



BUSINESS DATA ANALYSIS USING SQL

📌 DATA EXPLORATION & INSIGHTS FOR DECISION MAKING

2025

TABLE OF CONTENT

- Project Introduction
- Data Overview
- Framework
- Database Exploration
- Dimensions Exploration

- Date Range Analysis
- Key Metrics Analysis
- Magnitude Analysis
- Revenue Analysis
- Ranking Analysis
- Key Findings & Recommendations

PROJECT INTRODUCTION

This project aims to analyze business data using SQL to extract valuable insights and make data-driven decisions. The analysis included exploring the database structure, studying customers and products, analyzing time-based trends, and measuring performance through sales and revenue. The project also allowed me to apply my skills practically by writing complex queries, deeply exploring data, and understanding how to join tables to extract accurate information. It enhanced my expertise in data analysis, performance optimization, and using SQL in real-world scenarios to solve business problems efficiently.



DATA OVERVIEW

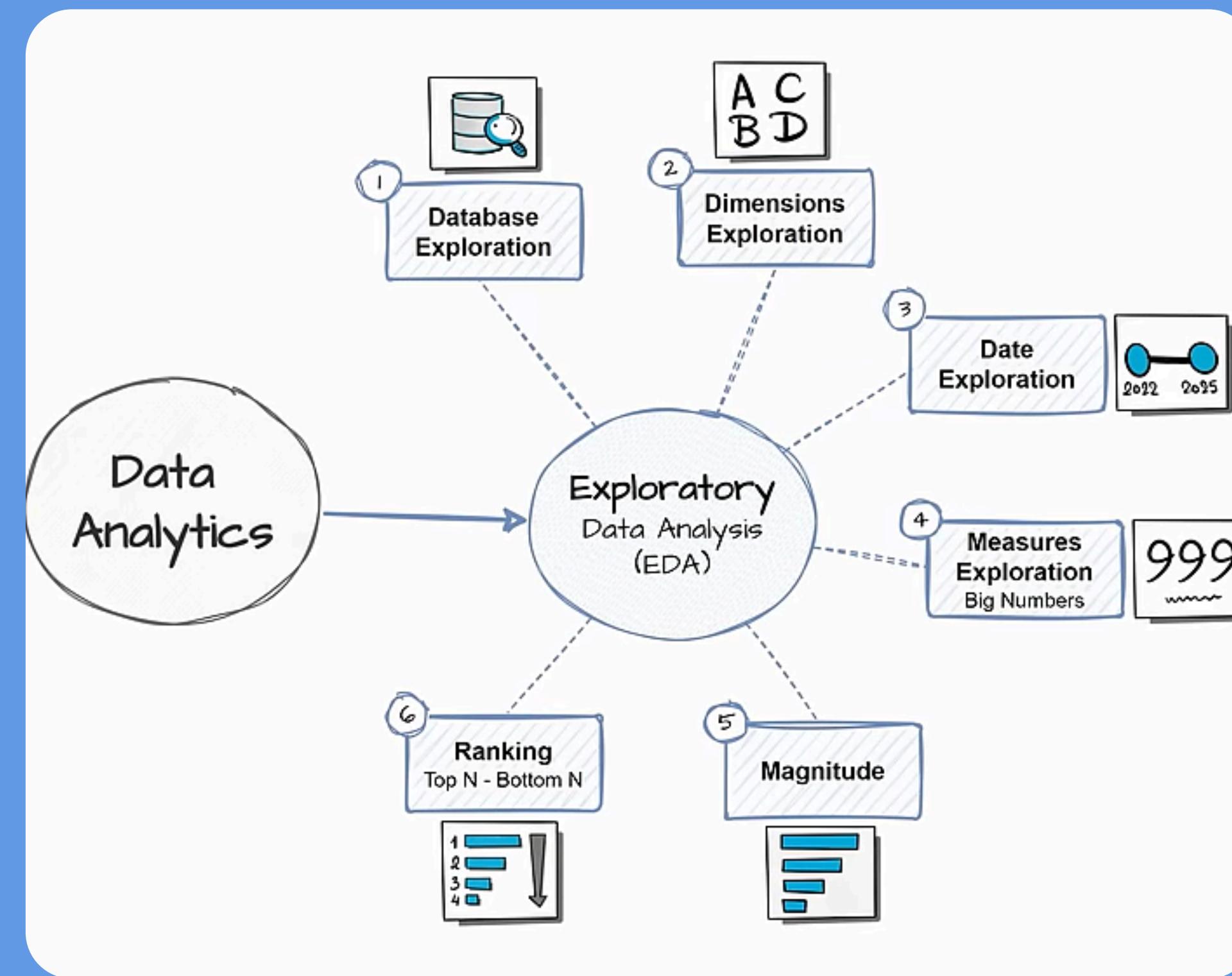
The project uses a star schema with one fact table (`fact_sales`) and two dimension tables (`dim_customers` and `dim_products`) to analyze sales, customers, and products.

