

Dimensional Model

Bhagyashri Avinash Pagar

NUID: 002310690

Dimensions:

DIM_DATE

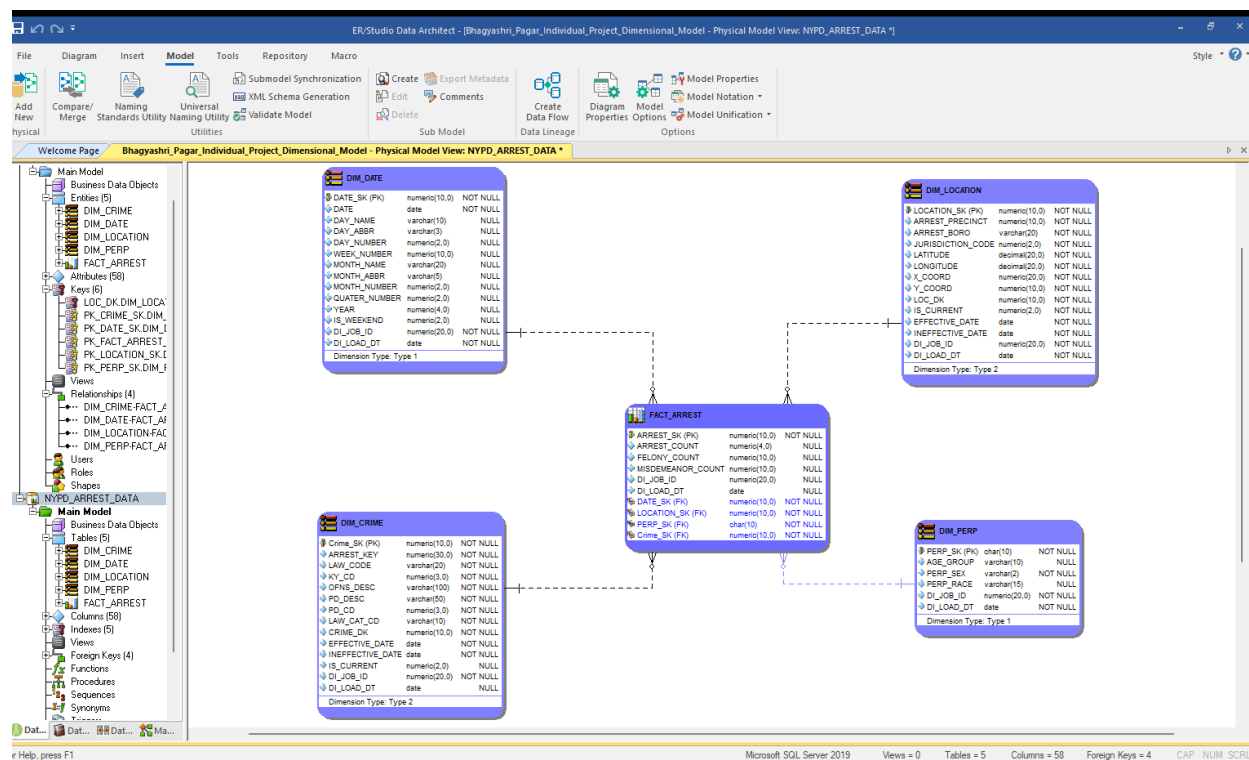
DIM_LOCATION

DIM_CRIME

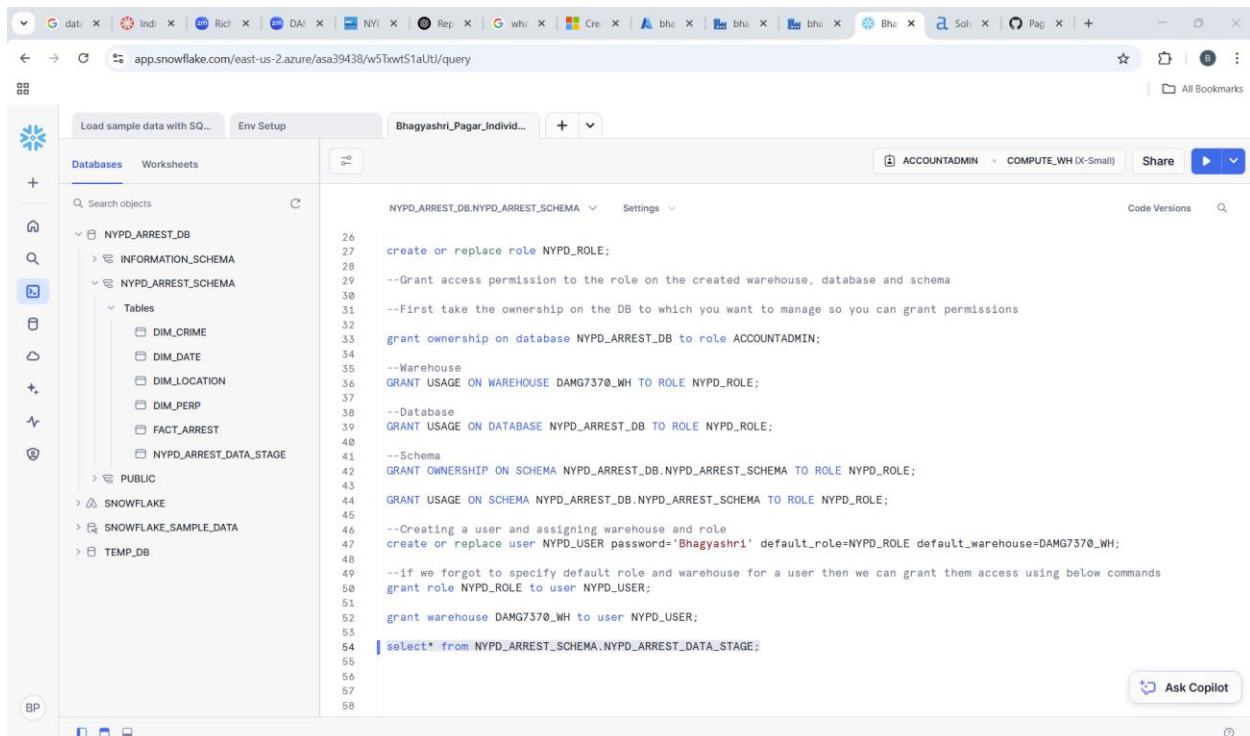
DIM_PERP

Fact:

FACT_ARREST



Tables for the above dimensions and facts are created in the snowflake:



SQL Queries for the given Business Requirements:

1. How many arrests occurred on any specific day, week, month, quarter, or year?

```
SELECT d.DATE, d.WEEK_NUMBER, d.MONTH_NAME, d.QUARTER_NUMBER,
d.YEAR,
```

```
SUM(f.ARREST_COUNT) AS Total_Arrests
```

```
FROM FACT_ARREST f JOIN DIM_DATE d ON f.DATE_SK = d.DATE_SK
```

```
GROUP BY d.DATE, d.WEEK_NUMBER, d.MONTH_NAME, d.QUARTER_NUMBER,
d.YEAR
```

```
ORDER BY d.YEAR, d.QUARTER_NUMBER, d.MONTH_NAME, d.WEEK_NUMBER,
d.DATE;
```

2. What are the peak days and months for arrests?

```
SELECT d.DAY_NAME, d.MONTH_NAME, SUM(f.ARREST_COUNT) AS Total_Arrests
```

```
FROM FACT_ARREST f
```

```
JOIN DIM_DATE d ON f.DATE_SK = d.DATE_SK
GROUP BY d.DAY_NAME, d.MONTH_NAME
ORDER BY Total_Arrests DESC
LIMIT 10;
```

3. What are the top 5 most frequently occurring crimes?

```
SELECT c.OFNS_DESC, COUNT(*) AS Crime_Count
FROM FACT_ARREST f
JOIN DIM_CRIME c ON f.Crime_SK = c.Crime_SK
GROUP BY c.OFNS_DESC
ORDER BY Crime_Count DESC
LIMIT 5;
```

4. Which crimes have increased or decreased the most over time?

```
SELECT d.YEAR, c.OFNS_DESC, COUNT(*) AS Crime_Count
FROM FACT_ARREST f
