

This data warehouse aims to address a pressing issue related to building use complaints and Department of Buildings (DOB) permit issuance in the city. The problem at hand revolves around the efficient handling of citizen complaints regarding building use violations and the subsequent permit issuance process by the Department of Buildings. Building use complaints, typically reported through the 311 system, are multifaceted and can range from concerns about unauthorized construction activities, structural safety issues, zoning violations, noise disturbances, to improper occupancy. These complaints serve as crucial indicators of potential violations and safety hazards within the community, reflecting the concerns and well-being of the city's residents.

On the other hand, DOB permit issuance data from NYC Open Data is essential for understanding the legal aspects of construction and renovation activities within the city. It includes information about permits issued for various building projects, such as new constructions, renovations, repairs, and demolitions. Integrating this dataset with the 311 building use complaints data provides a comprehensive perspective. By doing so, the group aims to establish meaningful correlations between citizen complaints and the subsequent regulatory response, evaluating the effectiveness and efficiency of the DOB's permit issuance process in addressing reported issues. This integrated dataset can unveil patterns and trends, enabling authorities to proactively identify problematic areas, streamline the permit issuance process, and enhance overall urban governance. Addressing this issue will not only improve the city's regulatory framework but also contribute significantly to the safety and quality of life for its residents, ensuring that building use complaints are addressed promptly and efficiently.

Data Sources:

<https://data.cityofnewyork.us/Social-Services/311-Service-Requests-from-2010-to-Present/erm2-nwe9>

<https://data.cityofnewyork.us/Housing-Development/DOB-Permit-Issuance/ipu4-2q9a>

[29033] (311 Building /Use)

Dimensions:

