

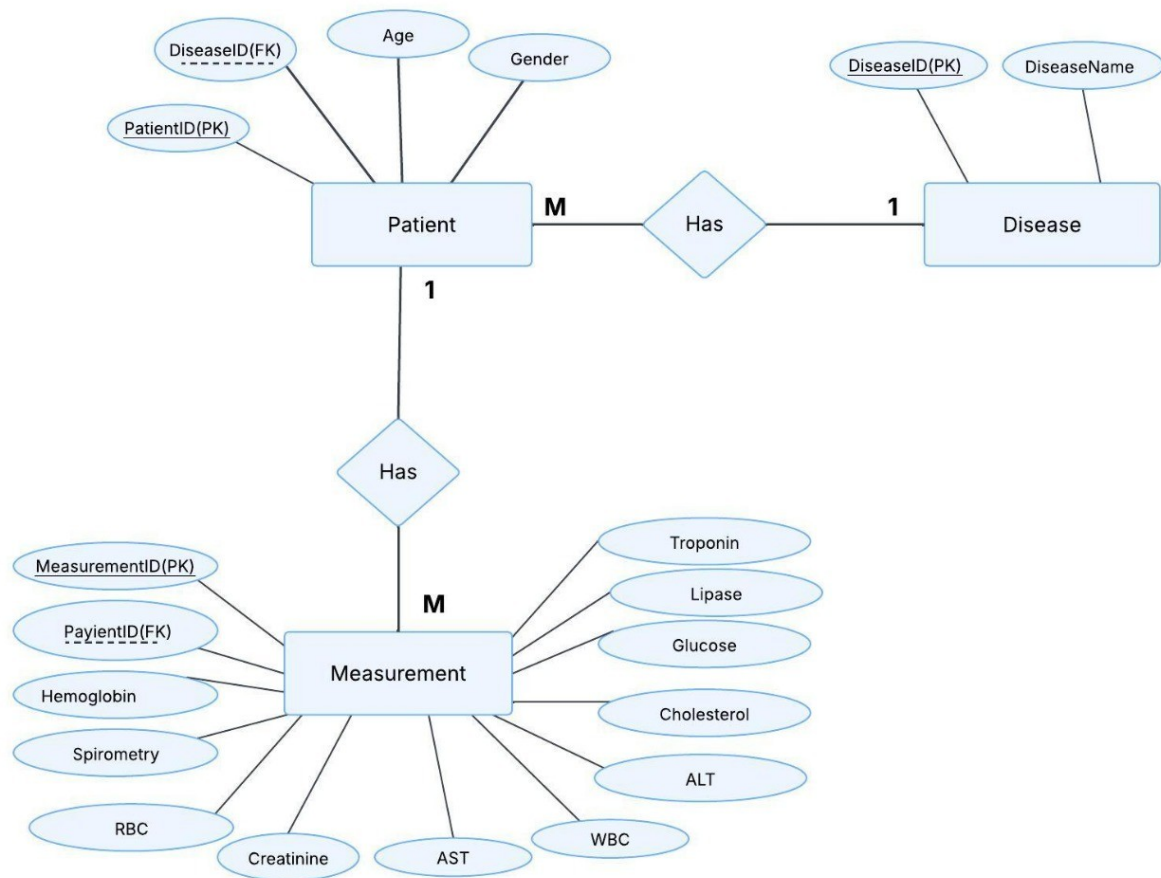
Data warehouse final project

DB source name: laboratory__data

Team members:

- | | | |
|---------------------------|----------|-----------------------------|
| • Rawan Elsayed Hisham | 42110142 | (42110142.Rawan@acu.edu.eg) |
| • Salma Abdrabo Abdelbary | 42110089 | |
| • Omar Esam Shokry | 42110192 | |
| • Yassa khalil naguip | 42110348 | |
| • Mustafa Safwat | 42110304 | |
| • Abdelwahab Ashraf | 42110375 | |

Source ERD:

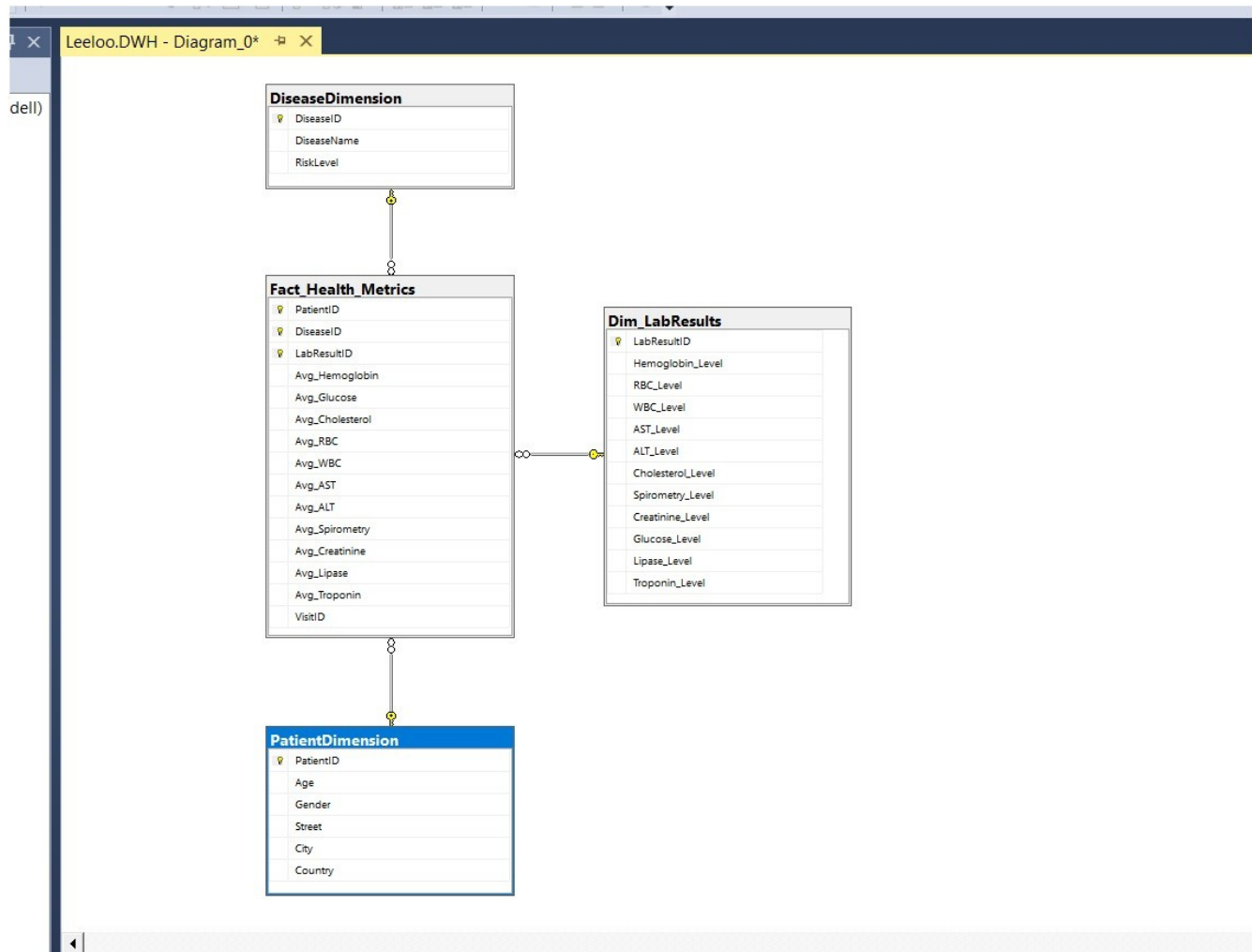


Motivation:

We are creating the Patient Health **Star Schema** & **Snowflake Schema** to analyze average medical measurements (such as Hemoglobin, Glucose, and Cholesterol) across patient demographics (Gender and Age), diseases (e.g., Anemia, Asthma), and raw measurement values. This will enable healthcare professionals to identify health trends, detect patterns in medical conditions, and assess risk factors for specific diseases, ultimately supporting better diagnosis and treatment decisions.

Star schema

Star schema model:



Schema description:

Fact Table (Fact_Health_Metrics)

- Contains aggregated health metrics (averages of lab results) for patients.
- Foreign keys:
 - PatientID (links to PatientDimension)
 - DiseaseID (links to DiseaseDimension)
 - LabResultID (links to Dim_LabResults)
- Metrics include averages for hemoglobin, glucose, cholesterol, RBC, WBC, AST, ALT, spirometry, creatinine, lipase, and troponin.

2. Dimension Tables:

- **DiseaseDimension**
 - Stores disease-related data (DiseaseID, DiseaseName, RiskLevel).

- **Dim_LabResults**
 - Contains raw lab test results (e.g., hemoglobin level, glucose level, cholesterol level).
- **PatientDimension**
 - Holds patient demographics (PatientID, Age, Gender, Street, City, Country).

Transformation process queries:

Query 1: Average Hemoglobin by Disease Risk Level

SQLQuery1.sql - (lo...(LEELOO\del (63)))* X Leeloo.DWH - Diagram_0*

```

SELECT
    DD.DiseaseName,
    DD.RiskLevel,
    AVG(FHM.AVG_Hemoglobin) AS AvgHemoglobinByDisease
FROM
    Fact_Health_Metrics FHM
JOIN
    DiseaseDimension DD ON FHM.DiseaseID = DD.DiseaseID
GROUP BY
    DD.DiseaseName, DD.RiskLevel
HAVING
    AVG(FHM.AVG_Hemoglobin) > 12
ORDER BY
    AvgHemoglobinByDisease DESC;

```

100 %

Results Messages

	DiseaseName	RiskLevel	AvgHemoglobinByDisease
1	Infection	Low	14.5510204081633
2	Kidney Disease	Medium	14.4317673378076
3	Diabetics	Medium	14.3597122302158
4	Liver Disease	Medium	14.3502581755594
5	Heart attack	High	14.3247588424437
6	Pancreatitis	High	14.3236282194849
7	Asthma	Low	14.3089005235602
8	Cardiovascular disease	High	14.2725968436155

Query 2: Patients with High Glucose or High Cholesterol

SQLQuery1.sql - (lo...(LEELOO\de11 (63))* X Leeloo.DWH - Diagram_0*

```

WHERE
    FHM.AVG_Glucose > 200
UNION
SELECT
    PD.PatientID,
    PD.Age,
    PD.Gender,
    FHM.AVG_Glucose,
    FHM.AVG_Cholesterol
FROM
    Fact_Health_Metrics FHM
JOIN
    PatientDimension PD ON FHM.PatientID = PD.PatientID
WHERE
    FHM.AVG_Cholesterol > 240
ORDER BY
    PD.PatientID;

```

100 %

Results Messages

	PatientID	Age	Gender	AVG_Glucose	AVG_Cholesterol
1	40	73	Female	70	250
2	69	66	Male	74	260
3	87	75	Female	76	260
4	112	20	Female	92	255
5	124	74	Female	96	266
6	143	83	Male	97	253
7	164	65	Female	74	262
8	184	84	Male	89	263
9	190	27	Male	87	270
10	198	41	Female	81	260
11	212	81	Male	88	263
12	247	41	Female	75	255
13	255	54	Male	85	256
14	273	32	Female	93	268
15	302	50	Male	92	268
16	311	52	Female	70	260
17	325	63	Male	97	271

