

# Bike Station Analytics (2022-2025)

## 1. Project Overview and Objective

This project focuses on analysing bike station operations using a provided dataset. It involves cleaning and transforming raw station-level data, followed by creating an interactive Power BI dashboard to visualize key performance indicators and operational metrics.

The objective is to:

- **Analyse bike station capacity and availability:** Study the distribution of docking capacity and available bikes across different contracts, identifying areas with high utilization or excess capacity.
  - **Evaluate station performance and status:** Assess how many stations are operational (OPEN) versus CLOSED, and measure the concentration of stations and bike stands across leading contracts.
  - **Understand geographic distribution and service coverage:** Examine how stations are clustered geographically using an interactive map, and analyse the rollout of key services like banking/card payments to identify coverage gaps.
  - **Compare contract-level performance:** Compare bike availability against total capacity for each contract, highlighting which contracts show a healthy balance and which may be under- or over-supplied.
  - **Visualize insights with interactive dashboards:** Build interactive Power BI reports that allow users to filter by station status (OPEN/CLOSED) and bonus offerings to explore trends, service levels, and geographic patterns in an intuitive way.
- 

## 2. Data Sources

- **Source Description and Timeline:** Data sourced from the Bike Station dataset containing station-level details such as capacity, availability, status, banking, and location, with records spanning from 2022 to 2025.
  - **Domain:** Transportation (Public Bike Sharing).
- 

## 3. Problem Statement

- To analyse bike station availability and capacity from 2022 to 2025 to identify key utilization trends.

- To assess station performance across different contracts to uncover areas with over-supply or under-supply of bikes.
- To study the geographic distribution of stations and banking service coverage to identify gaps in accessibility and customer convenience.
- To evaluate operational status (OPEN vs CLOSED) and the impact of bonus features on station usage patterns.
- To develop an interactive dashboard enabling exploration of station dynamics for strategic fleet management and infrastructure planning.

#### 4. Attribute (Column /Features) Details:

Attribute Name	Data Type	Description
Station No.	Integer	Unique identifier for each bike station (Station ID).
Contract Name	Categorical	Name of the city or contract area (e.g., Lyon, Valence).
Station Name	Categorical	Full name of the bike station.
Address	Categorical	Street address or location description of the station.
Latitude	Numeric (Decimal)	Geographic latitude coordinate of the station.
Longitude	Numeric (Decimal)	Geographic longitude coordinate of the station.
Banking	Boolean / Categorical	Indicates if the station supports bank card payments (TRUE/FALSE).
Bonus	Boolean / Categorical	Indicates if the station is a bonus station offering incentives (TRUE/FALSE).
Bike Stands	Numeric	Total capacity (number of docking slots) at the station.
Available Bike Stands	Numeric	Number of empty slots currently available for parking.
Available Bikes	Numeric	Number of bikes currently available for rent.
Status	Categorical	Operational status of the station (OPEN or CLOSED).

Last Update (Date)	Date	Year derived from the last update timestamp, used for trend analysis.
-----------------------	------	--

