

Project Deliverable 5 – Table Data Populations

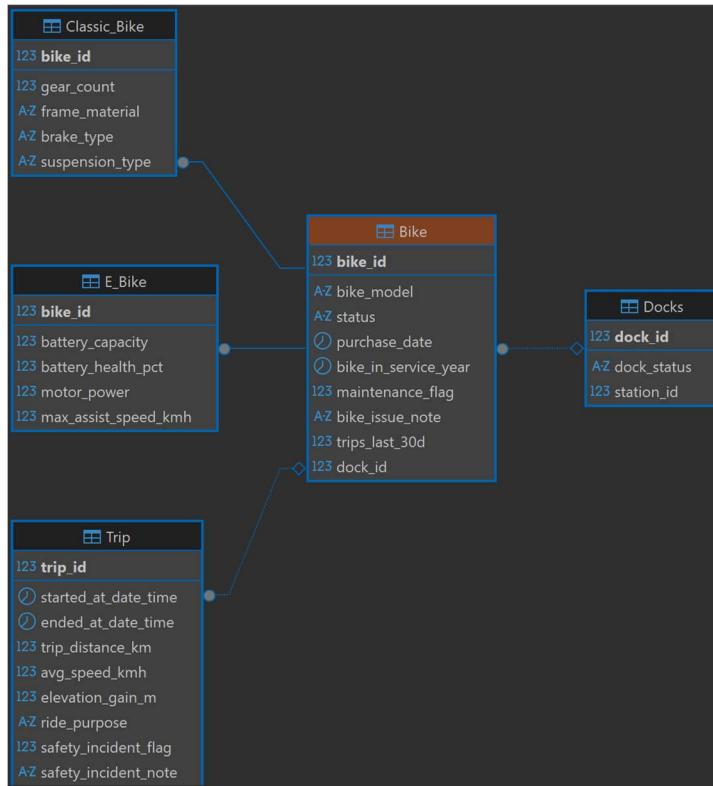
Group 3 - Sabereen Safa, Jordan Eberhart, James Eastwood

The work contained here is the work of Group 3 and Group 3 alone.

Properties		Data	Diagram
Table Name:	Bike	Engine:	InnoDB
Description:		Auto Increment:	0
Columns		Name	Value
		Row Count	60
		Avg Row Length	273
		Data Length	16K
		Max data length	0
		Data free	0
		Index length	32K
		Row format	Dynamic
		Create Time	2025-11-02 14:39:26.0
		Update time	2025-11-02 14:53:15.0
Statistics		Check time	
DDL			
Virtual			

Bike_ID was synthesized using AI as a sequential number with the leading 32 standing for Classic Bike and the leading 52 standing for Electric Bike, then zeroes before the sequence starting at 1.

Purchase_Date was synthesized data using AI based on the provided data of Bike_In_Service_Year.

