

Case Study

- Ashwin Harish P

1. Uploading CSV to Azure

Creating a new container “Retail Sales” in Azure

2. Ingest it into Snowflake using Snowpark,

Creating a new database and schema

RETAIL_SALES.RETAIL_SALES_SCHEMA Settings

Open in Workspaces Code Versions

```
1 // Creating Database and Schema
2 CREATE OR REPLACE DATABASE RETAIL_SALES;
3 CREATE OR REPLACE SCHEMA RETAIL_SALES.RETAIL_SALES_SCHEMA;
4
5
```

Results Chart

	status
1	Schema RETAIL_SALES_SCHEMA successfully created.

Query Details

Query duration 54ms

Rows 1

Creating a new integration which points to azure

```
5 // Creating Integration
6 CREATE OR REPLACE STORAGE INTEGRATION AZURE_INTEGRATION
7     TYPE = EXTERNAL_STAGE
8     STORAGE_PROVIDER = AZURE
9     ENABLED = TRUE
10    AZURE_TENANT_ID = '7540734b-e567-46c3-9ad3-ec9fb9e50140'
11    STORAGE_ALLOWED_LOCATIONS = ( 'azure://ashwinblobst.blob.core.windows.net/retail-sales' );
12
13
14 DESC INTEGRATION AZURE_INTEGRATION;
15
```

Results Chart

	property	property_type	property_value	property_default
1	ENABLED	Boolean	true	false
2	STORAGE_PROVIDER	String	AZURE	
3	STORAGE_ALLOWED_L	List	azure://ashwinblobst.blob	[]
4	STORAGE_BLOCKED_L	List		[]
5	USE_PRIVATELINK_EN	Boolean	false	false
6	AZURE_TENANT_ID	String	7540734b-e567-46c3-9	

Query Details

Query duration 34ms

Rows 9

Query ID 01bfe0a7-0001-672b-0...

Show more

Creating a file format

```
15 // Creating a File
16 CREATE OR REPLACE FILE FORMAT RETAIL_SALES.RETAIL_SALES_SCHEMA.CSV_FILE
17     TYPE=CSV
18     FIELD_DELIMITER=';',
19     SKIP_HEADER=1;
```

Results Chart

	status
1	File format CSV_FILE successfully created.

Query Details

Query duration 52ms

Creating a Stage

```
22 // Creating Stage
23 CREATE OR REPLACE STAGE RETAIL_SALES.RETAIL_SALES_SCHEMA.AZURE_STAGE
24     STORAGE_INTEGRATION = AZURE_INTEGRATION
25     URL='azure://ashwinblobst.blob.core.windows.net/retail-sales'
26     FILE_FORMAT = CSV_FILE;
```

ResultsChart

	status
1	Stage area AZURE_STAGE successfully created.

Listing files in Stage

```
--
17 -- List files in the Stage
18 LIST @RETAIL_SALES.RETAIL_SALES_SCHEMA.AZURE_STAGE;
19
```

ResultsChart

	name	size	md5	last_modified
1	azure://ashwinblobst.blob	3004	a38163e0b0e75a2fc23a	Wed, 22 Oct 2025 10:14:...

Creating Table

```
--
21 -- Creating new table
22 CREATE OR REPLACE TABLE RETAIL_SALES.RETAIL_SALES_SCHEMA.SALES_DATA (
23     OrderID STRING,
24     CustomerID STRING,
25     ProductID STRING,
26     Quantity INT,
27     Price FLOAT,
28     OrderDate DATE
29 );
30
```

ResultsChart

	status
1	Table SALES_DATA successfully created.

