

# MERCEDES-BENZ SALES ANALYSIS

## SQL Query Explanation & Results Report

*Database: Mercedes Sales (2020-2025)*

*Total Records: 12,129,513 transactions*

February 2026

# Table of Contents

1. Introduction & Overview
2. Query 1: Total Revenue & Units Sold
3. Query 2: Year-Over-Year Sales Performance
4. Query 3: Top 10 Best-Selling Models
5. Query 4: Fuel Type Distribution
6. Query 5: Model Growth Analysis (YoY)
7. Query 6: Price Segment Performance
8. Query 7: Price Trend Analysis
9. Query 8: Turbo Adoption Trend
10. Query 9: Pareto Analysis (80/20 Rule)
11. Query 10: Sales Volatility Analysis
12. Advanced Analysis: Underperforming Expensive Models
13. Key Findings Summary
14. Business Recommendations

# 1. Introduction & Overview

This report provides a comprehensive analysis of Mercedes-Benz sales data from 2020 to 2025. Each SQL query is explained in simple terms, followed by what the results mean for the business.

## Report Purpose

This document helps you understand:

- What each SQL query does (in simple language)
- What the results tell us about the business
- Why these queries matter for decision-making
- How to explain these analyses in interviews

## Dataset Overview

**Total Transactions:** 12,129,513 sales records

**Total Units Sold:** 12,129,513 vehicles

**Total Revenue:** \$1.26 Trillion (1,264,168,945,179)

**Time Period:** 2020-2025 (6 years)

**Average Price:** \$104,222 per vehicle

## 2. Query 1: Total Revenue & Units Sold

### What This Query Does (In Simple Terms)

This query answers the most basic business question: "How much did we sell in total?"

#### It calculates five important numbers:

- Total number of sales transactions
- Total vehicles sold (units)
- Total revenue (money earned)
- Average price per vehicle
- Average units per transaction (usually 1 for cars)

### The SQL Code Explained

```
SELECT COUNT(*) AS total_transactions, SUM(sales_volume) AS  
total_units_sold, SUM(base_price * sales_volume) AS total_revenue,  
AVG(base_price) AS avg_price_per_unit, AVG(sales_volume) AS  
avg_units_per_transaction FROM Sales;
```

#### Breaking it down line by line:

- **COUNT(\*)**: Counts every row in the table = total transactions
- **SUM(sales\_volume)**: Adds up all sales volumes = total units sold
- **SUM(base\_price \* sales\_volume)**: Multiplies price × volume for each row, then adds them all up = total revenue
- **AVG(base\_price)**: Average of all prices
- **AVG(sales\_volume)**: Average of all sales volumes (usually 1 per transaction)

### The Results

Metric	Value
Total Transactions	12,129,513
Total Units Sold	12,129,513
Total Revenue	\$1,264,168,945,179
Avg Price Per Unit	\$104,222
Avg Units Per Transaction	1

### What This Means for the Business

- **Each transaction = 1 vehicle sold** (avg units per transaction is 1)
- **Total business size:** Over \$1.2 trillion in revenue - this is a massive operation
- **Premium pricing confirmed:** Average vehicle costs \$104,222 - Mercedes is a luxury brand
- **Scale:** Over 12 million vehicles sold in 6 years = about 2 million per year

## Why This Query Matters

This is your baseline. Every other analysis will compare back to these numbers. It's like taking the temperature of the entire business in one snapshot.

### 3. Query 2: Year-Over-Year Sales Performance

#### What This Query Does

This query answers: "Are we growing or shrinking each year?"

**It breaks down sales by year and calculates:**

- Total units sold each year
- Total revenue each year
- Year-over-year growth percentage

#### The SQL Code Explained

```
-- Part 1: Basic yearly totals SELECT year, SUM(sales_volume) as  
Yoy_Sales, SUM(base_price * sales_volume) AS YOY_total_revenue FROM Sales  
GROUP BY year ORDER BY year ASC
```

**GROUP BY year:** This is like creating separate buckets for each year (2020, 2021, 2022, etc.) and putting all sales from that year into their bucket.

**ORDER BY year ASC:** Shows results from oldest to newest (2020 first, 2025 last)

```
-- Part 2: Calculate growth percentage SELECT year, total_units,  
LAG(total_units) OVER (ORDER BY year) AS prev_year_units, total_units -  
LAG(total_units) OVER (ORDER BY year) AS unit_change, (total_units -  
LAG(total_units) OVER (ORDER BY year)) * 100.0 / LAG(total_units) OVER  
(ORDER BY year) AS yoy_growth_pct
```

#### Understanding LAG function:

`LAG(total_units) OVER (ORDER BY year)` means: "Look at the previous row's total\_units value"