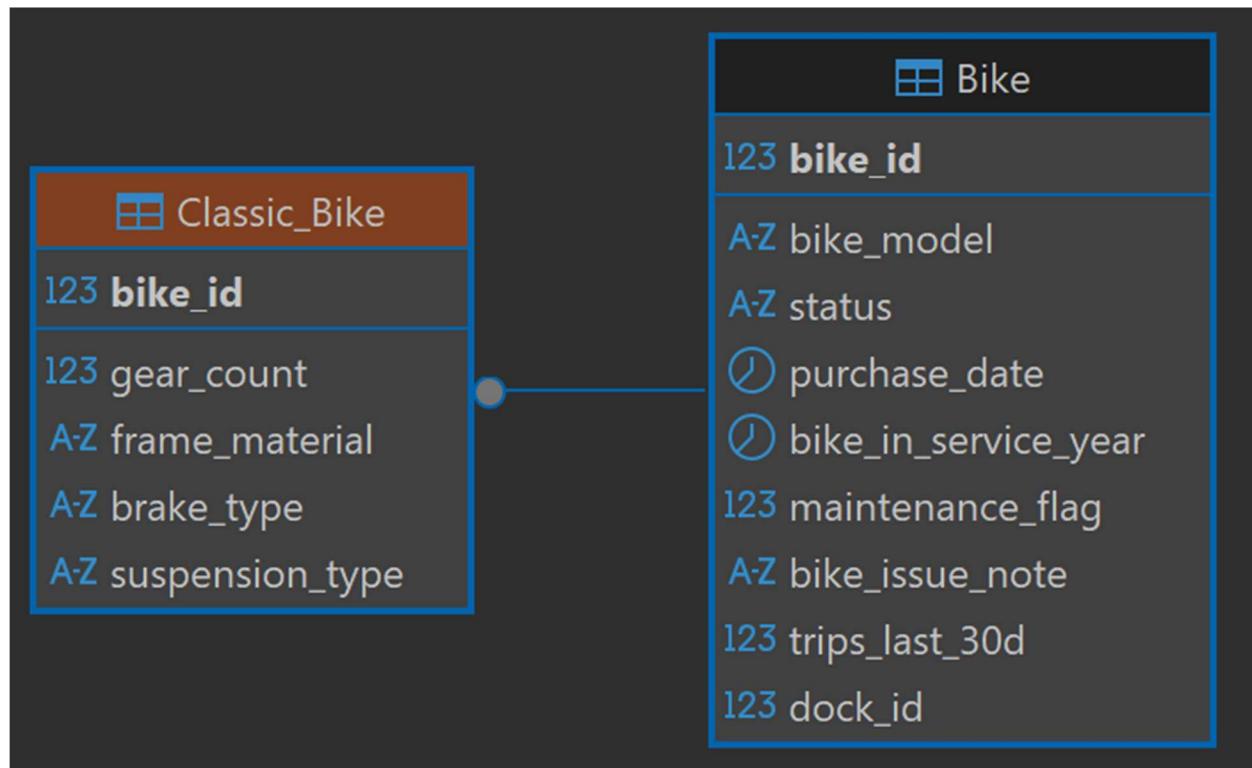


Properties Data Diagram

Table Name:	Classic_Bike	Engine:	InnoDB
Description:		Auto Increment:	0
		Charset:	utf8mb4
Description			
Columns	Name	Value	
	Row Count	30	
	Avg Row Length	546	
	Data Length	16K	
	Max data length	0	
	Data free	0	
	Index length	16K	
	Row format	Dynamic	
	Create Time	2025-10-26 19:54:46.0	
	Update time	2025-11-02 14:55:31.0	
Statistics	Check time		
DDL			
Virtual			

Gear_count was synthesized using AI falling within the constraint of the 25 gear max.

Frame_material, brake_type, suspension_type, were also synthesized using AI and randomly assigned to bike_ids.