- ◆ **gold.dim_customers:** Stores customer details like name, gender, country, birthdate, and marital status. Used for customer segmentation and behavior analysis.
- ◆ **gold.dim_products:** Contains product information such as category, subcategory, cost, and product lineage. Helps in product sales and pricing analysis.
- ◆ **gold.fact_sales:** Records transactional data including order details, sales amount, quantity, and prices. Used for revenue tracking and sales performance analysis.

Relationships:

- `fact_sales` links to `dim_customers` via `customer_key` and `dim_products` via `product_key`.

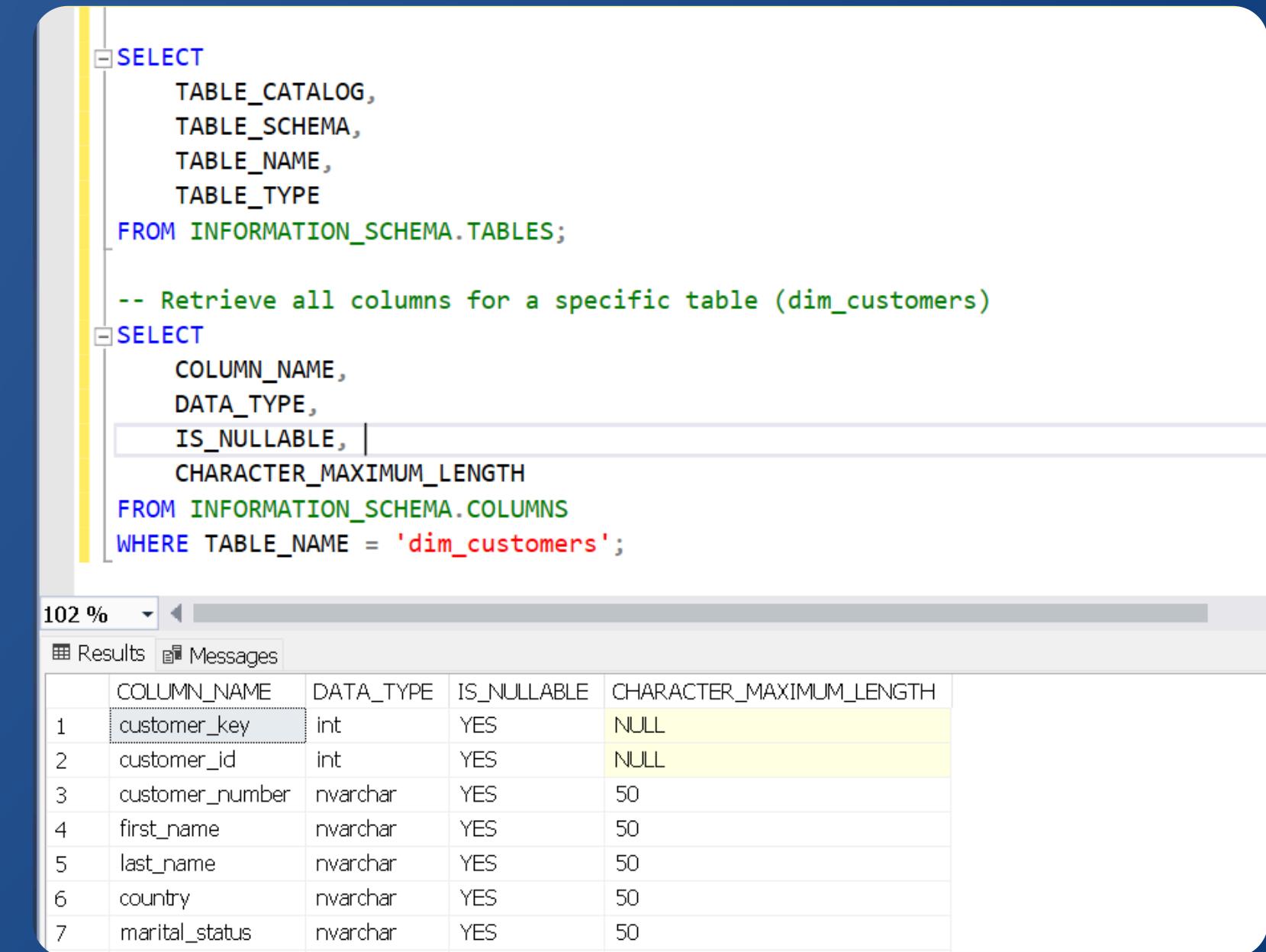
EXPLORATORY DATA ANALYSIS (EDA) FRAMEWORK



