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 3

(Aug 28 – Sep 1)

01 Case Study

02 Basics of BI

03 OLTP and OLAP

04 OLAP in Detail, Azure SQL

05 Azure VM and GitHub Actions

3.1 Case Study

Did a case study on "Streamline Solutions Inc. - Optimizing Software Deployment", and performed the following tasks:

1. Create and configure project in Azure DevOps with Epics, Sprints, Product and Sprint Backlog, User Stories, Tasks, Test Cases, and Defects
2. Created an Azure Virtual Machine and connected via SSH
3. Installed tools like Docker, Maven on the VM
4. Built a Spring Boot app using Maven and Dockerised it, published to Docker Hub, and launched it through an Azure Web App
5. Created a CI/CD workflow for automatically deploying the app on changes using GitHub Actions
6. Raised PR against the main repository on GitHub after adding features

3.2 Basics of Business Intelligence

Learnt about the following:

1. Key Resource Areas (KRI), and Key Performance Indicators (KPI)
2. Need for Business Intelligence (BI)
3. Parts of BI:
    1. Dimensional Modelling
    2. Data Warehousing
    3. Data Mining
4. Operational System Support (OSS) and Decision Support System (DSS)

3.3 OLTP and OLAP

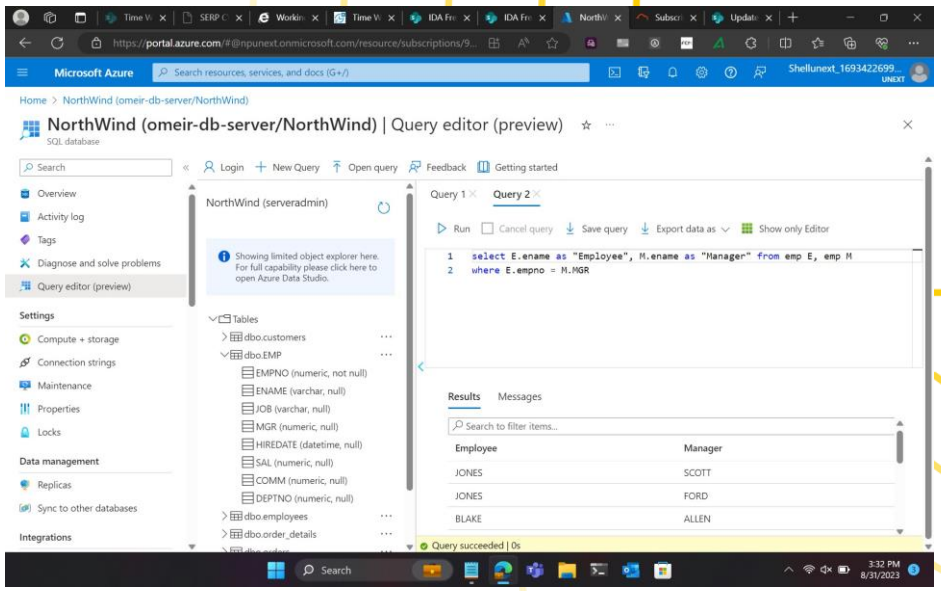
Learnt about the following:

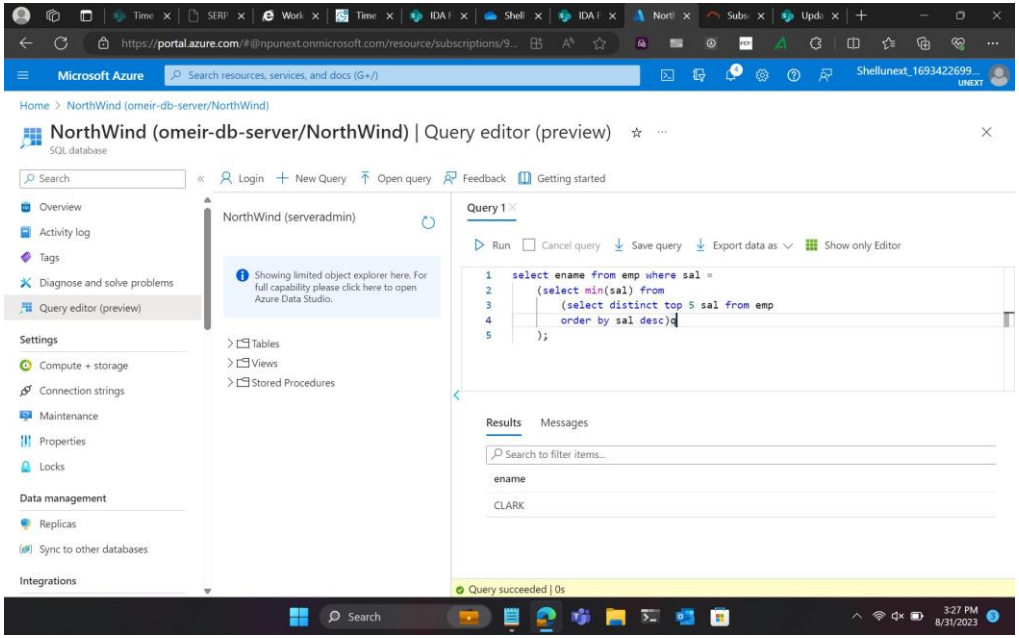
1. Online Transaction Processing (OLTP)
2. Online Analytical Processing (OLAP)
3. Shifting Data from OLTP to OLAP (ELT)
4. Difference between DB Backup and OLAP
5. OLAP vs Data Warehouse
6. Dimensional Modelling
7. Fact Tables
8. Slowly-Changing Dimensions