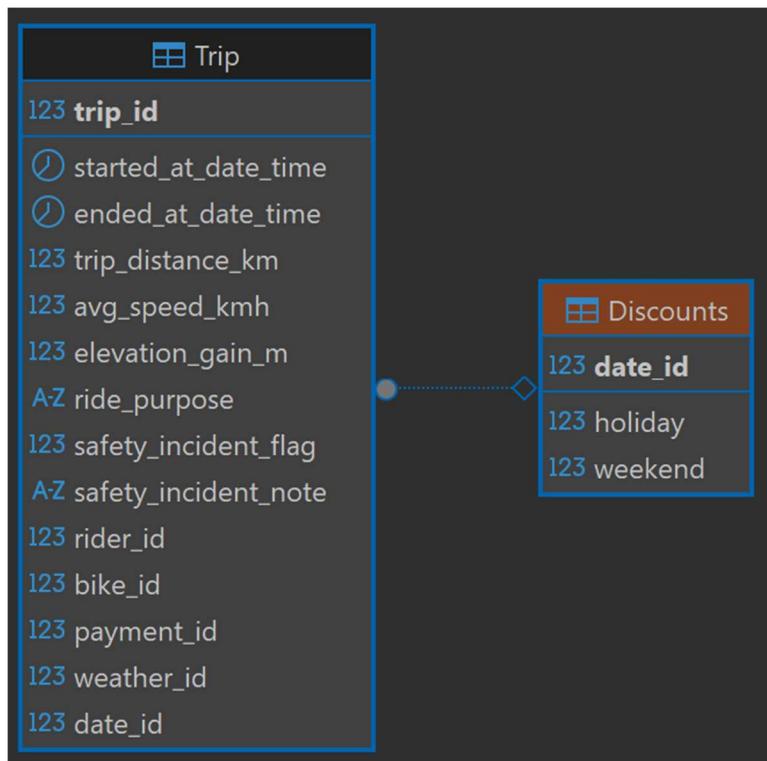


Properties Data Diagram

Table Name:	Discounts	Engine:	Inno
Description:		Auto Increment:	0
		Charset:	utf8
Columns	Name	Value	
Constraints	Row Count	365	
Foreign Keys	Avg Row Length	44	
References	Data Length	16K	
Triggers	Max data length	0	
Indexes	Data free	0	
Partitions	Index length	16K	
Statistics	Row format	Dynamic	
DDL	Create Time	2025-10-26 19:29:50.0	
Virtual	Update time	2025-11-02 14:29:47.0	
	Check time		

Date_id was synthesized with 49193 standing for DISC because of the name of the Table “Discounts.” Then preceding zeroes and then a sequential number starting at 1 and ending at 366 (since 2024 was a Leap Year) were added.

Holiday and Weekend are boolean variables which were synthesized based on whether they were holidays or weekends in the year of 2024.

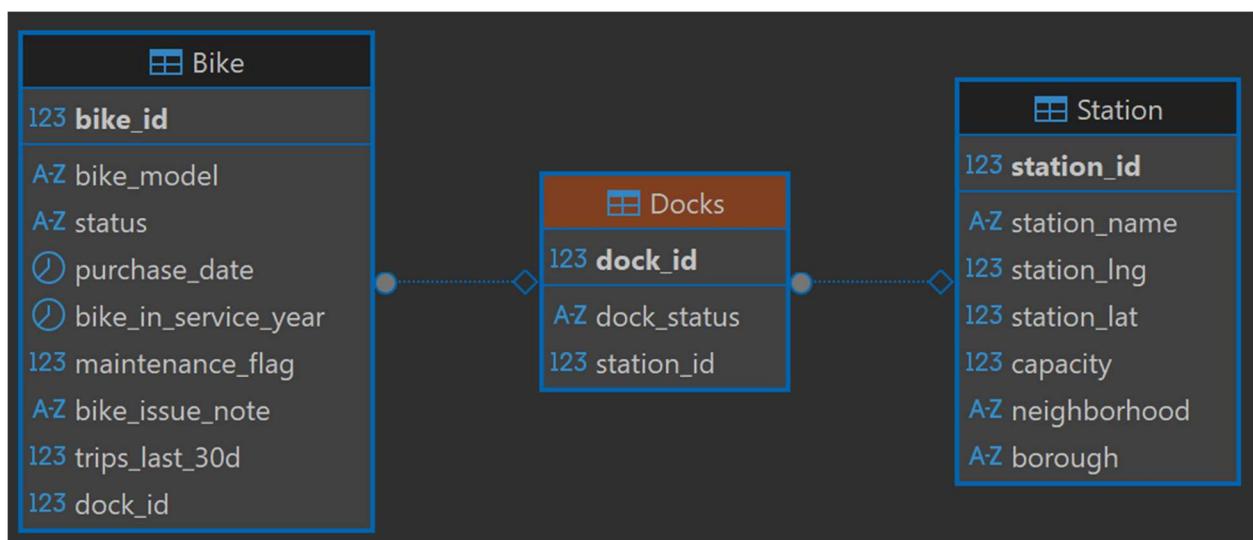


Properties Data Diagram

Table Name:	Docks	Engine:	InnoDB
Description:			
		Auto Increment:	0
		Charset:	utf8mb4
Columns	Name	Value	
Constraints	Row Count	45	
Foreign Keys	Avg Row Length	364	
References	Data Length	16K	
Triggers	Max data length	0	
Indexes	Data free	0	
Partitions	Index length	32K	
Statistics	Row format	Dynamic	
DDL	Create Time	2025-10-26 19:30:35.0	
Virtual	Update time	2025-11-02 13:33:25.0	
	Check time		

Dock_id was an AI synthesized sequential number with a leading 4 representing “D” for Dock and then zeroes to separate the actual dock_id.