DATABASE EXPLORATION

🔍 Understanding the database schema

- List all tables in the database
- Inspect table columns and metadata
- Validate stored data



The screenshot shows a database management interface with two SQL queries in the query editor and a results grid below.

```
SELECT
    TABLE_CATALOG,
    TABLE_SCHEMA,
    TABLE_NAME,
    TABLE_TYPE
FROM INFORMATION_SCHEMA.TABLES;

-- Retrieve all columns for a specific table (dim_customers)
SELECT
    COLUMN_NAME,
    DATA_TYPE,
    IS_NULLABLE,
    CHARACTER_MAXIMUM_LENGTH
FROM INFORMATION_SCHEMA.COLUMNS
WHERE TABLE_NAME = 'dim_customers';
```

The results grid displays the columns of the 'dim_customers' table:

	COLUMN_NAME	DATA_TYPE	IS_NULLABLE	CHARACTER_MAXIMUM_LENGTH
1	customer_key	int	YES	NULL
2	customer_id	int	YES	NULL
3	customer_number	nvarchar	YES	50
4	first_name	nvarchar	YES	50
5	last_name	nvarchar	YES	50
6	country	nvarchar	YES	50
7	marital_status	nvarchar	YES	50

DIMENSIONS EXPLORATION

Analyzing dimensional tables (Customers, Products, etc.)

- List unique customer countries
- Retrieve unique product categories and subcategories

```
-- Retrieve a list of unique countries from which customers originate
SELECT DISTINCT
    country
FROM gold.dim_customers
ORDER BY country;
```

102 %

Results Messages

	country
1	Australia
2	Canada
3	France
4	Germany
5	n/a
6	United Kingdom
7	United States

```
-- Retrieve a list of unique categories, subcategories, and products
SELECT DISTINCT
    category,
    subcategory,
    product_name
FROM gold.dim_products
ORDER BY category, subcategory, product_name;
```