---

## 5. Tools & Technologies

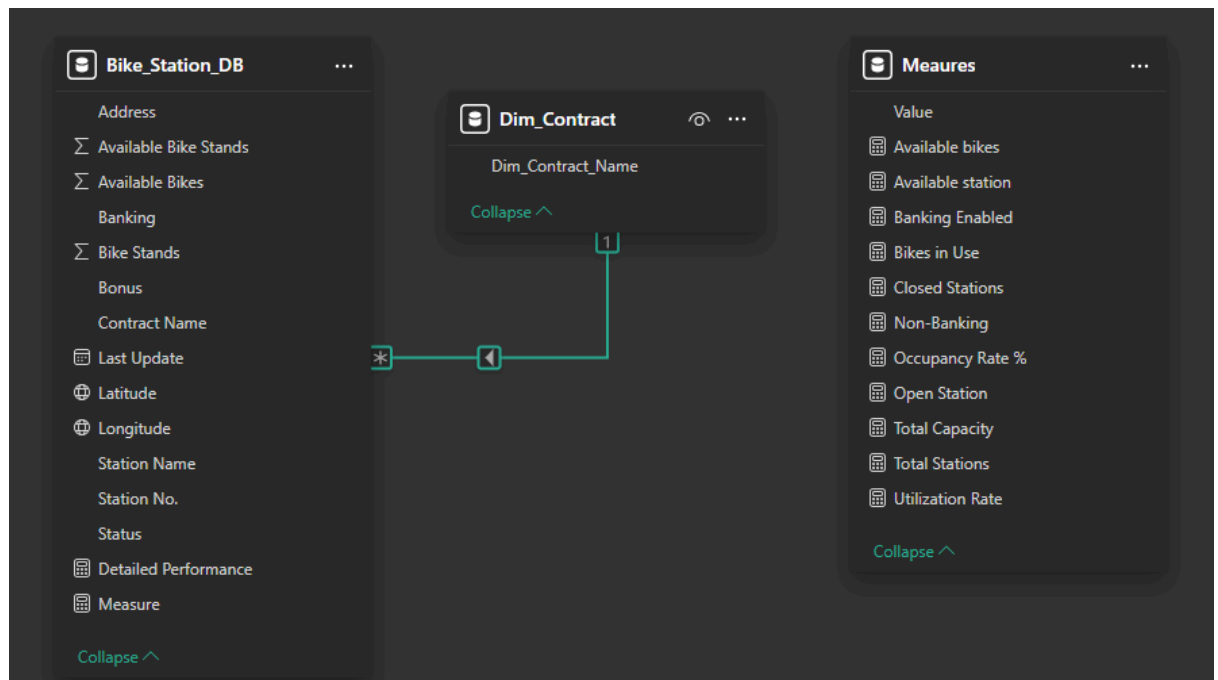
- Power Query: Data cleaning and transformation.
  - Power BI: Data modelling, DAX calculations, visualization, and interactive dashboard creation.
- 

## 6. Data Pre-Processing (Power Query)

- **Data Collection:** Imported the bike station dataset containing station details, capacity, availability, and service attributes.
  - **Data Cleaning:** Removed duplicates, handled missing or null values, corrected data types, and renamed columns (e.g., Station No., Contract Name) for clarity.
  - **Data Transformation:** Split Position into separate Latitude and Longitude columns, derived Last Update (Year), and created structured fields for status, banking, and bonus.
  - **Filtering & Sorting:** Filtered out invalid or inactive records where needed and sorted stations and contracts to support clear reporting and ranking.
  - **Data Integration:** Shaped the cleaned data into a single, analysis-ready table for use in Power BI modelling and visualization.
- 

## 7. Data Modelling and DAX (Power BI)

- **Data Model:** Implemented a simple star-style model with one main fact table containing all bike station details (station attributes, capacity, availability, banking, bonus, status, and year). A separate Measures table was created to centralize key DAX calculations such as capacity, availability, utilization, and banking metrics, making the model easier to manage and reuse across visuals.



- **Calculated Columns & DAX Measures :**

Total Capacity – sum of all bike stands across stations.

Available Bikes – sum of available bikes for the current filter context.

Available Stations – count of stations with at least one available bike stand.

Total Stations – distinct count of station numbers.

Open Station – count of stations where status = OPEN.

Closed Stations – count of stations where status = CLOSED.

Banking Enabled – count of stations with banking = TRUE.

Non-Banking – count of stations with banking = FALSE.

Bikes in Use – difference between total capacity and available bike stands.

Occupancy Rate % – ratio of bikes in use to total capacity, expressed as a percentage.

Utilization Rate – proportion of available bikes relative to total capacity or total bikes, used to assess how effectively stations are being used.

