

Project DOCUMENTATION

By : Abdelneem Alaref





Welcome to My Project

As a BI Engineer, passionate about data. I would like to share my journey through a challenging task.

The task was to analyze and transform retail invoice data collected monthly from February to May.



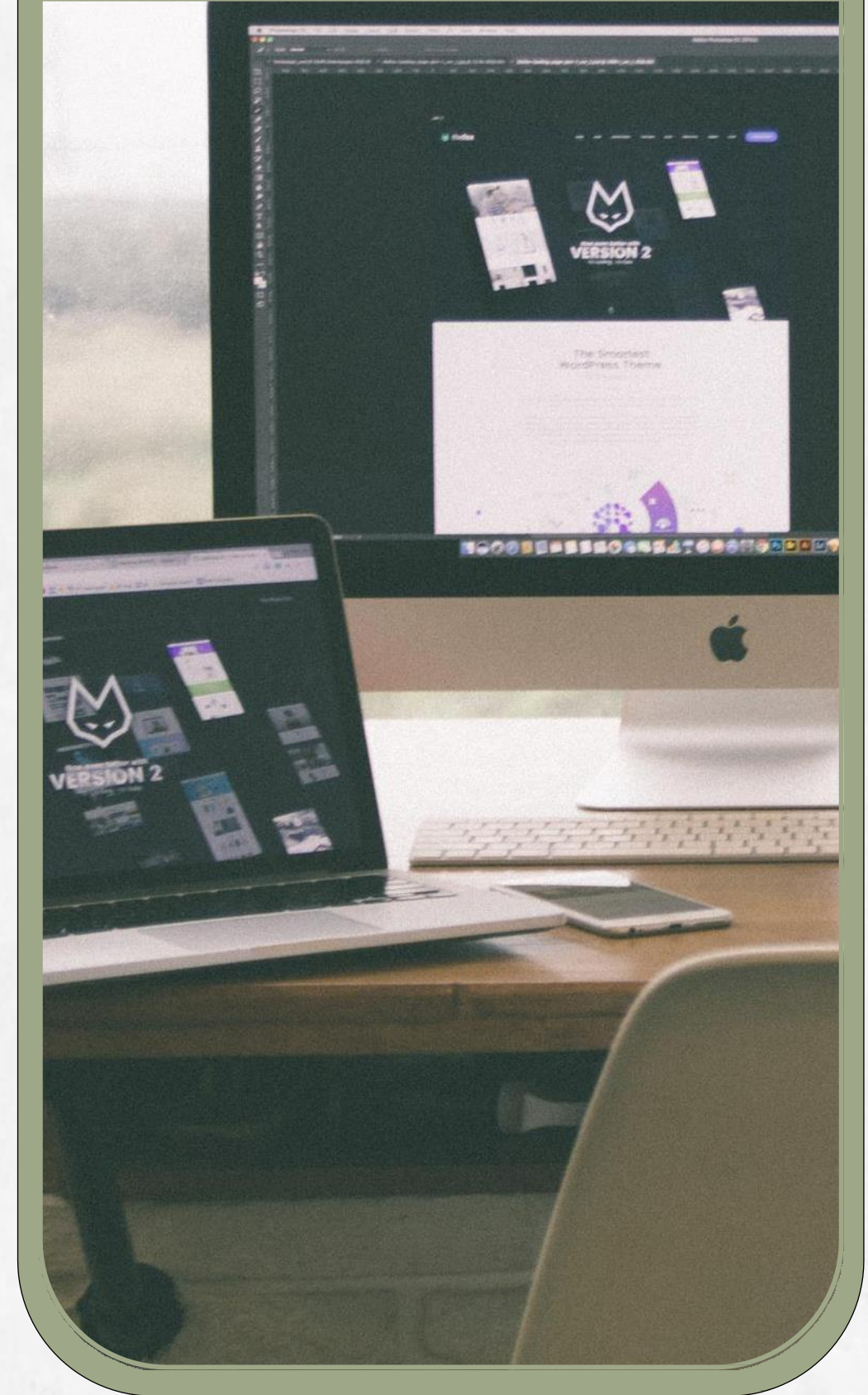
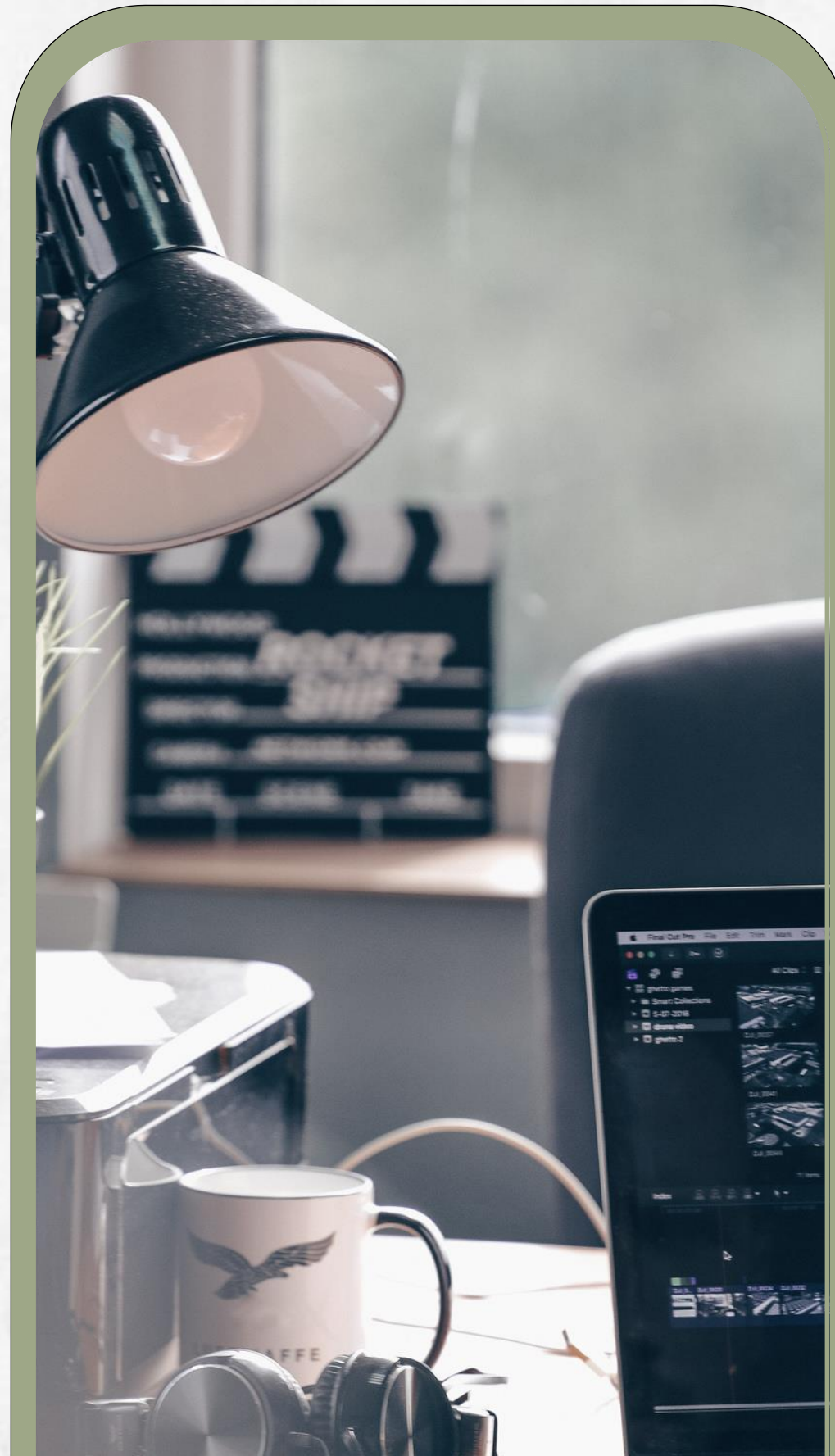
Welcome to The Data