102 %

Results Messages

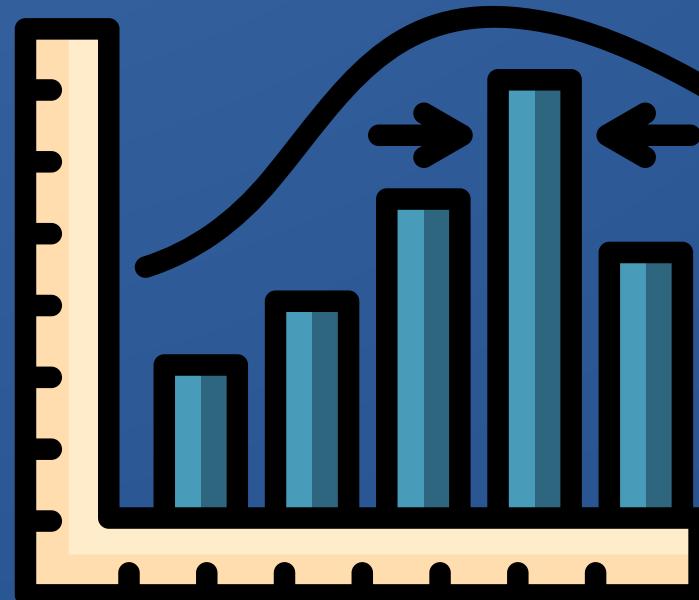
	category	subcategory	product_name
1	NULL	NULL	HL Mountain Pedal
2	NULL	NULL	HL Road Pedal
3	NULL	NULL	LL Mountain Pedal
4	NULL	NULL	LL Road Pedal
5	NULL	NULL	ML Mountain Pedal
6	NULL	NULL	ML Road Pedal
7	NULL	NULL	Touring Pedal
8	Accessories	Bike Racks	Hitch Rack - 4-Bike
9	Accessories	Bike Stands	All-Purpose Bike Stand
10	Accessories	Bottles and Cages	Mountain Bottle Cage
11	Accessories	Bottles and Cages	Road Bottle Cage
12	Accessories	Bottles and Cages	Water Bottle - 30 oz.
13	Accessories	Cleaners	Bike Wash - Dissolver
14	Accessories	Fenders	Fender Set - Mountain
15	Accessories	Helmet	Sport-100 Helmet- Black
16	Accessories	Helmet	Sport-100 Helmet- Blue

DATE RANGE ANALYSIS

JUL 17

Understanding historical data

- Identify the first and last order dates
- Calculate the total duration in months
- Analyze customer age distribution



```
--Find the date of the first and last order
--How many months of sales are available
select
    min(order_date) as first_order_date ,
    max(order_date) as last_order_date,
    DATEDIFF(MONTH,min(order_date),max(order_date)) as order_range_months
from gold.fact_sales

--Find the youngest and the oldest customer
select
    min(birthdate) as oldest_birthdate,
    DATEDIFF(YEAR,min(birthdate),GETDATE()) as oldest_age,
    max(birthdate) as youngest_birthdate,
    DATEDIFF(YEAR,max(birthdate),GETDATE()) as youngest_age
from gold.dim_customers
```

102 %

	first_order_date	last_order_date	order_range_months
1	2010-12-29	2014-01-28	37

	oldest_birthdate	oldest_age	youngest_birthdate	youngest_age
1	1916-02-10	109	1986-06-25	39

MEASURES EXPLORATION (KEY METRICS)

Evaluating business performance

- Total sales
- Total number of orders
- Average selling price



```
-- Generate a Report that shows all key metrics of the business
SELECT 'Total Sales' AS measure_name, SUM(sales_amount) AS measure_value FROM gold.fact_sales
UNION ALL
SELECT 'Total Quantity', SUM(quantity) FROM gold.fact_sales
UNION ALL
SELECT 'Average Price', AVG(price) FROM gold.fact_sales
UNION ALL
SELECT 'Total Orders', COUNT(DISTINCT order_number) FROM gold.fact_sales
UNION ALL
SELECT 'Total Products', COUNT(DISTINCT product_name) FROM gold.dim_products
UNION ALL
SELECT 'Total Customers', COUNT(customer_key) FROM gold.dim_customers;
```

102 %

	measure_name	measure_value
1	Total Sales	29356250
2	Total Quantity	60423
3	Average Price	486
4	Total Orders	27659
5	Total Products	295
6	Total Customers	18484

