

1. This dataset comprises several Excel files related to sales, including dimensions for customers, stock items, cities, dates, and employees, as well as a fact table for sales transactions. Create a Data Lakehouse in Microsoft fabric. Upload this dataset to the Microsoft fabric environment using the Upload Files option or Data Flow Gen2.

The top screenshot shows the Microsoft Fabric Lakehouse interface for 'Assignment_4_lakehouse'. The Explorer pane on the left shows a folder structure with 'Tables' and 'Files'. The main area displays a large paperclip icon and the text 'Add Content to this Folder'. Below this, it says 'You can create subfolders or upload files to this folder. You can also choose from other data sources.' and there is a 'Get data' button.

The bottom screenshot shows a GitHub repository named 'Data-Warehousing / Assignment 4'. The file list on the left shows several CSV files: DimCity.csv, DimCustomer.csv, DimDate.csv, DimEmployee.csv, DimStockitem.csv, and FactSale.csv. The main area shows a table of commit history for these files.

Name	Last commit message	Last commit date
..		
DimCity.csv	add csv file	12 hours ago
DimCustomer.csv	add csv file	12 hours ago
DimDate.csv	add csv file	12 hours ago
DimEmployee.csv	add csv file	12 hours ago
DimStockitem.csv	add csv file	12 hours ago
FactSale.csv	add csv file	12 hours ago

2. Apply the necessary preprocessing and transforming steps and create files in Data Lake Delta Tables. To do this ETL process use Data Flow Gen2 or Fabric Notebooks.

The image displays two screenshots of the Microsoft Fabric Dataflow Gen2 interface, showing the transformation of data from a source table to a target table.

Top Screenshot: The interface shows a dataflow named "Dataflow 1" with a query titled "Table.TransformColumnTypes". The query is applied to a table named "FactSale". The query settings show the "Entity type" as "FactSale" and the "Data destination" as "Lakehouse". The query is completed in 1.23 seconds, with 21 columns and 99+ rows.

Sale Key	City Key	Customer Key	Bill To Customer Key	Stock Item Key	Invoice Date Key	Delivery Date Key	Salesperson Key	WWI
49265	41568	0	0	204	10/22/2013	10/23/2013	83	
49372	48937	0	0	173	10/22/2013	10/23/2013	83	
50379	91464	0	0	174	10/28/2013	10/29/2013	70	
51852	48937	0	0	171	11/6/2013	11/7/2013	74	
52374	89925	0	0	173	11/6/2013	11/9/2013	39	
53463	72709	0	0	183	11/15/2013	11/16/2013	91	
53591	87633	0	0	202	11/16/2013	11/17/2013	91	
54792	77092	0	0	176	11/22/2013	11/23/2013	91	
55582	47692	0	0	175	11/28/2013	11/29/2013	19	
56531	47134	0	0	191	12/3/2013	12/4/2013	91	
57267	72894	0	0	188	12/7/2013	12/8/2013	81	
57869	91464	0	0	169	12/12/2013	12/13/2013	39	
57915	47692	0	0	166	12/12/2013	12/13/2013	86	
58792	48937	0	0	171	12/19/2013	12/20/2013	39	
49620	87481	0	0	182	10/24/2013	10/25/2013	19	

Bottom Screenshot: The interface shows a dataflow named "Dataflow 1" with a query titled "Table.ReplaceValue". The query is applied to a table named "DimStockItem". The query settings show the "Entity type" as "DimStockItem" and the "Data destination" as "Lakehouse". The query is completed in 1.29 seconds, with 21 columns and 99+ rows.