Query Details

Query duration140ms

Doing Transformations in Snowpark

Installing Necessary Packages

```
%pip install snowflake-connector-python
%pip install snowflake-snowpark-python

Downloading cffi-1.17.1-cp312-cp312-manylinux_2_17_aarch64.manylinux2014_aarch64.whl (478 kB)
Downloading cryptography-46.0.0-cp311-abi3-manylinux_2_34_aarch64.whl (4.3 MB)
----- 4.3/4.3 MB 78.5 MB/s eta 0:00:00

Installing collected packages: tzlocal, cffi, cryptography, snowflake-connector-python, snowflake-snowpark-python
Attempting uninstall: cffi
  Found existing installation: cffi 2.0.0
  Uninstalling cffi-2.0.0:
    Successfully uninstalled cffi-2.0.0
Attempting uninstall: cryptography
  Found existing installation: cryptography 46.0.3
  Uninstalling cryptography-46.0.3:
    Successfully uninstalled cryptography-46.0.3
Attempting uninstall: snowflake-connector-python
  Found existing installation: snowflake-connector-python 4.0.0
  Uninstalling snowflake-connector-python-4.0.0:
    Successfully uninstalled snowflake-connector-python-4.0.0
Successfully installed cffi-1.17.1 cryptography-46.0.0 snowflake-connector-python-3.18.0 snowflake-snowpark-python-1.40.0 tzlocal-5.3.1
Note: you may need to restart the kernel using %restart_python or dbutils.library.restartPython() to use updated packages
```

```
%restart_python
```

Importing Necessary Packages

```
from snowflake.snowpark import Session
from snowflake.snowpark.functions import col, avg, split, year, month
```

Importing Necessary Packages

▶ ✓ 1 minute ago (1s)

5

```
from snowflake.snowpark import Session
from snowflake.snowpark.functions import col, avg
```

Creating Connection with Snowflake and Creating Session

▶ ✓ 4 minutes ago (2s)

7

```
connection_parameters = {
    "account": "okmtjfr-rf98839",
    "user": "ASHWINHARISHP",
    "password": "Ashwinakhilesh21",
    "warehouse": "COMPUTE_WH",
    "database": "RETAIL_SALES",
    "schema": "RETAIL_SALES_SCHEMA"
}
session = Session.builder.configs(connection_parameters).create()
```

▶ ✓ 04:33 PM (1s)

9

```
df = session.read.option("header", True).csv("@RETAIL_SALES.RETAIL_SALES_SCHEMA.AZURE_STAGE/Retail_Sales.csv")
```

Doing Transformation

Renaming Column Names

```
df = df.rename({
    "c1": "OrderID",
    "c2": "OrderDate",
    "c3": "OrderMonth",
    "c4": "CustomerID",
    "c5": "CustomerName",
    "c6": "Country",
    "c7": "Region",
    "c8": "City",
    "c9": "Category",
    "c10": "SubCategory",
    "c11": "Quantity",
    "c12": "Discount",
    "c13": "Sales",
    "c14": "Profit"
})
```

Type Casting

```
df = df.with_column("Quantity", col("Quantity").cast("INTEGER")) \
    .with_column("Discount", col("Discount").cast("FLOAT")) \
    .with_column("Sales", col("Sales").cast("FLOAT")) \
    .with_column("Profit", col("Profit").cast("FLOAT")) \
    .with_column("OrderDate", col("OrderDate").cast("DATE"))
```

Injecting to a Snowflake Table

Ingesting to a Snowflake Table

▶

✓ 1 minute ago (2s)

19

```
df.write.save_as_table("RETAIL_SALES.RETAIL_SALES_SCHEMA.SALES_DATA", mode="overwrite")
```

Verifying Data in Snowflake

19
20 -- Verifying Data
21 SELECT * FROM SALES_DATA;

Results

Chart

	OrderID	OrderDate	MonthOfSale	CustomerID	CustomerName	Country	Region	City	Category
1	ORD-5F8D6FOC	2024-10-08	2024-10	CUST1000	Ananya Sharma	India	South	Mumbai	Office Supp
2	ORD-BF0078E4	2024-08-11	2024-08	CUST1001	Aarav Iyer	India	Central	Lucknow	Technolog
3	ORD-86CD58A3	2024-06-12	2024-06	CUST1002	Arjun Sharma	USA	East	Kolkata	Furniture
4	ORD-FB0CD2D9	2024-12-18	2024-12	CUST1003	Ananya Das	India	North	Kolkata	Office Supp
5	ORD-EF35596B	2024-10-27	2024-10	CUST1004	Ishaan Bhat	UK	Central	Chennai	Furniture

Query Details

...