Query executed successfully. (lo

Query 3: Lab Results Not Associated with High Risk Diseases

Execute

SQLQuery1.sql - (lo...(LEELOO\dell (63))* Leeloo.DWH - Diagram_0*

```
Dim_LabResults DLR
WHERE
    DLR.LabResultID IN (
        SELECT FHM.LabResultID
        FROM Fact_Health_Metrics FHM
        JOIN DiseaseDimension DD ON FHM.DiseaseID = DD.DiseaseID
        WHERE DD.RiskLevel != 'High'
    )
EXCEPT
SELECT
    DLR.LabResultID,
    DLR.Hemoglobin_LEVEL,
    DLR.RBC_LEVEL
FROM
    Dim_LabResults DLR
WHERE
    DLR.Hemoglobin_LEVEL < 10;
```

100 %

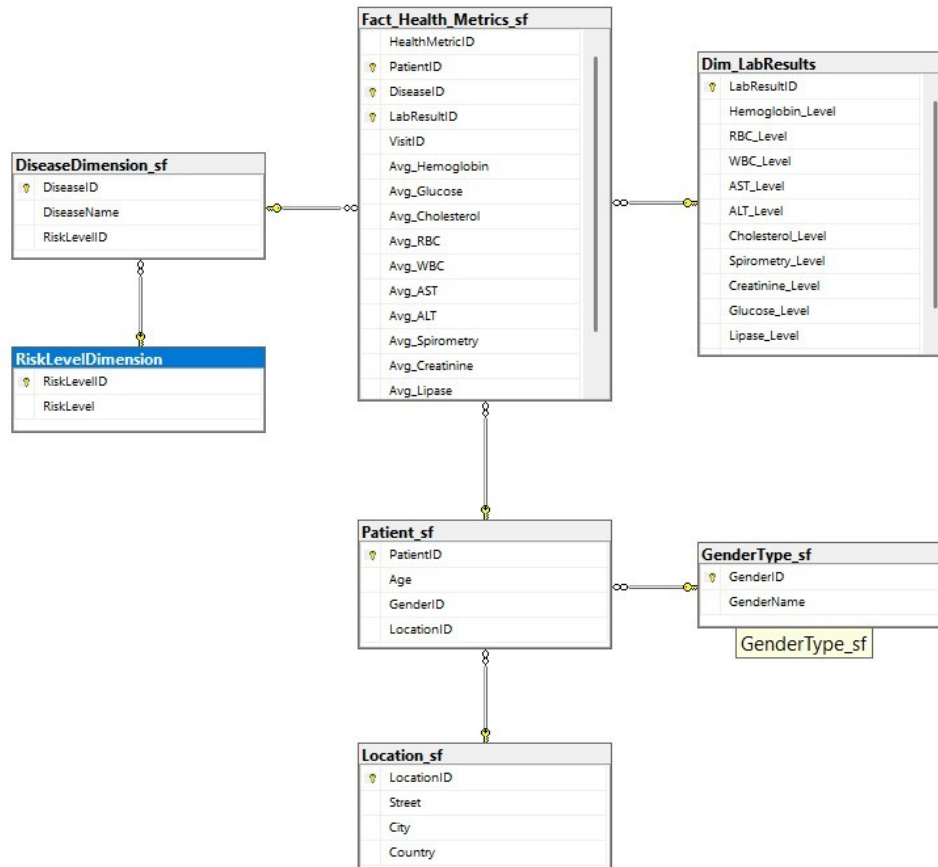
Results Messages

	LabResultID	Hemoglobin_LEVEL	RBC_LEVEL
1	2	40	20
2	3	14	3.54999995231628
3	4	16	4.94999980926514
4	5	16	3.96000003814697
5	6	15	4.30999994277954
6	8	34	24
7	10	16	4.5
8	11	15	4.46999979019165
9	12	14	4.84000015258789
10	13	13	4.51999998092651
11	14	13	4.82999992370605
12	16	16	4.53000020980835
13	17	13	4.71000003814697
14	21	15	4.8600001335144
15	22	13	3.52999997138977
16	23	16	3.57999992370605
17	24	13	5.05000019073486

Query executed successfully. (local) (16.0 I

Snowflake schema

Snowflake schema model:



Schema description:

Fact Table (Fact_Health_Metrics_sf)

- Contains aggregated health metrics (averages of lab results) linked to patients, diseases, and lab tests.
- Foreign keys:
 - PatientID → Patient_sf
 - DiseaseID → DiseaseDimension_sf
 - LabResultID → Dim_LabResults
 - VisitID (tracks patient visits)

- Metrics include averages for hemoglobin, glucose, cholesterol, RBC, WBC, AST, ALT, spirometry, creatinine, lipase, and troponin.

Dimension Tables (Normalized Structure):

1. **DiseaseDimension_sf**

- Stores disease data (DiseaseID, DiseaseName, RiskLevelID).
- Linked to **RiskLevelDimension** (RiskLevelID, RiskLevel) for risk categorization.

2. **Dim_LabResults**

- Contains raw lab test results (e.g., hemoglobin level, glucose level, cholesterol level).

3. **Patient_sf**

- Holds patient data (PatientID, Age, GenderID, LocationID).
- Linked to:
 - **GenderType_sf** (GenderID, GenderName) for gender classification.
 - **Location_sf** (LocationID, Street, City, Country) for address details.

Comparison to Star Schema :

- **Snowflake:** More normalized (e.g., Patient_sf links to GenderType_sf and Location_sf, DiseaseDimension_sf Linked to RiskLevelDimension).
- **Star:** Simpler, with denormalized dimensions (e.g., PatientDimension included gender and location directly, DiseaseDimension_sf included RiskLevel directly).