Stock Item Key	WWI Stock Item ID	Stock Item	Color	Selling Package	Buying Package	Brand	Size
1	219	Void fill 400 L bag (White) 400L	N/A	Each	Each	N/A	400L
2	218	Void fill 300 L bag (White) 300L	N/A	Each	Each	N/A	300L
3	217	Void fill 200 L bag (White) 200L	N/A	Each	Each	N/A	200L
4	216	Void fill 100 L bag (White) 100L	N/A	Each	Each	N/A	100L
5	215	Air cushion machine (Blue)	N/A	Each	Each	N/A	N/A
6	214	Air cushion film 200mmx200mm 325m	N/A	Each	Each	N/A	325m
7	213	Air cushion film 200mmx100mm 325m	N/A	Each	Each	N/A	325m
8	212	Large replacement blades 18mm	N/A	Each	Each	N/A	18mm
9	211	Small 9mm replacement blades 9mm	N/A	Each	Each	N/A	9mm
10	210	Packing knife with metal insert blade (Yellow) 18mm	N/A	Each	Each	N/A	18mm
11	209	Packing knife with metal insert blade (Yellow) 9mm	N/A	Each	Each	N/A	9mm
12	208	Permanent marker red 5mm nib (Red) 5mm	N/A	Each	Each	N/A	5mm
13	207	Permanent marker blue 5mm nib (Blue) 5mm	N/A	Each	Each	N/A	5mm
14	206	Permanent marker black 5mm nib (Black) 5mm	N/A	Each	Each	N/A	5mm
15	205	Tape dispenser (Blue)	N/A	Each	Each	N/A	N/A

Power BI interface showing a Dataflow query named "Table.Skip (#'Changed column type', 1)". The query is applied to a table with columns: Employee Key, WWI Employee ID, Employee, Preferred Name, Is Salesperson, Photo, Valid From, Valid To, and Lineage Key. The data is displayed in a table view with 16 rows and 9 columns. The query settings panel on the right shows the entity type as "DimEmployee" and the data destination as "Lakehouse".

Employee Key	WWI Employee ID	Employee	Preferred Name	Is Salesperson	Photo	Valid From	Valid To	Lineage Key
1	1	Lily Code	Lily	0	NULL	00:00:00	00:00:00	3
2	2	Isabella Rupp	Isabella	0	NULL	00:00:00	00:00:00	3
3	3	Ethan Onslow	Ethan	0	NULL	00:00:00	00:00:00	3
4	4	Amy Trell	Amy	1	NULL	00:00:00	00:00:00	3
5	5	Jai Shand	Jai	0	NULL	00:00:00	00:00:00	3
6	6	Anthony Grosse	Anthony	1	NULL	00:00:00	00:00:00	3
7	7	Taj Shand	Taj	1	NULL	00:00:00	00:00:00	3
8	8	Hudson Hollinworth	Hudson	1	NULL	00:00:00	00:00:00	3
9	9	Jack Potter	Jack	1	NULL	00:00:00	00:00:00	3
10	10	Piper Koch	Piper	0	NULL	00:00:00	00:00:00	3
11	11	Hudson Onslow	Hudson	1	NULL	00:00:00	00:00:00	3
12	12	Sophia Hinton	Sophia	1	NULL	00:00:00	00:00:00	3
13	13	Henry Forlonge	Henry	0	NULL	00:00:00	00:00:00	3
14	14	Stella Rosenhain	Stella	0	NULL	00:00:00	00:00:00	3
15	15	Kayla Wibodcock	Kayla	1	NULL	00:00:00	00:00:00	3
16	16	Katie Darwin	Katie	0	NULL	00:00:00	00:00:00	3

Power BI interface showing a Dataflow query named "Table.TransformColumnTypes (#'Promoted headers', {('Date', type date), ('Day Number', Int64.Type), ('Day', Int64.Type), ('Month', type date), ('Short Month', type date), ('Calendar Month Number', type date), ('Calendar Month Label', type date), ('Calendar Year', type date), ('Calendar Year Label', type date)}). The query is applied to a table with columns: Date, Day Number, Day, Month, Short Month, Calendar Month Number, Calendar Month Label, Calendar Year, and Calendar Year Label. The data is displayed in a table view with 16 rows and 9 columns. The query settings panel on the right shows the entity type as "DimDate" and the data destination as "Lakehouse".