Query duration 516ms

Rows 25

Query ID 01bfe0f8-0001-6723-0...

Show more

Generating Power BI Report

Loading Data from External Source-Snowflake

Get Data



All

Database

All



Snowflake

Certified Connectors

Template Apps

Connect

Cancel

Snowflake

Server

Warehouse

▸ Advanced options

OK

Cancel

Transforming Data

Navigator

Display Options ▾

OKMTJFR-RF98839.snowflakecomputing.co...

COPY_INT0

CRON_JOB_DB

CUSTOMER_DB

DATASHARING

MANAGED_DB

OUR_FIRST_DB

PRATICE_1

PROD_DB

RETAIL_SALES [3]

INFORMATION_SCHEMA

PUBLIC

RETAIL_SALES_SCHEMA [2]

☐ SALES

☒ SALES_DATA

SNOWFLAKE

SNOWFLAKE_LEARNING_DB

SNOWFLAKE_SAMPLE_DATA

SNOWPIPE

STREAM_DB

SALES_DATA

OrderID	OrderDate	MonthOfSale	CustomerID	CustomerName
ORD-5F8D6FOC	2024-10-08	2024-10	CUST1000	Ananya Sharma
ORD-BF0078E4	2024-08-11	2024-08	CUST1001	Aarav Iyer
ORD-86CD58A3	2024-06-12	2024-06	CUST1002	Arjun Sharma
ORD-FB0CD2D9	2024-12-18	2024-12	CUST1003	Ananya Das
ORD-EF35596B	2024-10-27	2024-10	CUST1004	Ishaan Bhat
ORD-60D1DA88	2024-08-26	2024-08	CUST1005	Neha Iyer
ORD-A5081404	2025-09-15	2025-09	CUST1006	Arjun Iyer
ORD-E1C9BE42	2024-02-27	2024-02	CUST1007	Priya Singh
ORD-4FCB3B05	2025-05-26	2025-05	CUST1008	Kabir Menon
ORD-921966C8	2025-03-14	2025-03	CUST1009	Arjun Chopra
ORD-E4A002F0	2024-06-12	2024-06	CUST1010	Ananya Patel
ORD-0944D71F	2025-07-20	2025-07	CUST1011	Sanjay Gupta
ORD-7E28FF54	2025-03-21	2025-03	CUST1012	Ananya Khan
ORD-55961D4C	2025-09-15	2025-09	CUST1013	Neha Mehta
ORD-6BE57CAD	2025-01-06	2025-01	CUST1014	Aarav Reddy
ORD-1D9DC086	2024-05-08	2024-05	CUST1015	Ishaan Bhat
ORD-9B484AF9	2025-02-12	2025-02	CUST1016	Rohan Khan
ORD-42167295	2025-08-25	2025-08	CUST1017	Kabir Sharma
ORD-A91119D6	2025-06-10	2025-06	CUST1018	Kabir Iyer
ORD-951CD78B	2024-04-06	2024-04	CUST1019	Sneha Menon
ORD-B6A75DA5	2025-02-16	2025-02	CUST1020	Ananya Verma
ORD-6D74C638	2024-09-12	2024-09	CUST1021	Arjun Gowda
ORD-93240C28	2024-08-30	2024-08	CUST1022	Aarav Iyer

Select Related Tables

Load

Transform Data

Cancel

Connection settings

You can choose how to connect to this data source. Import allows you to bring a copy of the data into Power BI. DirectQuery will connect live to this data source.