Date

Location

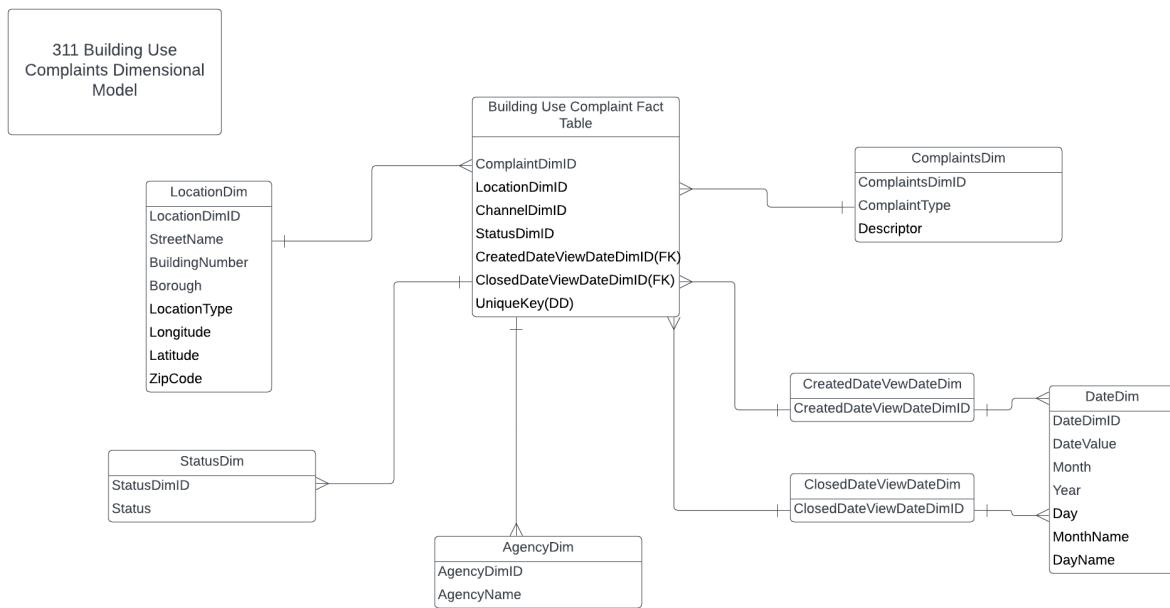
Complaint Type

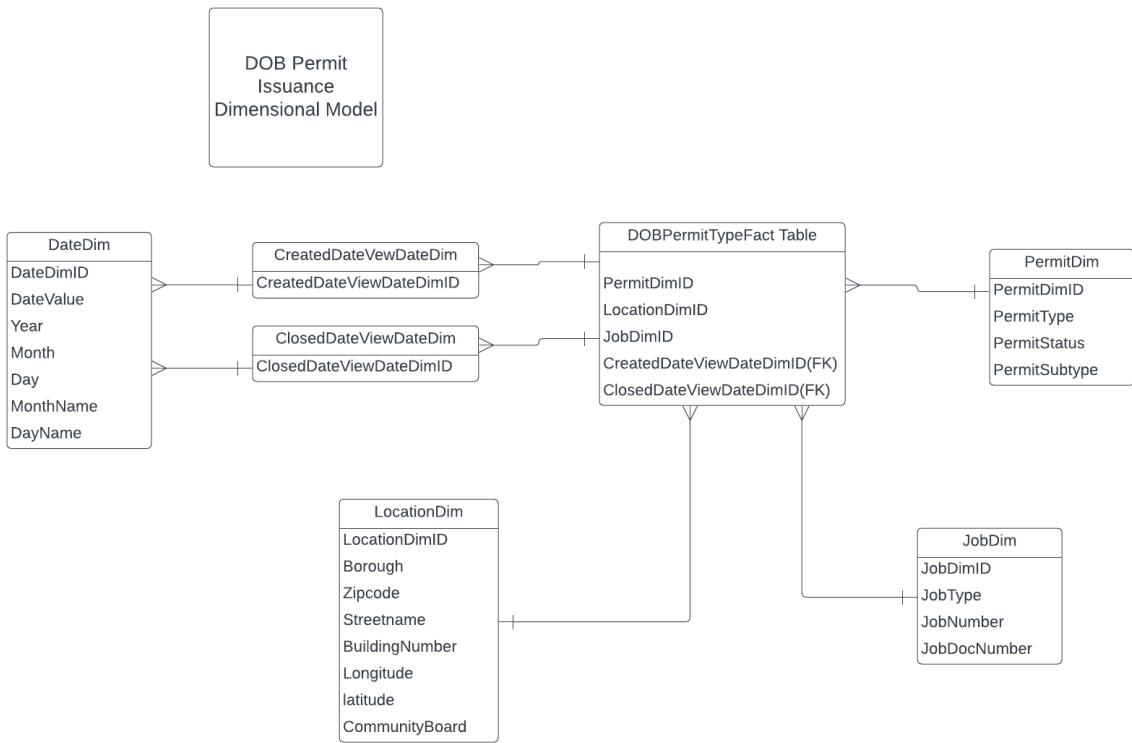
Permit Type

AgencyType

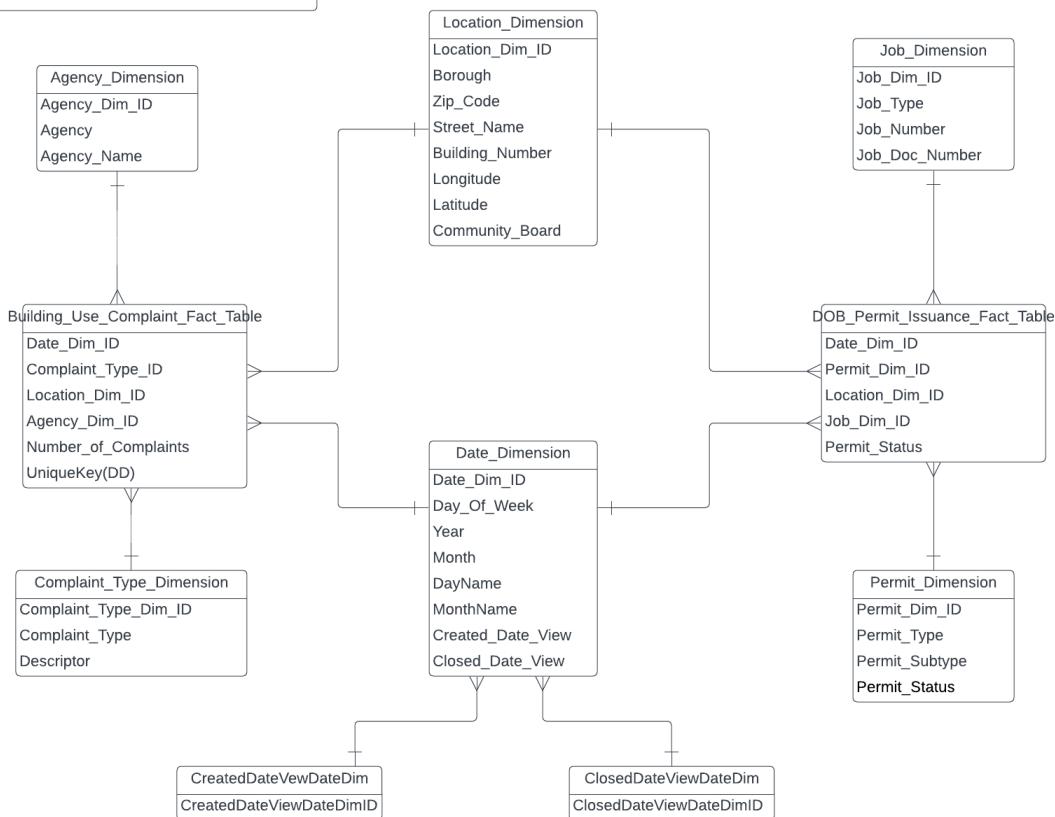
Job Type

3.