The data was divided into four sheets: Sales, Products, Reps, and Stores.

The goal was to create a comprehensive solution using various tools and present it effectively.

STEPS INDEX

- Excel
- SQL
- SSIS
- Power Query
- Power bi
- Presentation

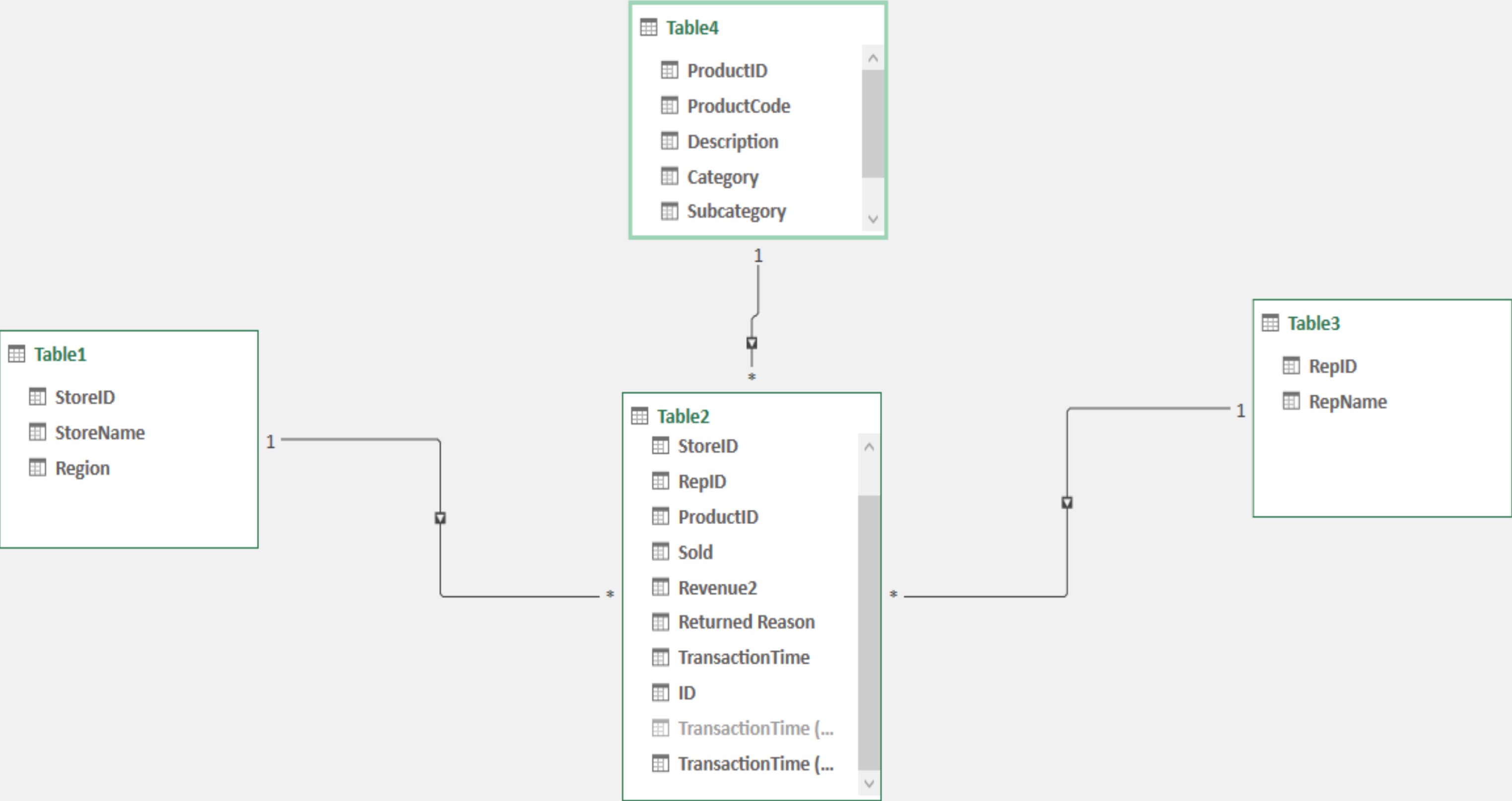


EXCEL

- Reading the data
- Cleaning Data
- Creating some calculations and Data Model



J12																		
1																		
2																		
3		Product ID	Sold Amount															
4		11	132															
5		386	127				Worest Brand	Sales			Top Brand	Sales		Store/Month	Sales			
6		392	97				BOBOO	\$123			Goo Jit Zu	\$47,035		StoreA				
7		462	1193				CRACKERS	\$73			L.O.L. Surprise	\$162,750		Feb	\$83,101			
8		465	286				Craftabelle	\$225			Match Attax	\$55,665		Mar	\$128,077			
9		470	227		RepName	Orders	CUP CAKE	\$80			Nilco	\$64,103		Apr	\$206,545			
10		471	183		Ahmed	909	Magic Mixies	\$0			Our Generation	\$53,507		May	\$3,910			
11		474	106		Alaa	534	POP CORN	\$183			Pokémon	\$160,569		StoreB				
12		475	138		Asya	409	Slice CAKE	\$40			Rainbocorns	\$109,388		Feb	\$17,305			
13		476	115		Bassem	1075	SNACK & GO	\$68			Rainbow High	\$78,239		Mar	\$65,618			
14		4028	421		Essam	1178	So Glow	\$0			Style 4 Ever	\$42,918		Apr	\$98,901			
15		4078	404		Nora	473	Sticks	\$20			X-Shot	\$129,223		May	\$2,718			
16		4079	325		Ola	369	Grand Total	\$810			Grand Total	\$903,395		StoreC				
17		4087	93		Omar	596								Feb	\$124,316			
18		4088	107		Shady	646								Mar	\$156,030			
19		4089	91		Yara	337	Store/Reps	Sales			Brand/Product	Sales		Apr	\$305,106			
20		4090	327		Grand Total	6526	StoreA				4M			May	\$2,569			
21		4091	176				Ahmed	\$31,266			Education & Learning	\$3,356		StoreD				
22		4092	188				Ali	-\$825			Akedo			Feb	\$35,859			
23		4202	100				Asmaa	-\$288			Games	\$1,925		Mar	\$66,953			
24		Grand Total	4836				Asya	\$68,149			Art Lab			Apr	\$151,314			
25							Bassem	\$20,833			Education & Learning	\$8,550		May	\$742			
26							Esraa	\$45,361			ASTRA			StoreE				
27							Loaa	\$39,215			F&B	\$495		Feb	\$17,561			
28							Mariam	\$55,414			Audi			Mar	\$24,420			
29							Nada	\$44,961			Season	\$705		Apr	\$32,975			





SQL

- Creating a Data Warehouse
- Hosting the stare schema and creating the modeling of DWH

Object Explorer

Connect

DESKTOP-1SKMSHS\SQLEXPRESS (SQL)

Databases

System Databases

Database Snapshots

AdventureWorks2014

AdventureWorksDW2017

AdventureWorksDW2019

AdventureWorksLT2019

CustomerChurn

EO_AdventureWorksDW2019

UrbanStores

SalesDWH

Database Diagrams

Tables

System Tables

FileTables

External Tables

Graph Tables

dbo.Dim_Product

dbo.Dim_Reps

dbo.Dim_Stores

dbo.Fact_Sales

Views

External Resources

Synonyms

Programmability

Query Store

Service Broker

Storage

SQLQuery8.sql - D...1SKMSHS\DELL (64))

DESKTOP-1SKMSHS...sDWH - Diagram_0*

SQLQuery7.sql - D...1SKMSHS\DELL (63))

SQLQuery6.sql - D...1SKMSHS\DELL (62))

SELECT TOP (1000) [OrderNumer]

, [StoreID]

, [RepID]

, [ProductID]

, [Sold]

, [Revenue2]

, [TransactionTime]

, [ID]

FROM [SalesDWH].[dbo].[Fact_Sales]