Transformation process queries:

Q*: This query calculates the average health metrics grouped by gender and disease:

SQLQuery7.sql - (lo...(LEELOO\dell (69))* X SQLQuery6.sql - (lo...(LEELOO\dell (62)) SQLQuery5.sql - (lo...(LEELOO\dell (95)) Leeloo.DWH -

```

gt.GenderName,
dd.DiseaseName,
AVG(fhm.Avg_Hemoglobin) AS Avg_Hemoglobin,
AVG(fhm.Avg_Glucose) AS Avg_Glucose,
AVG(fhm.Avg_Cholesterol) AS Avg_Cholesterol,
COUNT(*) AS Patient_Count
FROM
Fact_Health_Metrics_sf fhm
JOIN Patient_sf p ON fhm.PatientID = p.PatientID
JOIN GenderType_sf gt ON p.GenderID = gt.GenderID
JOIN DiseaseDimension_sf dd ON fhm.DiseaseID = dd.DiseaseID
GROUP BY
gt.GenderName,
dd.DiseaseName
HAVING
COUNT(*) > 1
ORDER BY
gt.GenderName, Avg_Glucose DESC;

```

100 %

Results Messages

	GenderName	DiseaseName	Avg_Hemoglobin	Avg_Glucose	Avg_Cholesterol	Patient_Count
1	Female	Diabetics	14.417518	135.097810	215.017518	685
2	Female	Infection	14.691637	87.909407	216.926829	574
3	Female	Asthma	14.332599	86.568281	214.585903	454
4	Female	Kidney Disease	14.433666	86.550641	215.503566	701
5	Female	Pancreatitis	14.256900	86.002123	214.891719	471
6	Female	Heart attack	14.462668	85.877600	214.440636	817
7	Female	Anemia	4.740625	85.558750	215.234375	1600
8	Female	Liver Disease	14.139573	84.965517	215.724137	609
9	Female	Cardiovascular disease	14.117449	83.510067	255.617449	298
10	Male	Diabetics	14.289752	133.807420	215.980565	566
11	Male	Cardiovascular disease	14.388471	87.308270	258.120300	399
12	Male	Pancreatitis	14.398104	86.345971	216.061611	422
13	Male	Heart attack	14.172086	85.939024	216.272357	738

Query executed successfully. (local) (16.0 RTM) | LEELOO\dell (69) | D

Ln 19 Col 37 Ch 37 INS

Search

Query 1: Average Health Metrics by Location and Disease

SQLQuery7.sql - (lo...(LEELOO\del (69))*)						
SQLQuery6.sql - (lo...(LEELOO\del (62))						
SQLQuery5.sql - (lo...(LEELOO\de						
<pre> SELECT l.City, dd.DiseaseName, AVG(fhm.Avg_Hemoglobin) AS Avg_Hemoglobin, AVG(fhm.Avg_Glucose) AS Avg_Glucose, AVG(fhm.Avg_Cholesterol) AS Avg_Cholesterol, COUNT(DISTINCT fhm.PatientID) AS Unique_Patient_Count FROM Fact_Health_Metrics_sf fhm JOIN Patient_sf p ON fhm.PatientID = p.PatientID JOIN Location_sf l ON p.LocationID = l.LocationID JOIN DiseaseDimension_sf dd ON fhm.DiseaseID = dd.DiseaseID GROUP BY l.City, dd.DiseaseName HAVING COUNT(DISTINCT fhm.PatientID) >= 2 ORDER BY Avg_Glucose DESC; </pre>						
100 %						
Results Messages						
	City	DiseaseName	Avg_Hemoglobin	Avg_Glucose	Avg_Cholesterol	Unique_Patient_Count
1	Los Angeles	Diabetics	14.422705	135.004830	215.236714	414
2	New York	Diabetics	14.242647	134.678921	215.259803	408
3	Chicago	Diabetics	14.410256	133.883449	215.846153	429
4	Los Angeles	Infection	14.455470	87.020356	216.923664	393
5	New York	Infection	14.739240	86.817721	216.835443	395
6	Chicago	Infection	14.456185	86.783505	217.221649	388
7	Los Angeles	Asthma	14.396226	86.468553	214.795597	318
8	New York	Pancreatitis	14.223684	86.394736	215.332236	304
9	Chicago	Asthma	14.184126	86.266666	214.695238	315
10	Chicago	Cardiovascular disease	14.369098	86.158798	256.798283	233
11	New York	Heart attack	14.220272	86.142300	215.337231	513
12	Los Angeles	Pancreatitis	14.351973	86.065789	215.486842	304
13	New York	Kidney Disease	14.386877	86.065610	215.900452	442
Query executed successfully. (local) (16.						
Ln 19 Col 22						

Query 2: High-Risk Patients by Gender

WITH HighRiskPatients AS (

SELECT

p.PatientID,

gt.GenderName,

dd.DiseaseName,

fhm.Avg_Hemoglobin

FROM

Fact_Health_Metrics_sf fhm

JOIN Patient_sf p ON fhm.PatientID = p.PatientID

JOIN GenderType_sf gt ON p.GenderID = gt.GenderID

JOIN DiseaseDimension_sf dd ON fhm.DiseaseID = dd.DiseaseID

JOIN RiskLevelDimension rl ON dd.RiskLevelID = rl.RiskLevelID

```
WHERE
    rl.RiskLevel = 'High'
)
SELECT
    GenderName,
    DiseaseName,
    COUNT(DISTINCT PatientID) AS HighRisk_Patient_Count
FROM
    HighRiskPatients
WHERE
    PatientID IN (
        SELECT PatientID
        FROM HighRiskPatients
        EXCEPT
        SELECT PatientID
        FROM HighRiskPatients
        WHERE Avg_Hemoglobin BETWEEN 12 AND 16 -- Normal range for hemoglobin
    )
GROUP BY
    GenderName,
    DiseaseName
ORDER BY
    HighRisk_Patient_Count DESC;
```

SQLQuery7.sql - (lo...(LEELOO\deli (69))* X SQLQuery6.sql - (lo...(LEELOO\deli (62)) SQLQuery5.sql - (lo...(LEELOO\deli (95)) Leeloo

```

)
SELECT
    GenderName,
    DiseaseName,
    COUNT(DISTINCT PatientID) AS HighRisk_Patient_Count
FROM
    HighRiskPatients
WHERE
    PatientID IN (
        SELECT PatientID
        FROM HighRiskPatients
        EXCEPT
        SELECT PatientID
        FROM HighRiskPatients
        WHERE Avg_Hemoglobin BETWEEN 12 AND 16 -- Normal range for hemoglobin
    )
GROUP BY
    GenderName,
    DiseaseName
ORDER BY
    HighRisk_Patient_Count DESC;