☐ Import

☒ DirectQuery

[Learn more about DirectQuery](#)

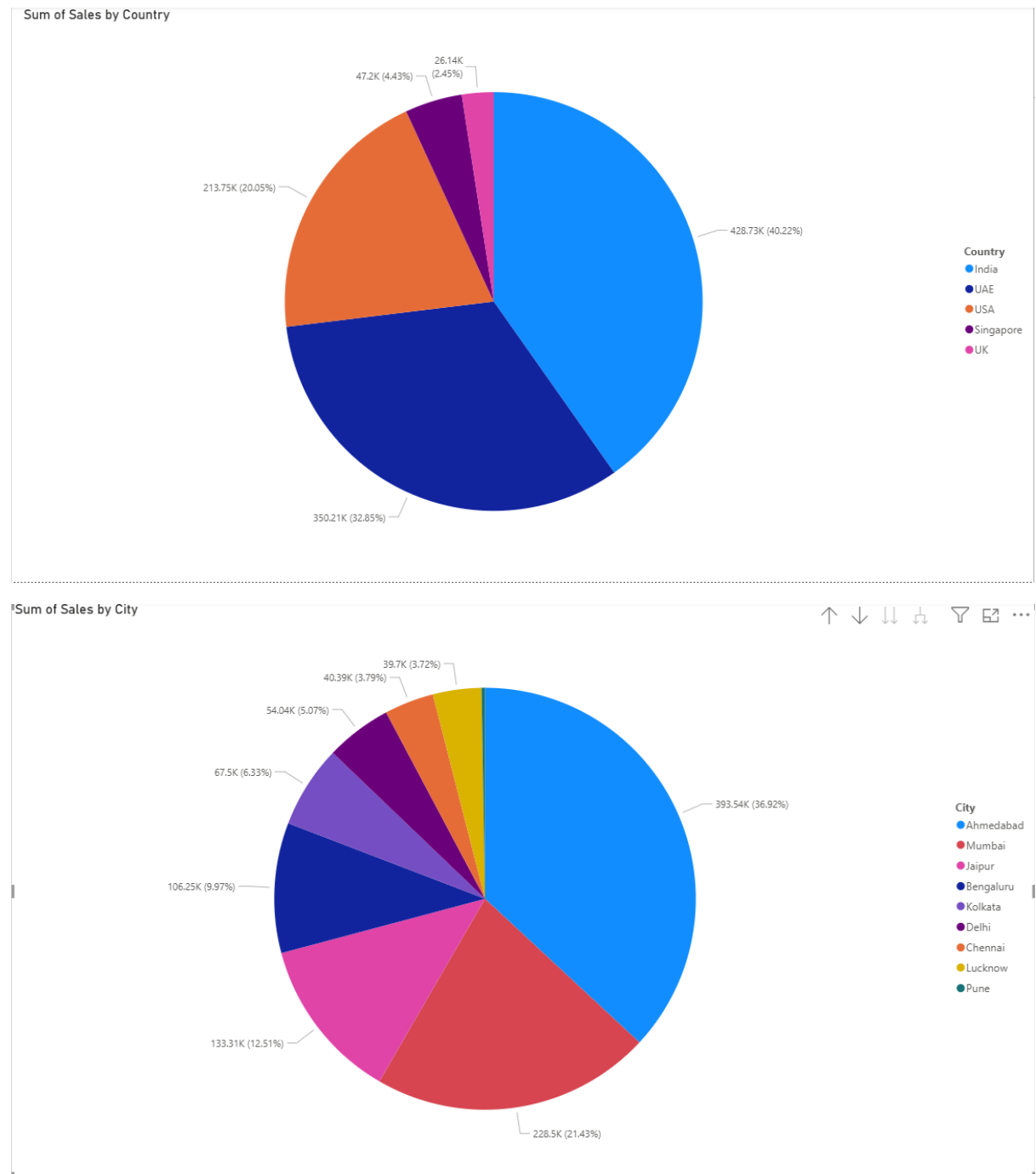
OK