100 %

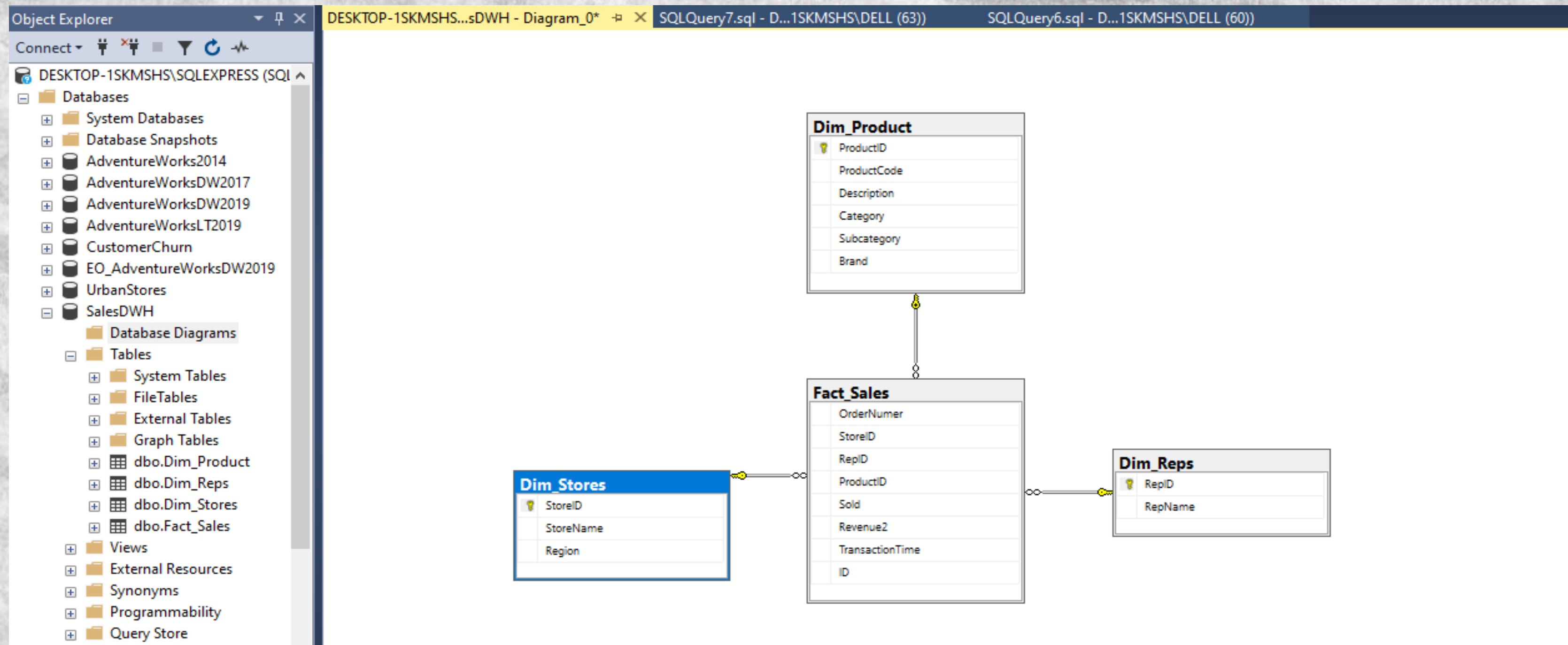
Results

Messages

	OrderNumer	StoreID	RepID	ProductID	Sold	Revenue2	TransactionTime	ID
1	55	3	30	3532	1	200	2023-02-02 20:17:00.0000000	3-55
2	60	3	30	687	1	750	2023-02-02 21:55:00.0000000	3-60
3	61	3	30	3510	1	213	2023-02-02 22:03:00.0000000	3-61
4	61	3	30	3509	1	213	2023-02-02 22:03:00.0000000	3-61
5	62	3	30	678	1	188	2023-02-02 22:06:00.0000000	3-62
6	69	3	30	672	1	150	2023-02-02 22:40:00.0000000	3-69
7	74	3	30	3360	1	975	2023-02-03 00:05:00.0000000	3-74