## 8. Analysis and Visualizations (Power BI)

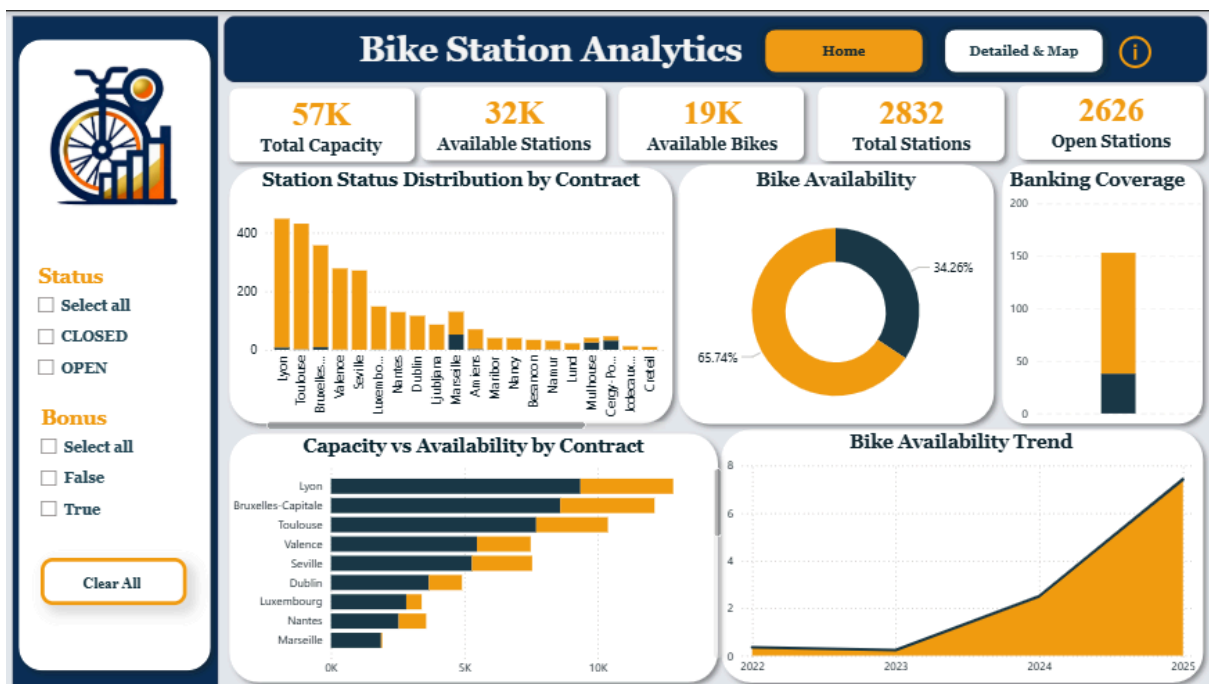
### Dashboard Features:

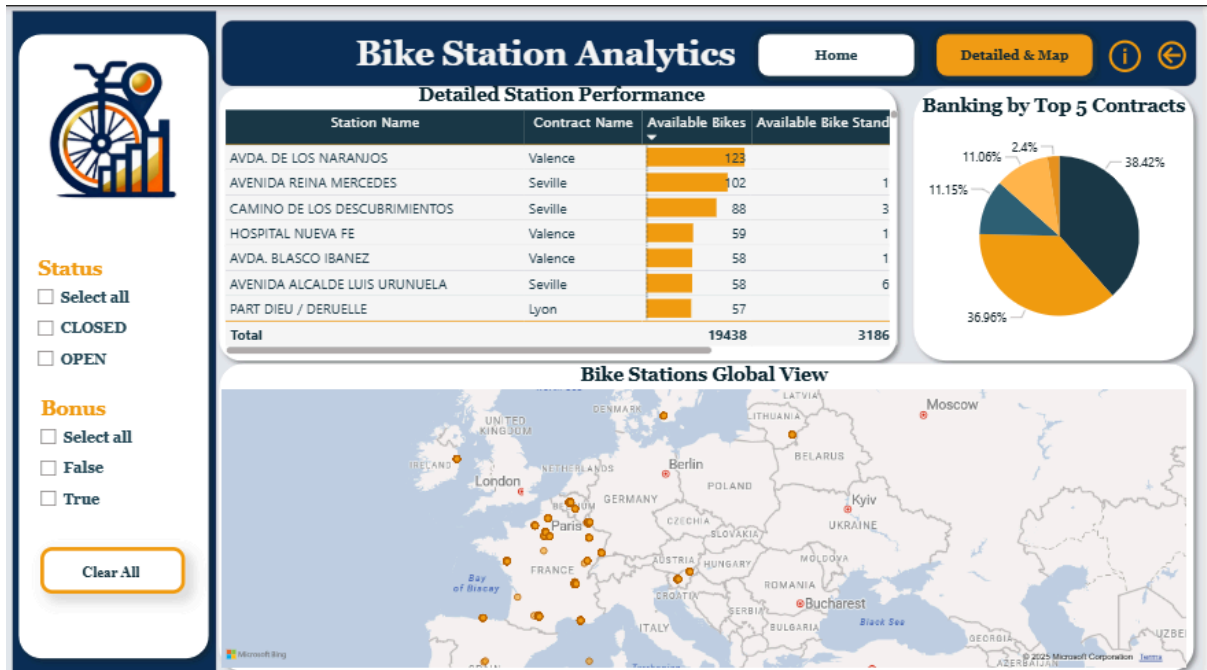
Developed a set of visuals aligned with the problem statement, including bar charts, stacked bars, donut charts, KPI cards, tables, and a map to clearly communicate bike station insights. Designed the report with interactive slicers (Status, Bonus) and navigation buttons (Home, Detailed & Map, Info) for smooth, granular