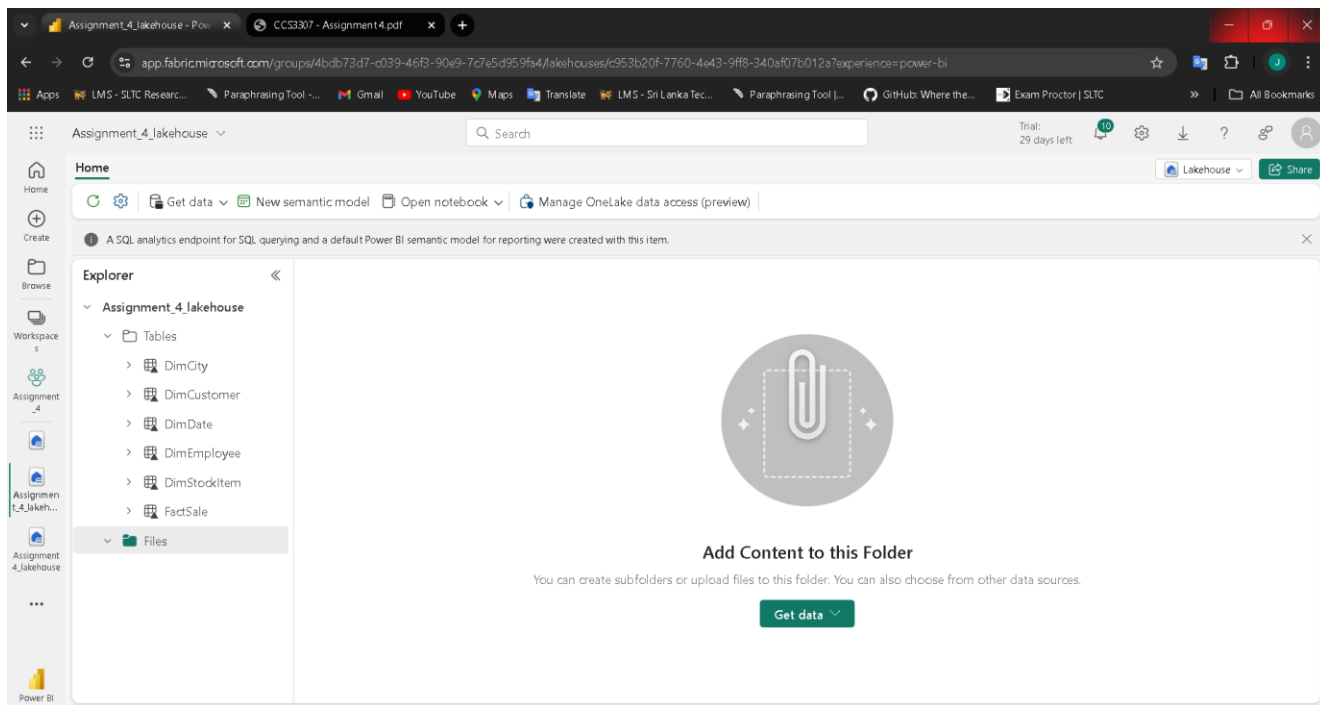
Date	Day Number	Day	Month	Short Month	Calendar Month Number	Calendar Month Label	Calendar Year	Calendar Year Label
1/1/2013	1	1	January	Jan	1	CY2013-Jan	2013	CY2013
1/2/2013	2	2	January	Jan	1	CY2013-Jan	2013	CY2013
1/3/2013	3	3	January	Jan	1	CY2013-Jan	2013	CY2013
1/4/2013	4	4	January	Jan	1	CY2013-Jan	2013	CY2013
1/5/2013	5	5	January	Jan	1	CY2013-Jan	2013	CY2013
1/6/2013	6	6	January	Jan	1	CY2013-Jan	2013	CY2013
1/7/2013	7	7	January	Jan	1	CY2013-Jan	2013	CY2013
1/8/2013	8	8	January	Jan	1	CY2013-Jan	2013	CY2013
1/9/2013	9	9	January	Jan	1	CY2013-Jan	2013	CY2013
1/10/2013	10	10	January	Jan	1	CY2013-Jan	2013	CY2013
1/11/2013	11	11	January	Jan	1	CY2013-Jan	2013	CY2013
1/12/2013	12	12	January	Jan	1	CY2013-Jan	2013	CY2013
1/13/2013	13	13	January	Jan	1	CY2013-Jan	2013	CY2013
1/14/2013	14	14	January	Jan	1	CY2013-Jan	2013	CY2013
1/15/2013	15	15	January	Jan	1	CY2013-Jan	2013	CY2013