8	79	1	15	182	1	138	2023-02-02 16:40:00.0000000	1-79
9	79	3	27	3479	1	875	2023-02-03 13:38:00.0000000	3-79
10	79	3	27	4078	1	50	2023-02-03 13:38:00.0000000	3-79
11	79	3	27	4078	1	50	2023-02-03 13:38:00.0000000	3-79

Star Schema

It enhances analytical reporting by providing faster query response times and easier data navigation due to its simplified design.



SSIS



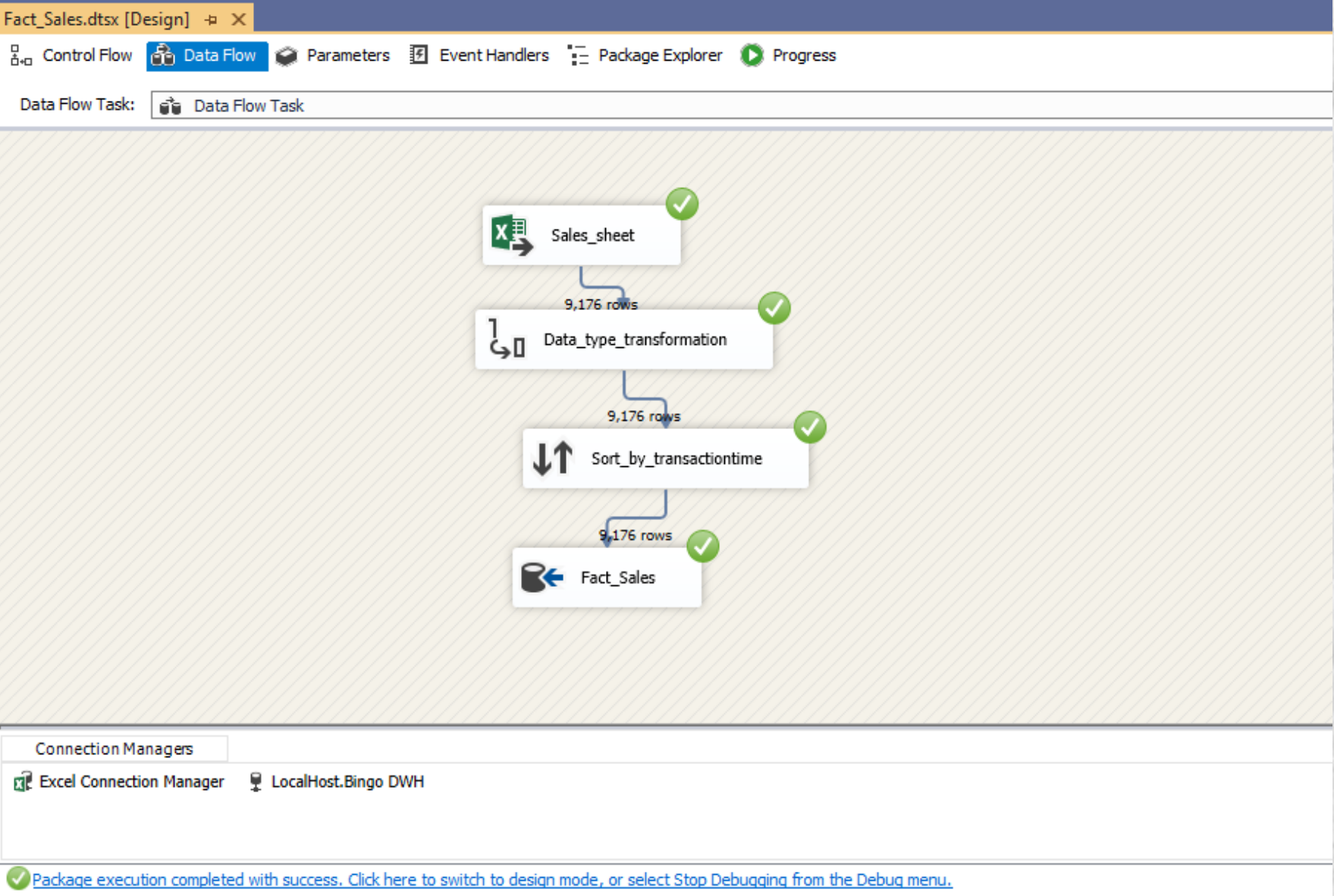
Creating data types transformation and sorting the data