exploration of the data. Used a consistent layout, clear chart titles, and custom icons to improve readability and user experience across the two report pages.

### Highlights:

- **Capacity & Availability:** Station Status Distribution and Capacity vs Availability by Contract charts show how stations and docking capacity are distributed across contracts and how many bikes are available versus total stands.
- **Service Coverage:** Banking Coverage column chart and Banking by Top 5 Contracts donut highlight how banking-enabled stations are spread across contracts and which contracts contribute most to banking capacity.
- **Trend Analysis:** Bike Availability Trend line/area chart tracks changes in bike availability over years, helping identify growth and shifts in supply.
- **Station-Level Details:** Detailed Station Performance table surfaces top stations by available bikes and stands, while the Bike Stations Global View map shows where stations are located geographically for quick spatial context.





## 9. Insights & Conclusions

### Key Findings

- Descriptive:** Capacity and station counts are heavily concentrated in a few contracts such as Lyon, Valence, Bruxelles Capitale, Toulouse, and Seville, while many smaller contracts operate with modest capacity. Bike availability is generally healthy, with roughly two-thirds of docks either filled or bikes available, and banking services are dominated by the top five contracts.

- **Diagnostic:** Capacity vs Availability visuals show that some large contracts run close to full occupancy while others have many empty stands, indicating imbalances in where bikes are placed. Banking coverage gaps in several contracts suggest that payment convenience may be limiting usage in those areas.
- **Predictive:** The strong upward trend in bike availability from 2023 to 2025 points to continued expansion, which could lead to under-used capacity in lower-demand contracts if rebalancing is not performed. Stations with persistently high spare capacity are likely candidates for resizing or relocation.
- **Prescriptive:** Operations teams should regularly rebalance bikes from low-utilization contracts and stations toward high-demand ones, and prioritize enabling banking at busy non-banking stations. Reviewing bonus stations and underperforming locations for targeted incentives, marketing, or infrastructure changes can further improve overall performance.

### Performance Insights

- High-capacity contracts contribute most of the total bikes and stands, creating clear “hub” cities that drive overall performance.
- Some contracts show strong utilization with limited spare stands, while others exhibit excess capacity, highlighting opportunities to optimize placement.
- Top stations in Valence and Seville carry substantial available bikes and stands, acting as key hubs but also potential sources of redeploy able capacity.
- Banking-enabled capacity is concentrated in a few contracts, suggesting that expanding card-payment support elsewhere could unlock additional demand.
- The interactive dashboards support continuous monitoring of capacity, availability, and service coverage, helping stakeholders make data-driven decisions on planning and operations.

---

## 10. Conclusions

The use of Power Query and Power BI enabled a complete analysis of bike station capacity, availability, and service features across multiple contracts. The project highlights how capacity and stations are concentrated in a few key contracts, where banking coverage is strong or limited, and which locations show signs of over- or under-utilization. By transforming raw station data into interactive KPIs, charts, and maps, the Bike Station Analytics dashboard gives stakeholders a clear, visual way to monitor performance, rebalance bikes and capacity, and plan future service improvements.

---