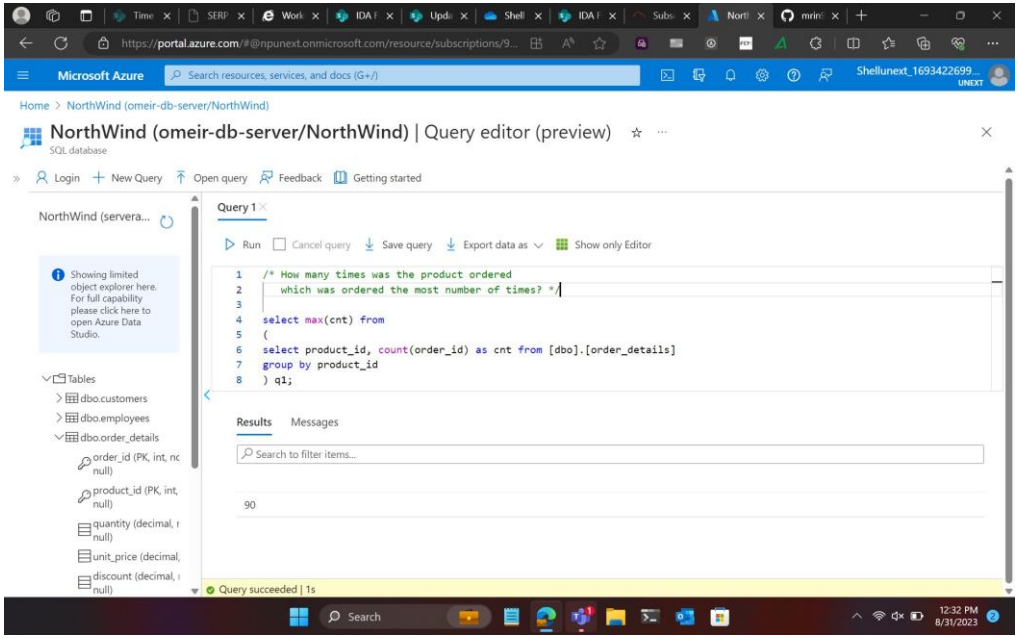
3.4 OLAP in Detail, Azure SQL

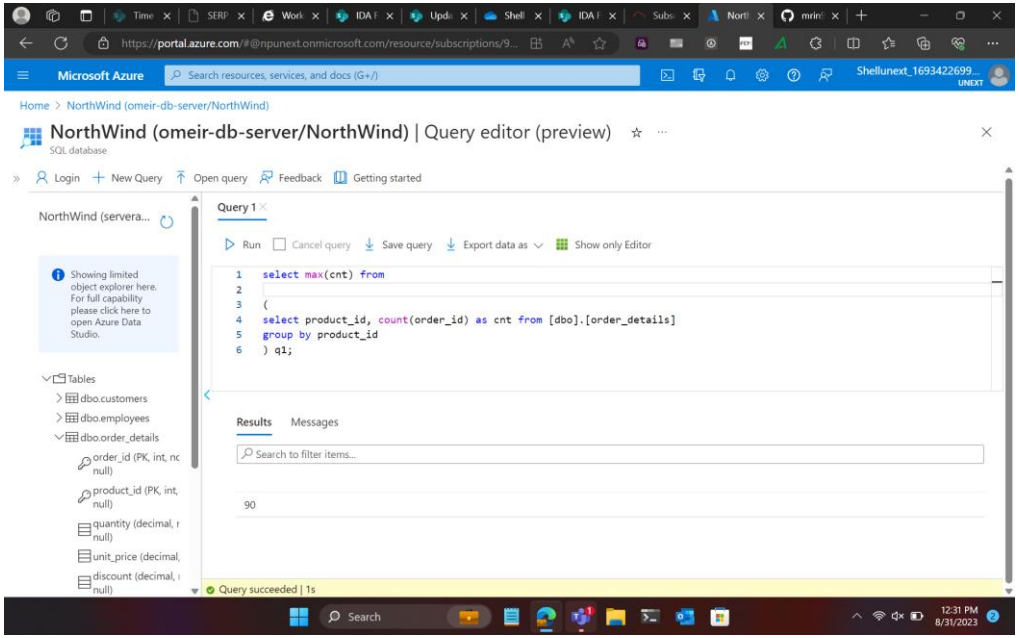
Learnt about the following:

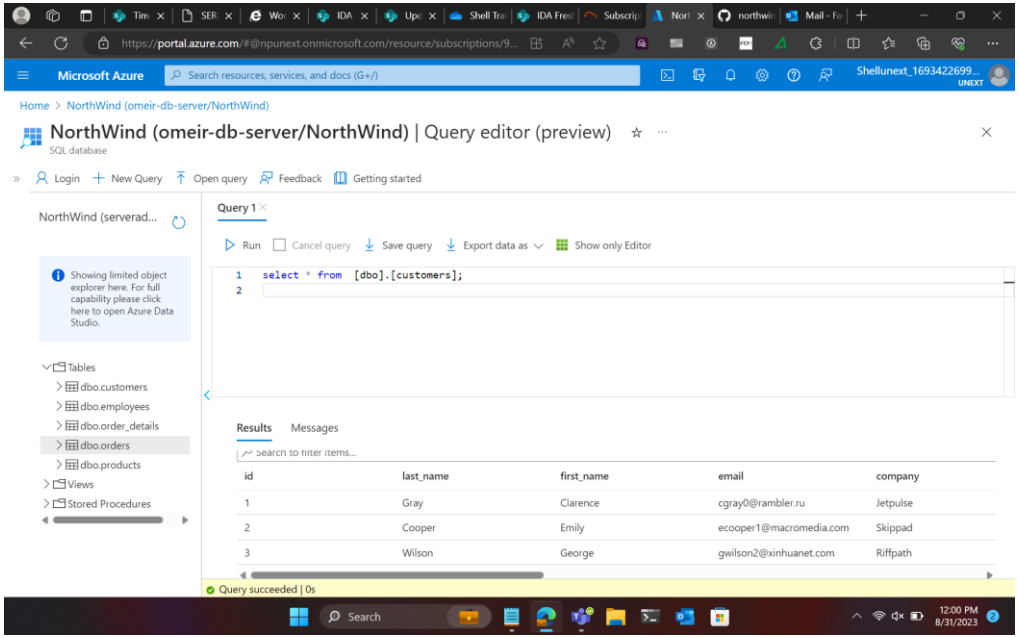
1. Operational Data Store (ODS)
2. Types of Fact Table Schemas: Star, Galaxy, Snowflake, Aggregate
3. Data Cubes
4. Types of OLAPs:
    1. MOLAP
    2. ROLAP: Relational
    3. HOLAP: Hybrid
    4. COLAP/DOLAP: Client/Desktop
5. Created a DB on Azure SQL and ran queries




