



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

A large, stylized sunburst graphic in shades of yellow and orange, centered on the left side of the slide. It consists of multiple concentric, wavy rings that create a sense of depth and radiance.

Week 5

(Sep 12 – Sep 15)

01 Azure Synapse Analytics

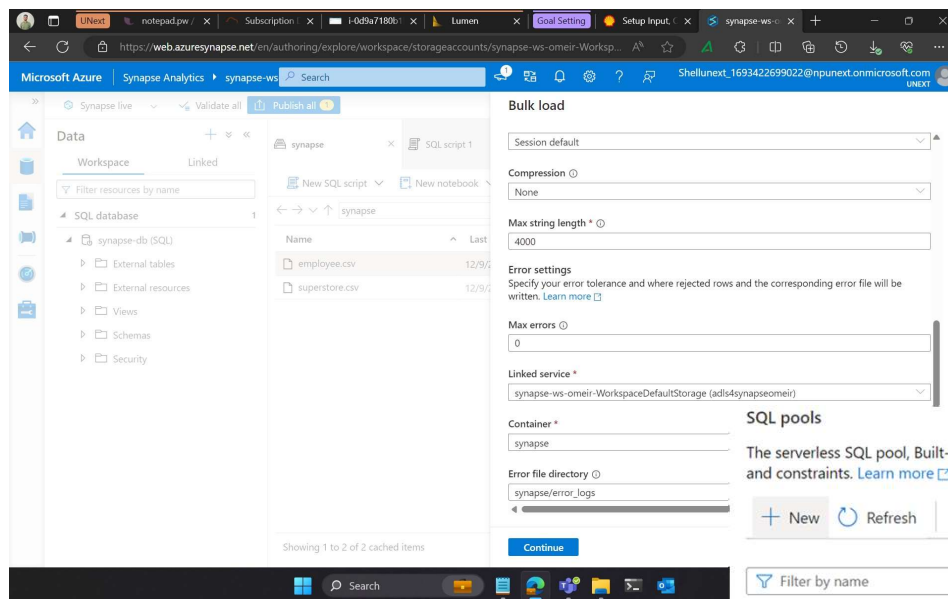
02 Power BI Visualisation

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



SQL pools

The serverless SQL pool, Built-in, is immediately available for your workspace. Dedicated SQL pools can be configured to adapt to team or organizational requirements and constraints. [Learn more](#)

[+ New](#) [Refresh](#)

Filter by name
Showing 1-2 of 2 items (1 Serverless, 1 Dedicated)

Name	Type	Status	Size
Built-in	Serverless	Online	Auto
bulkloaddedicatedpool	Dedicated	Online	DW100c

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

Navigator

Display Options ▾

AdventureWorks Sales.xlsx [14]

- Customer
- Date
- Product
- Reseller
- Sales
- SalesOrder
- SalesTerritory
- Customer_data
- Date_data
- Product_data**
- Reseller_data
- Sales Order_data
- Sales Territory_data
- Sales_data

Product_data

ProductKey	SKU	Product
210	FR-R92B-58	HL Road Frame - Black, 58
211	FR-R92R-58	HL Road Frame - Red, 58
212	HL-U509-R	Sport-100 Helmet, Red
213	HL-U509-R	Sport-100 Helmet, Red
214	HL-U509-R	Sport-100 Helmet, Red
215	HL-U509	Sport-100 Helmet, Black
216	HL-U509	Sport-100 Helmet, Black
217	HL-U509	Sport-100 Helmet, Black
218	SO-B909-M	Mountain Bike Socks, M
219	SO-B909-L	Mountain Bike Socks, L
220	HL-U509-B	Sport-100 Helmet, Blue
221	HL-U509-B	Sport-100 Helmet, Blue
222	HL-U509-B	Sport-100 Helmet, Blue
223	CA-1098	AWC Logo Cap
224	CA-1098	AWC Logo Cap
225	CA-1098	AWC Logo Cap
226	LI-0192-S	Long-Sleeve Logo Jersey, S
227	LI-0192-S	Long-Sleeve Logo Jersey, S
228	LI-0192-S	Long-Sleeve Logo Jersey, S
229	LI-0192-M	Long-Sleeve Logo Jersey, M

employee csv

COMMISSION_PCT
DEPARTMENT_ID
EMAIL
Σ EMPLOYEE_ID
FIRST_NAME
HIRE_DATE
JOB_ID
LAST_NAME
MANAGER_ID

department csv

DEPARTMENT_ID
DEPARTMENT_NAME
LOC

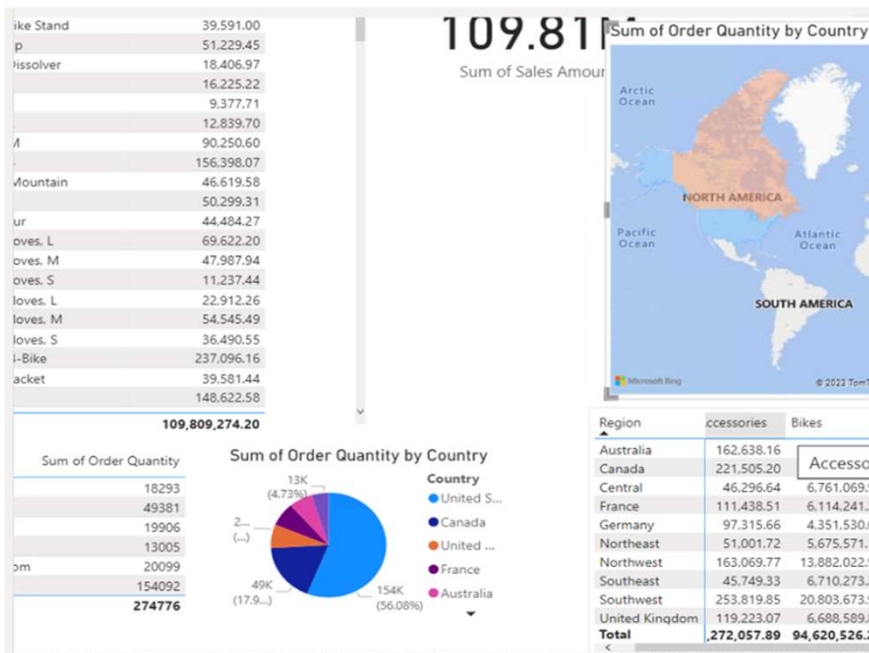
[Collapse ^](#)

[Collapse ^](#)

5.3 Power BI Features

Learnt about the following:

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



Create relationship

Select tables and columns that are related.

Date_data

DateKey	Date	Fiscal Year	Fiscal Quarter	Month	Full Date	Mont
20190701	Monday, July 1, 2019	FY2020	FY2020 Q1	Monday, July 1, 2019	Monday, July 1, 2019	
20190702	Tuesday, July 2, 2019	FY2020	FY2020 Q1	Monday, July 1, 2019	Tuesday, July 2, 2019	
20190703	Wednesday, July 3, 2019	FY2020	FY2020 Q1	Monday, July 1, 2019	Wednesday, July 3, 2019	

Sales_data

SalesOrderLineKey	ResellerKey	CustomerKey	ProductKey	OrderDateKey	DueDateKey	ShipDateKey
43663001	510	-1	322	20170707	20170717	20170714
43666001	511	-1	330	20170709	20170719	20170716
43666006	511	-1	334	20170709	20170719	20170716

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