Power BI interface showing a Power Query transformation. The query is named "Table.TransformColumnTypes(*Replaced value 1*, {{"Credit Limit", Currency.Type}})". The data table has 16 rows and 12 columns. The columns are: Customer Key, WW1 Customer ID, Customer, Bill To Customer, Category, Buying Group, Primary Contact, and Postal C. The data is transformed from a source table to a target table.

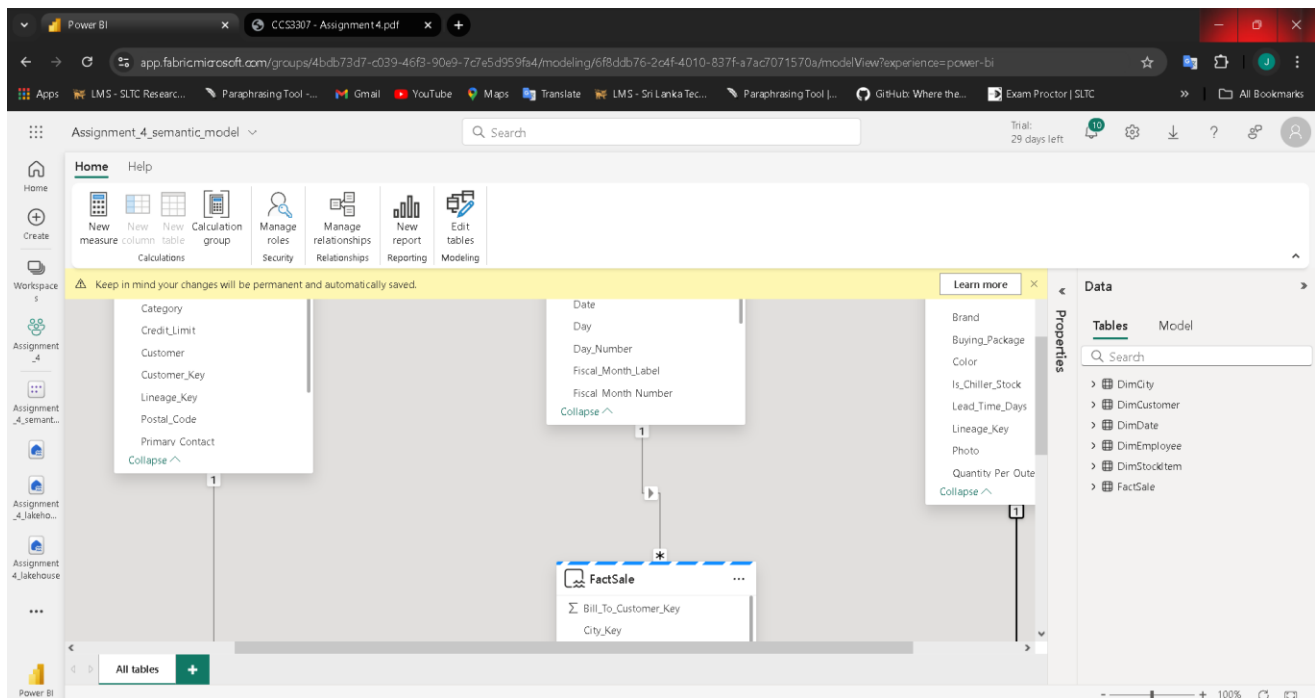
Customer Key	WW1 Customer ID	Customer	Bill To Customer	Category	Buying Group	Primary Contact	Postal C
1	1	Talpin Toys (Head Office)	Talpin Toys (Head Office)	Novelty Shop	Talpin Toys	Waldemar Fisar	90410
2	2	Talpin Toys (Sylvanite, MT)	Talpin Toys (Head Office)	Novelty Shop	Talpin Toys	Lorena Cindric	90216
3	3	Talpin Toys (Peoples Valley, AZ)	Talpin Toys (Head Office)	Novelty Shop	Talpin Toys	Bhaargav Rambhatla	90205
4	4	Talpin Toys (Medicine Lodge, KS)	Talpin Toys (Head Office)	Novelty Shop	Talpin Toys	Daniel Roman	90152
5	5	Talpin Toys (Gaspert, NY)	Talpin Toys (Head Office)	Novelty Shop	Talpin Toys	Johanna Huting	90261
6	6	Talpin Toys (Jessie, ND)	Talpin Toys (Head Office)	Novelty Shop	Talpin Toys	Bismajet Thakur	90298
7	7	Talpin Toys (Frankewing, TN)	Talpin Toys (Head Office)	Novelty Shop	Talpin Toys	Kalidas Nadar	90761
8	8	Talpin Toys (Bow Mar, CO)	Talpin Toys (Head Office)	Novelty Shop	Talpin Toys	Kanti Kotadia	90484
9	9	Talpin Toys (Netcong, NJ)	Talpin Toys (Head Office)	Novelty Shop	Talpin Toys	Sonitu Aalto	90129
10	10	Talpin Toys (Wimbleton, ND)	Talpin Toys (Head Office)	Novelty Shop	Talpin Toys	Siddhartha Parkar	90061
11	11	Talpin Toys (Devault, PA)	Talpin Toys (Head Office)	Novelty Shop	Talpin Toys	Eliaz Javan	90185
12	12	Talpin Toys (Biscay, MN)	Talpin Toys (Head Office)	Novelty Shop	Talpin Toys	Helosa Fernandes	90054
13	13	Talpin Toys (Stonfort, IL)	Talpin Toys (Head Office)	Novelty Shop	Talpin Toys	Razena Hosseini	90685
14	14	Talpin Toys (Long Meadow, MD)	Talpin Toys (Head Office)	Novelty Shop	Talpin Toys	Tereza Valentova	90633
15	15	Talpin Toys (Batson, TX)	Talpin Toys (Head Office)	Novelty Shop	Talpin Toys	Filips Jaunzents	90631
16	16	Talpin Toys (Head Office)	Talpin Toys (Head Office)	Novelty Shop	Talpin Toys	Talpin Toys	90631

