

### **Individual Project 1: NYPD Arrest Data**

In this project, Alteryx and Azure Data Factory (ADF) will be used to create a Dimensional Model for NYPD arrest data analysis and to develop a Data Staging Pipeline. Supporting analytical inquiries about NYC crime trends is the main objective to assist the public and city officials in comprehending arrest trends.

-Business Requirements:

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

Relevant - The DIM\_DATE table provides DAY\_NUM, WEEK\_NUM, MONTH\_NUM, QTR\_NUM, and YEAR\_NUM for time-based analysis.

QUERY:

```
SELECT YEAR_NUM, MONTH_NUM, WEEK_NUM, DAY_NUM, COUNT(*) AS  
total_arrests  
  
FROM CRIME_FACT CF  
  
JOIN DIM_DATE DD ON CF.DATE_SK = DD.DATE_SK  
  
GROUP BY YEAR_NUM, MONTH_NUM, WEEK_NUM, DAY_NUM  
  
ORDER BY YEAR_NUM DESC, MONTH_NUM DESC, WEEK_NUM DESC,  
DAY_NUM DESC;
```

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

Relevant - DIM\_DATE allows us to identify the most common days and months.

QUERY:

```
SELECT MONTH_NUM, COUNT(*) AS total_arrests  
  
FROM CRIME_FACT CF  
  
JOIN DIM_DATE DD ON CF.DATE_SK = DD.DATE_SK  
  
GROUP BY MONTH_NUM  
  
ORDER BY total_arrests DESC
```

LIMIT 5;

To analyze peak days:

```
SELECT DAY_NUM, COUNT(*) AS total_arrests FROM CRIME_FACT CF JOIN  
DIM_DATE DD ON CF.DATE_SK = DD.DATE_SK GROUP BY DAY_NUM ORDER BY  
total_arrests DESC LIMIT 5;
```

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

Relevant - The DIM\_CRIMES table contains PD\_DESC (precinct description) and OFFENSE\_CAT (offense category).

QUERY:

```
SELECT PD_DESC, COUNT(*) AS total_occurrences  
FROM CRIME_FACT CF  
JOIN DIM_CRIMES DC ON CF.CRIMES_SK = DC.CRIMES_SK  
GROUP BY PD_DESC  
ORDER BY total_occurrences DESC  
LIMIT 5;
```

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

Relevant - The DIM\_DATE table allows us to analyze trends over time.

QUERY:

```
SELECT YEAR_NUM, PD_DESC, COUNT(*) AS total_arrests  
FROM CRIME_FACT CF  
JOIN DIM_CRIMES DC ON CF.CRIMES_SK = DC.CRIMES_SK  
JOIN DIM_DATE DD ON CF.DATE_SK = DD.DATE_SK  
GROUP BY YEAR_NUM, PD_DESC  
ORDER BY YEAR_NUM DESC, total_arrests DESC;
```

5. Are there specific precincts with higher felony arrests compared to misdemeanors? (Hint: A precinct is a police district within a city.)

Relevant - The DIM\_ARRESTS table contains ARREST\_PRECINCT, and DIM\_CRIMES has LAW\_CATEGORY (which can indicate felony vs. misdemeanor).

QUERY:

```
SELECT DA.ARREST_PRECINCT,  
       DC.LAW_CATEGORY,  
       COUNT(*) AS total_arrests  
  FROM CRIME_FACT CF  
  JOIN DIM_ARRESTS DA ON CF.ARRESTS_SK = DA.ARRESTS_SK  
  JOIN DIM_CRIMES DC ON CF.CRIMES_SK = DC.CRIMES_SK  
 GROUP BY DA.ARREST_PRECINCT, DC.LAW_CATEGORY  
 ORDER BY DA.ARREST_PRECINCT, total_arrests DESC;
```

6. Which borough has the highest number of arrests? (Hint: A borough is a large administrative division in NYC, such as Manhattan (M), Brooklyn (K), Queens (Q), The Bronx (B), and Staten Island (S).)

Relevant - DIM\_ARRESTS contains ARREST\_BOROUGH.

QUERY:

```
SELECT ARREST_BOROUGH, COUNT(*) AS total_arrests  
  FROM CRIME_FACT CF  
  JOIN DIM_ARRESTS DA ON CF.ARRESTS_SK = DA.ARRESTS_SK  
 GROUP BY ARREST_BOROUGH  
 ORDER BY total_arrests DESC;
```

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

Relevant - The DIM\_CRIMINAL table contains AGE, SEX, and RACE.

QUERY:

```
SELECT AGE, SEX, RACE, COUNT(*) AS total_arrests  
FROM CRIME_FACT CF  
JOIN DIM_CRIMINAL DC ON CF.CRIMINAL_SK = DC.CRIMINAL_SK  
GROUP BY AGE, SEX, RACE  
ORDER BY total_arrests DESC;
```

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

Partially Relevant - DIM\_LOCATION contains LATITUDE, LONGITUDE, and GEOREFERENCE, which help identify crime hotspots. However, this dataset does not include crime severity, time of day, or socioeconomic factors necessary for a full predictive analysis.

```
SELECT DL.LATITUDE, DL.LONGITUDE, COUNT(*) AS total_arrests  
FROM CRIME_FACT CF  
JOIN DIM_LOCATION DL ON CF.LOCATION_SK = DL.LOCATION_SK  
GROUP BY DL.LATITUDE, DL.LONGITUDE  
ORDER BY total_arrests DESC  
LIMIT 10;
```

-Identifying the Grain:

Each record in the dataset represents a single arrest event, which is the grain of our data.

Dimension Tables:

- DIM\_DATE – Stores date-related information (day, month, quarter, year).
- DIM\_LOCATION – Stores geographic data (latitude, longitude, borough, precinct).

- DIM\_CRIMES – Stores crime-related details (crime type, law code, offense category).
- DIM\_ARRESTS – Stores arrest-related attributes (arrest ID, precinct, borough, jurisdiction).
- DIM\_CRIMINAL – Stores information about arrestees (age, race, gender).

