

|  |
| --- |
| **Data Analytics engineering Mentoring program AU25** |



Contents

[1 Business Description 3](#_Toc62212630)

[1.1 Business background 3](#_Toc62212631)

[1.2 Problems. Current Situation 3](#_Toc62212632)

[1.3 The benefits of implementing a database. Project Vision 3](#_Toc62212633)

[2 Model description 3](#_Toc62212634)

[2.1 Definitions & Acronyms 3](#_Toc62212635)

[2.2 Logical Scheme 4](#_Toc62212636)

[2.3 Objects 4](#_Toc62212637)

# 

# Business Description

## Business background

This database supports the operations of a metropolitan subway system, including trains, stations, ticketing, infrastructure, and scheduling. Its objective is to centralize all operational data, enabling efficient planning of routes, sales, maintenance, and passenger services.

## Problems. Current Situation

Before database integration, information was scattered across different tools and spreadsheets. This caused duplication of data, inconsistent train schedules, difficulties tracking maintenance, and slow access to reliable operational reports.

## the Benefits of implementing a database. Project Vision

The database provides a single source of truth for metro operations. It improves data consistency, automates scheduling, ensures accurate ticketing, supports planning and maintenance, and increases overall operational reliability and customer satisfaction.

# Model description

## Definitions & Acronyms

**Logical Scheme**

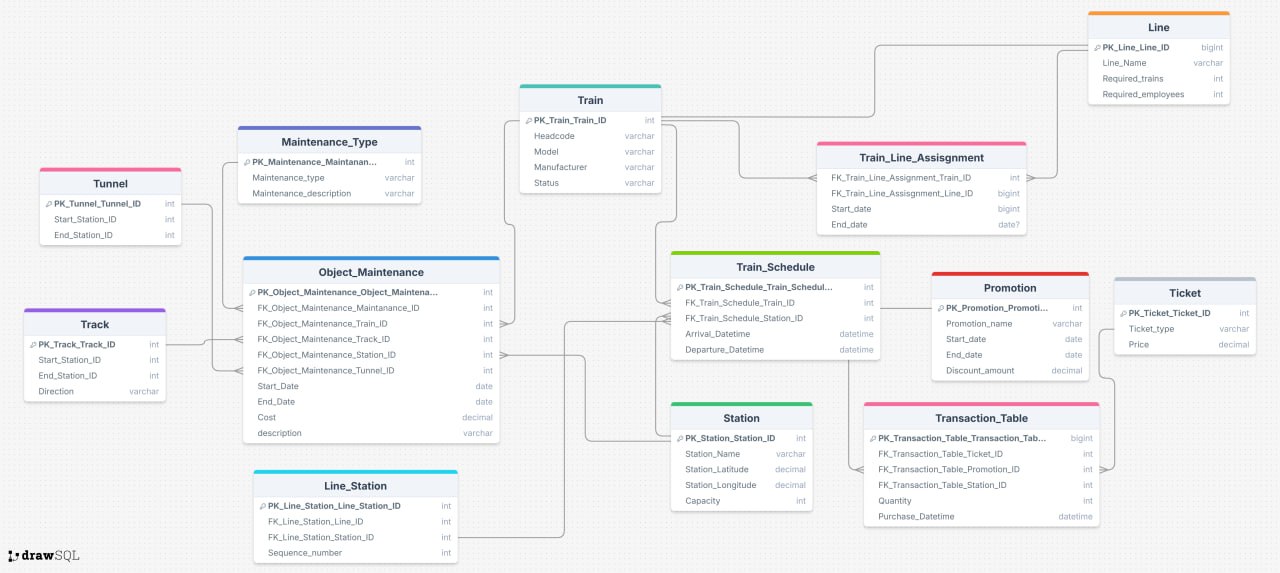
The model supports core metro processes:

* train movement and scheduling
* ticket sales and promotions
* tracking stations along multiple lines
* maintaining physical infrastructure (tracks, tunnels, trains)

**Key Design Decisions**

* **M:N Relationship (Train ↔️ Line):** Because a train can serve multiple lines over time and each line can have multiple trains, a junction table (Train\_Line\_Assignment) is used.
* **Historical accuracy:** Start\_Date and End\_Date allow tracking changes over time.
* **3NF Applied:** Each table stores one concept, all non-key attributes depend on the key, and no transitive dependencies exist.

## Logical Scheme



## Objects

**Table Description**

**OPERATIONS TABLES**

|  |  |
| --- | --- |
| Table Name | Description |
| Train | **Stores information about each train, including its model, headcode, manufacturer, status, and the line it currently operates on. Each train runs on one line at a time.** |
| Train\_Schedule | **Defines the arrival and departure timetable of trains at stations. It connects a train to a specific station with arrival and departure datetimes and references the sequence of that station on its line.** |
| Line | **Represents a railway line. Each line has a name and required resource information such as number of trains and employees needed to operate it.** |
| **Line\_Station** | **A junction table that defines the ordered sequence of stations along a railway line. It ensures a line has multiple stations in the correct route order.** |
| **Station** | **Stores station details such as name, coordinates (latitude & longitude), capacity, and the line it belongs to.** |

**INFRASTRUCTURE TABLES**

|  |  |
| --- | --- |
| Table Name | Description |
| Track | Defines an above-ground track segment between two stations, including direction information (e.g., Northbound/Southbound). |
| Tunnel | Defines a tunnel segment between two stations, similar to tracks, but for underground routes. |
| Object\_Maintenance | Tracks all maintenance performed on infrastructure or trains. It links to trains, tracks, stations, or tunnels and includes start/end dates, cost, and a description. |
| Maintanance\_Type | Lists types of maintenance (e.g., electrical, engine service, track repair) along with a description of each maintenance category. |

**TICKETING & PROMOTION TABLES**

|  |  |
| --- | --- |
| Table Name | Description |
| Ticket | Represents a ticket type (e.g., standard, express, VIP), including its base price. |
| Promotion | Defines discounts available during specific dates (e.g., holiday sale, weekend offers). Contains name, validity period, and discount amount. |
| Transaction\_Table | Logs completed ticket purchases. It records which ticket was bought, which promotion (if any) was applied, which station the purchase occurred at, total price, quantity, and purchase timestamp. |

**Table detailed description**

|  |  |  |  |
| --- | --- | --- | --- |
| Table Name | Field name | Field Description | Data Type |
| Train | PK\_Train\_Line\_Assignment\_Train\_ID | References the Train table (Train\_ID), Identifies which train is assigned to the line, FK | Int |
| PK\_Train\_Line\_Assignment\_Line\_Line\_ID | References the Line Table (Line\_ID), identifies the railway line that the train is assigned to, FK | Bigint |
| Start\_Date | The date when the train starts operating on the assigned line, used to track historical changes and prevent overlapping assignments. | Date |
| End\_Date | The date when the train stops operating on that line. If Null, it indicates the train is currently assigned to that line. | Date (Nullabe) |

**Comments on table relationships**

A single **Train** can serve multiple Lines over time, and a **Line** can operate multiple Trains. Because this is a many-to-many relationship, it is modeled using the **Train\_Line\_Assignmnet** table.

Example with data

|  |  |  |  |
| --- | --- | --- | --- |
| PK\_Train\_Line\_Assignment\_Train\_ID | PK\_Train\_Line\_Assignment\_Line\_ID | Start\_Date | End\_Date |
| 12 | 3 | 2025-01-10 | 2025-05-30 |
| 12 | 5 | 2025-06-01 | NULL |

Meaning: Train 12 served Line 3 until May, then switched to Line 5 and is still currently in service on that line.