	22.0	9.2	10.6	7.0	8.1	72	186	1.0	1.0	0.32	0.37	41	84	70	94	6	14
30027	RIVER YAMUNA AT PONTA SAHIB (H.P)	HIMACHAL PRADESH	20.0	24.3	9.0	10.5	8.0	8.4	134	518	1.0	9.0	0.49	1.32	7	7800	170	33000	4	2400
1553	RIVER YAMUNA U/S PAONTA SAHIB	HIMACHAL PRADESH	16.0	23.0	6.3	8.4	7.0	8.4	264	702	1.0	2.5	0.54	0.90	2	94	49	220	2	2
1554	RIVER YAMUNA D/S PAONTA SAHIB	HIMACHAL PRADESH	16.0	22.0	6.3	8.4	7.7	8.3	243	666	1.6	3.6	0.54	0.90	2	220	94	540	2	2
4439	RIVER YAMUNA U/S SUN PHARMACEUTICALS	HIMACHAL PRADESH	15.0	23.0	6.4	8.5	7.4	8.3	257	587	1.3	2.9	0.54	0.92	40	240	170	350	2	2
4440	RIVER YAMUNA D/S SUN PHARMACEUTICALS	HIMACHAL PRADESH	16.0	24.0	6.4	8.4	7.7	8.4	301	550	1.2	3.5	0.00	0.92	46	350	280	920	2	2
30028	RIVER YAMUNA AT BURIYA U/S JAGADHARI, MANDOLI, HARYANA	HARYANA	17.5	27.5	7.9	8.8	8.0	8.4	211	358	1.0	15.0	0.53	0.89	110	4900	790	92000	2	490
30035	RIVER YAMUNA AT HATHNIKUND	HARYANA	14.0	20.0	7.6	9.8	7.8	8.3	208	390	1.0	6.0	0.79	1.07	14	1300	79	13000	5	330
1117	RIVER YAMUNA AT HATHNIKUND, YAMUNANAGAR	HARYANA			6.8	7.8	7.3	8.2	181	953	1.0	4.2	0.30	0.30	2	200	200	900	2	90
1496	RIVER YAMUNA AT KALANAUR, YAMUNA NAGAR	HARYANA			6.0	7.9	7.1	8.2	210	954	1.0	6.2	0.30	0.30	100	400	300	1200	2	200
4914	RIVER YAMUNA AT MANGLAURA, KARNAL	HARYANA			5.2	7.8	6.9	8.1	238	864	1.2	30.0	0.00	0.30	100	1500	400	4200	2	580
30032	RIVER YAMUNA AT SONAULI ROAD, SHAMLI BORDER , PANIPAT (HARYANA)	HARYANA			6.8	10.2	7.7	8.6	221	413	1.0	6.0	0.52	1.66	45	4900	700	11000	2	1100
10004	RIVER YAMUNA AT KHOJKIPUR PANIPAT	HARYANA			3.8	7.4	6.9	7.8	189	989	1.2	6.8	0.30	0.30	100	600	600	1700	2	200
30029	RIVER YAMUNA AT SONIPAT, BAGHPAT ROAD(HARYANA)	HARYANA			6.3	8.6	7.7	8.7	299	620	1.0	8.0	0.66	1.99	92	2300	270	4900	11	1400
1119	RIVER YAMUNA AT SONEPAT	HARYANA			6.0	7.8	6.1	8.1	195	350	1.0	6.2	0.30	0.32	2	1000	200	10000	2	200
10005	RIVER YAMUNA AT PALLA,SONEPAT	HARYANA			5.2	7.8	5.5	8.2	235	1080	1.8	6.8	0.00	0.30	100	1100	400	3400	2	200
10006	RIVER YAMUNA AT PALLA	DELHI	16.0	16.0	11.2	11.2	7.5	7.5	1172	1172	13.0	13.0	2.00	2.00	1400	1400	2200	2200	33	33
30026	RIVER YAMUNA AT PALLA (DELHI)	DELHI			6.2	9.2	7.4	8.2	346	1265	1.0	16.0	0.32	2.31	200	2600000	54000	7000000	230	17000
1120	RIVER YAMUNA AT PALLA, DELHI	DELHI			6.3	11.3	7.1	7.5			1.0	3.0	0.00	0.00	180	560				

Figure 1: Sample CPCB PDF table used as a source for station-wise water-quality data.

