Data analytics reference guide

Juraj Grešo | 08.08.2025

Step 0: Database exploration

Explore tables and columns: INFORMATION_SCHEMA.TABLES, INFORMATION_SCHEMA.COLUMNS

Step 1: Exploratory data analysis

Step	Focus	Formula	Functions	Example
1.1	Dimensions	DISTINCT [Dimension]	DISTINCT ORDER BY	DISTINCT country ORDER BY country
1.2	Dates	MIN/MAX [Date Dimension]	MIN/MAX DATEDIFF	DATEDIFF(YEAR, MAX(birthdate), GETDATE())
1.3	Measures	Σ [Measure]	AVG COUNT SUM	AVG(price) COUNT(DISTINCT order_number) SUM(quantity)
1.4	Magnitude	Σ [Measure] by [Dimension]	AVG COUNT SUM GROUP BY ORDER BY	AVG(cost) COUNT(customer_key) SUM(sales) AS total_sales GROUP BY customer_key ORDER BY total_sales
1.5	Ranking	Σ [Measure] by [Dimension]	DENSE_RANK() RANK() ROW_NUMBER() TOP GROUP BY ORDER BY	SELECT TOP 5 product_name, SUM(sales) AS total_sales FROM fact LEFT JOIN dim ON key GROUP BY product_name ORDER BY total_sales DESC

Step 2: Advanced data analysis

Step	Focus	Formula	Functions	Example
2.1	Change-over-Time	Σ [Measure] by [Date Dimension]	AVG COUNT SUM DATEPART() DATETRUNC() FORMAT	DATETRUNC(month, order_date) AS order_date FORMAT(order_date, 'yyyy-MMM') AS order_date
2.2	Cumulative	Σ [Cumulative Measure] by [Date Dimension]	SUM() OVER() AVG() OVER()	SUM(total_sales) OVER (ORDER BY order_date) AS running_total_sales
2.3	Performance	Current [Measure] - Target [Measure]	AVG() OVER() LAG() CASE	Current - Average Current year - Previous year Current - Highest
2.4	Part-to-Whole	([Measure] / [Total Measure]) * 100 by [Dimension]	AVG SUM SUM() OVER()	SUM(sales) AS total_category SUM(total_category) OVER() AS total_sales
2.5	Segmentation	[Measure] by [Measure]	CASE GROUP BY	CASE WHEN price <= 100 THEN 'Low' ELSE 'High' END AS price_flag

Step 3: Reports

Create reports for customer, product, geography, sales channel, and more or less as needed.

```
CREATE VIEW report AS
WITH base AS (fact + dim),
aggregation AS (SUM + COUNT FROM base)
SELECT aggregation + segmentation + KPIs)
```