#### -Data Profiling Using Alteryx:

The NYPD Arrest dataset was profiled using Alteryx Designer, revealing several data quality issues:

- Missing Values:
  - PD\_CD: 8 records are null
  - KY\_CD: 32 records are null
  - LAW\_CAT\_CD: 1390 records are null
  - LATITUDE: 4 records are null
  - LONGITUDE: 4 records are null
  - NEW\_GEOREFERENCED\_COLUMN: 4 records are null
- Data Formatting Issues:
  - Date format is MMDDYYYY, but should be converted to YYYYMMDD.
  - Data types need standardization for PD\_CD, KY\_CD, ARREST\_PRECINCT, JURISDICTION\_CODE, LATITUDE, and LONGITUDE.

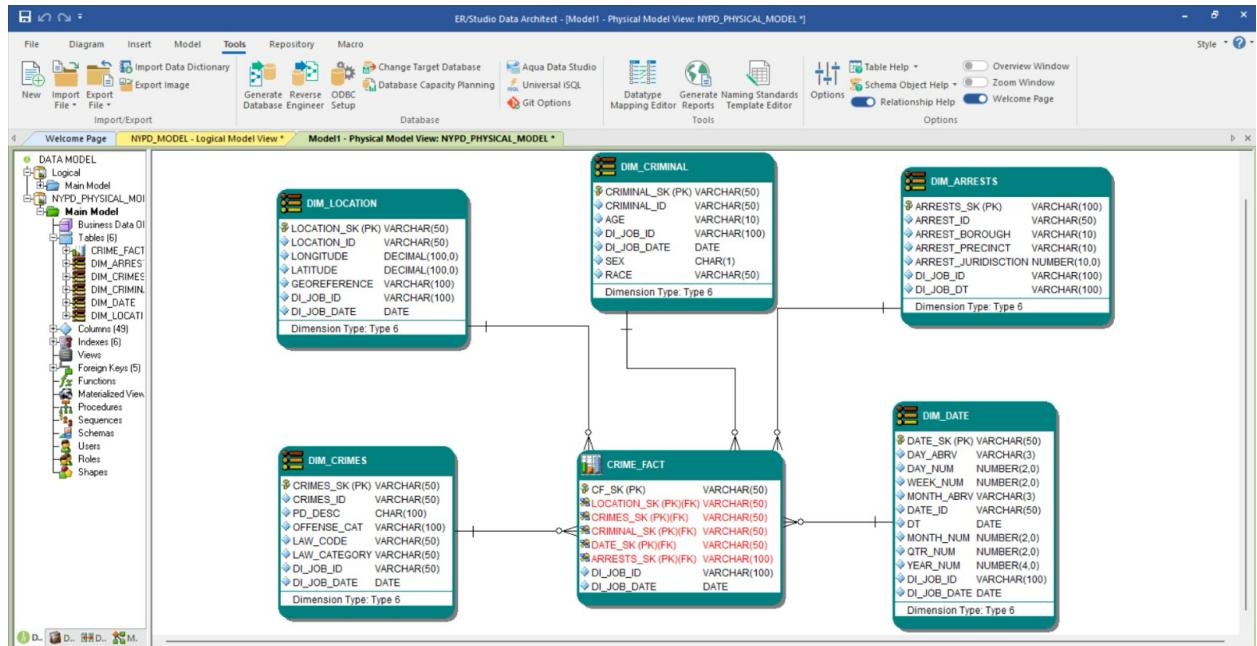
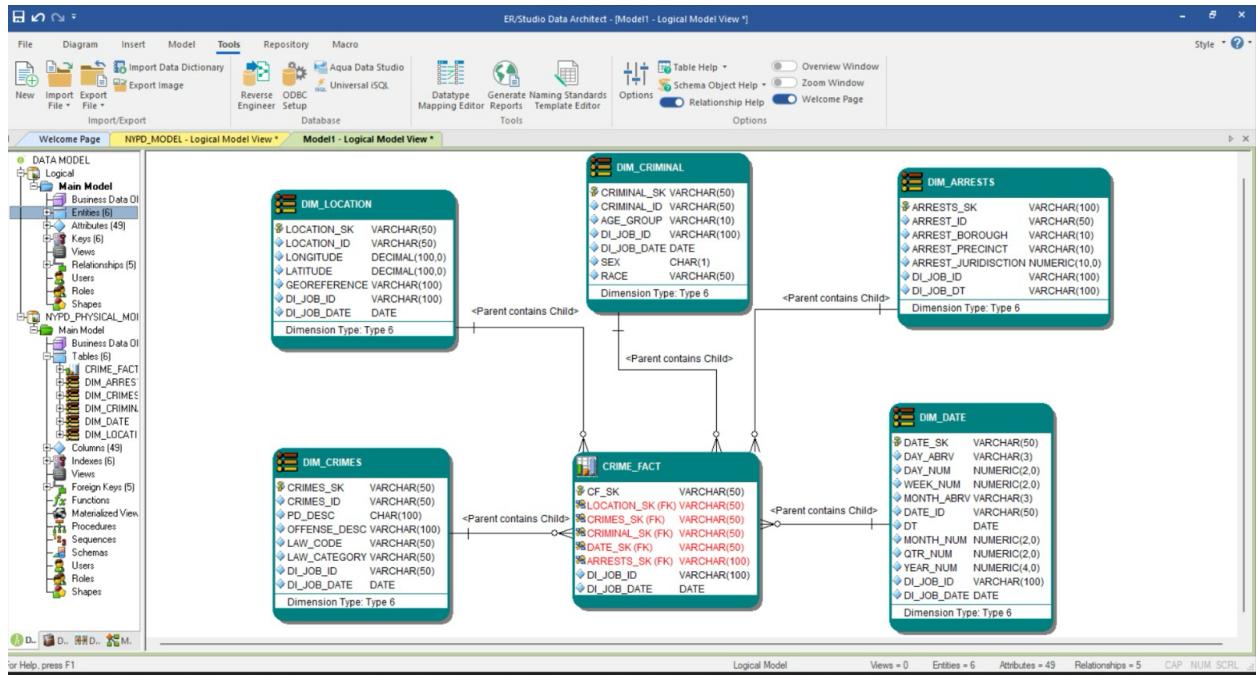
#### Data Cleansing Approach:

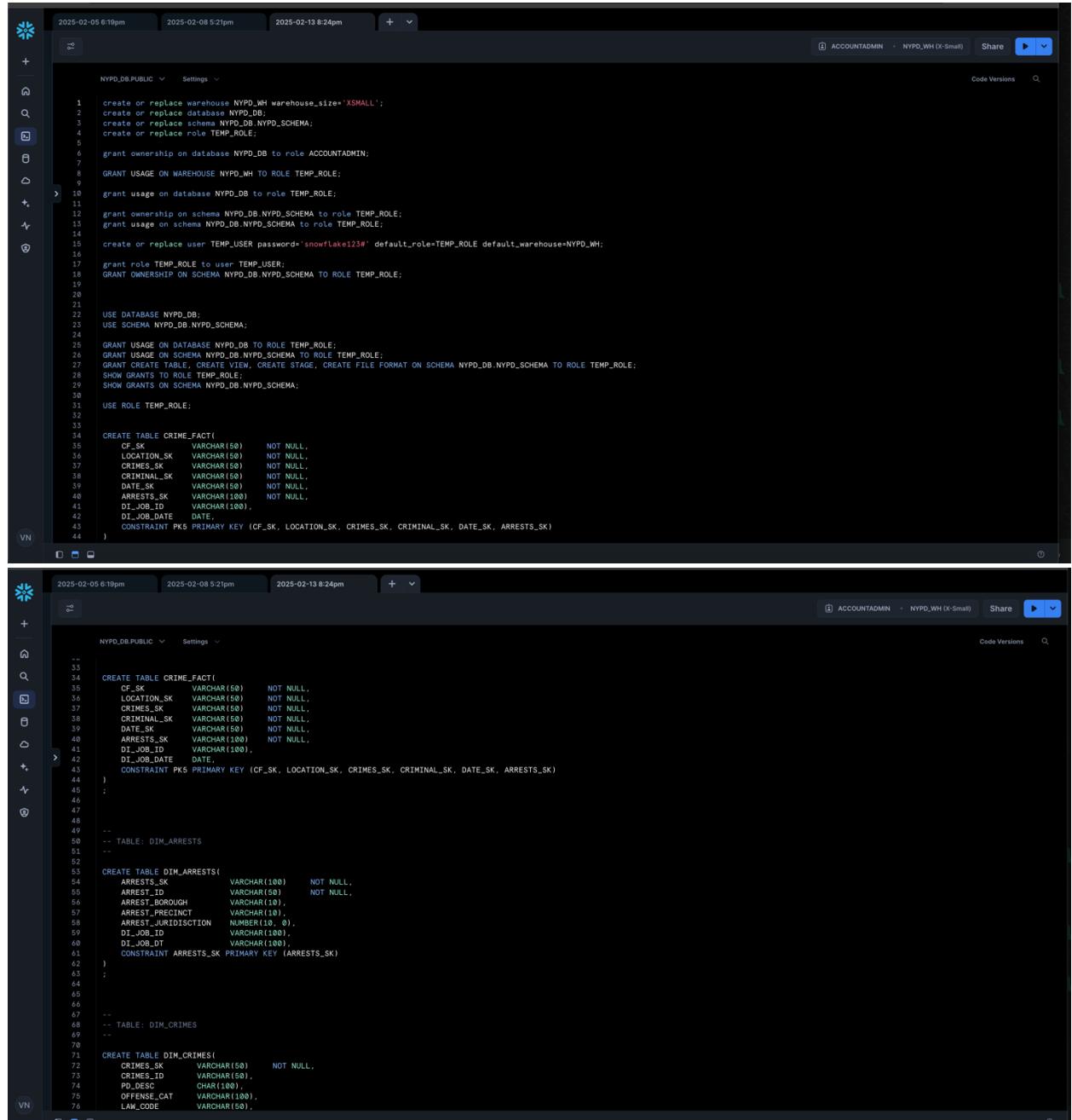
- Replace null values with appropriate placeholders or remove invalid records.
- Convert date format for consistency.
- Standardize data types to maintain integrity in Snowflake.

#### Transformations:

- NULLCLEAR: Handle missing values.
- DateFormat: Convert date format to YYYYMMDD.
- Standardization: Ensure data types match schema requirements.