6 Data Processing Steps

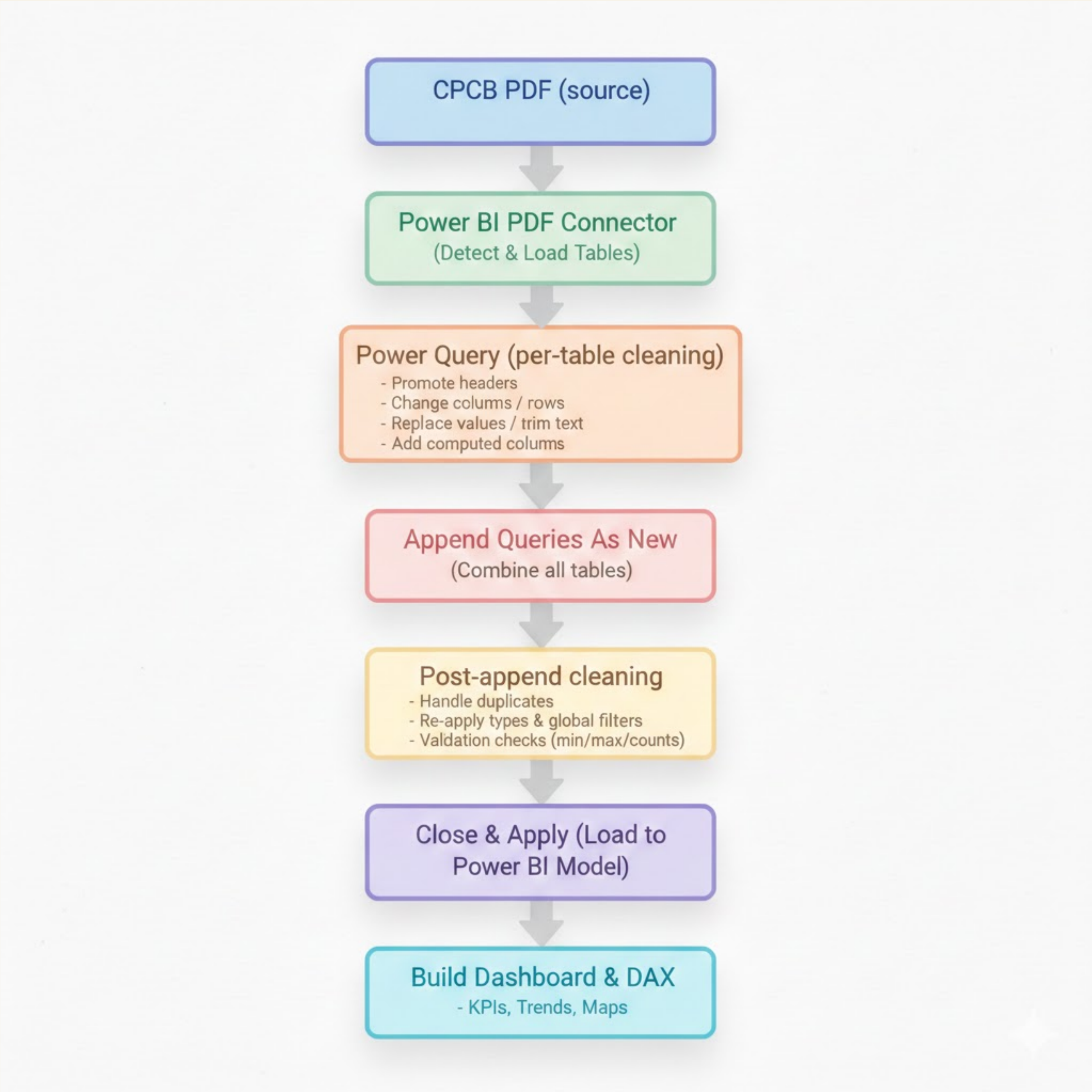


Figure 2: Data processing flowchart for the Yamuna Canal Rejuvenation dashboard.

The data preprocessing workflow included:

- Importing multiple PDF tables into Power BI,
- Cleaning column names, handling missing values, and unifying units,
- Creating calculated columns (e.g., year, station category),
- Building relationships and star-schema style modeling,
- Defining DAX measures for KPIs (DO, BOD, Total Coliform, pH, compliance percentages).

7 Analytical Graphs & Interpretation

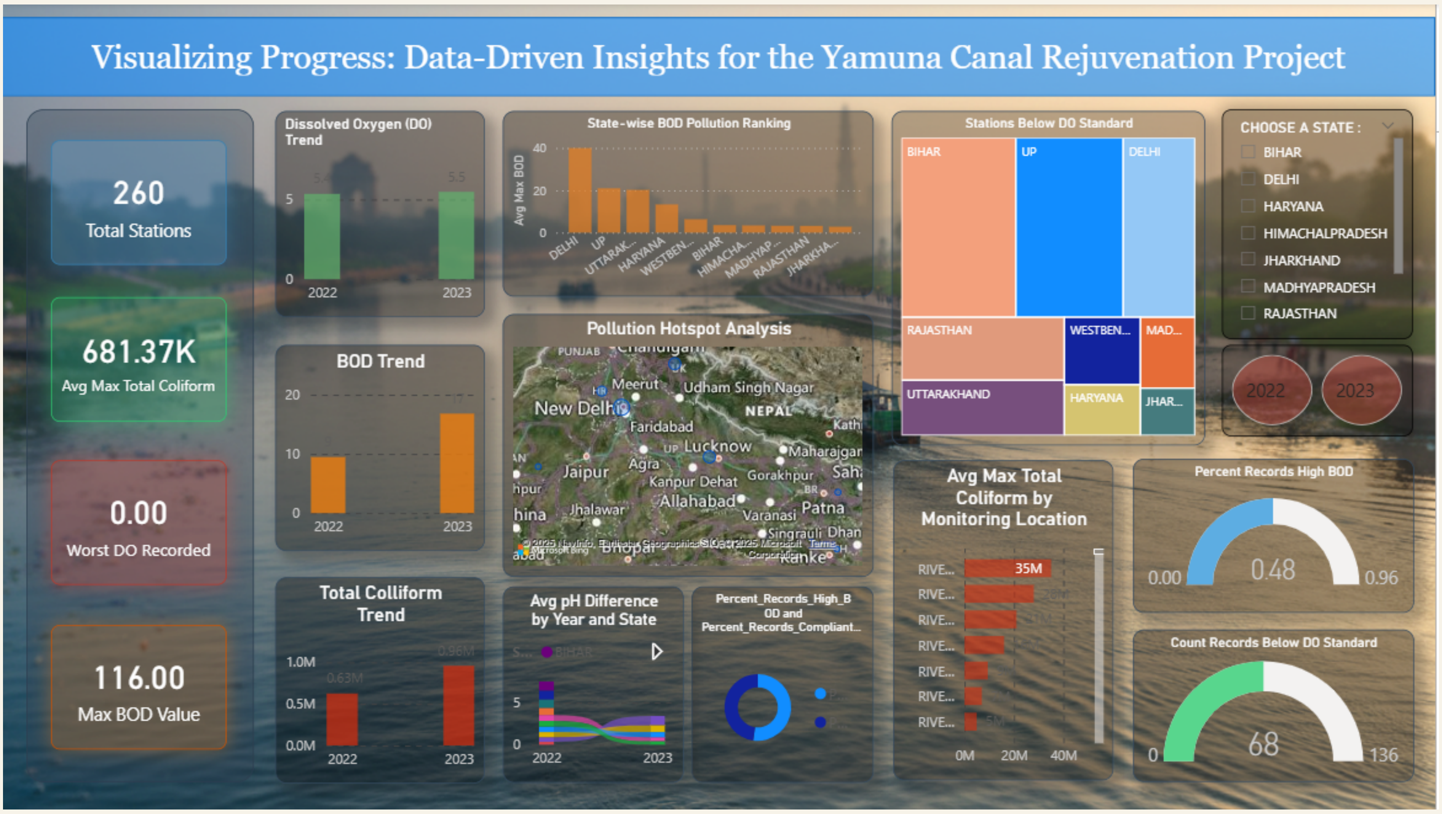
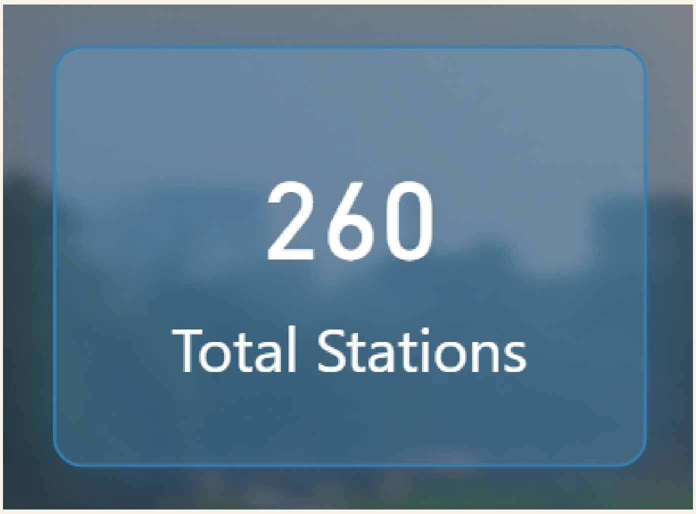