Dock_status was randomly selected between “available,” “occupied,” or “out of service.”

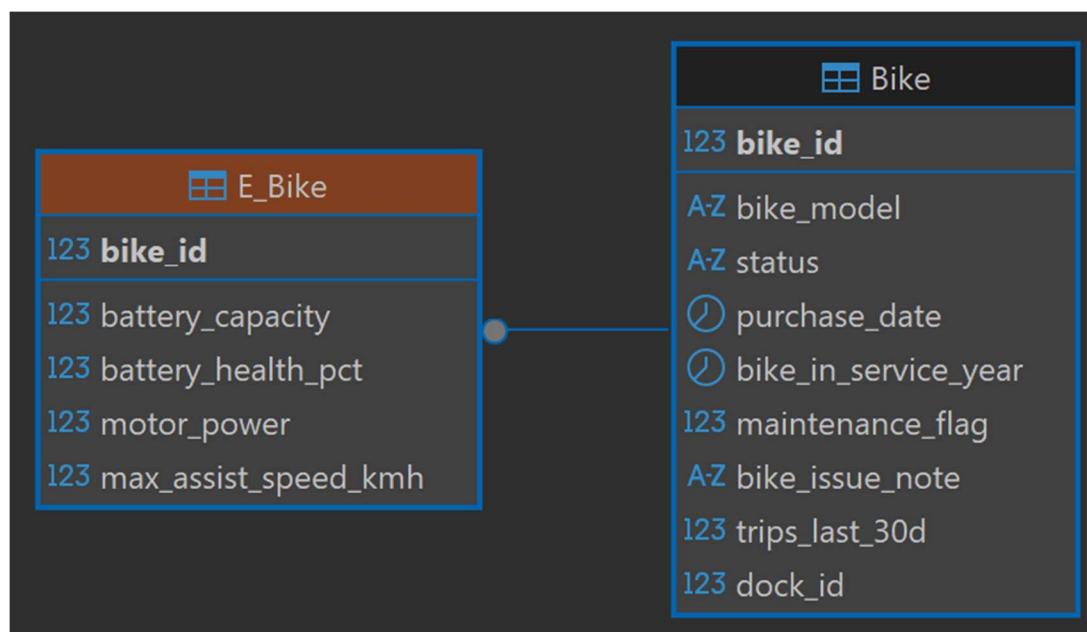


Properties **Data** **Diagram**

Table Name:	E_Bike	Engine:	Inno
Description:		Auto Increment:	0
		Charset:	utf8r

	Name	Value
Columns	Row Count	30
Constraints	Avg Row Length	546
Foreign Keys	Data Length	16K
References	Max data length	0
Triggers	Data free	0
Indexes	Index length	16K
Partitions	Row format	Dynamic
Statistics	Create Time	2025-10-26 19:58:04.0
	Update time	2025-11-02 14:54:48.0
	Check time	
DDL		
Virtual		

Battery_capacity, battery_health_pct, motor_power, and max_assist_speed_kmh were all randomly AI synthesized.