Defining the fact tables and dimension tables

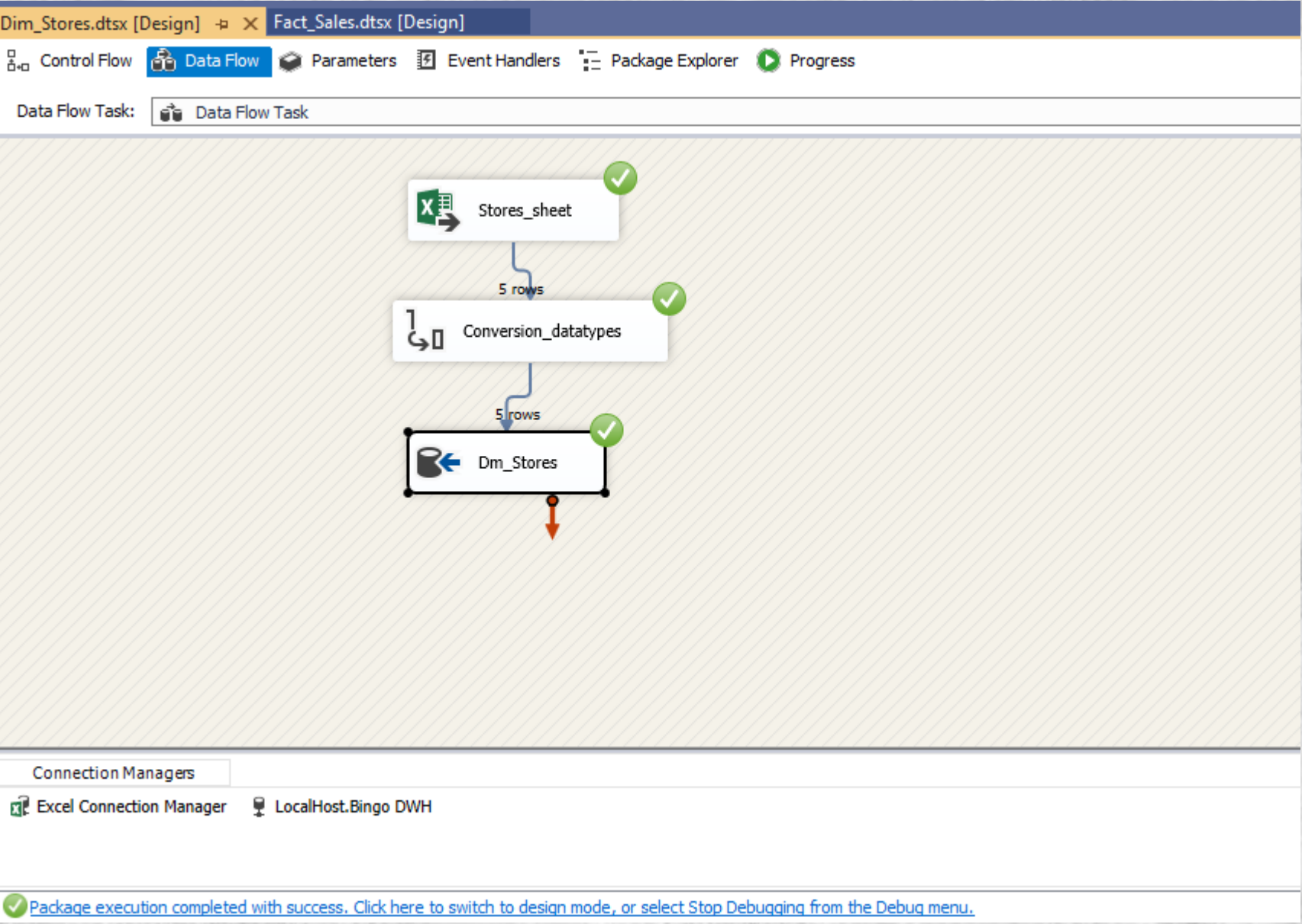


Creating ETL pipelines to populate the data into the DWH

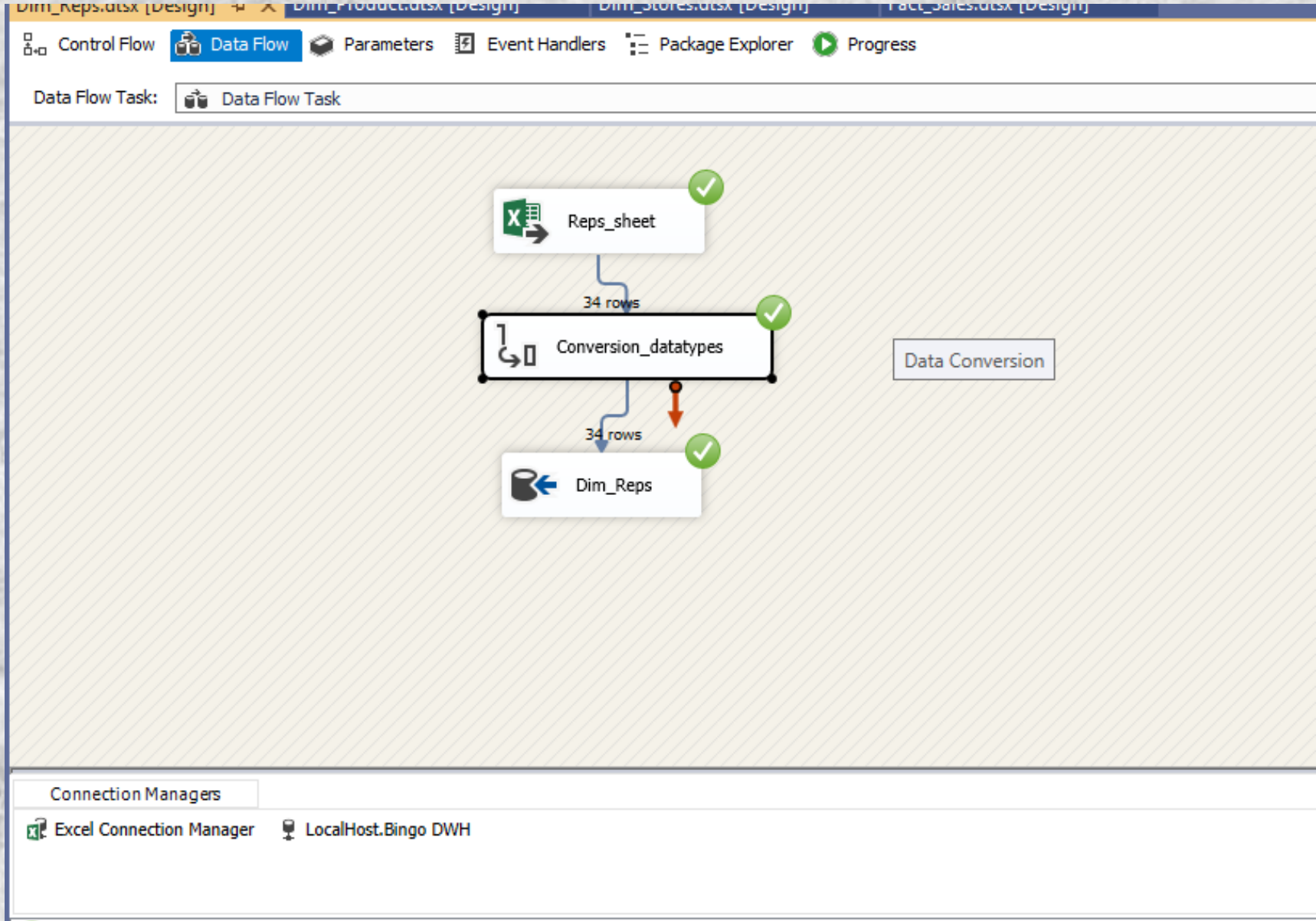


Fact Sales SSIS Package

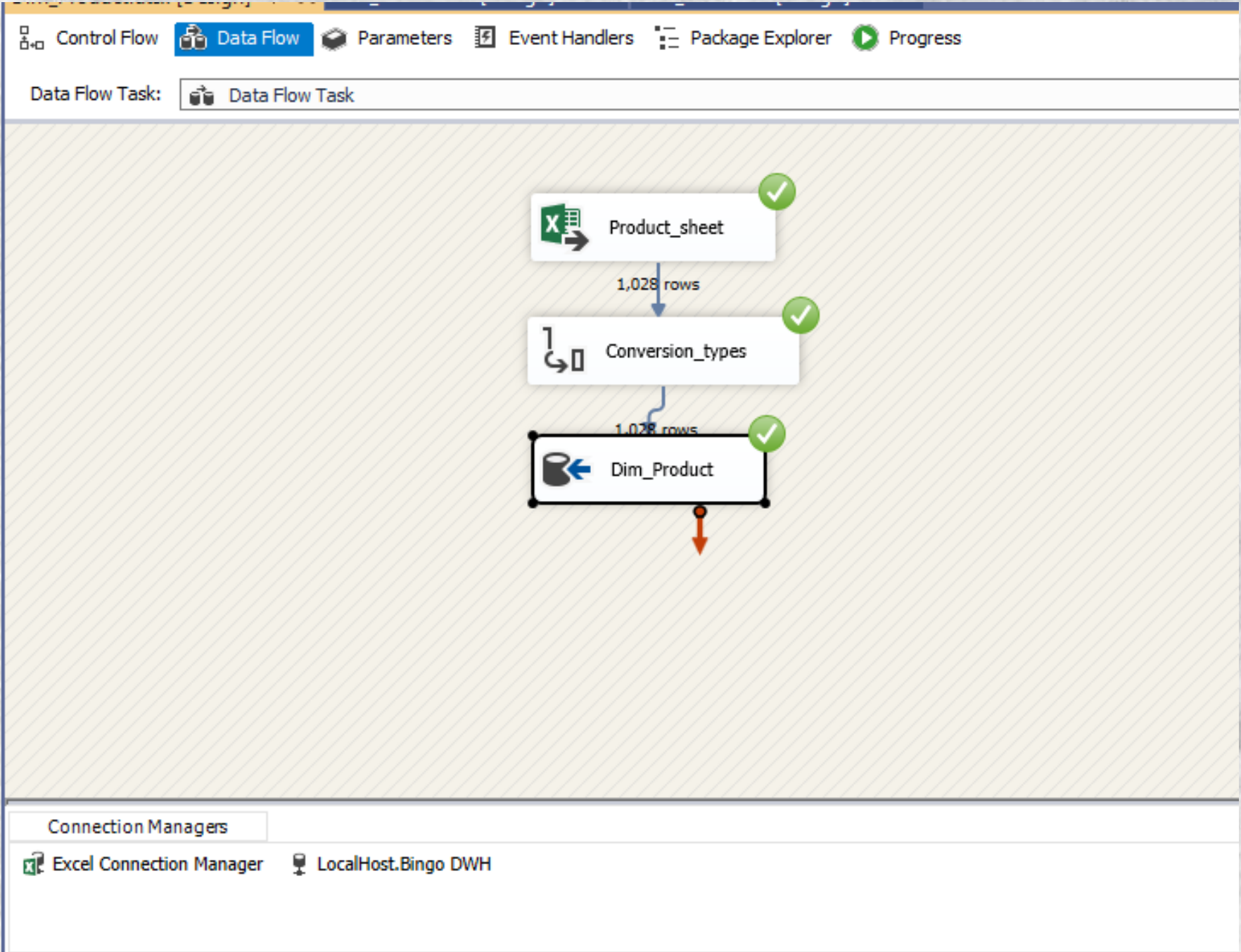
Store Dimension SSIS Package



Product Dimension SSIS Package



Reps Dimension SSIS Package





Power BI

- Explaining our process of thinking and working during the task
- Providing a brief presentation and dashboard