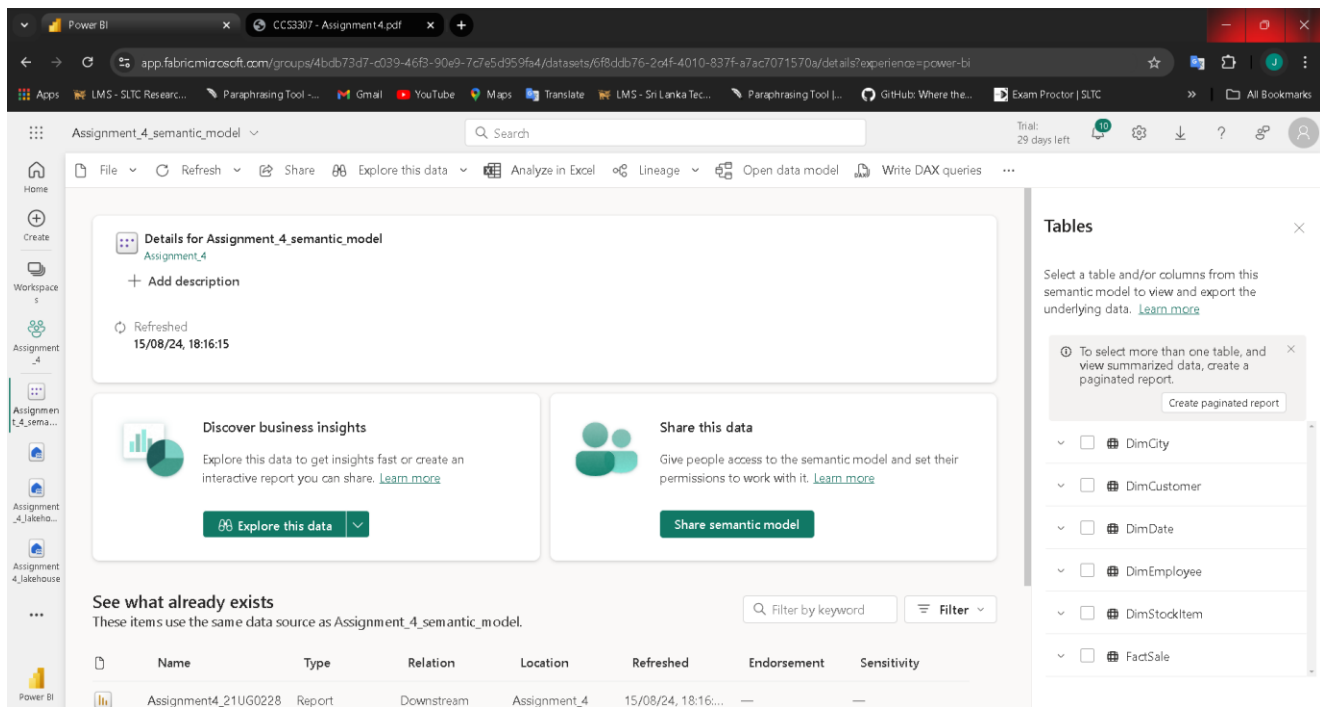
Power BI interface showing a Power Query transformation. The query is named "Table.TransformColumnTypes(*Promoted headers*, {{"City Key", Int64.Type}, {"City", type text}, {"State Province", type text}, {"Country", type text}, {"Sales Territory", type text}, {"Location", type text}, {"Latest Recorded Population", type text}})". The data table has 16 rows and 7 columns. The columns are: City Key, City, State Province, Country, Sales Territory, Location, and Latest Recorded Population. The data is transformed from a source table to a target table.