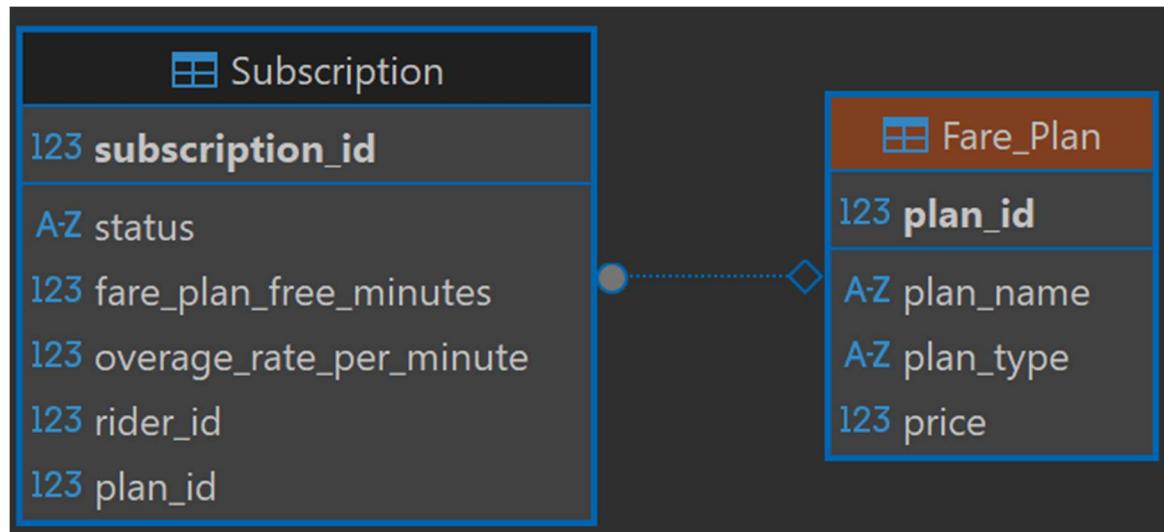


Properties Data Diagram

Table Name:	Fare_Plan	Engine:	In
Description:		Auto Increment:	0
		Charset:	utf8mb4
Columns	Name	Value	
Constraints	Row Count	12	
Foreign Keys	Avg Row Length	1,365	
References	Data Length	16K	
Triggers	Max data length	0	
Indexes	Data free	0	
Partitions	Index length	16K	
Statistics	Row format	Dynamic	
DDL	Create Time	2025-10-26 20:01:34.0	
Virtual	Update time	2025-11-02 13:40:18.0	
	Check time		

Plan_id was an AI synthesized random number with a lead 616 for “F” and “P” for the Fare_Plan table. Then there were separating zeroes and the sequential number of all of the plans, of which there are only 12 due to there not being any additional reasonable fare plan options.

Plan_name, plan_type, and price were all AI generated and chosen as the best options to differentiate between fare plans with data that was reasonable and realistic.

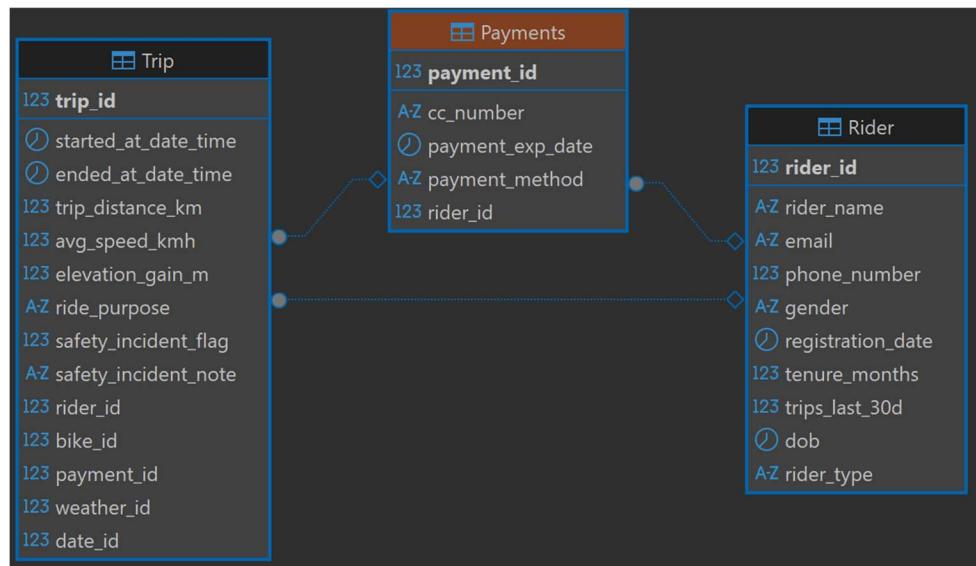


Properties Data Diagram

Table Name:	Payments	Engine:	InnoDB
Description:		Auto Increment:	0
		Charset:	utf8mb4
Columns	Name	Value	
	Row Count	60	
Constraints	Avg Row Length	273	
Foreign Keys	Data Length	16K	
References	Max data length	0	
Triggers	Data free	0	
Indexes	Index length	32K	
Partitions	Row format	Dynamic	
Statistics	Create Time	2025-10-26 19:32:55.0	
	Update time	2025-11-02 20:19:54.0	
	Check time		
DDL			
Virtual			

Payment_id was an AI synthesized, random number with a leading 16 for “P” for Payments, then a separator zero, and then the random number for a grand total of 10 digits.