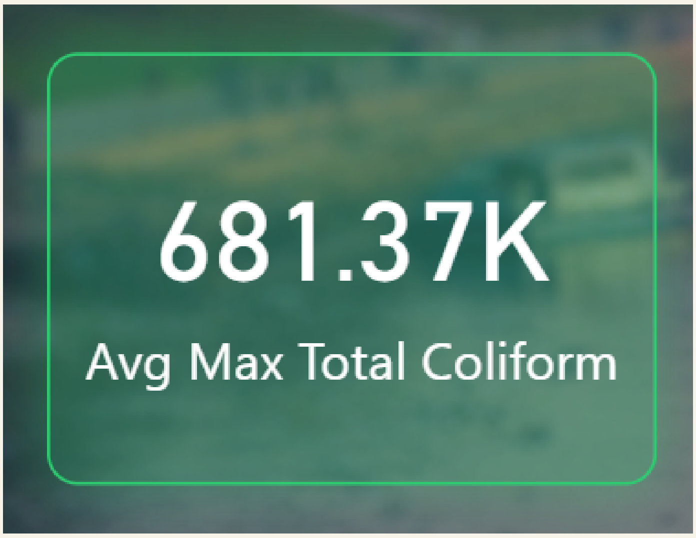
Figure 3: Enter Caption

Figure 4: Complete Power BI dashboard — Yamuna Canal Rejuvenation overview.

The figures below present key visuals from the dashboard. Images are arranged two per row for comparative analysis.



(a) KPI — Total monitoring stations (260)



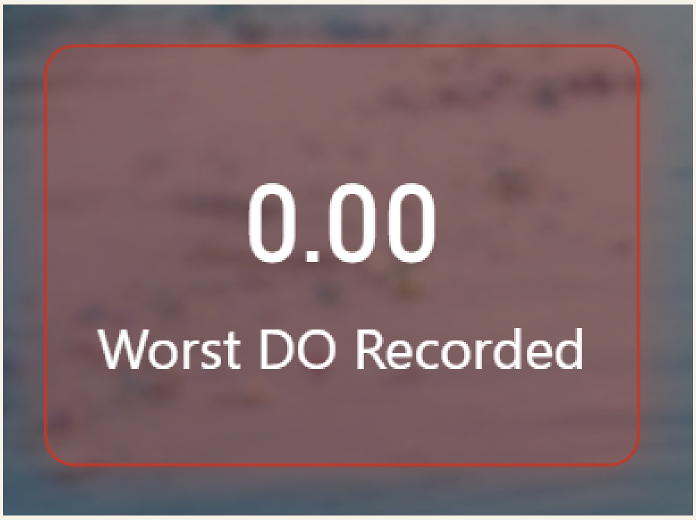
(b) KPI — Average maximum Total Coliform (e.g., 681.37K)

Visual 1 — Total Stations (KPI Card)

This KPI indicates the total number of monitoring stations (260) included in the analysis. A dense monitoring network improves spatial resolution and helps identify localized pollution sources across states and canal stretches.

Visual 2 — Average Maximum Total Coliform (KPI Card)

Total Coliform is a key microbial contamination indicator. An average maximum value in the range of $\sim$ 681K MPN/100 mL indicates severe fecal contamination at multiple locations, driven by domestic sewage, open drains, and untreated wastewater entering the canal.



(a) KPI — Worst Dissolved Oxygen recorded (e.g., 0.00 mg/L)

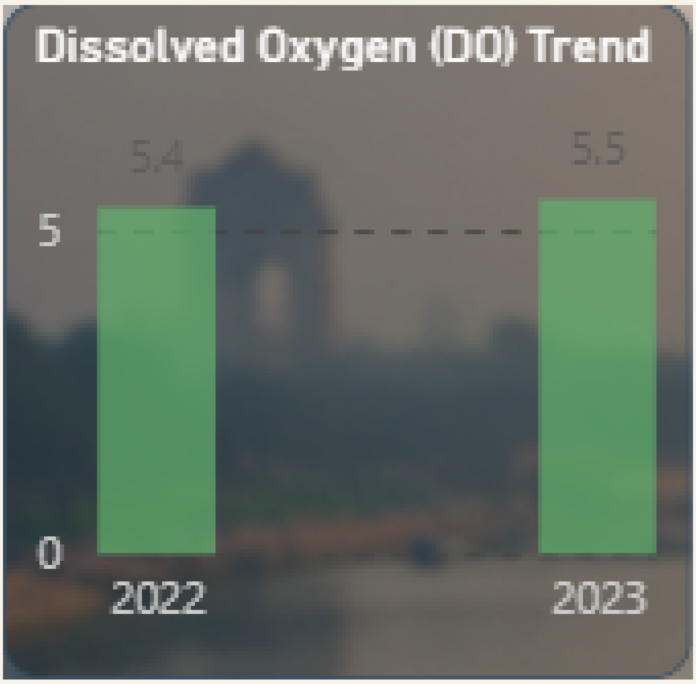
(b) KPI — Maximum BOD value (e.g., 116 mg/L)

Visual 3 — Worst Dissolved Oxygen (KPI Card)