JOIN DIM_DATE d ON f.DATE_SK = d.DATE_SK
JOIN DIM_CRIME c ON f.Crime_SK = c.Crime_SK
GROUP BY d.YEAR, c.OFNS_DESC
ORDER BY d.YEAR, Crime_Count DESC;
```

5. Are there specific precincts with higher felony arrests compared to misdemeanors?

```
SELECT l.ARREST_PRECINCT,
       SUM(f.FELONY_COUNT) AS Total_Felonies,
       SUM(f.MISDEMEANOR_COUNT) AS Total_Misdemeanors
FROM FACT_ARREST f
JOIN DIM_LOCATION l ON f.LOCATION_SK = l.LOCATION_SK
GROUP BY l.ARREST_PRECINCT
```

```
HAVING SUM(f.FELONY_COUNT) > SUM(f.MISDEMEANOR_COUNT)
ORDER BY Total_Felonies DESC;
```

6. Which borough has the highest number of arrests?

```
SELECT l.ARREST_BORO, SUM(f.ARREST_COUNT) AS Total_Arrests
FROM FACT_ARREST f
JOIN DIM_LOCATION l ON f.LOCATION_SK = l.LOCATION_SK
GROUP BY l.ARREST_BORO
ORDER BY Total_Arrests DESC
LIMIT 1;
```

7. What is the distribution of arrestees by age, race, and gender?

```
SELECT p.AGE_GROUP, p.PERP_RACE, p.PERP_SEX, COUNT(*) AS Arrestee_Count
FROM FACT_ARREST f
JOIN DIM_PERP p ON f.PERP_SK = p.PERP_SK
GROUP BY p.AGE_GROUP, p.PERP_RACE, p.PERP_SEX
ORDER BY Arrestee_Count DESC;
```

8. Can we predict high-crime areas based on past arrest data?

Yes

```
SELECT l.LOCATION_SK, l.ARREST_PRECINCT, l.ARREST_BORO, COUNT(*) AS
Arrest_Count
FROM FACT_ARREST f
JOIN DIM_LOCATION l ON f.LOCATION_SK = l.LOCATION_SK
GROUP BY l.LOCATION_SK, l.ARREST_PRECINCT, l.ARREST_BORO ORDER BY
Arrest_Count DESC
LIMIT 10;
```