CC_number, payment_exp_date, and payment_method were all AI synthesized numbers where the Visa, Mastercard, and Discover payment_method have 16 digit CC_numbers while American Express have 15 digits. Payment_exp_dates were generated to be 2 to 6 years into the future.



Properties		Data	Diagram
Table Name:	Rider	Engine:	Innodb
Description:		Auto Increment:	0
		Charset:	utf8mb4
Columns	Name	Value	
Constraints	Row Count	60	
Foreign Keys	Avg Row Length	273	
References	Data Length	16K	
Triggers	Max data length	0	
Indexes	Data free	0	
Partitions	Index length	16K	
Statistics	Row format	Dynamic	
DDL	Create Time	2025-10-26 20:06:46.0	
Virtual	Update time	2025-11-02 16:41:31.0	
Statistics	Check time		
DDL			
Virtual			

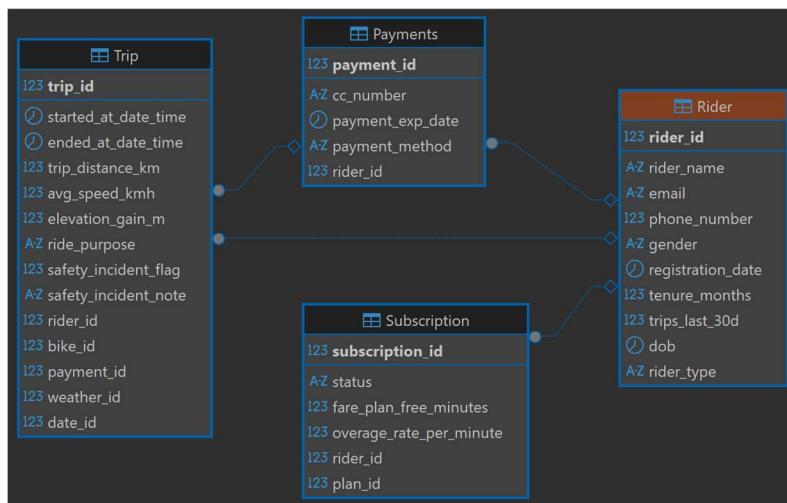
Rider_id was an AI synthesized number with a leading 18 for “R” for Rider, then a separator zero, and then the random number for a grand total of 10 digits.

Rider_name, email, phone_number, and dob were all randomly generated by AI.

Gender was randomly chosen from “M,” “F,” “Other,” or “Prefer Not to Say,” with a bias put towards M and F if the name of the rider was a generally M or F name.

Registration_date was based on tenure_months which was a provided datapoint in the original dataset. If tenure_months was 6, then the registration date would have been a randomly generated date in a 30 day window surrounding the ride date, 6 months in the past.

Tenure_months and trips_last_30d were data that was all provided in the original dataset.



Properties Data Diagram

Table Name: Station Engine:

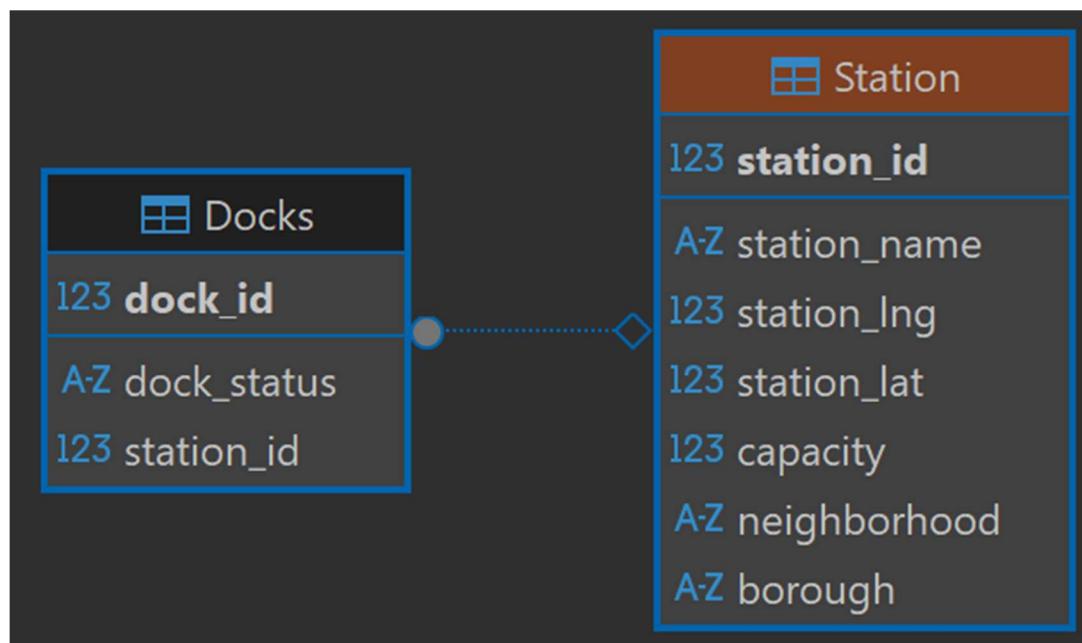
Description: Auto Increment:

Charset:

Name	Value
Row Count	9
Avg Row Length	1,820
Data Length	16K
Max data length	0
Data free	0
Index length	16K
Row format	Dynamic
Create Time	2025-10-26 20:09:10.0
Update time	2025-10-28 20:15:31.0
Check time	

Columns Constraints Foreign Keys References Triggers Indexes Partitions Statistics DDL Virtual

Station_id was an ID number provided in the original dataset, along with every other attribute and value. As the only table that needed no data generated, if we just left it alone, even though it only has 9 records, it made sense to leave it untouched.



Properties Data Diagram

Table Name: Subscription Engine: InnoDB

Description:

Auto Increment: 0 Charset: utf8mb4

Name	Value
Row Count	60
Avg Row Length	273
Data Length	16K
Max data length	0
Data free	0
Index length	48K
Row format	Dynamic
Create Time	2025-10-26 20:12:28.0
Update time	2025-11-02 17:01:25.0

Statistics

Check time

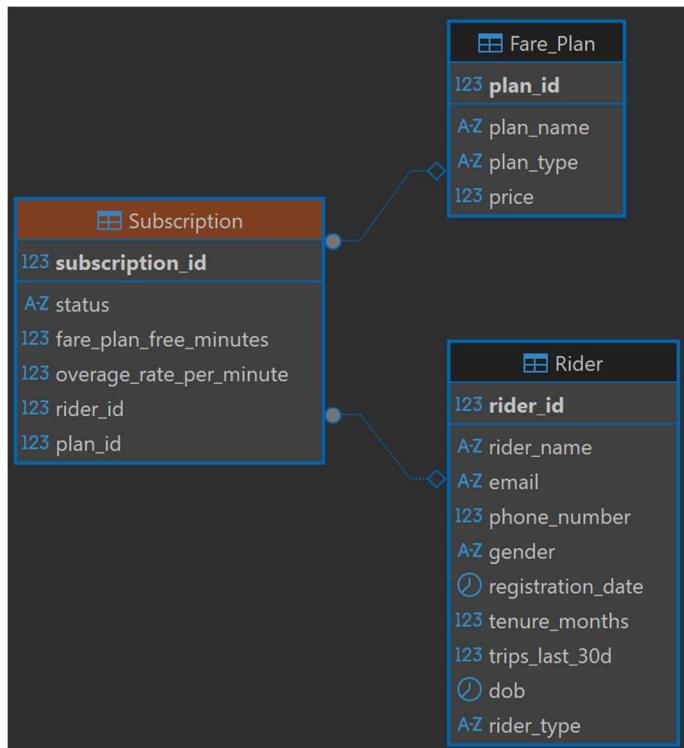
DDL

Virtual

Subscription_id was an AI synthesized, random number. The number is led by 5 digits 19212 for “S,” “U,” “B” for SUB in Subscription, then a separator zero, and then the random number for a total of 10 digits.