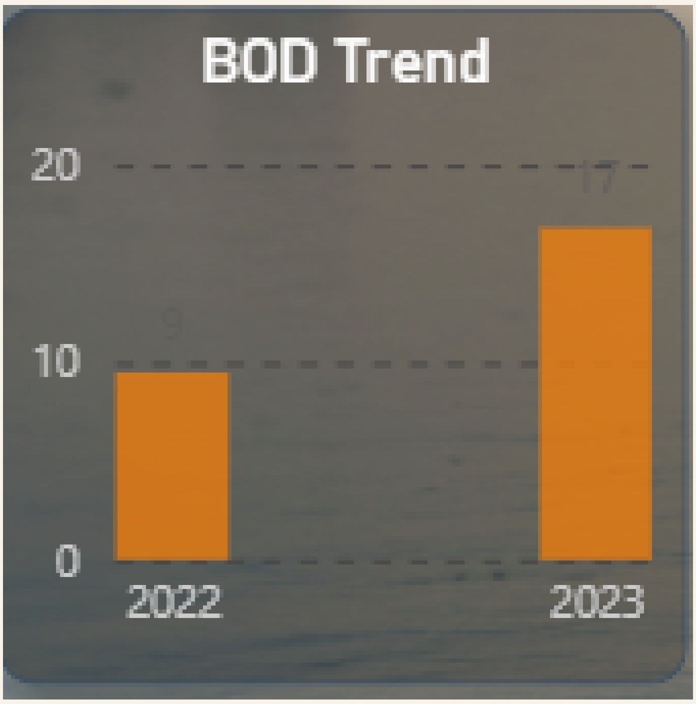
A DO value of 0 mg/L indicates extreme oxygen depletion, making the water unsuitable for aquatic life. Such conditions are commonly associated with high organic pollution, stagnant stretches, and sustained pollutant loading.

Visual 4 — Maximum BOD Value (KPI Card)

For good-quality water, BOD should typically be below 3 mg/L. A peak value of 116 mg/L is alarming, showing heavy organic loads from industrial discharge and untreated sewage. High BOD usually correlates with low DO, explaining the degraded ecological health observed in hotspots.



(a) Column chart — DO trend across 2022 and 2023



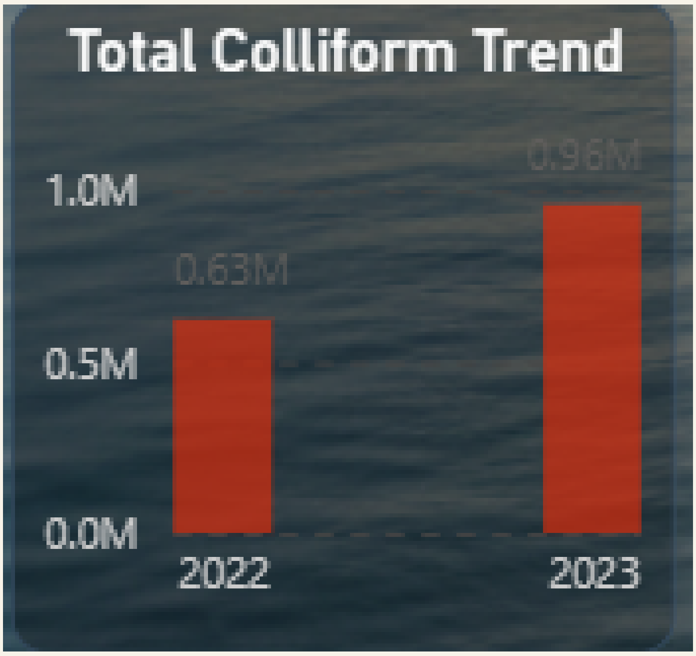
(b) Column chart — BOD trend across 2022 and 2023

Visual 5 — Dissolved Oxygen (DO) Trend

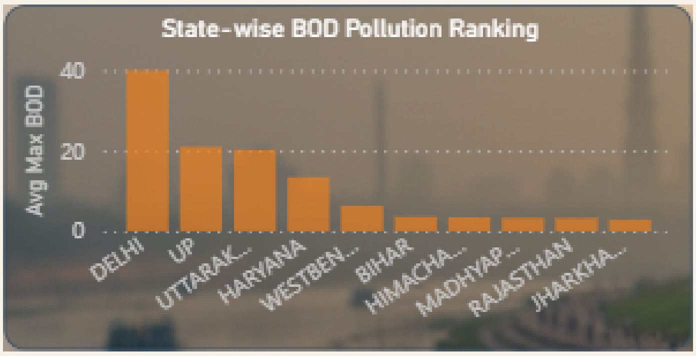
- A slight rise in average DO in 2023 suggests minor improvement in oxygen conditions.
- However, DO values remain far below the recommended 5.0 mg/L threshold for Class-B waters, so the canal remains unfit for aquatic life in many stretches.
- Small improvements may reflect seasonal effects, increased flow, or partial treatment upgrades.

Visual 6 — BOD Trend

- BOD increases from ~ 10 mg/L to ~ 20 mg/L between 2022 and 2023.
- This indicates worsening organic pollution over time.
- The pattern is consistent with continued inflow of untreated industrial and domestic wastewater.



(a) Column chart — Total Coliform trend



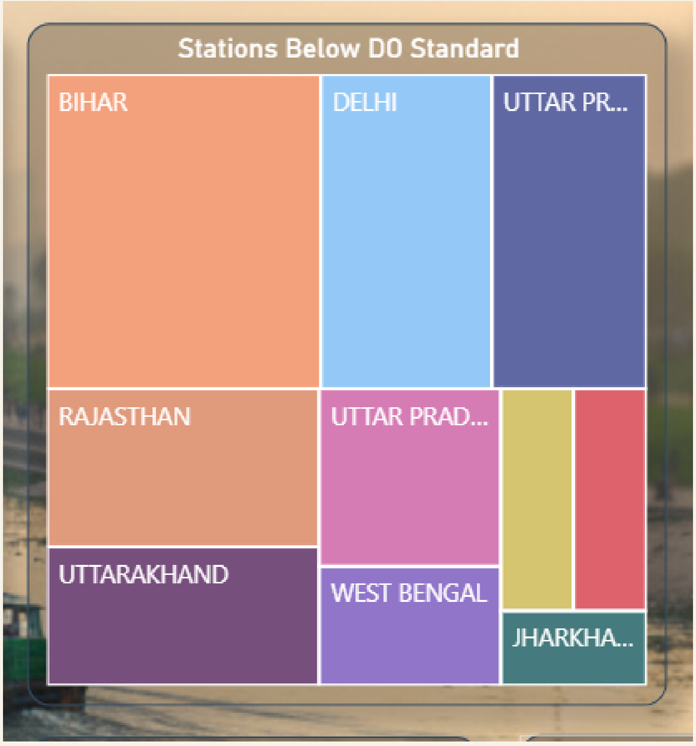
(b) State-wise average BOD ranking (bar chart)

