Supply Chain Performance Analysis and Optimization

- Team Members: -Kholoud Hagag -Reham Sobeeh

-Mahmoud Seera -Ahmed Hussein

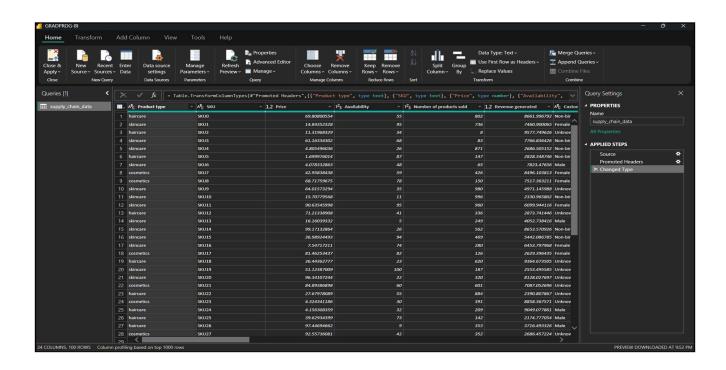
- Team Leader: Kholoud Hagag

- Date Submitted: 3/10/2025

Checking the Data (Assumptions & Data Quality):

What has been checked	What we found	What we did	
One row header	As a CSV file, so it's text	Promoted headers	
Missing values	No missing values		
Duplicates	No duplicates		
Totals	No totals		
Column types	Was all text	Changed columns types	
Columns were fully	24 columns are too hard to	Created a dictionary defining each column	
understood	get by only looking through		
	the data		
Primary key	Unique column SKU is already		
	in the data		

Applied steps:



First: promoted headers→from trasnform →use the first row as headers.

Second: changed data types → from transform → detect data type.

Data dictionary:

1	Column Name	Description	Data Type
2	Product type	The category of the product ("cosmatics", "haircare," "skincare").	Text
3	SKU	Stock Keeping Unit, a unique code for each product.	Text
4	Price	The price of a single unit of the product.	Decimal Number
5	Availability	The number of products currently in stock and ready to be sold.	Whole Number
6	Number of products sold	The total quantity of products sold.	Whole Number
7	Revenue generated	The total revenue earned from the sale of the product.	Decimal Number
8	Customer demographics	Information about the customer, such as gender or age group.	Text
9	Stock levels	The current inventory level for the product.	Whole Number
10	Lead times	The duration from order placement to delivery.	Whole Number
11	Order quantities	The number of units ordered in a single transaction.	Whole Number
12	Shipping times	The time taken to ship the product to the customer.	Whole Number
13	Shipping carriers	The company responsible for transporting the products ("carrier A", "carrier B", "carrier C").	Text
14	Shipping costs	The cost associated with shipping the product.	Decimal Number
15	Supplier name	The name of the product's supplier.	Text
16	Location	The geographic location of the supply chain node.	Text
17	Lead time	The time it takes for a supplier to deliver products.	Whole Number
18	Production volumes	The number of units produced.	Whole Number
19	Manufacturing lead time	The time it takes to manufacture the product.	Whole Number
20	Manufacturing costs	The cost of producing the product.	Decimal Number
21	Inspection results	The outcome of the quality inspection ("Pass," "Fail," "Pending").	Text
22	Defect rates	The percentage of products with defects.	Decimal Number
23	Transportation modes	The method of transport used ("Road," "Rail," "Air").	Text
24	Routes	The specific path or route used for transportation.	Text
25	Costs	The overall cost associated with the product.	Decimal Number

Finally:-

- We obtained clean data with 24 columns and 100 rows, with each row having its unique primary key.
- Created a dictionary describing every column.
- Having a model that contains one table ready for to next step of calculating.