Cancel

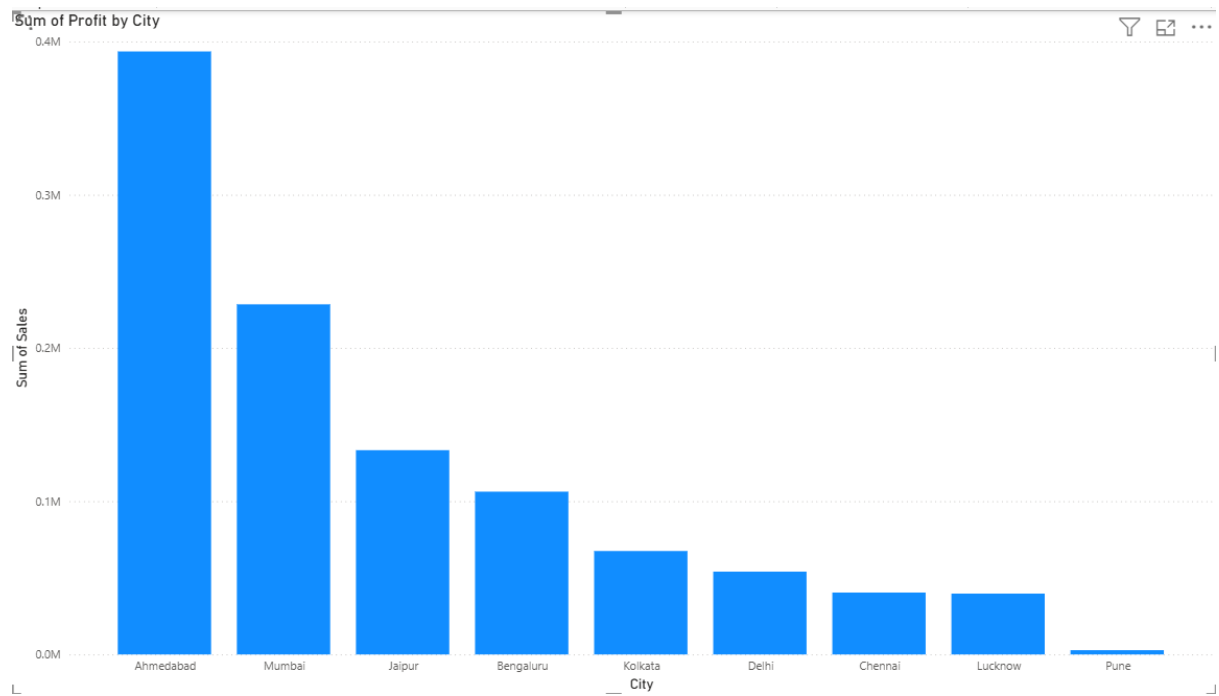
After making some transformation, I loaded the data