Visual 7 — Total Coliform Trend

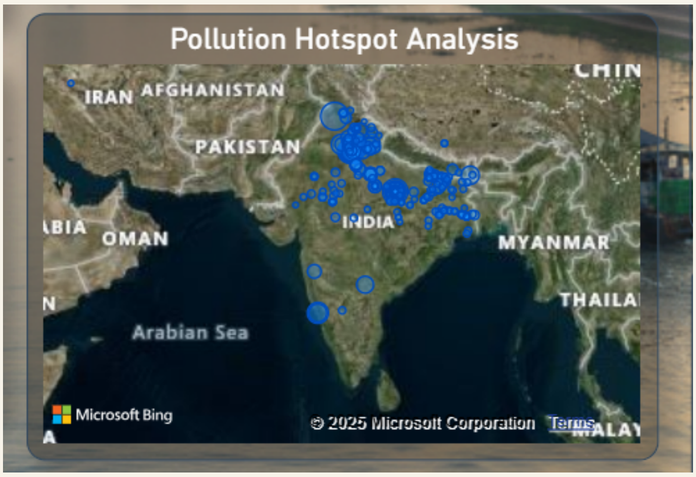
- Total Coliform trends from $\sim 600{,}000$ to nearly 1,000,000 MPN/100 mL.
- This sharp increase reflects escalating microbial contamination, a major public health concern.
- It indicates limited progress in sanitation coverage and wastewater management.

Visual 8 — State-wise BOD Pollution Ranking

- States like Delhi, Haryana, and Uttar Pradesh show the highest average BOD values.
- Pollution intensity generally decreases upstream towards Himalayan regions, where flows are cleaner and settlements sparser.
- These high-BOD states have dense populations, large sewage loads, and significant industrial activity.



(a) Treemap — stations below DO standard



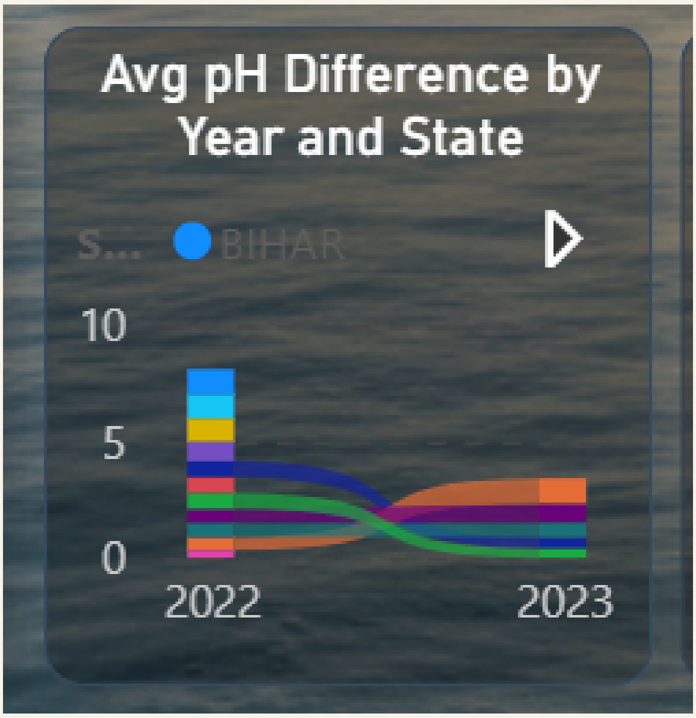
(b) Geospatial map — pollution hotspots across monitoring stations

Visual 9 — Stations Below DO Standard (Treemap)

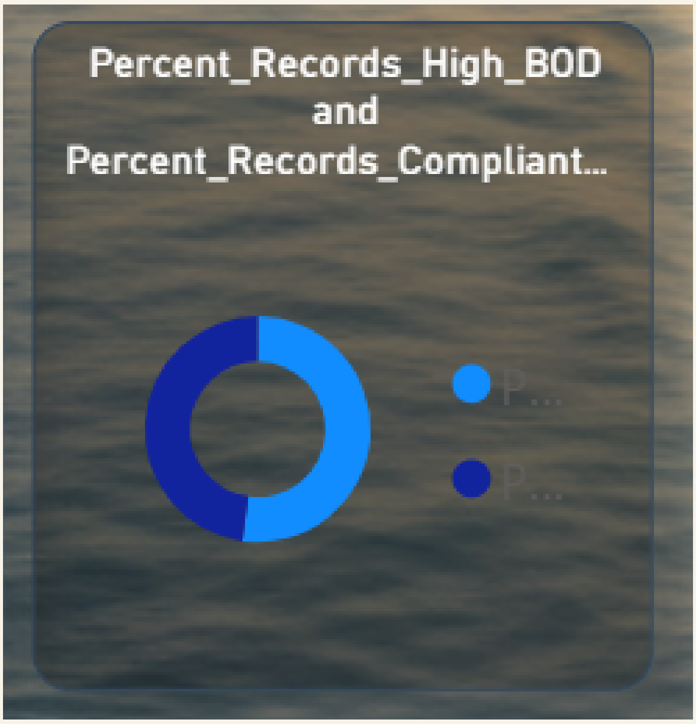
- States such as Bihar, Delhi, Rajasthan, and Uttarakhand contain large blocks in the treemap, indicating many stations with DO below standard.
- This reveals chronic low-oxygen zones where remediation should be prioritized.
- The treemap helps quickly rank states by failure count or severity.

Visual 10 — Pollution Hotspot Analysis (Map)

- Each point represents a monitoring station; darker or larger markers indicate higher pollution intensity.
- Hotspots cluster in the northern plains, especially around Delhi, Haryana, and parts of Uttar Pradesh.
- Upstream Himalayan stations appear comparatively cleaner, underscoring the role of urbanization and industrial corridors.