```

100 %

Results Messages

	GenderName	DiseaseName	HighRisk_Patient_Count
1	Male	Heart attack	18
2	Female	Heart attack	17
3	Female	Cardiovascular disease	12
4	Male	Cardiovascular disease	10
5	Female	Pancreatitis	8
6	Male	Pancreatitis	8

Query 3: Compare Lab Metrics for Young vs. Older Patients

```

SELECT
    'Young (Age <= 40)' AS Age_Group,
    dd.DiseaseName,
    AVG(fhm.Avg_Glucose) AS Avg_Glucose,
    AVG(fhm.Avg_Cholesterol) AS Avg_Cholesterol,
    COUNT(*) AS Patient_Count
FROM
    Fact_Health_Metrics_sf fhm
    JOIN Patient_sf p ON fhm.PatientID = p.PatientID
    JOIN DiseaseDimension_sf dd ON fhm.DiseaseID = dd.DiseaseID
WHERE
    p.Age <= 40
GROUP BY
    dd.DiseaseName
UNION
SELECT
    'Older (Age > 40)' AS Age_Group,
    dd.DiseaseName,
    AVG(fhm.Avg_Glucose) AS Avg_Glucose,
    AVG(fhm.Avg_Cholesterol) AS Avg_Cholesterol,
    COUNT(*) AS Patient_Count
FROM
    Fact_Health_Metrics_sf fhm
    JOIN Patient_sf p ON fhm.PatientID = p.PatientID
    JOIN DiseaseDimension_sf dd ON fhm.DiseaseID = dd.DiseaseID
WHERE
    p.Age > 40
GROUP BY
    dd.DiseaseName

```

```

COUNT(*) AS Patient_Count
FROM
    Fact_Health_Metrics_sf fhm
    JOIN Patient_sf p ON fhm.PatientID = p.PatientID
    JOIN DiseaseDimension_sf dd ON fhm.DiseaseID = dd.DiseaseID
WHERE
    p.Age > 40
GROUP BY
    dd.DiseaseName
ORDER BY
    DiseaseName, Age_Group;

```

SQLQuery7.sql - (lo...(LEELOO\del1 (69)))* - X SQLQuery6.sql - (lo...(LEELOO\del1 (62)) SQLQuery5.sql - (lo...(LEELOO\del1 (9

```

JOIN Patient_sf p ON fhm.PatientID = p.PatientID
JOIN DiseaseDimension_sf dd ON fhm.DiseaseID = dd.DiseaseID
WHERE
    p.Age <= 40
GROUP BY
    dd.DiseaseName
UNION
SELECT
    'Older (Age > 40)' AS Age_Group,
    dd.DiseaseName,
    AVG(fhm.Avg_Glucose) AS Avg_Glucose,
    AVG(fhm.Avg_Cholesterol) AS Avg_Cholesterol,
    COUNT(*) AS Patient_Count
FROM
    Fact_Health_Metrics_sf fhm
    JOIN Patient_sf p ON fhm.PatientID = p.PatientID
    JOIN DiseaseDimension_sf dd ON fhm.DiseaseID = dd.DiseaseID
WHERE
    p.Age > 40
GROUP BY
    dd.DiseaseName
ORDER BY
    DiseaseName, Age_Group;

```

100 %

Results Messages

	Age_Group	DiseaseName	Avg_Glucose	Avg_Cholesterol	Patient_Count
1	Older (Age > 40)	Anemia	84.955404	215.351429	2063
2	Young (Age <= 40)	Anemia	85.237991	215.861353	916
3	Older (Age > 40)	Asthma	86.861356	214.523598	678
4	Young (Age <= 40)	Asthma	84.270758	214.501805	277
5	Older (Age > 40)	Cardiovascular disease	85.529292	257.589898	495
6	Young (Age <= 40)	Cardiovascular disease	86.064356	255.727722	202
7	Older (Age > 40)	Diabetics	134.223121	215.804624	865
8	Young (Age <= 40)	Diabetics	135.165803	214.665803	386
9	Older (Age > 40)	Heart attack	85.854205	215.438317	1070
10	Young (Age <= 40)	Heart attack	86.022680	215.026804	485
11	Older (Age > 40)	Infection	87.958748	217.163584	703
12	Young (Age <= 40)	Infection	85.262156	216.737843	473

Query executed successfully. (local) (16.0 RT (local) (16.0 Encryption:

Ln 31 Col 28

Q4: Dangerous diseases only

SQLQuery7.sql - (lo...(LEELOO\deII (69))* SQLQuery6.sql - (lo...(LEELOO\deII (62)) SQLQuery5.sql - (lo...(LEELOO\deII (61))

```
SELECT
    d.DiseaseName,
    r.RiskLevel
FROM DiseaseDimension_sf d
JOIN RiskLevelDimension r ON d.RiskLevelID = r.RiskLevelID
WHERE r.RiskLevel = 'High';
```

100 %

Results Messages

	DiseaseName	RiskLevel
1	Cardiovascular disease	High
2	Pancreatitis	High
3	Heart attack	High

Query executed successfully. (loc

Q5: Patients number in each city

SQLQuery7.sql - (lo...(LEELOO\dell (69))* SQLQuery6.sql - (lo...(LEELOO\dell (62)) SQLQuery5.sql - (lo...(L

```
SELECT
    l.City,
    COUNT(DISTINCT p.PatientID) AS NumberOfPatients
FROM Patient_sf p
JOIN Location_sf l ON p.LocationID = l.LocationID
GROUP BY l.City
ORDER BY NumberOfPatients DESC;
```