Visualizations

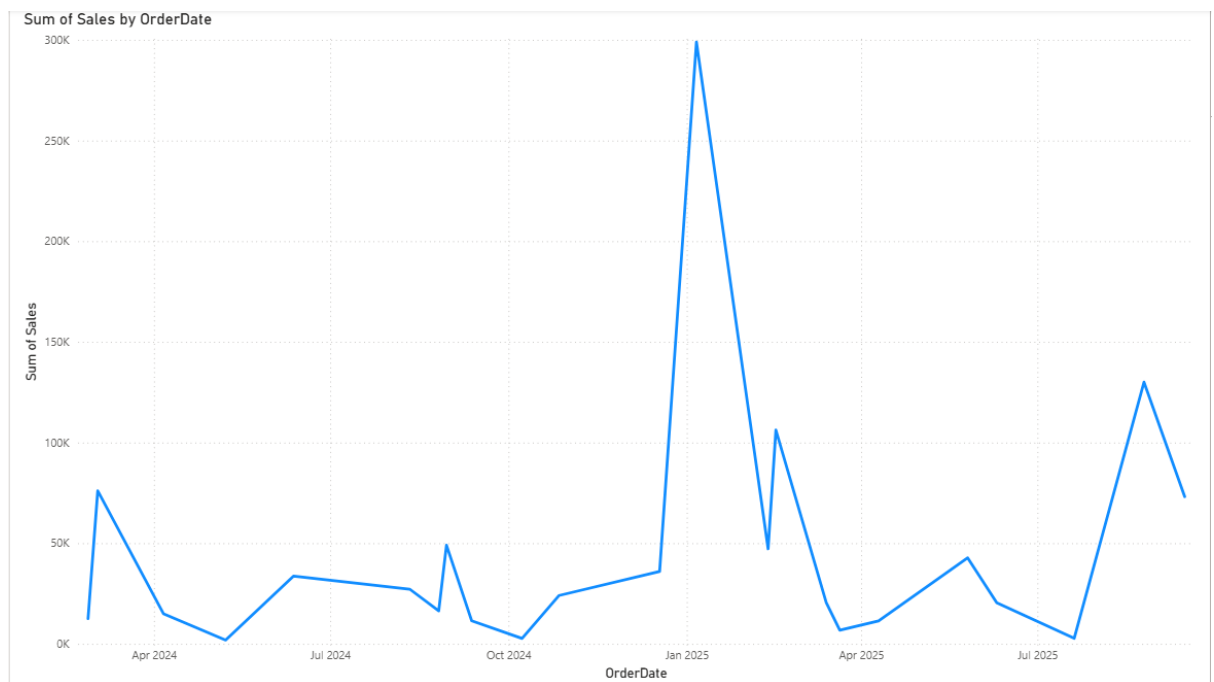
1. Sales of Each City and Country using Drill Through