Fare_plan_free_minutes, and overage_rate_per_minute are actual data included in the original dataset.

Status was an AI randomly chosen selection between “active,” “cancelled,” and “paused.”



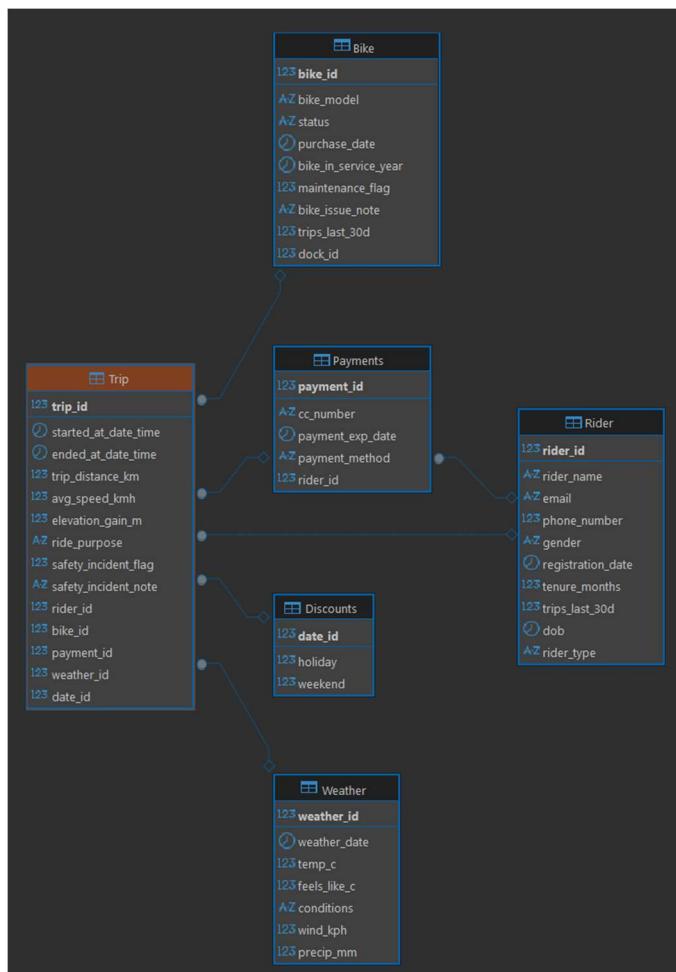
Properties Data Diagram

Table Name: Trip Engine: InnoDB
Description: Auto Increment: 0
Charset: utf8mb4

	Name	Value
Columns	Row Count	60
Constraints	Avg Row Length	273
Foreign Keys	Data Length	16K
References	Max data length	0
Triggers	Data free	0
Indexes	Index length	96K
Partitions	Row format	Dynamic
Statistics	Create Time	2025-10-26 20:15:29.0
DDL	Update time	2025-11-02 20:33:34.0
Virtual	Check time	

Trip_id was an AI synthesized sequential number with a leading 20 for “T” in Trip, and then a lot of separator zeroes, followed by the sequence starting at 1, with a grand total of 10 digits.

Every other column, aside from the Foreign Keys are data that was provided in the original dataset.



Properties Data Diagram

Table Name: Weather Engine:

Description: Auto Increment: 0

Charset: u

	Name	Value
Columns	Row Count	60
Constraints	Avg Row Length	273
Foreign Keys	Data Length	16K
References	Max data length	0
Triggers	Data free	0
Indexes	Index length	16K
Partitions	Row format	Dynamic
Statistics	Create Time	2025-11-03 08:18:39.0
	Update time	2025-11-03 08:19:05.0
Check time		
DDL		
Virtual		

Weather_id is an AI synthesized number led by 2321 for “W,” and “U” because when we first generated the weather_id data, we mistakenly generated a varchar identifier with a leading WU, which I then converted with 2321. Then there were separator zeroes and a sequential number starting at 1 for a grand total of 10 digits.

All other data in the table was provided by the original dataset.