Power Query Processes

FileHomeTransformAdd ColumnViewToolsHelp

Close & Apply

Close

New Source

New Query

Recent Sources

New Query

Enter Data

Data source settings

Data Sources

Manage Parameters

Parameters

Refresh Preview

Query

Properties

Advanced Editor

Choose Columns

Manage Columns

Remove Columns

Manage Columns

Keep Rows

Reduce Rows

Remove Rows

Reduce Rows

Sort

Split Column

Group By

Data Type: Text

Use First Row as Headers

Replace Values

Merge Queries

Combine

Append Queries

Combine

Combine Files

Combine

Text Analytics

AI Insights

Vision

AI Insights

Azure Machine Learning

AI Insights

Queries [4]

Dim_Product

Dim_Reps

Dm_Stores

Fact_Sales

TransactionTime

ID

Dim_Product

Dim_Reps

Dm_Stores

1

ull

2/2/2023 4:40:32 PM

1-79

Value

Value

Value

2

ull

2/2/2023 6:43:17 PM

1-94

Value

Value

Value

3

ull

2/2/2023 8:17:06 PM

3-55

Value

Value

Value

4

ull

2/2/2023 9:30:43 PM

1-105

Value

Value

Value

5

ull

2/2/2023 9:30:43 PM

1-105

Value

Value

Value

6

ull

2/2/2023 9:55:35 PM

3-60

Value

Value

Value

7

ull

2/2/2023 10:03:29 PM

3-61

Value

Value

Value

8

ull

2/2/2023 10:03:29 PM

3-61

Value

Value

Value

9

ull

2/2/2023 10:06:17 PM

3-62

Value

Value

Value

10

ull

2/2/2023 10:40:06 PM

3-69

Value

Value

Value

11

ull

2/3/2023 12:05:53 AM

3-74

Value

Value

Value

12

ull

2/3/2023 12:47:38 PM

1-112

Value

Value

Value

13

ull

2/3/2023 12:49:39 PM

1-113

Value

Value

Value

14

ull

2/3/2023 1:38:18 PM

3-79

Value