MAGNITUDE ANALYSIS

📌 Understanding data distribution

- Customer count by country
- Number of products in each category
- Average product cost per category

```
-- Find total customers by countries
SELECT
    country,
    COUNT(customer_key) AS total_customers
FROM gold.dim_customers
GROUP BY country
ORDER BY total_customers DESC;
```

102 %

	country	total_customers
1	United States	7482
2	Australia	3591
3	United Kingdom	1913
4	France	1810
5	Germany	1780
6	Canada	1571
7	n/a	337

```
SELECT
    category,
    COUNT(product_key) AS total_products
FROM gold.dim_products
GROUP BY category
ORDER BY total_products DESC;
```

102 %

	category	total_products
1	Components	127
2	Bikes	97
3	Clothing	35
4	Accessories	29
5	NULL	7

```
-- What is the average costs in each category?
SELECT
    category,
    AVG(cost) AS avg_cost
FROM gold.dim_products
GROUP BY category
ORDER BY avg_cost DESC;
```

102 %

	category	avg_cost
1	Bikes	949
2	Components	264
3	NULL	28
4	Clothing	24
5	Accessories	13

REVENUE ANALYSIS

\$ Measuring revenue performance

- Total revenue by category
- Revenue per customer

```
-- What is the total revenue generated for each category?  
SELECT  
    p.category,  
    SUM(f.sales_amount) AS total_revenue  
FROM gold.fact_sales f  
LEFT JOIN gold.dim_products p  
    ON p.product_key = f.product_key  
GROUP BY p.category  
ORDER BY total_revenue DESC;
```

102 %

Results Messages

	category	total_revenue
1	Bikes	28316272
2	Accessories	700262
3	Clothing	339716

```
-- What is the total revenue generated by each customer?  
SELECT  
    c.customer_key,  
    c.first_name,  
    c.last_name,  
    SUM(f.sales_amount) AS total_revenue  
FROM gold.fact_sales f  
LEFT JOIN gold.dim_customers c  
    ON c.customer_key = f.customer_key  
GROUP BY  
    c.customer_key,  
    c.first_name,  
    c.last_name  
ORDER BY total_revenue DESC;
```

102 %

Results Messages

	category	total_revenue
1	Bikes	28316272
2	Accessories	700262
3	Clothing	339716

RANKING ANALYSIS

🏆 Identifying top & low-performing products and customers

- Top 3 highest revenue-generating category
- The 3 customers with the fewest orders placed

```
-- Which 3 category Generating the Highest Revenue?  
-- Simple Ranking  
select top 3  
p.category , sum(f.sales_amount) total_revenue  
from gold.dim_products p right join gold.fact_sales f  
on p.product_key = f.product_key  
group by p.category  
order by total_revenue desc
```

102 %

Results Messages

	category	total_revenue
1	Bikes	28316272
2	Accessories	700262
3	Clothing	339716

```
-- The 3 customers with the fewest orders placed  
SELECT TOP 3  
c.customer_key,  
c.first_name,  
c.last_name,  
COUNT(DISTINCT order_number) AS total_orders  
FROM gold.fact_sales f  
LEFT JOIN gold.dim_customers c  
ON c.customer_key = f.customer_key  
GROUP BY  
c.customer_key,  
c.first_name,  
c.last_name  
ORDER BY total_orders ;
```

102 %

Results Messages

	customer_key	first_name	last_name	total_orders
1	21	Jordan	King	1
2	41	Jesse	Murphy	1
3	37	Jennifer	Russell	1



KEY FINDINGS & RECOMMENDA- TIONS

01

✓ Enhance Marketing Strategies – Focus on promoting underperforming products through targeted campaigns, discounts, and improved visibility to boost sales.

02

✓ Expand in High-Value Markets – Identify and target regions with high-spending customers by tailoring offers, optimizing distribution channels, and improving customer engagement.

03

✓ Optimize Pricing for Profitability – Analyze pricing trends, adjust product prices based on demand and competition, and implement dynamic pricing strategies to maximize profit margins.



THANK YOU



01009218738



amrmashaly935@gmail.com