City Key	City	State Province	Country	Sales Territory	Location	Latest Recorded Population
18	Carson	California	United States	Far West	0x6100000010C345A588168EA4040AC1BEF...	91714
19	Carson	Washington	United States	Far West	0x6100000010C8B98CB8BD9DC4640CDBFA...	2279
22	Carson City	Nevada	United States	Far West	0x6100000010C8D919158F7944340B67526...	55274
27	Cartago	California	United States	Far West	0x6100000010C168156050F294240936FB6...	92
36	El Centro	California	United States	Far West	0x6100000010C4C37894160654040E430BD...	42598
37	El Cerrito	California	United States	Far West	0x6100000010C0746698F3754240EFAE0ED...	23549
43	El Dorado Hills	California	United States	Far West	0x6100000010C5F155E38C6574340DE7360...	42108
46	El Granada	California	United States	Far West	0x6100000010C0E4C6E1459C04240B9420E...	5467
52	El Monte	California	United States	Far West	0x6100000010C102A2790C80841405ABD1E...	113475
55	El Nido	California	United States	Far West	0x6100000010C77E4FD9459142403C7A0D...	330
62	El Paso de Robles	California	United States	Far West	0x6100000010CC72C7B12D8D1414051871...	29793
69	El Portal	California	United States	Far West	0x6100000010CAE9921FASAD6A240A1E4CA...	474
75	El Rio	California	United States	Far West	0x6100000010CF4F928232E1E4140A306D3...	7198
78	El Segundo	California	United States	Far West	0x6100000010CFED1DCAFA7F54040A7BA7...	16654
79	El Sobrante	California	United States	Far West	0x6100000010CC1880A1813FD4240ABC59...	12669
81	El Toro	California	United States	Far West	0x6100000010C278BD05669D04040C7B8F...	0