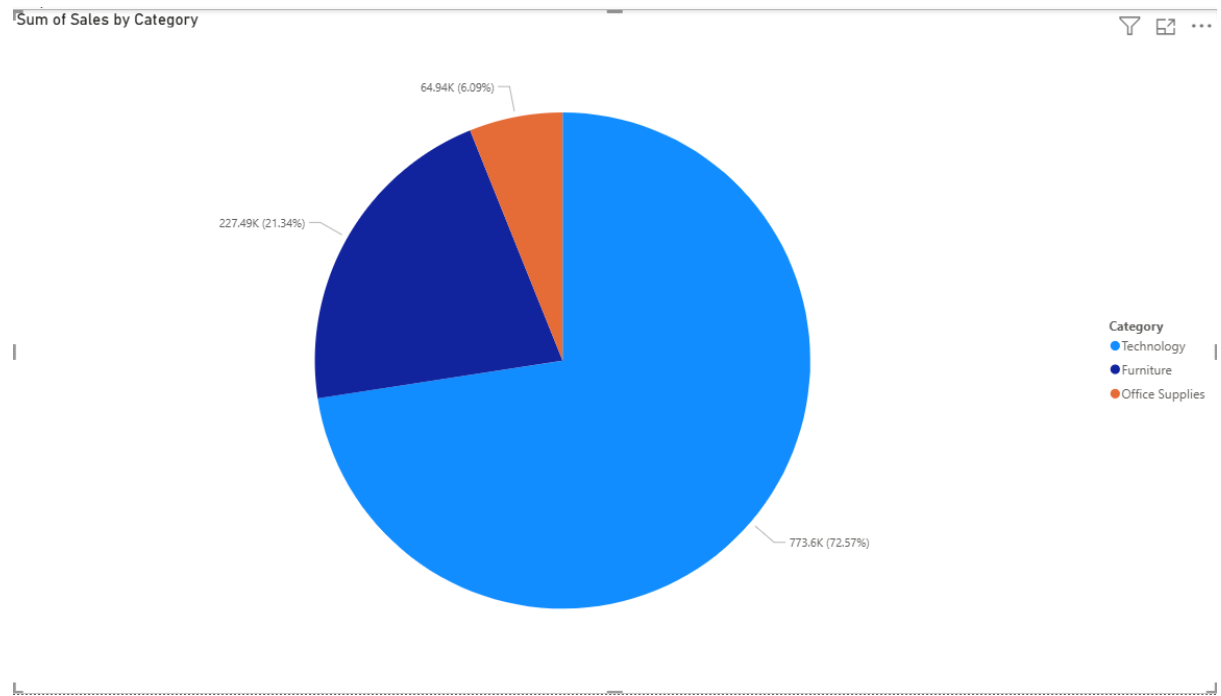
2. Profit of Each City



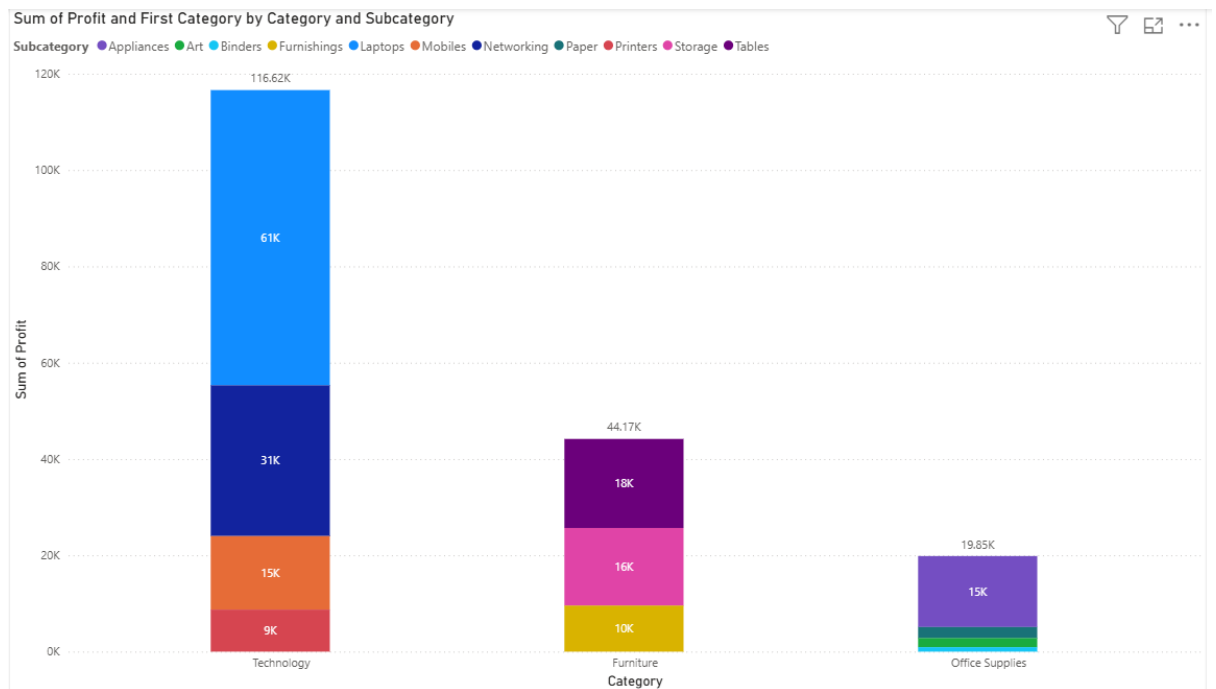
3. Sales by Month



4. Sales by Category and Sub Category using Drill Though



5. Profit by Category and Sub Category



6. Quantity Sold by Sub-Category