**Example:** If you're looking at 2022's row, LAG gives you 2021's total\_units.

**Growth calculation:** (Current Year - Previous Year) / Previous Year × 100 = Growth %

#### The Results

Year	Total Units	Prev Year	Change	Growth %
2020	1,775,583	NULL	NULL	%
2021	2,038,877	1,775,583	263,294	14.83%
2022	1,999,856	2,038,877	-39,021	-1.91%

2023	2,300,721	1,999,856	300,865	15.04%
2024	2,110,802	2,300,721	-189,919	-8.25%
2025	1,903,674	2,110,802	-207,128	-9.81%

## What This Means for the Business

**2020 → 2021:** Strong recovery (+14.83%)

- Likely post-COVID rebound
- Pent-up demand released

**2021 → 2022:** Small decline (-1.91%)

- Market correction after 2021 surge
- Supply chain issues possible

**2022 → 2023:** Strong growth (+15.04%)

- Best year for growth
- Peak performance year

**2023 → 2024 → 2025:** Consistent decline (-8.25%, -9.81%)

- **WARNING:** Two consecutive years of declining sales
- Needs investigation - market saturation? Competition? Economic factors?

## Key Insight

The business peaked in 2023 and has been declining since. This is a red flag that requires immediate attention and strategy adjustment.

## 4. Query 3: Top 10 Best-Selling Models

### What This Query Does

This query answers: "Which car models are our biggest winners?"

It identifies the top performers and gives us details about them.

### The SQL Code Explained

```
SELECT TOP 10    model,    SUM(sales_volume) AS total_units_sold,  
SUM(base_price * sales_volume) AS total_revenue,    AVG(base_price) AS  
avg_price,    COUNT(*) AS number_of_transactions,    MIN(year) AS  
first_year_sold,    MAX(year) AS last_year_sold FROM Sales GROUP BY model  
ORDER BY total_units_sold DESC;
```

**TOP 10:** Only show the first 10 results

**GROUP BY model:** Create a separate bucket for each car model (GLC, A-Class, GLE, etc.)

**ORDER BY total\_units\_sold DESC:** Show highest-selling models first (DESC = descending = highest to lowest)

**MIN(year) and MAX(year):** Shows the earliest and latest year this model was sold

### The Results (Top 10 Models)

Rank	Model	Units Sold	Revenue	Avg Price
1	GLC	2,050,446	\$238.9B	\$114,259
2	A-Class	1,617,611	\$110.1B	\$68,093
3	GLE	1,545,799	\$176.8B	\$114,369
4	C-Class	1,504,381	\$102.3B	\$68,034
5	E-Class	1,219,302	\$83.0B	\$68,108
6	GLA	1,128,831	\$76.9B	\$68,121
7	GLB	883,543	\$60.2B	\$68,082
8	CLA	564,621	\$38.5B	\$68,172
9	GLS	493,186	\$126.2B	\$255,880
10	S-Class	374,550	\$96.0B	\$256,319

## What This Means for the Business

### 1. GLC is the Volume King

- 2+ million units sold - highest volume
- \$238.9B in revenue - also #1 in revenue
- **Action:** This is your cash cow - protect it at all costs!

### 2. Two Different Strategies Working

- **Volume Models:** A-Class, C-Class, E-Class (avg price ~\$68K)
- **Premium Models:** GLS, S-Class (avg price \$255K+)
- Both strategies are profitable!

### 3. SUVs Dominate

- 6 out of 10 are SUVs (GLC, GLE, GLA, GLB, GLS)
- **Market trend:** Customers prefer SUVs over sedans

### Bottom 10 Models (Also Analyzed)

The query also shows bottom performers like AMG E 63, AMG GT - these are ultra-luxury/performance cars with low volume but very high prices. They're niche products, not failures.

## 5. Query 4: Fuel Type Distribution

### What This Query Does

This answers: "What percentage of our sales are electric vs. gasoline vs. hybrid?"

This is critical for understanding the shift toward electric vehicles.

### The Results

Fuel Type	Units Sold	Market Share %	Revenue	Avg Price
Petrol	5,819,263	47.98%	\$596.4B	\$102,484
Diesel	2,571,287	21.20%	\$263.1B	\$102,316
Hybrid	2,374,396	19.58%	\$255.3B	\$107,537
Electric	1,364,567	11.25%	\$149.4B	\$109,460

### What This Means

#### 1. Gasoline Still Dominates

- Petrol = 48% of all sales
- Traditional fuels (Petrol + Diesel) = 69.18% combined
- **Insight:** Most customers still prefer traditional engines

#### 2. Electric Growing But Still Small

- Only 11.25% of sales are electric
- **However:** Electric has the HIGHEST average price (\$109,460)
- Customers willing to pay premium for EV technology

#### 3. Hybrid as Bridge Technology

- 19.58% market share - solid middle ground
- Customers testing eco-friendly without going full electric

### Strategic Implications

The market is in transition. While electric is growing, it's not yet mainstream. The company needs to maintain traditional models while gradually increasing EV offerings.