## -ER/STUDIO:



**-SQL SCRIPT:**


```

NYPD_DB.PUBLIC  Settings

1  create or replace warehouse NYPD_WH warehouse_size='xSMALL';
2  create or replace database NYPD_DB;
3  create or replace schema NYPD_DB.NYPD_SCHEMA;
4  create or replace role TEMP_ROLE;
5
6  grant ownership on database NYPD_DB to role ACCOUNTADMIN;
7
8  GRANT USAGE ON WAREHOUSE NYPD_WH TO ROLE TEMP_ROLE;
9
10 grant usage on database NYPD_DB to role TEMP_ROLE;
11
12 grant ownership on schema NYPD_DB.NYPD_SCHEMA to role TEMP_ROLE;
13 grant usage on schema NYPD_DB.NYPD_SCHEMA to role TEMP_ROLE;
14
15 create or replace user TEMP_USER password='snowLake123#' default_role=TEMP_ROLE default_warehouse=NYPD_WH;
16
17 grant role TEMP_ROLE to user TEMP_USER;
18
19 GRANT OWNERSHIP ON SCHEMA NYPD_DB.NYPD_SCHEMA TO ROLE TEMP_ROLE;
20
21
22 USE DATABASE NYPD_DB;
23 USE SCHEMA NYPD_DB.NYPD_SCHEMA;
24
25 GRANT USAGE ON DATABASE NYPD_DB TO ROLE TEMP_ROLE;
26 GRANT USAGE ON SCHEMA NYPD_DB.NYPD_SCHEMA TO ROLE TEMP_ROLE;
27 GRANT CREATE TABLE, CREATE VIEW, CREATE STAGE, CREATE FILE FORMAT ON SCHEMA NYPD_DB.NYPD_SCHEMA TO ROLE TEMP_ROLE;
28 SHOW GRANTS TO ROLE TEMP_ROLE;
29 SHOW GRANTS ON SCHEMA NYPD_DB.NYPD_SCHEMA;
30
31 USE ROLE TEMP_ROLE;
32
33
34 CREATE TABLE CRIME_FACT(
35   CF_SK          VARCHAR(50)    NOT NULL,
36   LOCATION_SK   VARCHAR(50)    NOT NULL,
37   CRIMES_SK     VARCHAR(50)    NOT NULL,
38   CRIMINAL_SK   VARCHAR(50)    NOT NULL,
39   DATE_SK        VARCHAR(50)    NOT NULL,
40   ARRESTS_SK    VARCHAR(100)   NOT NULL,
41   DI_JOB_ID     VARCHAR(100),
42   DI_JOB_DATE   DATE,
43
44   CONSTRAINT PK5 PRIMARY KEY (CF_SK, LOCATION_SK, CRIMES_SK, CRIMINAL_SK, DATE_SK, ARRESTS_SK)
45 )
46
47
48
49 -- TABLE: DIM_ARRESTS
50
51
52 CREATE TABLE DIM_ARRESTS(
53   ARRESTS_SK    VARCHAR(100)   NOT NULL,
54   ARREST_ID     VARCHAR(50)    NOT NULL,
55   ARREST_BOROUGH VARCHAR(10),
56   ARREST_PRECINCT VARCHAR(10),
57   ARREST_NEIGHBORHOOD VARCHAR(10),
58   ARREST_DISTRICT NUMBER(10, 0),
59   DI_JOB_ID     VARCHAR(100),
60   DI_JOB_DT     VARCHAR(100),
61
62   CONSTRAINT ARRESTS_SK PRIMARY KEY (ARRESTS_SK)
63 )
64
65
66
67 -- TABLE: DIM_CRIMES
68
69
70 CREATE TABLE DIM_CRIMES(
71   CRIMES_SK     VARCHAR(50)    NOT NULL,
72   CRIMES_ID     VARCHAR(50),
73   PD_DESC       CHAR(100),
74   OFFENSE_CAT   VARCHAR(100),
75   LAW_CODE      VARCHAR(50),
76
77 )

```

# VARANA NAVADIYA-002308207

The image shows two side-by-side screenshots of a database editor interface, likely Oracle SQL Developer, displaying SQL code for creating dimension tables. Both tabs are titled 'NYPD\_DB.PUBLIC' and have the same timestamp: 2025-02-05 6:19pm.

**Top Tab (Left):**

```
67  --
68  -- TABLE: DIM_CRIMES
69  --
70
71  CREATE TABLE DIM_CRIMES(
72      CRIMES_SK          VARCHAR(50)      NOT NULL,
73      CRIMES_ID          VARCHAR(50),
74      PD_DESC            CHAR(100),
75      OFFENSE_CAT        VARCHAR(100),
76      LAW_CODE           VARCHAR(50),
77      LAW_CATEGORY       VARCHAR(50),
78      DI_JOB_ID          VARCHAR(50),
79      DI_JOB_DATE        DATE,
80      CONSTRAINT PK_CRIMES_SK PRIMARY KEY (CRIMES_SK)
81  )
82  ;
83
84
85
86  --
87  -- TABLE: DIM_CRIMINAL
88  --
89
90  CREATE TABLE DIM_CRIMINAL(
91      CRIMINAL_SK         VARCHAR(50)      NOT NULL,
92      CRIMINAL_ID         VARCHAR(50),
93      AGE                 VARCHAR(10),
94      DI_JOB_ID           VARCHAR(100),
95      DI_JOB_DATE         DATE,
96      SEX                 CHAR(1),
97      RACE                VARCHAR(50),
98      CONSTRAINT PK3 PRIMARY KEY (CRIMINAL_SK)
99  )
100 ;
101
102
103
104  --
105  -- TABLE: DIM_DATE
106  --
107
108  CREATE TABLE DIM_DATE(
109      DATE_SK             VARCHAR(50)      NOT NULL,
110      DAY_ABRV            VARCHAR(3),
111      DAY_NUM              NUMBER(2, 0),
112      WEEK_NUM             NUMBER(2, 0),
113      MONTH_ABRV           VARCHAR(3),
114      DATE_ID              VARCHAR(50),
115      DT                  DATE,
116      MONTH_NUM            NUMBER(2, 0),
117      QTR_NUM              NUMBER(2, 0),
118      YEAR_NUM             NUMBER(4, 0),
119      DI_JOB_ID            VARCHAR(100),
120      DI_JOB_DATE          DATE,
121      CONSTRAINT PK4 PRIMARY KEY (DATE_SK)
122  )
123 ;
124
125
126
127  --
```