Microsoft Azure
Search resources, services, and docs (G+/)
Home > NorthWind (omeir-db-server/NorthWind)
NorthWind (omeir-db-server/NorthWind) | Query editor (preview)
SQL database
Search
Login
New Query
Open query
Feedback
Getting started
Overview
Activity log
Tags
Diagnose and solve problems
Query editor (preview)
Settings
Compute + storage
Connection strings
Maintenance
Properties
Locks
Data management
Replicas
Sync to other databases
Integrations
NorthWind (serveradmin)
Showing limited object explorer here. For full capability please click here to open Azure Data Studio.
Tables
Views
Stored Procedures
Query 1
Run
Cancel query
Save query
Export data as
Show only Editor
select ename from emp where sal =
(select min(sal) from
(select distinct top 5 sal from emp
order by sal desc)a
);
Results
Messages
Search to filter items...
ename
CLARK
Query succeeded | 0s
Search
3:27 PM
8/31/2023


Microsoft Azure
Search resources, services, and docs (G+/)
Home > NorthWind (omeir-db-server/NorthWind)
NorthWind (omeir-db-server/NorthWind) | Query editor (preview)
SQL database
Login
New Query
Open query
Feedback
Getting started
NorthWind (servera...
Showing limited object explorer here. For full capability please click here to open Azure Data Studio.
Tables
dbo.customers
dbo.employees
dbo.order_details
order_id (PK, int, nc null)
product_id (PK, int, null)
quantity (decimal, r null)
unit_price (decimal,
discount (decimal, r null)
Query 1
Run
Cancel query
Save query
Export data as
Show only Editor
/* How many times was the product ordered
which was ordered the most number of times? */
select max(cnt) from
(
select product_id, count(order_id) as cnt from [dbo].[order_details]
group by product_id
) q1;
Results
Messages
Search to filter items...
90
Query succeeded | 1s
Search
12:32 PM
8/31/2023





3.5 Unstructured Data and Big Data

Learnt about the following:

1. Types of NoSQL Databases:
    1. Key-Value: Redis
    2. Document: MongoDB
    3. Wide-Column: Cassandra
    4. Graph-Based: Neo4j
2. Data Lake and Data Lakehouse
3. Big Data and its Tools:
    1. Apache Hadoop
    2. Apache MapReduce (not popular anymore)
    3. Apache Spark
4. Advanced SQL:
    1. Chasm, Fan, and Loop Traps
    2. Coalesce and Custom Functions in SQL
    3. UNION, MINUS, INTERSECT

Query 1 | Query 2

Run | Cancel query | Save query | Export data as | Show only Editor

```
1  select ename as "Employee", sal + coalesce(comm, 0) as "Income" from emp
2
```

Query 1

Run | Cancel query | Save query | Export data as | Show only Editor

```
1  ALTER FUNCTION Manager_Name (@empno  int)
2  RETURNS varchar(10) AS
3  BEGIN
4  DECLARE @mgname varchar(10);
5      select @mgname=ename from emp where mgr=@empno;
6      RETURN @mgname;
7  END;
8
```