00 %

Results Messages

	City	NumberOfPatients
1	Chicago	4003
2	Los Angeles	4003
3	New York	4003

Q6: Patients with ALT levels higher than the general average

```

SELECT
    f.PatientID,
    l.ALT_Level
FROM Fact_Health_Metrics_sf f
JOIN Dim_LabResults l ON f.LabResultID = l.LabResultID
WHERE l.ALT_Level > (
    SELECT AVG(ALT_Level)
    FROM Dim_LabResults
);

```

100 %

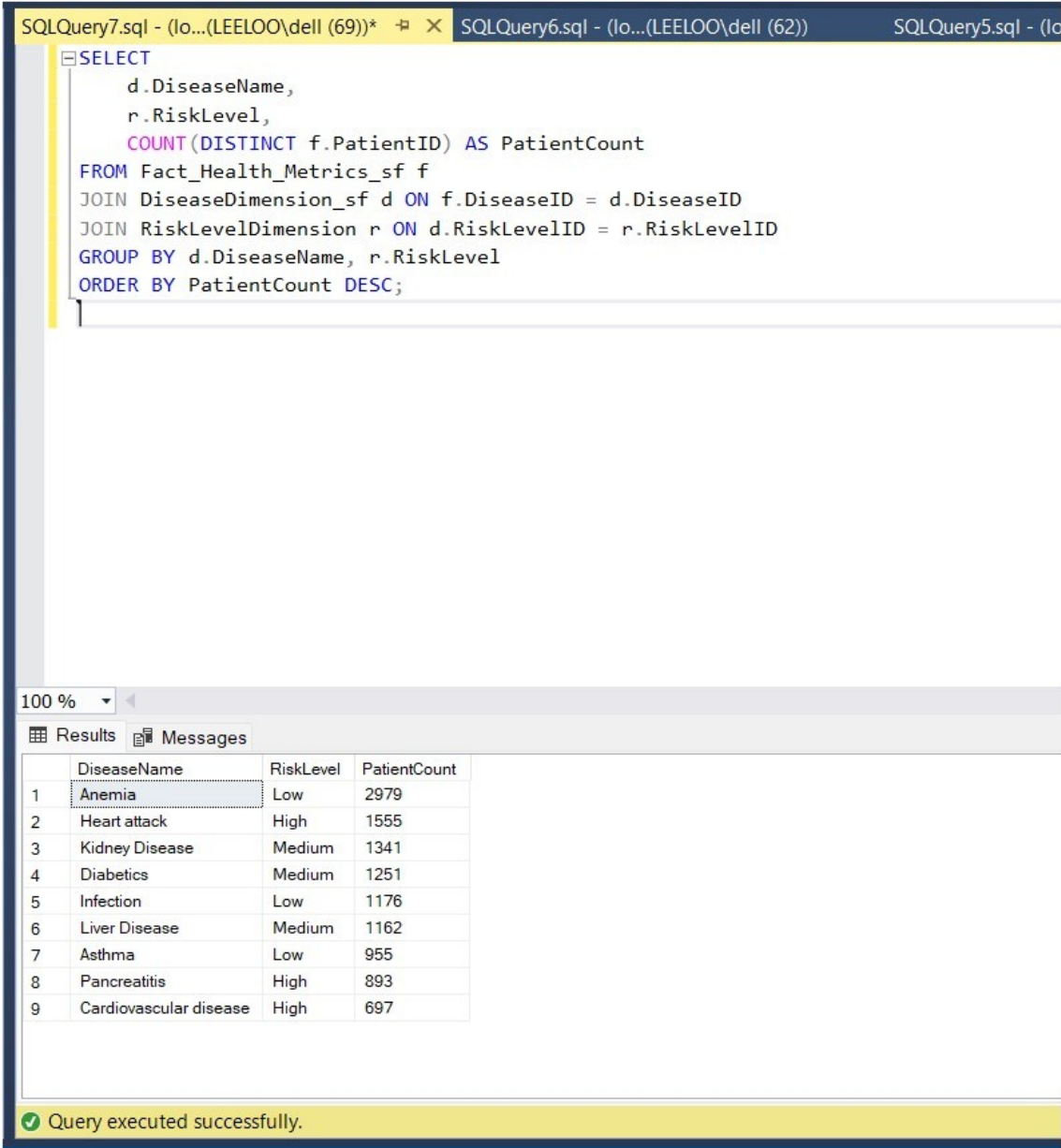
Results Messages

	PatientID	ALT_Level
1	4	62
2	6	35
3	9	32
4	10	63
5	17	61
6	18	35
7	20	34
8	22	63
9	27	65
10	28	33
11	29	32
12	31	33

✓ Query executed successfully.

(lc

Q7: Analysis of the number of patients by disease type and severity



The screenshot shows a SQL Server Enterprise Manager window with three tabs: 'SQLQuery7.sql - (lo...(LEELOO\delI (69)))*', 'SQLQuery6.sql - (lo...(LEELOO\delI (62))', and 'SQLQuery5.sql - (lo...'. The active tab is 'SQLQuery7.sql'. The query editor displays the following SQL code:

```
SELECT
    d.DiseaseName,
    r.RiskLevel,
    COUNT(DISTINCT f.PatientID) AS PatientCount
FROM Fact_Health_Metrics_sf f
JOIN DiseaseDimension_sf d ON f.DiseaseID = d.DiseaseID
JOIN RiskLevelDimension r ON d.RiskLevelID = r.RiskLevelID
GROUP BY d.DiseaseName, r.RiskLevel
ORDER BY PatientCount DESC;
```

Below the query editor, the 'Results' tab is selected, showing a table with 9 rows and 4 columns: 'DiseaseName', 'RiskLevel', and 'PatientCount'. The data is as follows:

	DiseaseName	RiskLevel	PatientCount
1	Anemia	Low	2979
2	Heart attack	High	1555
3	Kidney Disease	Medium	1341
4	Diabetics	Medium	1251
5	Infection	Low	1176
6	Liver Disease	Medium	1162
7	Asthma	Low	955
8	Pancreatitis	High	893
9	Cardiovascular disease	High	697

At the bottom of the window, a yellow status bar indicates 'Query executed successfully.'