**Bottom Tab (Right):**

```
84
85
86  --
87  -- TABLE: DIM_CRIMINAL
88  --
89
90  CREATE TABLE DIM_CRIMINAL(
91      CRIMINAL_SK         VARCHAR(50)      NOT NULL,
92      CRIMINAL_ID         VARCHAR(50),
93      AGE                 VARCHAR(10),
94      DI_JOB_ID           VARCHAR(100),
95      DI_JOB_DATE         DATE,
96      SEX                 CHAR(1),
97      RACE                VARCHAR(50),
98      CONSTRAINT PK3 PRIMARY KEY (CRIMINAL_SK)
99  )
100 ;
101
102
103
104  --
105  -- TABLE: DIM_DATE
106  --
107
108  CREATE TABLE DIM_DATE(
109      DATE_SK             VARCHAR(50)      NOT NULL,
110      DAY_ABRV            VARCHAR(3),
111      DAY_NUM              NUMBER(2, 0),
112      WEEK_NUM             NUMBER(2, 0),
113      MONTH_ABRV           VARCHAR(3),
114      DATE_ID              VARCHAR(50),
115      DT                  DATE,
116      MONTH_NUM            NUMBER(2, 0),
117      QTR_NUM              NUMBER(2, 0),
118      YEAR_NUM             NUMBER(4, 0),
119      DI_JOB_ID            VARCHAR(100),
120      DI_JOB_DATE          DATE,
121      CONSTRAINT PK4 PRIMARY KEY (DATE_SK)
122  )
123 ;
124
125
126
127  --
```

# VARANA NAVADIYA-002308207

The image displays two vertically stacked PostgreSQL code editors, likely from a tool like pgAdmin or similar, showing the creation of database tables and the addition of foreign key constraints.

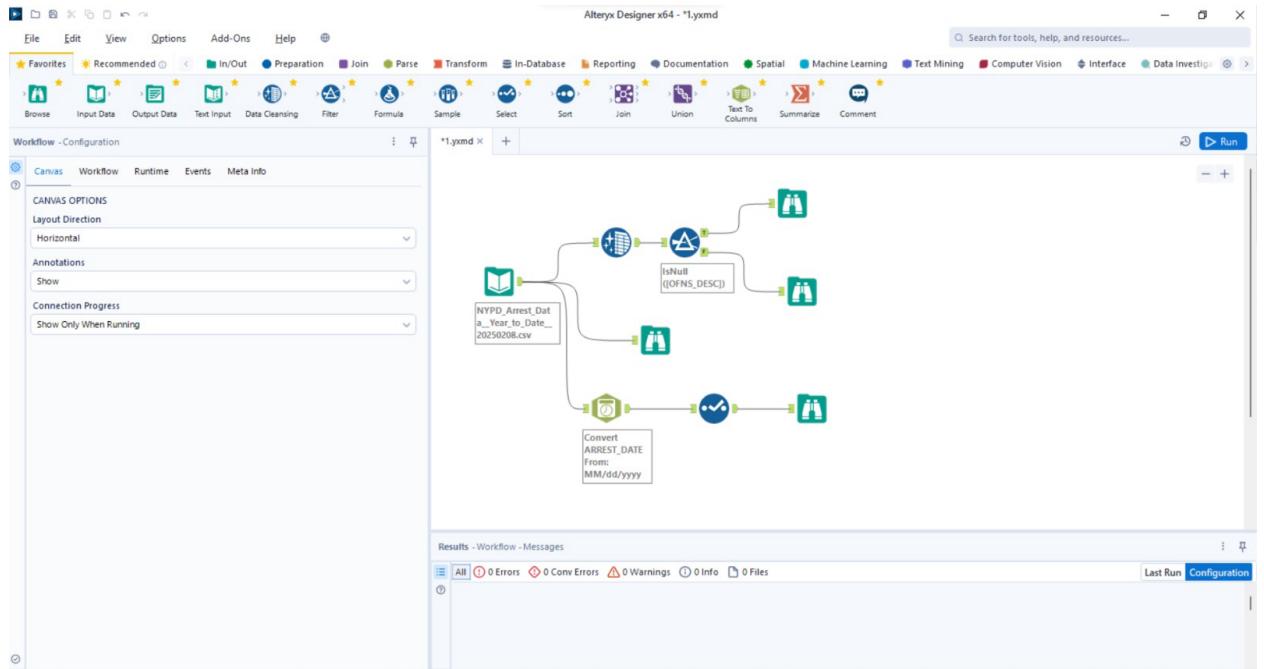
**Top Editor (Schema Creation):**

```
105 -- TABLE: DIM_DATE
106 --
107
108 CREATE TABLE DIM_DATE(
109     DATE_SK          VARCHAR(50)      NOT NULL,
110     DAY_ABRV         VARCHAR(5),
111     DAY_NUM          NUMBER(2, 0),
112     WEEK_NUM         NUMBER(2, 0),
113     MONTH_ABRV       VARCHAR(5),
114     DATE_ID          VARCHAR(50),
115     DT               DATE,
116     MONTH_NUM        NUMBER(2, 0),
117     QTR_NUM          NUMBER(2, 0),
118     YEAR_NUM         NUMBER(4, 0),
119     DI_JOB_ID        VARCHAR(100),
120     DI_JOB_DATE      DATE,
121     CONSTRAINT PK4 PRIMARY KEY (DATE_SK)
122 )
123 :
124 :
125 :
126 :
127 --
128 -- TABLE: DIM_LOCATION
129 --
130
131 CREATE TABLE DIM_LOCATION(
132     LOCATION_SK       VARCHAR(50)      NOT NULL,
133     LOCATION_ID       VARCHAR(50),
134     LONGITUDE         DECIMAL(100, 0),
135     LATITUDE          DECIMAL(100, 0),
136     GEOREFERENCE     VARCHAR(100),
137     DI_JOB_ID        VARCHAR(100),
138     DI_JOB_DATE      DATE,
139     CONSTRAINT PK2 PRIMARY KEY (LOCATION_SK)
140 )
141 :
142 :
143 :
144 :
145 --
146 -- TABLE: CRIME_FACT
147 --
148
```