(a) Multi-line chart — average pH difference by year and state



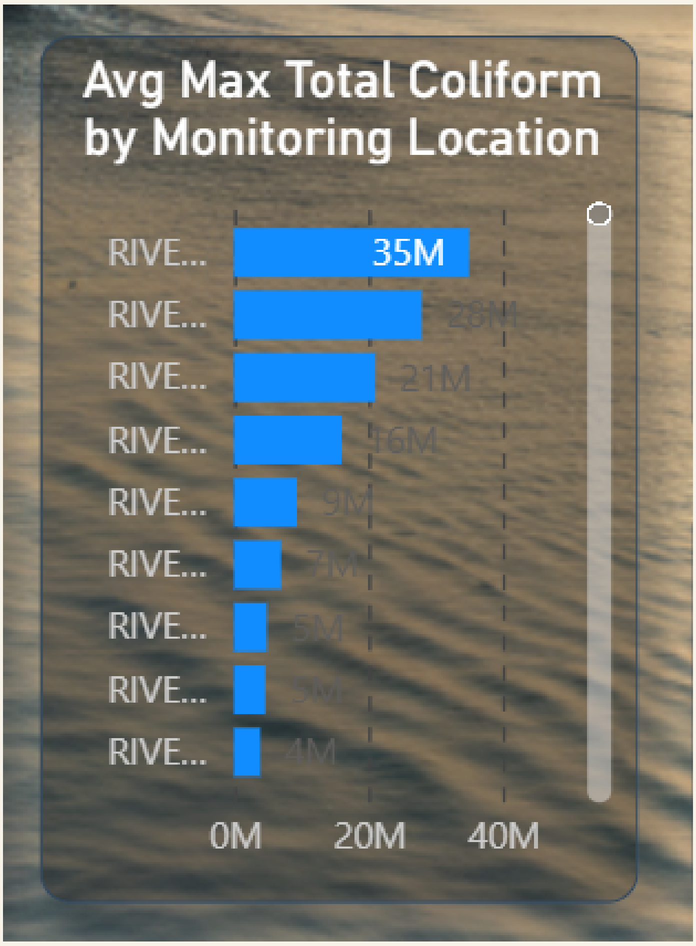
(b) Donut chart — percentage of BOD-compliant vs high-BOD records

Visual 11 — Average pH Difference by Year and State

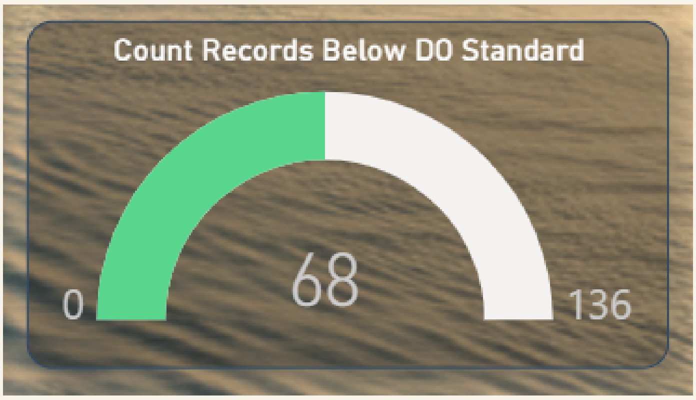
- pH mostly remains within the acceptable 6.5–8.5 range.
- Slight fluctuations across states are normal and do not suggest widespread chemical pollution.
- pH is therefore less critical compared to BOD and Coliform for this dataset.

Visual 12 — Percent High BOD and Compliant Records

- The donut chart indicates that roughly 48% of records are compliant while about 52% have high BOD.
- More than half of the measurements exceed recommended limits, signaling poor waste-management and treatment coverage.
- This metric is useful as a single, easy-to-understand KPI for decision-makers.



(a) Horizontal bar chart — average maximum Total Coliform by monitoring location



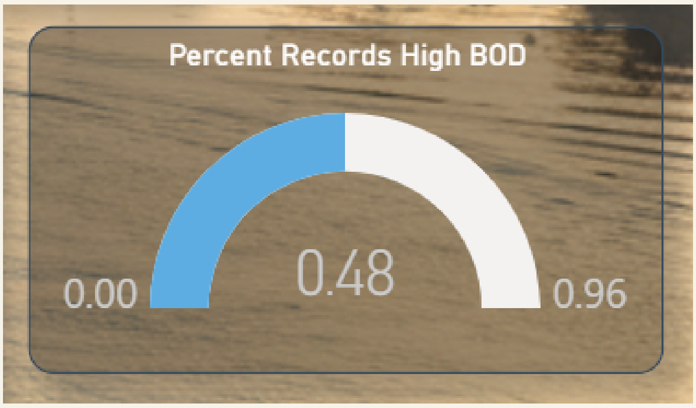
(b) Gauge — count of records below DO standard

Visual 13 — Average Maximum Total Coliform by Location

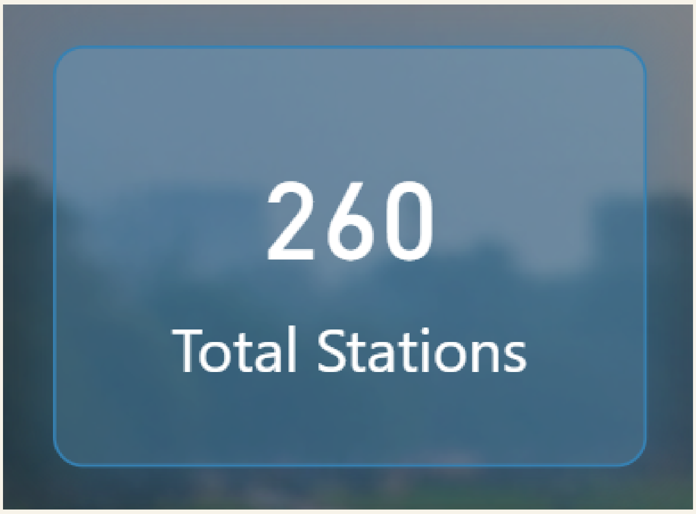
- Stations such as Palla, Panipat, and Hathnikund show extremely high coliform counts.
- These locations likely act as direct discharge points for untreated or partially treated wastewater.
- The chart highlights urgent need for decentralized wastewater treatment units near these hotspots.

Visual 14 — Count of Records Below DO Standard

- The gauge summarizes how many readings had DO values below the selected standard (e.g., 68 out of 136).
- More than 50% of samples falling below standard indicates systemic oxygen depletion.
- This is consistent with high organic loads and insufficiently treated sewage.



(a) Gauge — percent of records in high-BOD category



(b) KPI context — total stations supporting gauge analysis

Visual 15 — Percent Records with High BOD

- This gauge focuses specifically on the share of records exceeding the BOD standard.
- Nearly half or more of the dataset shows high BOD, indicating widespread failure to meet environmental quality norms.
- When combined with the total stations KPI, this supports station-wise risk prioritization.

Overall Interpretation and Policy Recommendations

- **Summary of findings:** Combined visuals show widespread organic and microbial pollution (high BOD and Total Coliform) and significant oxygen depletion (low DO), especially in the northern plains (Delhi, Haryana, Uttar Pradesh).
- **Short-term recommendations:** Deploy targeted wastewater treatment at identified hotspots, enforce industrial effluent regulations, and implement emergency aeration or flow-restoration measures at critically low-DO sites.
- **Long-term actions:** Upgrade sewage-collection and treatment infrastructure, strengthen continuous monitoring and public dashboards, and invest in decentralized sanitation in peri-urban and rural fringe regions.
- **Monitoring KPIs:** Track DO compliance percentage, state-wise average BOD, and a top-10 list of highest Coliform locations as recurring KPIs to measure progress over time (see References [1]–[3], [10], [17]–[19]).