Query 8: Analyzes average measurements by age group and disease with risk level

```
SELECT
CASE
    WHEN p.Age <= 30 THEN 'Young (<=30)'
    WHEN p.Age BETWEEN 31 AND 50 THEN 'Middle (31-50)'
    ELSE 'Senior (>50)'
END AS Age_Group,
dd.DiseaseName,
rl.RiskLevel,
AVG(fhm.Avg_Hemoglobin) AS Avg_Hemoglobin,
AVG(fhm.Avg_Glucose) AS Avg_Glucose,
```

```

    AVG(fhm.Avg_Cholesterol) AS Avg_Cholesterol,
    COUNT(DISTINCT p.PatientID) AS Patient_Count,
    SUM(CASE WHEN fhm.Avg_Hemoglobin < 10 THEN 1 ELSE 0 END) AS
Critical_Hemoglobin_Count
FROM
    Fact_Health_Metrics_sf fhm
    JOIN Patient_sf p ON fhm.PatientID = p.PatientID
    JOIN DiseaseDimension_sf dd ON fhm.DiseaseID = dd.DiseaseID
    JOIN RiskLevelDimension rl ON dd.RiskLevelID = rl.RiskLevelID
WHERE
    dd.DiseaseName IN ('Anemia', 'Asthma')
GROUP BY
    CASE
        WHEN p.Age <= 30 THEN 'Young (≤30)'
        WHEN p.Age BETWEEN 31 AND 50 THEN 'Middle (31-50)'
        ELSE 'Senior (>50)'
    END,
    dd.DiseaseName,
    rl.RiskLevel
HAVING
    SUM(CASE WHEN fhm.Avg_Hemoglobin < 10 THEN 1 ELSE 0 END) > 0
ORDER BY
    Critical_Hemoglobin_Count DESC, Age_Group;

```

SQLQuery7.sql - (lo...(LEELOO\dell (69))* SQLQuery6.sql - (lo...(LEELOO\dell (62)) SQLQuery5.sql - (lo...(LEELOO\dell (95)) Leefoo.DWH - Diagram_1*

```

AVG(fhm.Avg_Glucose) AS Avg_Glucose,
AVG(fhm.Avg_Cholesterol) AS Avg_Cholesterol,
COUNT(DISTINCT p.PatientID) AS Patient_Count,
SUM(CASE WHEN fhm.Avg_Hemoglobin < 10 THEN 1 ELSE 0 END) AS Critical_Hemoglobin_Count
FROM
Fact_Health_Metrics_sf fhm
JOIN Patient_sf p ON fhm.PatientID = p.PatientID
JOIN DiseaseDimension_sf dd ON fhm.DiseaseID = dd.DiseaseID
JOIN RiskLevelDimension rl ON dd.RiskLevelID = rl.RiskLevelID
WHERE
dd.DiseaseName IN ('Anemia', 'Asthma')
GROUP BY
CASE
WHEN p.Age <= 30 THEN 'Young (<30)'
WHEN p.Age BETWEEN 31 AND 50 THEN 'Middle (31-50)'
ELSE 'Senior (>50)'
END,
dd.DiseaseName,
rl.RiskLevel
HAVING
SUM(CASE WHEN fhm.Avg_Hemoglobin < 10 THEN 1 ELSE 0 END) > 0
ORDER BY
Critical_Hemoglobin_Count DESC, Age_Group;

```

100 %

Results Messages

	Age_Group	DiseaseName	RiskLevel	Avg_Hemoglobin	Avg_Glucose	Avg_Cholesterol	Patient_Count	Critical_Hemoglobin_Count
1	Senior (>50)	Anemia	Low	4.660046	85.738901	215.304906	1712	1641
2	Middle (31-50)	Anemia	Low	4.883963	83.172099	215.775749	767	738
3	Young (<30)	Anemia	Low	4.922000	85.526000	215.794000	500	482
4	Senior (>50)	Asthma	Low	14.260794	87.295336	214.934369	579	11
5	Young (<30)	Asthma	Low	14.036764	85.073529	215.500000	136	2
6	Middle (31-50)	Asthma	Low	14.579166	83.837500	212.954166	240	1

Query executed successfully. (local) (16.0 RTM) LEELOO\dell (69) DWH 00:00:00 6 rows

Query 9: Identify Patients with Abnormal Measurements Excluding Low-Risk Diseases

WITH AbnormalPatients AS (

SELECT

p.PatientID,

p.Age,

gt.GenderName,

dd.DiseaseName,

fhm.Avg_Hemoglobin

FROM

Fact_Health_Metrics_sf fhm

JOIN Patient_sf p ON fhm.PatientID = p.PatientID

JOIN GenderType_sf gt ON p.GenderID = gt.GenderID

JOIN DiseaseDimension_sf dd ON fhm.DiseaseID = dd.DiseaseID

JOIN RiskLevelDimension rl ON dd.RiskLevelID = rl.RiskLevelID

WHERE

fhm.Avg_Hemoglobin NOT BETWEEN 12 AND 16

)

SELECT

GenderName,

DiseaseName,

```
AVG(Age) AS Avg_Age,
COUNT(DISTINCT PatientID) AS Abnormal_Patient_Count,
MIN(Avg_Hemoglobin) AS Min_Hemoglobin,
MAX(Avg_Hemoglobin) AS Max_Hemoglobin
FROM
    AbnormalPatients
WHERE
    PatientID IN (
        SELECT PatientID
        FROM AbnormalPatients
        EXCEPT
        SELECT PatientID
        FROM AbnormalPatients ap
        JOIN DiseaseDimension_sf dd ON ap.DiseaseName = dd.DiseaseName
        JOIN RiskLevelDimension rl ON dd.RiskLevelID = rl.RiskLevelID
        WHERE rl.RiskLevel = 'Low'
    )
GROUP BY
    GenderName,
    DiseaseName
ORDER BY
    Abnormal_Patient_Count DESC;
```

