Data Catalog for Gold Layer

Overview

The Gold Layer is the business-level data representation, structured to support analytical and reporting use cases. It consists of **dimension tables** and **fact tables** for specific business metrics.

---

1. gold.dim_vehicle

- **Purpose:** Stores reference data about vehicle types used in bookings.
- Columns:

Column Name	Data Type	Descriptioin
vehicle_key	INT	Surrogate key uniquely identifying each vehicle type.
vehicle_type	NVARCHAR(50)	The type of vehicle selected by the customer (e.g., Car, Bike, Auto).

2. gold.dim_location

- **Purpose:** Stores distinct pickup and drop-off locations.
- Columns:

Column Name	Data Type	Descriptioin
location_key	INT	Surrogate key uniquely identifying each location.
location_name	NVARCHAR(50)	The location name, either pickup or drop-off, as recorded in the bookings.

3. gold.dim_date

- **Purpose:** Provides calendar details extracted from booking datetime.
- Columns:

Column Name	Data Type	Descriptioin
date_key	INT	Surrogate key in YYYYMMDD format.
full_date	DATE	Full date in standard format (YYYY-MM-DD)..
day	INT	Day of the month (1–31).
month	INT	Month of the year (1–12).
year	INT	Year of the booking (e.g., 2025).

4. gold.dim_payment

- **Purpose:** Captures payment method information used in bookings.
- Columns:

Column Name	Data Type	Descriptioin
payment_key	INT	Surrogate key uniquely identifying each payment method.
payment_method	NVARCHAR(50)	Payment method used (e.g., Cash, Credit Card, Wallet).

5. gold.dim_cancellation_reason

- **Purpose:** Stores standardized reasons for booking cancellations.
- Columns:

Column Name	Data Type	Descriptioin
cancel_key	INT	Surrogate key uniquely identifying each cancellation reason.
cancellation_reason	NVARCHAR(50)	Reason provided for cancellation (customer or driver initiated).

6. gold.fact_booking
- **Purpose:** Stores all booking transactions with references to dimensions.
- Columns:

Column Name	Data Type	Descriptioin
booking_id	NVARCHAR(50)	Unique identifier for the booking.
date_key	INT	Foreign key to gold.dim_date.
vehicle_key	INT	Foreign key to gold.dim_vehicle.
pickup_location_key	INT	Foreign key to gold.dim_location (pickup location).
drop_location_key	INT	Foreign key to gold.dim_location (drop location).
payment_key	INT	Foreign key to gold.dim_payment.
booking_status	NVARCHAR(50)	Status of the booking (e.g., Completed, Cancelled).
ride_distance	FLOAT	Distance of the ride in kilometers.
booking_value	FLOAT	Total fare or value of the booking.
driver_rating	FLOAT	Rating given to the driver for the booking.

7. gold.fact_cancellation
- **Purpose:** Captures details of cancelled bookings.
- Columns:

Column Name	Data Type	Descriptioin
cancellation_id	BIGINT	Surrogate key uniquely identifying each cancellation event.
booking_id	BIGINT	Identifier linking to the cancelled booking.
date_key	INT	Foreign key to gold.dim_date.
cancelled_by	NVARCHAR(50)	Specifies who cancelled the ride (Customer or Driver).
cancel_key	INT	Foreign key to gold.dim_cancellation_reason.

8. gold.fact_revenue

- **Purpose:** Stores revenue details for completed rides.
- Columns:

Column Name	Data Type	Descriptioin
revenue_id	BIGINT	Surrogate key uniquely identifying each revenue event.
booking_id	BIGINT	Identifier linking to the booking.
date_key	INT	Foreign key to gold.dim_date.
vehicle_key	INT	Foreign key to gold.dim_vehicle.
payment_key	INT	Foreign key to gold.dim_payment..
revenue	FLOAT	Total revenue generated for the booking.

9. gold.report_success_rate

- **Purpose:** Summarizes booking performance by completion vs cancellation.
- Columns:

Column Name	Data Type	Descriptioin
date_key	INT	Reference to date (daily/hourly granularity).
total_booking	INT	Total number of bookings.
completed_booking	INT	Total number of completed bookings.
cancelled_booking	INT	Total number of cancelled bookings.
success_rate	FLOAT	Percentage of successful rides (completed / total * 100).

10. gold.report_revenue_distribution

- **Purpose:** Provides revenue distribution by vehicle type and payment method.
- Columns:

Column Name	Data Type	Descriptioin
date_key	INT	Reference to date (daily/hourly granularity).
vehicle_key	INT	Foreign key to gold.dim_vehicle.
payment_key	INT	Foreign key to gold.dim_payment.
total_revenue	FLOAT	Total revenue generated for the grouping.
ride_count	INT	Total number of rides for the grouping.

11. gold.report_cancellation_reason

- **Purpose:** Analyzes reasons for booking cancellations.
- Columns:

Column Name	Data Type	Descriptioin
date_key	INT	Reference to date (daily/hourly granularity).
cancel_key	INT	Foreign key to gold.dim_cancellation_reason.
cancel_count	INT	Number of bookings cancelled with this reason.