8 Socio-Economic Impact of Canal Rejuvenation

Socio-economic impact of improved water quality:

- Improved water quality reduces water-borne diseases and lowers the healthcare burden for nearby communities.
- Cleaner canal water supports safer irrigation, leading to better crop yield, improved soil health, and higher farmer income.
- Enhanced canal conditions boost local livelihoods, especially for fishing communities, small vendors, and nearby settlements dependent on the waterway.
- Reduction in pollution strengthens industrial compliance, preventing regulatory fines and encouraging sustainable operations.
- Rejuvenation efforts create employment opportunities in monitoring, treatment, waste management, and community-based initiatives (e.g., Yamuna Mitra groups).
- Overall improvement leads to a healthier ecosystem, better quality of life, and long-term economic growth along the canal corridor.

These socio-economic co-benefits highlight why environmental restoration is not only an ecological priority but also a critical development and public-health investment.

9 Proposed Solution: Yamuna Mitra Initiative

To address the problem of pollution and contamination in the Yamunanagar Canal, an integrated, multi-level solution is proposed based on the "Vocal for Local" principle. Under the "Namami Yamuna: Yamunanagar Canal Rejuvenation" project, this will be implemented as the **Yamuna Mitra Initiative**.

How the Solution Works

The solution operates across three major pillars:

1. Local Technical Interventions
2. Community Participation
3. Sustainable Economic Incentives

Local Technical Interventions

Decentralized waste treatment:
Low-cost, locally managed natural treatment systems (such as reed-bed systems and small-scale bioremediation tanks) can be installed at sewage inlets and drain outlets. These systems partially treat wastewater before it enters the canal, reducing BOD and pathogen loads.
Industrial waste monitoring:
A locally developed "Green Audit" and smart monitoring system can be implemented for industries. Trained local technicians (Yamuna Mitras) perform regular water-quality checks for industrial effluents, with data integrated back into dashboards.
Waste collection infrastructure:
Smart dustbins capable of sending fill-level alerts can be installed at high-pollution and sensitive locations along the canal. Scheduled waste collection will be enforced to prevent litter accumulation and plastic leakage.

Community Participation Model (Yamuna Mitra)

Formation of local volunteer groups:
Residents, youth, and retired individuals living near the canal are organized as **Yamuna Mitras**. They serve as the first line of monitoring and awareness.
Daily monitoring and awareness:
Yamuna Mitras conduct cleanliness drives, discourage open dumping, raise awareness about water pollution, and report illegal connections or irregular industrial discharges to authorities.
School engagement programs:
"River Guardian Clubs" in schools and colleges sensitise students to ecology and water quality, encouraging citizen science projects and periodic canal clean-up events.

Sustainable Economic Incentives

Waste-to-wealth initiative:
Collected plastic and recyclable waste is purchased by local Self-Help Groups (SHGs). SHGs convert waste into handicrafts and other local products, creating income opportunities and supporting the "Vocal for Local" principle (see References [10], [11]).
Cleanliness rating system:
Villages and colonies along the canal receive a cleanliness rating. High-performing areas are eligible for administrative grants, recognition, or small tax concessions, thereby rewarding collective action.

Beneficiaries of the Solution

- **Local residents:** Cleaner surroundings, fewer disease outbreaks, and improved living conditions.
- **Farmers:** Access to safer irrigation water, improved crop quality, and higher yields.

- **Local economy:** Employment generation through SHGs, volunteers, and service providers involved in monitoring, treatment, and waste management.
- **Industries:** Clear pollution-control guidelines and real-time monitoring to prevent penalties and disputes.
- **Environment:** Improved aquatic ecosystems, better groundwater recharge quality, and overall ecological rejuvenation of the Yamunanagar region.

10 Further Scope of Work

While the current analysis provides a comprehensive understanding of pollution levels in the Yamuna Canal, several opportunities exist to expand the scope and analytical depth of this project in the future.

Integration of Additional Data Sources

Future work can incorporate more diverse and real-time datasets including:

- IoT-based continuous water-quality sensor readings,
- Satellite imagery for vegetation index, land-use change and flood mapping,
- Meteorological data such as rainfall, temperature and river flow,
- Industrial discharge records and sewage treatment plant performance logs.

Integrating such datasets will allow for more accurate correlation between pollution triggers and environmental conditions (References [1]–[3], [10]).

Predictive and Forecasting Models

Machine learning and statistical techniques can be introduced to forecast future pollution levels and risk zones. Models such as ARIMA, Random Forest, LSTM, and time-series regression can be used to predict BOD, DO, and Coliform trends, seasonal variations, and hotspot formation.

Python-Based Advanced Analytics

Python scripts can enhance analytical capability through:

- Outlier detection using statistical tests,
- Correlation and dependency modeling,
- Clustering of stations based on pollution profiles,
- Automated anomaly detection for real-time alerts.

These insights can be integrated into Power BI or Tableau dashboards through embedded Python visuals or API connectors.

Enhanced Dashboard Platforms

The solution can be extended to multiple analytical tools such as:

- Tableau for advanced geospatial visualizations,
- IBM Cognos Analytics for enterprise-level reporting,
- Streamlit or Dash for Python-based web applications.

Automated Data Pipeline and Cloud Integration

Future versions can include:

- Automated ETL pipelines (Azure Data Factory, AWS Glue),
- Cloud storage solutions for scalable data handling,
- Scheduled refreshes for near real-time reporting.

Mobile App and Citizen Participation

A mobile-friendly dashboard or app can enable:

- Citizen reporting of pollution incidents,
- Live cleanliness scores for canal stretches,

- Alerts for sudden drops in water quality.

Policy Simulation and Scenario Modeling

Simulation tools can be developed to test "what-if" scenarios such as:

- Impact of installing new sewage treatment plants,
- Reduction in industrial discharge load,
- Effects of seasonal variations on canal pollution.

Such models help policymakers evaluate and compare the effectiveness of various intervention strategies.