SQLQuery7.sql - (lo...(LEELOO\deli (69))* X SQLQuery6.sql - (lo...(LEELOO\deli (62)) SQLQuery5.sql - (lo...(LEELOO\deli (95))

```

WITH AbnormalPatients AS (
    SELECT
        p.PatientID,
        p.Age,
        gt.GenderName,
        dd.DiseaseName,
        fhm.Avg_Hemoglobin
    FROM
        Fact_Health_Metrics_sf fhm
    JOIN Patient_sf p ON fhm.PatientID = p.PatientID
    JOIN GenderType_sf gt ON p.GenderID = gt.GenderID
    JOIN DiseaseDimension_sf dd ON fhm.DiseaseID = dd.DiseaseID
    JOIN RiskLevelDimension rl ON dd.RiskLevelID = rl.RiskLevelID
    WHERE
        fhm.Avg_Hemoglobin NOT BETWEEN 12 AND 16
)
SELECT
    GenderName,
    DiseaseName,
    AVG(Age) AS Avg_Age,
    COUNT(DISTINCT PatientID) AS Abnormal_Patient_Count,
    MIN(Avg_Hemoglobin) AS Min_Hemoglobin,
    MAX(Avg_Hemoglobin) AS Max_Hemoglobin

```

100 %

Results Messages

	GenderName	DiseaseName	Avg_Age	Abnormal_Patient_Count	Min_Hemoglobin	Max_Hemoglobin
1	Male	Heart attack	57	18	3.00	47.00
2	Female	Heart attack	56	17	4.00	42.00
3	Female	Diabetics	59	16	2.00	34.00
4	Male	Diabetics	50	16	3.00	44.00
5	Male	Kidney Disease	56	16	3.00	42.00
6	Female	Kidney Disease	55	13	3.00	23.00
7	Female	Liver Disease	51	12	3.00	6.00
8	Female	Cardiovascular disease	49	12	3.00	28.00
9	Male	Cardiovascular disease	62	10	3.00	6.00
10	Male	Liver Disease	54	9	4.00	27.00
11	Female	Pancreatitis	65	8	3.00	36.00
12	Male	Pancreatitis	47	8	3.00	39.00

Query executed successfully. (local) (16.0 RTM) LEELC

Ln 41 Col 33 Ch 33

Query 10: Average Measurements by Age Group and Disease with Risk Level

```

SELECT
    CASE
        WHEN p.Age <= 30 THEN 'Young (<=30)'
        WHEN p.Age BETWEEN 31 AND 50 THEN 'Middle (31-50)'
        ELSE 'Senior (>50)'
    END AS Age_Group,
    dd.DiseaseName,
    rl.RiskLevel,
    AVG(fhm.Avg_Hemoglobin) AS Avg_Hemoglobin,
    AVG(fhm.Avg_Glucose) AS Avg_Glucose,
    AVG(fhm.Avg_Cholesterol) AS Avg_Cholesterol,
    COUNT(DISTINCT p.PatientID) AS Patient_Count
FROM

```

```

Fact_Health_Metrics_sf fhm
JOIN Patient_sf p ON fhm.PatientID = p.PatientID
JOIN DiseaseDimension_sf dd ON fhm.DiseaseID = dd.DiseaseID
JOIN RiskLevelDimension rl ON dd.RiskLevelID = rl.RiskLevelID
WHERE
    dd.DiseaseName IN ('Anemia', 'Asthma')
GROUP BY
    CASE
        WHEN p.Age <= 30 THEN 'Young (<=30)'
        WHEN p.Age BETWEEN 31 AND 50 THEN 'Middle (31-50)'
        ELSE 'Senior (>50)'
    END,
    dd.DiseaseName,
    rl.RiskLevel
ORDER BY
    Age_Group, Avg_Hemoglobin DESC;

```

SQLQuery7.sql - (lo...(LEELOO\dell (69))* X SQLQuery6.sql - (lo...(LEELOO\dell (62)) SQLQuery5.sql - (lo...(LEELOO\dell (95)) Le

```

dd.DiseaseName,
rl.RiskLevel,
AVG(fhm.Avg_Hemoglobin) AS Avg_Hemoglobin,
AVG(fhm.Avg_Glucose) AS Avg_Glucose,
AVG(fhm.Avg_Cholesterol) AS Avg_Cholesterol,
COUNT(DISTINCT p.PatientID) AS Patient_Count
FROM
Fact_Health_Metrics_sf fhm
JOIN Patient_sf p ON fhm.PatientID = p.PatientID
JOIN DiseaseDimension_sf dd ON fhm.DiseaseID = dd.DiseaseID
JOIN RiskLevelDimension rl ON dd.RiskLevelID = rl.RiskLevelID
WHERE
    dd.DiseaseName IN ('Anemia', 'Asthma')
GROUP BY
    CASE
        WHEN p.Age <= 30 THEN 'Young (<=30)'
        WHEN p.Age BETWEEN 31 AND 50 THEN 'Middle (31-50)'
        ELSE 'Senior (>50)'
    END,
    dd.DiseaseName,
    rl.RiskLevel
ORDER BY
    Age_Group, Avg_Hemoglobin DESC;

```

100 %

Results Messages

	Age_Group	DiseaseName	RiskLevel	Avg_Hemoglobin	Avg_Glucose	Avg_Cholesterol	Patient_Count
1	Middle (31-50)	Asthma	Low	14.579166	83.837500	212.954166	240
2	Middle (31-50)	Anemia	Low	4.883963	83.172099	215.775749	767
3	Senior (>50)	Asthma	Low	14.260794	87.295336	214.934369	579
4	Senior (>50)	Anemia	Low	4.660046	85.738901	215.304906	1712
5	Young (<=30)	Asthma	Low	14.036764	85.073529	215.500000	136
6	Young (<=30)	Anemia	Low	4.922000	85.526000	215.794000	500

Query executed successfully. (local) (16.0 RTM) LEELOO