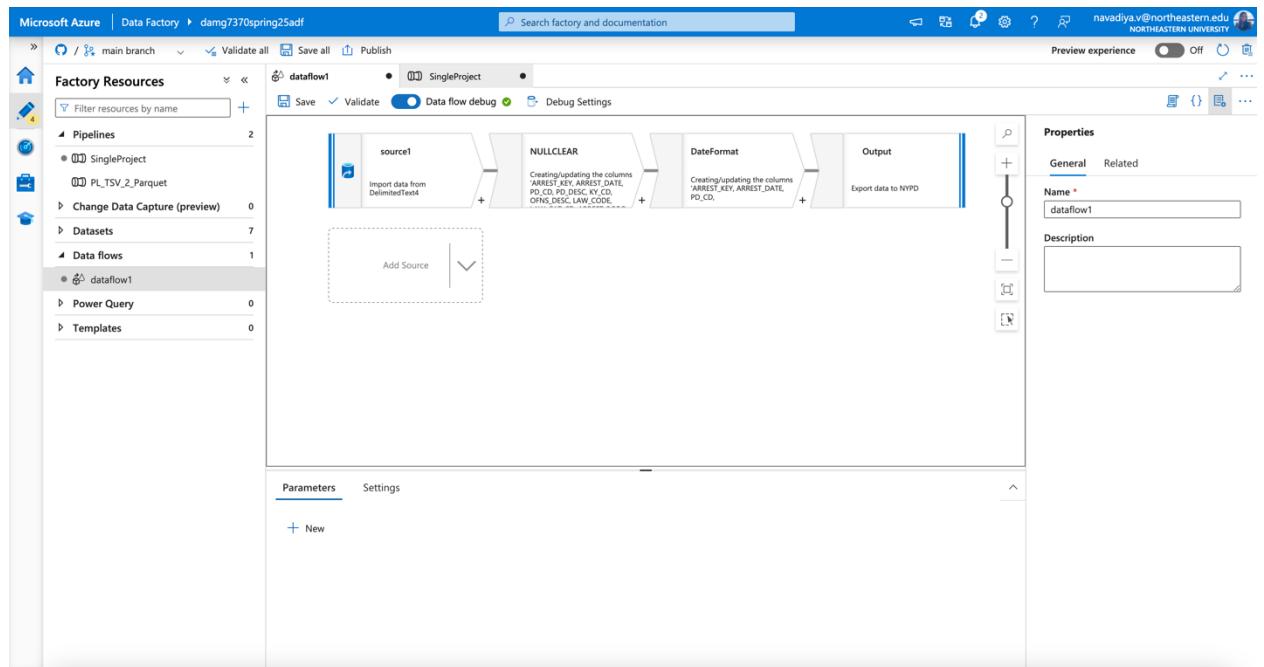
**Bottom Editor (Constraint Addition):**

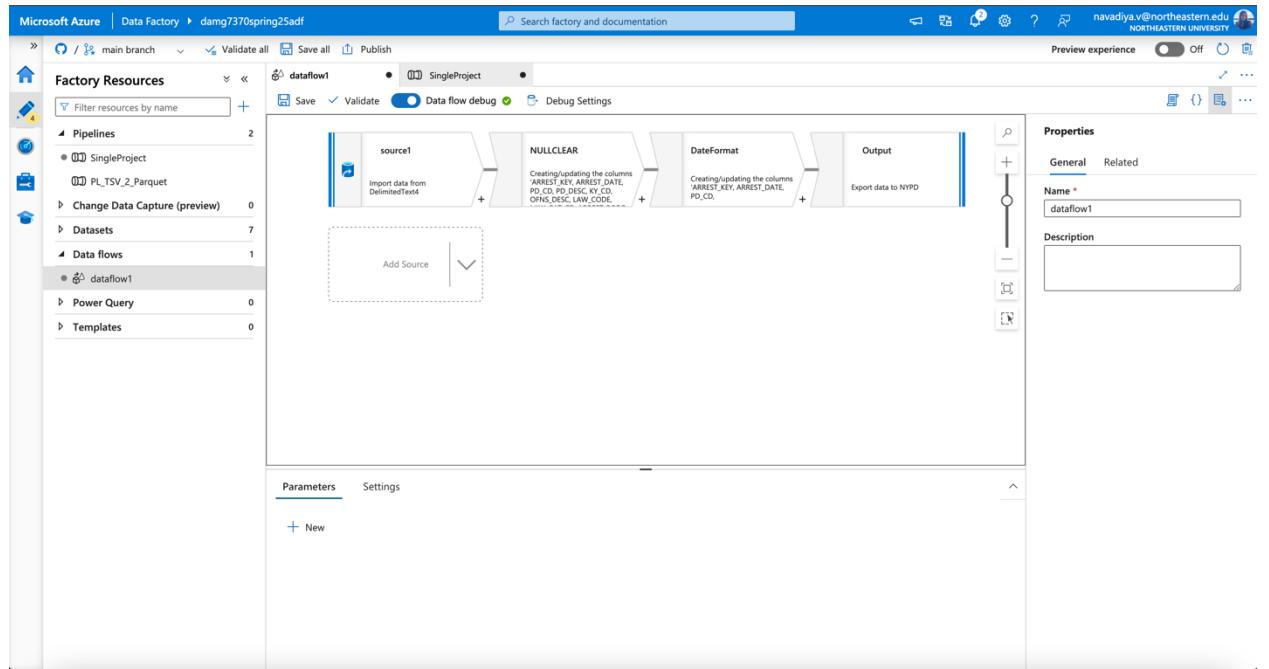
```
150
151 CREATE TABLE DIM_LOCATION(
152     LOCATION_SK       VARCHAR(50)      NOT NULL,
153     LOCATION_ID       VARCHAR(50),
154     LONGITUDE         DECIMAL(100, 0),
155     LATITUDE          DECIMAL(100, 0),
156     GEOREFERENCE     VARCHAR(100),
157     DI_JOB_ID        VARCHAR(100),
158     DI_JOB_DATE      DATE,
159     CONSTRAINT PK2 PRIMARY KEY (LOCATION_SK)
160 )
161 :
162 :
163 :
164 :
165 --
166 -- TABLE: CRIME_FACT
167 --
168
169 ALTER TABLE CRIME_FACT ADD CONSTRAINT RefDIM_LOCATION1
170     FOREIGN KEY (LOCATION_SK)
171     REFERENCES DIM_LOCATION(LOCATION_SK)
172 :
173
174 ALTER TABLE CRIME_FACT ADD CONSTRAINT RefDIM_CRIMES2
175     FOREIGN KEY (CRIMES_SK)
176     REFERENCES DIM_CRIMES(CRIMES_SK)
177 :
178
179 ALTER TABLE CRIME_FACT ADD CONSTRAINT RefDIM_CRIMINAL3
180     FOREIGN KEY (CRIMINAL_SK)
181     REFERENCES DIM_CRIMINAL(CRIMINAL_SK)
182 :
183
184 ALTER TABLE CRIME_FACT ADD CONSTRAINT RefDIM_DATE4
185     FOREIGN KEY (DATE_SK)
186     REFERENCES DIM_DATE(DATE_SK)
187 :
188
189 ALTER TABLE CRIME_FACT ADD CONSTRAINT RefDIM_ARRESTS5
190     FOREIGN KEY (ARRESTS_SK)
191     REFERENCES DIM_ARRESTS(ARRESTS_SK)
192 :
```