11 Conclusion

The Yamuna Canal Rejuvenation dashboard provides a comprehensive, data-driven assessment of the environmental condition of the Yamunanagar canal system. By integrating CPCB water-quality data with Power BI analytics, the project highlights critical pollution patterns and identifies the stations and states requiring immediate attention. The findings show that Dissolved Oxygen (DO) levels remain significantly below acceptable standards at many monitoring locations, while Biological Oxygen Demand (BOD) and Total Coliform concentrations demonstrate widespread non-compliance. These observations reflect severe organic and microbial pollution, driven largely by untreated sewage, industrial discharge, and inadequate waste-management systems (References [1], [4]–[7], [17]–[19]).

State-wise and location-wise analyses reveal that regions with dense populations and industrial clusters contribute most to the degradation of canal water quality. Visual tools such as KPIs, line charts, bar charts, treemaps, geographic maps, and gauge indicators enabled clear interpretation of trends, hotspots, and compliance levels. These insights can guide policymakers, environmental agencies, and local administrations in prioritizing interventions and designing targeted rejuvenation strategies.

The project also reinforces essential learning outcomes in data collection, extraction, cleaning, transformation, and modeling. The use of Power BI and DAX enhanced the analytical depth and interactivity of the dashboard, demonstrating the importance of modern data-analysis tools in environmental monitoring. The overall learning experience helped develop skills in end-to-end data processing, analytical storytelling, and domain-driven visualization.

In conclusion, this project lays a strong foundation for future enhancements such as predictive modeling, integration of additional data sources, IoT-driven real-time monitoring, and expanded analytical capabilities. The insights generated contribute meaningfully to ongoing efforts toward restoring the ecological health of the Yamuna Canal and advancing India's broader environmental and sustainable development objectives.

12 Annexure

12.1 A. Raw Data Snapshots

The CPCB PDF reports used as sources contain columns such as station name, state, year, DO, BOD, pH, and Total Coliform. Figure 1 (earlier) shows an example of a raw table extracted before cleaning and transformation.

Downloadable PDF Sources

PDF 1: CPCB Water Quality Report (Year 2022)
Click here to open

PDF 2: CPCB Water Quality Report (Year 2023)
Click here to open

12.2 B. DAX Queries Used in Power BI

To support KPI calculations, trend analysis, and compliance evaluation, several DAX measures were developed:

```
-- Query 1: Total Stations Total Stations = DISTINCTCOUNT('Data - Yamuna'[Station Code]) -- Query
```

```
-- Query 1:  Total Stations Total Stations = DISTINCTCOUNT('Data - Yamuna'[Station Code])
-- Query 2:  Avg Max BOD Avg Max BOD = AVERAGE('Data - Yamuna'[Max BOD (Mg/L)])
-- Query 3:  Avg Min DO Avg Min DO = AVERAGE('Data - Yamuna'[min Dissolved Oxygen (mg/L)])
-- Query 4:  Avg Max Total Coliform Avg Max Total Coliform = AVERAGE('Data - Yamuna'[max Total
Coliform (MPN/100mL)])
-- Query 5:  Avg Min Temp Avg Min Temp = AVERAGE('Data - Yamuna'[Min Temperature (C)])
-- Query 6:  Avg pH Difference Avg pH Difference = AVERAGEX( 'Data - Yamuna', 'Data - Yamuna'[max
ph] - 'Data - Yamuna'[min pH] )
-- Query 7:  Count_Records_Below_DO_Standard Count_Records_Below_DO_Standard =
CALCULATE(COUNTROWS('Data - Yamuna'), FILTER('Data - Yamuna', 'Data -
Yamuna'[min Dissolved Oxygen (mg/L)] < 4))
-- Query 8:  Max BOD Value Max BOD Value = MAX('Data - Yamuna'[Max BOD (Mg/L)])
-- Query 9:  Overall Min DO Overall Min DO = MIN('Data - Yamuna'[min Dissolved Oxygen (mg/L)])
-- Query 10:  Percent_Records_High_BOD Percent_Records_High_BOD = DIVIDE(CALCULATE(COUNTROWS('Data -
Yamuna'), FILTER('Data - Yamuna', 'Data - Yamuna'[Max BOD (Mg/L)] > 3)), COUNTROWS('Data -
Yamuna'))
-- Query 11:  Percent_Records_Compliant_BOD Percent_Records_Compliant_BOD = 1 - [Percent_Records_High_BOD]
```

13 References

1. Central Pollution Control Board (CPCB). *Water Quality Data Portal.* Ministry of Environment, Forest and Climate Change (MoEFCC), Government of India.
   https://cpcb.nic.in/water-quality-data/
2. Central Pollution Control Board (CPCB). *National Water Quality Monitoring Programme (NWQMP).*
   https://cpcb.nic.in/wqm/
3. National Mission for Clean Ganga (NMCG). *Ganga Basin Water Quality Reports.*
   https://nmcg.nic.in/
4. National River Conservation Directorate (NRCD). *Water Quality Standards and Environmental Regulations.* MoEFCC, Government of India.
5. Delhi Pollution Control Committee (DPCC). *River and Canal Water Quality Reports.* Government of Delhi.
6. Haryana State Pollution Control Board (HSPCB). *Water Quality Monitoring Data.* Government of Haryana.
7. Uttar Pradesh Pollution Control Board (UPPCB). *Water Quality Reports.* Government of Uttar Pradesh.
8. APHA. *Standard Methods for the Examination of Water and Wastewater.* American Public Health Association.
9. Chapman, D. (UNESCO). *Water Quality Assessment: A Guide to the Use of Biota, Sediments and Water in Environmental Monitoring.*
10. World Bank. (2021). *Quality Unknown: The Invisible Water Crisis.*
11. NITI Aayog. *Composite Water Management Index.* Government of India.
12. IIT Kanpur. *Ganga Basin River Water Quality Research Publications.*
13. Microsoft Documentation. *DAX Reference Guide.*
    https://learn.microsoft.com/power-bi/dax/
14. SQLBI (Marco Russo & Alberto Ferrari). *DAX Patterns and Best Practices.*
    https://www.sqlbi.com/
15. Microsoft Documentation. *Power Query M Language Specification.*
16. Microsoft Documentation. *Power BI Desktop — Data Modeling Concepts.*
17. Sharma, S., & Bhatt, P. *Assessment of Water Quality Using DO–BOD Indicators in Indian Rivers. Environmental Monitoring and Assessment.*

18. Kumar, A., & Singh, R. *Urbanization Impact on River Water Quality in India. Journal of Hydrology and Wastewater Research.*

19. Mishra, S. *Impact of Sewage and Industrial Discharge on Yamuna River. Indian Journal of Environmental Protection.*