Journey PPT

Shell Training Bootcamp (August 14 to October 6, 8 weeks)

Omeir Fawaz

Associate Data Engineer

Safety and Health On Site

Ensure confidential discussions are not overheard

Make sure your workspace is ergonomically sound

Ensure adequate lighting in the room when you work

Clean surfaces frequently

Have an emergency and evacuation plan in place

Ensure understanding of fire safety

- Know what the fire alarm sounds like
- Make sure that you can hear the fire alarm
- Make sure your smoke alarms work
- Maintain clear walkways and fire exits

On the Move

Do not take this call, or any other call, while driving – ever

Do not use any hands-free device – Bluetooth, built-in, etc. – whilst driving

Continue to follow COVID guidelines

In the event of any kind of emergency, please leave the call – promptly and safely

Week 5

(Sep 12 – Sep 15)

01 Azure Synapse Analytics

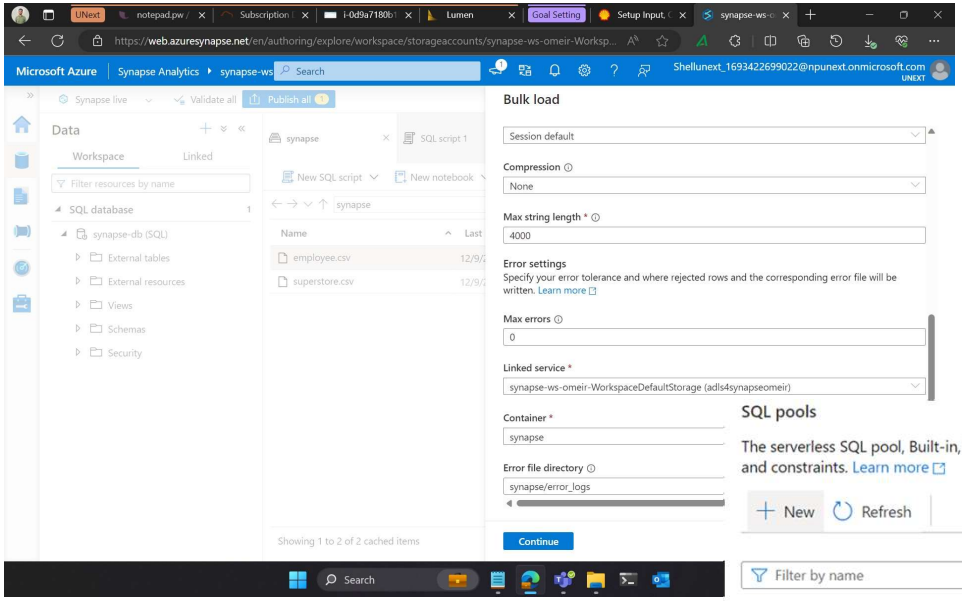
02 Power BI Visualisation

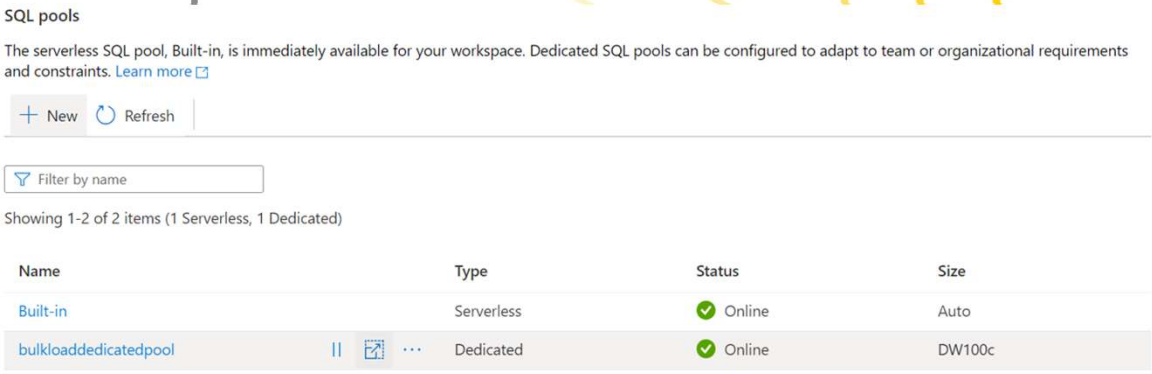
03 Python

5.1 Azure Synapse Analytics

Learnt about the following:

1. Synapse Workspace, Data Warehouse, distributed arch. and MPP
2. Synapse Mapping Data Flow, Pipeline; data ingestion, transformation, and analysis
3. SQL pools (built-in and dedicated), data warehouse interface and query editor