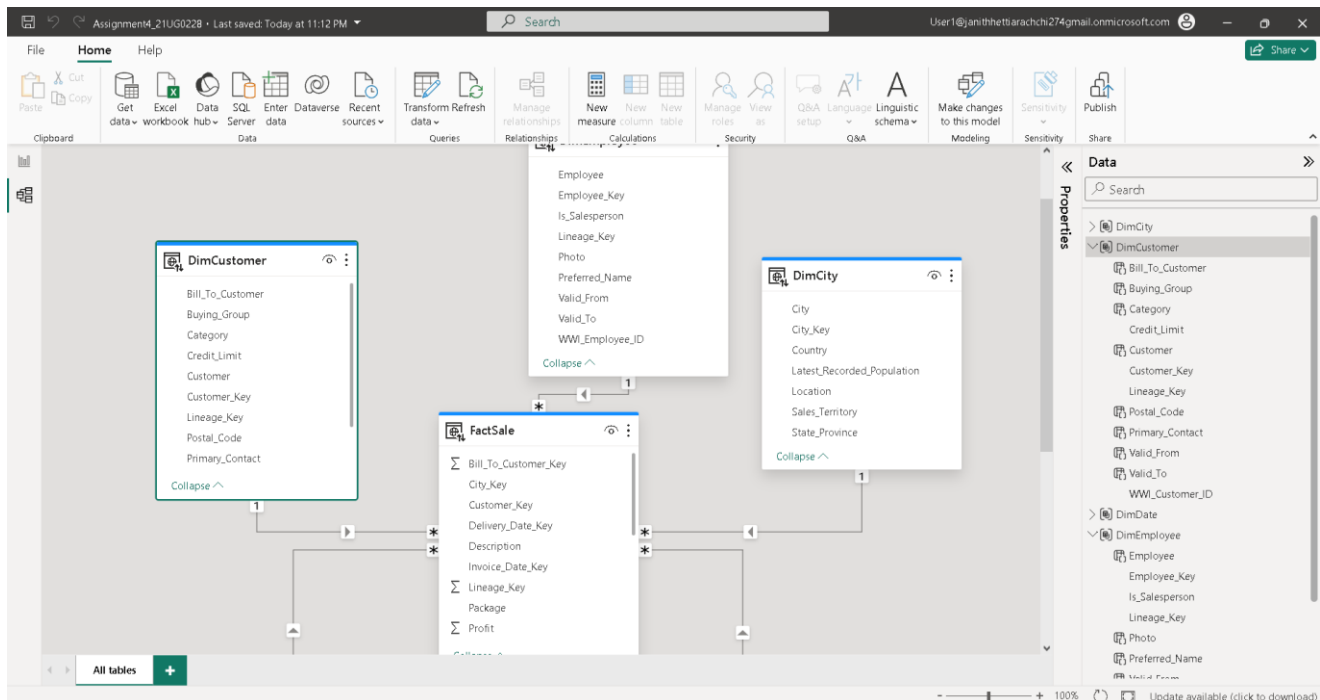


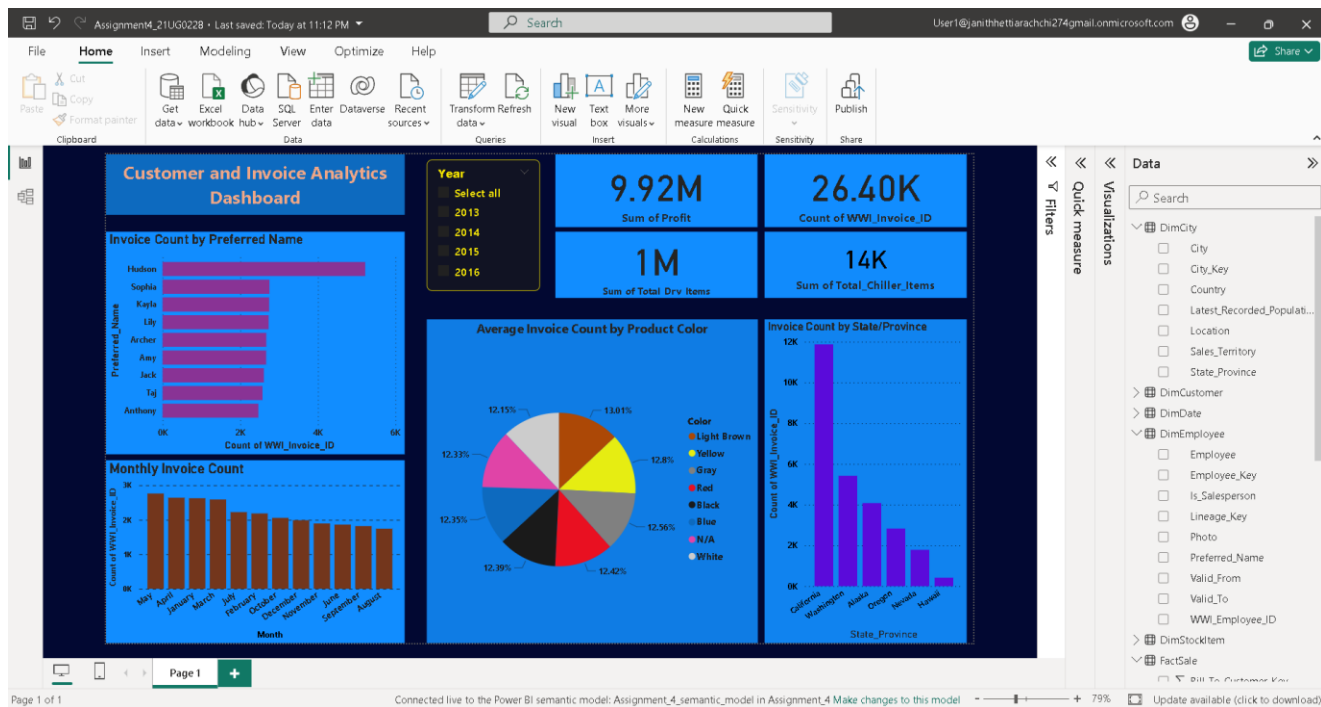
3. After Completing the ETL process and Delta Table Creation part on your data lake, create a semantic Model using the “New Semantic Model” option. Make sure to identify Keys in every table while creating the semantic model.





- Connect the semantic model to Microsoft Power BI and Create a Dashboard to visualize important KPIs according to given Sales Data.





- Publish your Power BI dashboard to Power BI service enabling a free trial and share a Link as a submission. Create a Lab Report adding necessary screenshots. Attach a Published link to the Lab Report. Make sure to rename the report using your student Index number.

Name	Type	Task	Owner	Refreshed	Next refresh	Endorsement	Sensitivity	Included in app
Assignment4_21UG0228	Report	—	Assignment_4	15/08/24, 18:16...	—	—	—	No
Assignment_4_lakehouse	Lakehouse	—	User1	—	—	—	—	—
Assignment_4_lakehouse	Semantic mo...	—	Assignment_4	15/08/24, 11:4...	N/A	—	—	—
Assignment_4_lakehouse	SQL analytics...	—	Assignment_4	15/08/24, 23:29...	N/A	—	—	—
Assignment_4_semantic_model	Semantic mo...	—	Assignment_4	15/08/24, 18:1...	N/A	—	—	—
Dataflow 1	Dataflow Gen2	—	User1	15/08/24, 17:5...	N/A	—	—	—

