```
Use Cars24;

-- Convert `Engine Capacity` and `Price(INR)` to numeric values
ALTER TABLE `dataset_used cars`
MODIFY COLUMN `Engine Capacity` INT,
MODIFY COLUMN `Price(INR)` INT;

-- Fill null values in `Varient` and `Details` with 'Unknown' or
appropriate placeholder
SET SQL_SAFE_UPDATES = 0;
UPDATE `dataset_used cars`
SET Varient = 'Unknown'
WHERE Varient IS NULL;

UPDATE `dataset_used cars`
SET Details = 'Unknown'
WHERE Details IS NULL;

SET SQL_SAFE_UPDATES = 1;

-- Calculate summary statistics for Distance Travelled
SELECT
    AVG(`Distance Travelled`) AS Avg_Distance,
    MIN(`Distance Travelled`) AS Min_Distance,
    MAX(`Distance Travelled`) AS Max_Distance,
    STDDEV(`Distance Travelled`) AS StdDev_Distance
FROM `dataset_used cars`;

-- Calculate summary statistics for Price(INR)
SELECT
    AVG(`Price(INR)`) AS Avg_Price,
    MIN(`Price(INR)`) AS Min_Price,
    MAX(`Price(INR)`) AS Max_Price,
    STDDEV(`Price(INR)`) AS StdDev_Price
FROM `dataset_used cars`;

SELECT ROUND(AVG(YEAR(CURDATE()) - Model)) AS Avg_Car_Age
FROM `dataset_used cars`;

-- Count cars by Manufacturer
SELECT Manufacturer, COUNT(*) AS Car_Count
FROM `dataset_used cars`
GROUP BY Manufacturer
ORDER BY Car_Count DESC;

-- Count cars by Fuel Type
SELECT `Fuel Type`, COUNT(*) AS Fuel_Count
FROM `dataset_used cars`
GROUP BY `Fuel Type`
ORDER BY Fuel_Count DESC;

-- Count cars by Transmission
SELECT Transmission, COUNT(*) AS Transmission_Count
FROM `dataset_used cars`
GROUP BY Transmission
ORDER BY Transmission_Count DESC;

SELECT Model, COUNT(*) AS Model_Count
```

```
FROM `dataset_used cars`
GROUP BY Model
ORDER BY Model DESC;

-- Average price by Manufacturer
SELECT Manufacturer, ROUND(AVG(`Price(INR)`),2) AS Avg_Price
FROM `dataset_used cars`
GROUP BY Manufacturer
ORDER BY Avg_Price DESC;

-- Average price by Fuel Type
SELECT `Fuel Type`, ROUND(AVG(`Price(INR)`),2) AS Avg_Fuel_Price
FROM `dataset_used cars`
GROUP BY `Fuel Type`
ORDER BY Avg_Fuel_Price DESC;

-- Average price by Transmission
SELECT Transmission, ROUND(AVG(`Price(INR)`),2) AS Avg_Transmission_Price
FROM `dataset_used cars`
GROUP BY Transmission
ORDER BY Avg_Transmission_Price DESC;

-- Most expensive car by Manufacturer
SELECT Manufacturer, Model, `Price(INR)`
FROM `dataset_used cars`
ORDER BY `Price(INR)` DESC
LIMIT 1;

-- Least expensive car by Manufacturer
SELECT Manufacturer, Model, `Price(INR)`
FROM `dataset_used cars`
ORDER BY `Price(INR)` ASC
LIMIT 1;

-- Average mileage by Manufacturer
SELECT Manufacturer, AVG(`Distance Travelled`) AS Avg_Distance
FROM `dataset_used cars`
GROUP BY Manufacturer
ORDER BY Avg_Distance DESC;

-- Average mileage by Fuel Type
SELECT `Fuel Type`, AVG(`Distance Travelled`) AS Avg_Fuel_Distance
FROM `dataset_used cars`
GROUP BY `Fuel Type`
ORDER BY Avg_Fuel_Distance DESC;

-- Average mileage by Transmission
SELECT Transmission, AVG(`Distance Travelled`) AS
Avg_Transmission_Distance
FROM `dataset_used cars`
GROUP BY Transmission
ORDER BY Avg_Transmission_Distance DESC;

-- Correlation between Engine Capacity and Price(INR)
SELECT
    (COUNT(*) * SUM(`Engine Capacity` * `Price(INR)`) - SUM(`Engine
Capacity`) * SUM(`Price(INR)`)) /
```

```
    (SQRT(COUNT(*) * SUM(POW(`Engine Capacity`, 2)) - POW(SUM(`Engine
Capacity`), 2)) *
     SQRT(COUNT(*) * SUM(POW(`Price(INR)`, 2)) - POW(SUM(`Price(INR)`),
2))) AS Correlation
FROM `dataset_used cars`;

-- Correlation between Distance Travelled and Price(INR)
SELECT
    (COUNT(*) * SUM(`Distance Travelled` * `Price(INR)`) - SUM(`Distance
Travelled`) * SUM(`Price(INR)`)) /
    (SQRT(COUNT(*) * SUM(POW(`Distance Travelled`, 2)) -
POW(SUM(`Distance Travelled`), 2)) *
     SQRT(COUNT(*) * SUM(POW(`Price(INR)`, 2)) - POW(SUM(`Price(INR)`),
2))) AS Correlation
FROM `dataset_used cars`;
```