### -ALTERYX:



### -DATAFLOW:



**-ADF PIPELINE:****-DATA POPULATED IN SNOWFLAKE DATABASE:**

```

create or replace table NYPD_SCHEMA.NYPD_STAGE_TABLE(
    ARREST_KEY VARCHAR(100),
    ARREST_DATE VARCHAR(100),
    PD_CD VARCHAR(100),
    PD_DESC VARCHAR(100),
    KY_CD VARCHAR(100),
    OFFENSE_DESCRIPTION VARCHAR(100),
    LAW_CODE VARCHAR(100),
    LAW_CAT_CD VARCHAR(100),
    BOROUGH VARCHAR(100),
    PRECINCT VARCHAR(100),
    JURISDICTION_CODE VARCHAR(100),
    AGE_GROUP VARCHAR(100),
    SEX VARCHAR(100),
    RACE VARCHAR(100),
    X_COORD VARCHAR(100),
    Y_COORD VARCHAR(100),
    LATITUDE VARCHAR(100),
    LONGITUDE VARCHAR(100),
    NEW_GEOREFERENCED_ID VARCHAR(100),
    dt_job_id STRING(100),
    dt_job_date DATE
);

```

The screenshot shows the DBeaver 24.3.3 interface. The top menu bar includes File, Edit, Navigate, Search, SQL Editor, Database, Window, and Help. The database navigation pane shows connections to 'NYPD.ARREST' and 'NYPD\_DB'. The current connection is 'NYPD\_DB'. The central area displays a SQL script for creating a table 'NYPD\_STAGE\_TABLE' with various columns. The bottom status bar shows the query count as 23:3:671 and the number of rows as 0. The system tray at the bottom indicates the user is connected to 'ENG' and has a battery level of 'IN'.

# VARANA NAVADIYA-002308207

DBeaver 24.3.3 - <NYPD\_DB> Script-6

File Edit Navigate Search SQL Editor Database Window Help

Database Navigator Project

\*<NYPD\_ARREST> Script-4 <NYPD\_DB> Script-5 <NYPD\_DB> Script-6

Enter a part of object name here

> NYPD\_ARREST tc74912.canada-central.azure.snowflakecomputing.com

> NYPD\_DB tc74912.canada-central.azure.snowflakecomputing.com:443

> NYPD\_DB

- INFORMATION\_SCHEMA
- NYPD\_SCHEMA
- SNOWFLAKE
- SNOWFLAKE\_SAMPLE\_DATA
- TEMP\_DB tc74912.canada-central.azure.snowflakecomputing.com:443

```

SEA VARCHAR(100),
RACE VARCHAR(100),
X_COORD VARCHAR(100),
Y_COORD VARCHAR(100),
LATITUDE VARCHAR(100),
LONGITUDE VARCHAR(100),
NEW_GEOREFERENCED_ID VARCHAR(100),
dt_job_id STRING(100),
dt_job_date DATE
);

SELECT * FROM NYPD_SCHEMA.NYPD_STAGE_TABLE;

```

NYPD\_STAGE\_TABLE 1 ×

Enter a SQL expression to filter results (use Ctrl+Space)

ARREST_KEY	ARREST_DATE	PD_CD	PD_DESC	KY_CD	OFFENSE_DESCRIPTION	LAW_CODE
295711043	2024-10-20	101	ASSAULT 3	344	ASSAULT 3 & RELATED OFFENSES	PL 1200001 M
295538544	2024-10-27	203	TRESPASS 3, CRIMINAL	352	CRIMINAL TRESPASS	PL 140100A M
294216589	2024-10-02	175	SEXUAL ABUSE 3.2	233	SEX CRIMES	PL 1305201 M
294326407	2024-10-04	511	CONTROLLED SUBSTANCE, POSSESS	235	DANGEROUS DRUGS	PL 2200300 M
293545116	2024-09-19	739	FRAUD,UNCLASSIFIED-FELONY	112	THEFT-FRAUD	PL 1908303 F
294681827	2024-10-10	439	LARCENY,GRAND FROM OPEN AREAS, UNATTENDED	109	GRAND LARCENY	PL 1533001 F
294879404	2024-10-15	109	ASSAULT 2.1,UNCLASSIFIED	106	FELONY ASSAULT	PL 1200502 F
295756651	2024-10-31	779	PUBLIC ADMINISTRATION,UNCLAS	126	MISCELLANEOUS PENAL LAW	PL 2155108 F
294057612	2024-09-29	969	TRAFFIC,UNCLASSIFIED INFRACTIO	881	OTHER TRAFFIC INFRACTION	VTL05101A M
293808940	2024-09-24	782	WEAPONS, POSSESSION, ETC	236	DANGEROUS WEAPONS	PL 2650102 M
293653580	2024-09-20	916	LEAVING SCENE-ACCIDENT-PERSONA	348	VEHICLE AND TRAFFIC LAWS	VTL06000A M
295066415	2024-10-18	223	BURGLARY,RESIDENCE,NIGHT	107	BURGLARY	PL 1402502 F
295873027	2024-11-03	101	ASSAULT 3	344	ASSAULT 3 & RELATED OFFENSES	PL 1200001 M
295712934	2024-10-30	339	LARCENY,PETIT FROM OPEN AREAS,	341	PETIT LARCENY	PL 1552500 M
294742532	2024-10-11	244	BURGLARY,UNCLASSIFIED,UNKNOWN	107	BURGLARY	PL 1402501 F
295567071	2024-10-28	101	ASSAULT 3	344	ASSAULT 3 & RELATED OFFENSES	PL 1200001 M
294622207	2024-10-09	705	FORGERY,ETC-MISD.	358	OFFENSES INVOLVING FRAUD	PL 1702000 M
295469405	2024-10-25	268	CRIMINAL MIS 2 & 3	121	CRIMINAL MISCHIEF & RELATED O	PL 1450902 F
294622199	2024-10-09	792	CRIMINAL POSSESSION WEAPON	118	DANGEROUS WEAPONS	PL 2650318 F