Value

Value

15

ull

2/3/2023 1:38:18 PM

3-79

Value

Value

Value

16

ull

2/3/2023 1:38:18 PM

3-79

Value

Value

Value

17

ull

2/3/2023 1:38:18 PM

3-79

Value

Value

Value

18

ull

2/3/2023 2:28:31 PM

3-90

Value

Value

Value

19

ull

2/3/2023 2:28:31 PM

3-90

Value

Value

Value

20

ull

2/3/2023 2:43:16 PM

1-127

Value

Value

Value

21

ull

2/3/2023 4:46:45 PM

1-146

Value

Value

Value

22

ull

2/3/2023 4:51:36 PM

1-147

Value

Value

Value

23

ull

2/3/2023 4:52:22 PM

3-110

Value

Value

Value

24

Fact_Sales

Properties

Name

Fact_Sales

All Properties

Applied Steps

source

Navigation

Changed Type

Added Conditional Column

Added Conditional Column1

Reordered Columns

Changed Type1

Removed Columns

Added Conditional Column2

Changed Type2

Removed Columns1

Inserted Week of Month

Reordered Columns1

Removed Columns2

Inserted Year

Inserted Month Name

Inserted Week of Month1

Inserted Day

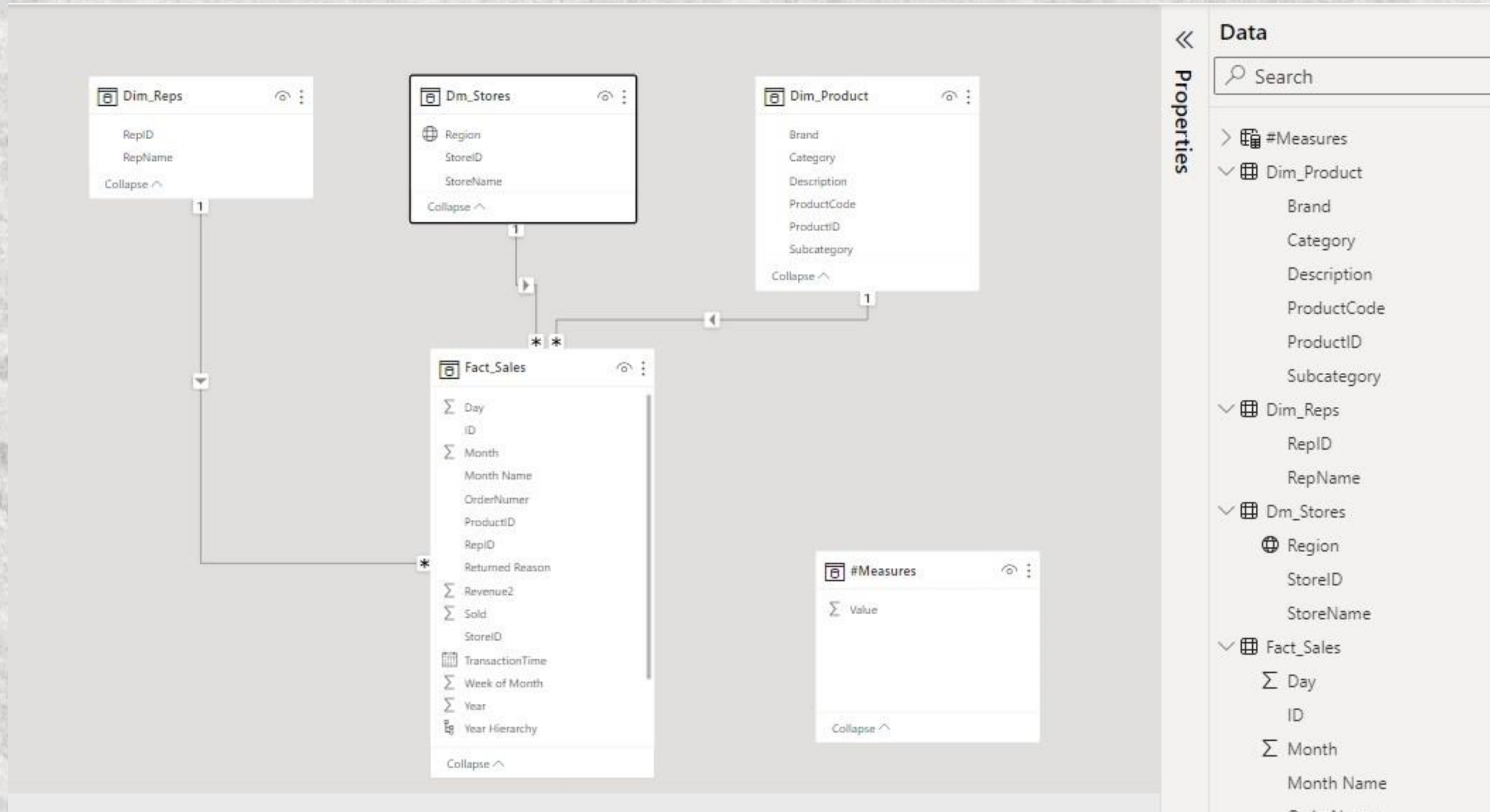
Inserted Month

17 COLUMNS, 999+ ROWS

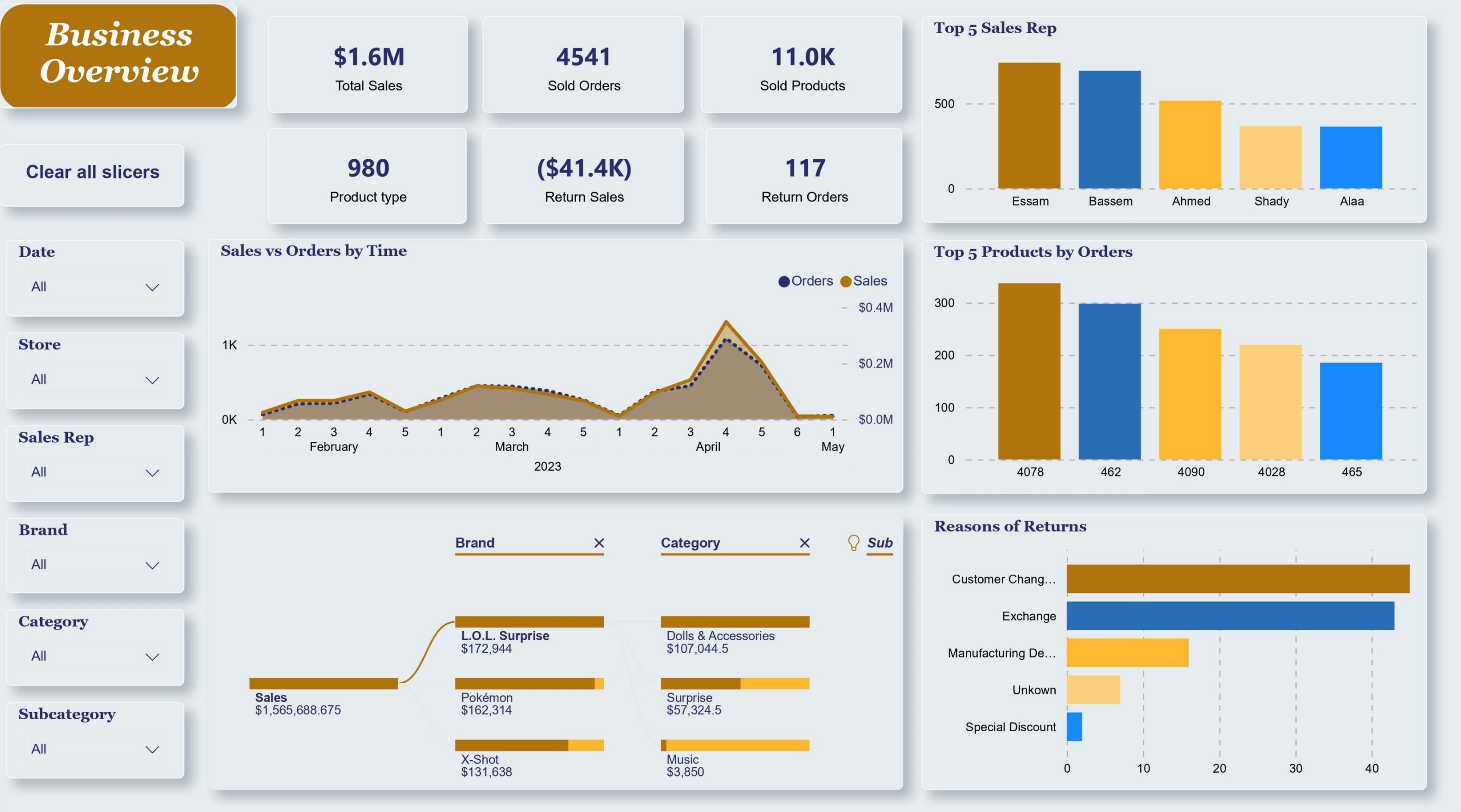
Column profiling based on top 1000 rows

PREVIEW DOWNLOADED AT 12:06 PM

Data Modeling Processes



Power BI Dashboard





INSIGHTS

- What is the Total Sales?
- What is the Total Orders?
- How much are the products sold?
- How much are the products returned?



INSIGHTS

- What are the Top Products sold?
- Who are the top sales representatives?
- How are the sales and orders by time?
- What are the most common reasons for returns?



THANKYOU

Abdelnaeem Alaref