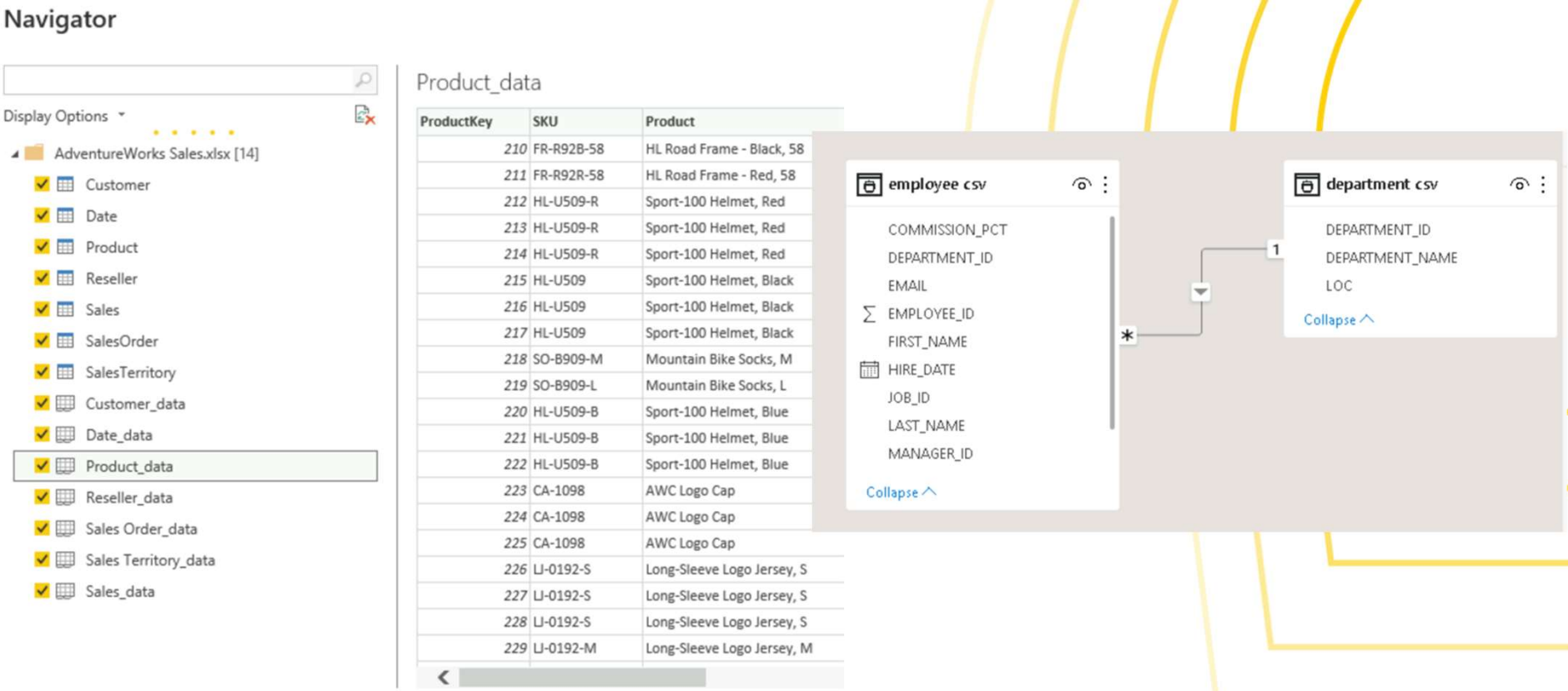




5.2 Basics of Power BI

Learnt about the following:

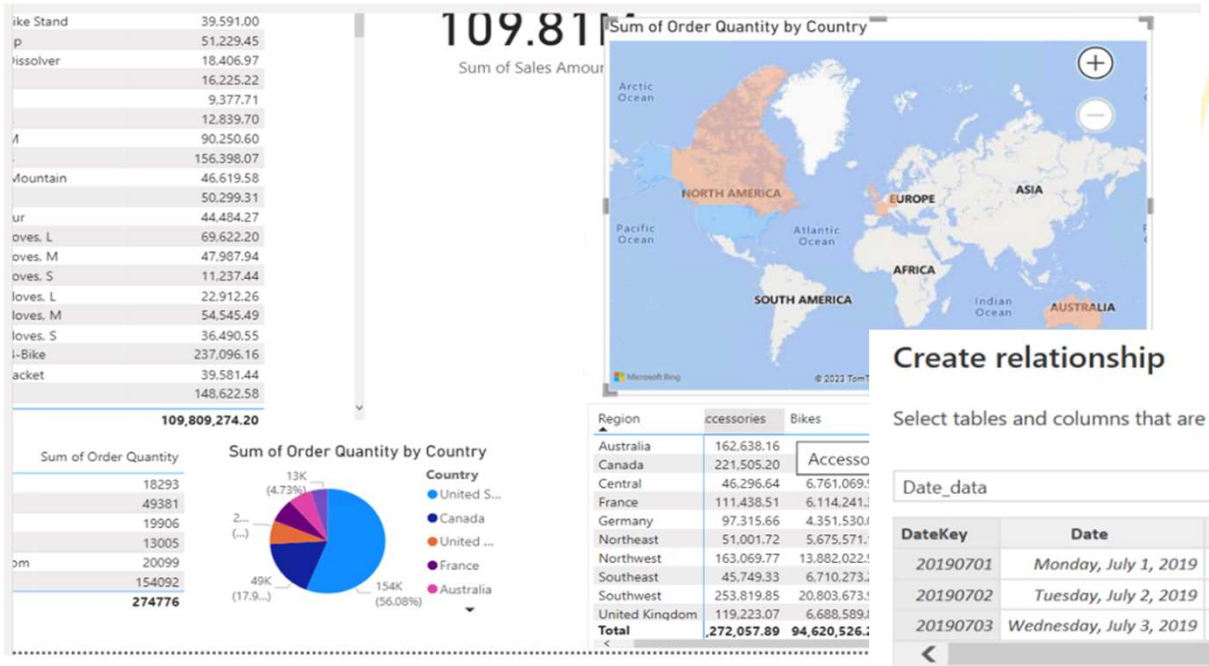
1. Importing data from multiple sources: CSV files, SQL server
2. Data loading, transformation, modelling and cleansing (removing duplicates, errors)
3. Calculated columns, row level security, aggregated functions

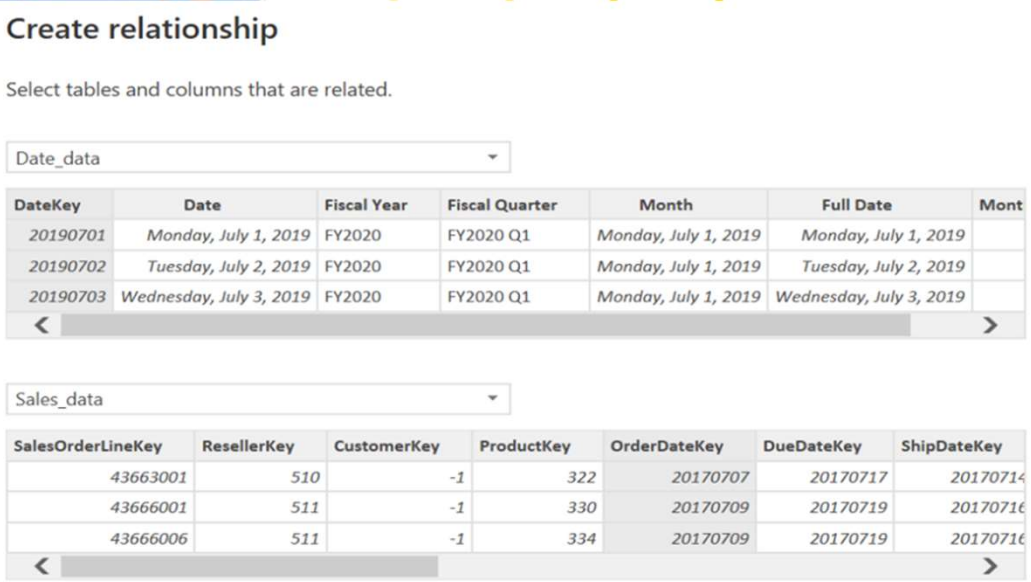


5.3 Power BI Features

Learnt about the following:

1. Creating visualizations, reports, and dashboards
2. Custom relationships between tables, DAX functions





5.4 Python

Learnt about the following:

1. Jupyter notebooks, data types in Python
2. Logical operations, zip() and enumerate() functions

```
In [11]: dates = ['2023-09-01', '2023-09-02', '2023-09-03', '2023-09-04', '2023-09-05']
         temperature = [78, 82, 79, 85, 88]
         humidity = [60, 65, 62, 58, 57]
         precipitation = [0.0, 0.2, 0.0, 0.0, 0.0]


         weather_data = list(zip(dates, temperature, humidity, precipitation))
         print(weather_data)

         hotdays=[]
         for idx,temp in enumerate(temperature):
             if temp>80:
                 hotdays.append(dates[idx])

         print(hotdays)
         weather_forecast=[]
         for temp in temperature:
             if temp>85:
                 forecast="Hot and sunny"
```

```
In [17]: a=0
         for i in temperature:
             a=a+i
         avg=a/len(temperature)
         print (avg)

         82.4
```