Project - General

Record

Refresh Save Cancel Export data 200 200+ 200 row(s) fetched - 0.397s (0.002s fetch), on 2025-02-13 at 21:59:13

EST en Writable Smart Insert 25 : 43 Set: 43 | 1

## -ROW COUNTS USING SQL QUERY:

DBeaver 24.3.3 - <NYPD\_DB> Script-6

File Edit Navigate Search SQL Editor Database Window Help

Database Navigator Project

\*<NYPD\_ARREST> Script-4 <NYPD\_DB> Script-5 <NYPD\_DB> Script-6

Enter a part of object name here

> NYPD\_ARREST tc74912.canada-central.azure.snowflakecomputing.com

> NYPD\_DB tc74912.canada-central.azure.snowflakecomputing.com:443

> NYPD\_DB

- INFORMATION\_SCHEMA
- NYPD\_SCHEMA
- SNOWFLAKE
- SNOWFLAKE\_SAMPLE\_DATA
- TEMP\_DB tc74912.canada-central.azure.snowflakecomputing.com:443

```

ARREST_DATE VARCHAR(100),
PD_CD VARCHAR(100),
PD_DESC VARCHAR(100),
KY_CD VARCHAR(100),
OFFENSE_DESCRIPTION VARCHAR(100),
LAW_CODE VARCHAR(100),
LAW_CAT_CD VARCHAR(100),
BOROUGH VARCHAR(100),
PRECINCT VARCHAR(100),
JURISDICTION_CODE VARCHAR(100),
AGE_GROUP VARCHAR(100),
SEX VARCHAR(100),
RACE VARCHAR(100),
X_COORD VARCHAR(100),
Y_COORD VARCHAR(100),
LATITUDE VARCHAR(100),
LONGITUDE VARCHAR(100),
NEW_GEOREFERENCED_ID VARCHAR(100),
dt_job_id STRING(100),
dt_job_date DATE
);

SELECT * FROM NYPD_SCHEMA.NYPD_STAGE_TABLE;
SELECT COUNT(*) FROM NYPD_SCHEMA.NYPD_STAGE_TABLE;

```

Results 1 ×

Enter a SQL expression to filter results (use Ctrl+Space)

COUNT(*)
260503

Project - General

Record

Refresh Save Cancel Export data 200 200+ 200 row(s) fetched - 0.397s (0.002s fetch), on 2025-02-13 at 21:59:13

EST en Writable Smart Insert 25 : 43 Set: 43 | 1

VARANA NAVADIYA-002308207

-GITHUB:

The image shows two screenshots of a GitHub repository page for 'DAMG7370SPRING25'.

**Top Screenshot (Repository Overview):**

- Repository Name:** DAMG7370SPRING25 (Public)
- Branches:** main (selected), 2 Branches
- Tags:** 0 Tags
- Commits:** 26 Commits (by varana2001)
- Activity:** Last commit was 1 minute ago.
- Statistics:** 0 stars, 0 forks, 1 watching.
- Releases:** No releases published. Create a new release.
- Packages:** No packages published. Publish your first package.

**Bottom Screenshot (Commit History):**

- Repository Name:** DAMG7370SPRING25
- Branch:** main
- Commits:**
  - Commits on Feb 13, 2025:
    - Adding dataflow: dataflow1 (5b09750) - committed 1 minute ago.
    - Adding linkedService: NYPD\_OUPUT (6bad85a) - committed 1 hour ago.
  - Commits on Feb 6, 2025:
    - Updating pipeline: PL\_TSV\_2\_Parquet (da7e005) - committed last week.
    - Updating DelimitedText: DelimitedText2 (bf8644c) - committed last week.
    - Updating pipeline: PL\_TSV\_2\_Parquet (6655184) - committed last week.
    - Updating pipeline: PL\_TSV\_2\_Parquet (bb4e4a6) - committed last week.
    - Updating linkedService: AzureBlobStorage1 (cb5e81c) - committed last week.
    - Updating pipeline: PL\_TSV\_2\_Parquet (911ddfd) - committed last week.
    - Updating pipeline: PL\_TSV\_2\_Parquet (7bedfe0) - committed last week.

VARANA NAVADIYA-002308207

GitHub Link: <https://github.com/varana2001/DAMG7370SPRING25>

DataFlow Link: <https://github.com/varana2001/DAMG7370SPRING25/tree/main/dataflow>

Pipeline Link: <https://github.com/varana2001/DAMG7370SPRING25/tree/main/pipeline>