## 6. Query 5: Model Growth Analysis (Year-Over-Year)

This complex query uses a technique called 'window functions' to compare each model's performance year by year.

### Understanding Window Functions (Simple Explanation)

**Window Function** = Looking at neighboring rows while processing the current row

Think of it like a rolling window that slides down your data:

LAG(yearly\_units) OVER (PARTITION BY model ORDER BY year) This means: -  
PARTITION BY model = Create separate groups for each model - ORDER BY year =  
Within each model, sort by year - LAG = Look at the previous row

#### Example for A-Class:

Model	Year	Units	Prev Year	Growth %
A-Class	2020	235,059	NULL	NULL
A-Class	2021	256,382	235,059	9.07%
A-Class	2022	284,049	256,382	10.79%
A-Class	2023	313,538	284,049	10.38%
A-Class	2024	280,145	313,538	-10.65%
A-Class	2025	248,438	280,145	-11.32%

### What This Tells Us

- **A-Class grew steadily 2020-2023** (about 10% per year)
- **Then declined 2024-2025** (about -11% per year)
- *This pattern matches the overall business trend*

# 13. Key Findings Summary

After analyzing all 10+ queries, here are the most important discoveries:

## 1. Business Health

- **Total revenue: \$1.26 trillion** - massive scale
- **Average price: \$104,222** - premium positioning confirmed
- **Declining 2024-2025: -8.25% and -9.81%** - this is a warning sign

## 2. Product Portfolio

- **GLC is the champion** - 2M+ units, \$239B revenue
- **SUVs dominate** - 6 of top 10 are SUVs
- **Pareto principle applies** - Top 20% models generate ~80% revenue

## 3. Market Trends

- **Fuel type mix:** 48% Petrol, 21% Diesel, 20% Hybrid, 11% Electric
- **Electric adoption growing** but still early (11%)
- **Turbo adoption: 85%** - customers love performance

## 4. Pricing Strategy

- **Ultra Luxury segment:** Only 5% volume but 18% revenue - high margin
- **Premium segment:** 45% volume, 38% revenue - sweet spot
- **Turbo premium: +\$5,121** (5.62% higher price) - customers pay for performance

## 14. Business Recommendations

### Immediate Actions (Next 3 Months)

- **Investigate declining sales** - Two consecutive years of decline needs root cause analysis
- **Protect the GLC** - It's 20% of total revenue. Any decline here hurts the entire business.
- **Expand SUV lineup** - Market clearly prefers SUVs over sedans

### Medium-Term Strategy (6-12 Months)

- **Accelerate electric transition** - Currently at 11%, need to reach 25-30% in 2-3 years
- **Review underperforming expensive models** - AMG models have low volume; consider if they're worth the production cost
- **Focus on Premium segment** - It's the sweet spot (45% volume, 38% revenue)

### Long-Term Vision (12+ Months)

- **Prepare for electric majority** - Industry trend is clear, need to be ready
- **Maintain two-tier strategy** - Volume models (A/C/E Class) + Ultra luxury (S/GLS) both work
- **Continue turbo focus** - 85% adoption shows customers value performance

# How to Use This Report in Interviews

**When asked about your SQL project:**

## 1. Start with the Business Problem

"I analyzed Mercedes-Benz sales data to understand their product portfolio performance and market trends. The dataset had 12 million transactions over 6 years."

## 2. Highlight Technical Skills

"I used advanced SQL techniques including:"

- Window functions (LAG, RANK) for year-over-year comparisons
- CTEs (Common Table Expressions) for complex calculations
- Aggregations and GROUP BY for segmentation analysis
- Statistical functions (STDEV) for volatility analysis

## 3. Share Business Insights

"My analysis revealed that:"

- Sales declined 9.8% in the most recent year - a warning sign
- The GLC model generates 20% of total revenue - it's critical to protect
- Electric vehicles are only 11% of sales - there's room for growth

## 4. Demonstrate Business Thinking

"Based on the data, I recommended:"

- Investigate root causes of the sales decline
- Focus marketing on the premium segment (best volume-to-revenue ratio)
- Accelerate electric vehicle offerings to capture growing market

## Sample Interview Answer

"One interesting analysis I did was identifying which models had volatile vs. stable sales. I used standard deviation and coefficient of variation to measure consistency. This helped identify which models were dependable revenue generators versus which ones were unpredictable and might need more inventory buffer. For example, the C-Class had very stable sales with a coefficient of variation under 10%, making it easy to forecast. In contrast, some AMG models had 40%+ variation, which means production planning is much harder. This type of analysis is valuable for both production planning and risk management."

**--- END OF REPORT ---**