311 Building Use / DOB Permit Issuance
Integrated Dimensional Model



4. DATA MODELING CODE

BUILDING FACT TABLE

```

WITH buildingtable AS (
    SELECT unique_key, agency, agency_name, Descriptor,
    status, Complaint_Type, Created_Date, Longitude, Latitude,
    Street_Name,
    COALESCE(City, 'NA') AS City,
    COALESCE(CAST(Zip_Codes AS STRING), 'NA') AS Zip_Codes,
    COALESCE(Borough, 'NA') AS Borough
    FROM `cis-4400-group-
project-403921.permit2.building_use` 
    WHERE FORMAT_DATE("%Y", Created_Date) = '2023'
),
    created_date_dimension AS (

```

```

        SELECT *
        FROM {{ ref('created_dateView') } }
) , 

closed_date_dimension AS (
    SELECT *
    FROM {{ ref('closed_dateView') } }
) , 

LocationDim AS (
    SELECT *
    FROM {{ ref('LocationDim') } }
) , 

agencydim AS (
    SELECT *
    FROM {{ ref('agencyDim') } }
) , 

statusdim AS (
    SELECT *
    FROM {{ ref('status') } }
) , 

complaintdim AS (
    SELECT *
    FROM {{ ref('Complaint') } }
)

SELECT
    dcl.ClosedDateViewDim,
    dcr.CreatedDateViewDim,
    l.Location_DIM_id,
    a.agency_DIM_id,
    s.Status_DIM_id,
    c.Complaints_DIM_id,
    b.unique_key
FROM
    buildingtable AS b
LEFT JOIN
    closed_date_dimension AS dcl ON dcl.full_date =
b.Created_Date
LEFT JOIN

```

```

    created_date_dimension AS dcr ON dcr.full_date =
b.Created_Date
LEFT JOIN
    LocationDim AS l ON (l.Latitude=b.Latitude AND
l.Longitude=b.Longitude) OR (l.Street_Name = b.Street_Name
AND l.Borough = b.Borough AND l.Zip_Codes = b.Zip_Codes AND
b.Latitude is NULL AND b.Longitude is NULL)
LEFT JOIN
    agencydim AS a ON a.agency = b.agency AND a.agency_name
= b.agency_name
LEFT JOIN
    statusdim AS s ON s.Status = b.Status
LEFT JOIN
    complaintdim AS c ON c.Complaint_Type =
b.Complaint_Type AND c.Descriptor = b.Descriptor

```

ClosedDateVie...	CreatedDateVie...	Location_DIM_id	agency_DIM_id	Status_DIM_id	Complaints_DI...	unique
41582	41582	30057	1	1	4	582310
1795	1795	1256	1	2	1	566489
37525	37525	27002	1	2	1	570444
12619	12619	9060	1	2	1	573504
40333	40333	29067	1	2	1	589522
4872	4872	3366	1	3	2	592083
38378	38378	27634	1	2	1	581861
51847	51847	38036	1	2	1	567989

- This our 311 Building / Use fact table consisting of our dimensions.

DIMENSIONS

AGENCY

```

WITH agencydim AS (SELECT
row_number() OVER () AS agency_DIM_id, *
FROM( SELECT DISTINCT agency, agency_name

```

```

FROM `cis-4400-group-
project-403921.permit2.building_use` 
))

SELECT *
FROM agencydim

```

agency_DIM_id	agency	agency_name
1	DOB	Department of Buildings

- This is our agency dimension which includes agency name and an abbreviation of the agency

LOCATION

```

WITH locationdim AS (
  SELECT
    ROW_NUMBER() OVER () AS Location_DIM_id,
    *
  FROM (
    SELECT DISTINCT
      Street_Name,
      Latitude,
      Longitude,
      COALESCE(City, 'NA') AS City,
      COALESCE(STRING(CAST(Zip_Codes AS STRING)), 'NA') AS Zip_Codes,
      COALESCE(Borough, 'NA') AS Borough
    FROM `cis-4400-group-
project-403921.permit2.building_use` 
  )
)

SELECT * FROM locationdim

```

Location_DIM_id	Street_Name	Latitude	Longitude	City	Zip_Codes	Borough
1	9 ROAD	40.78877837	-73.84725081	COLLEGE POINT	14191	QUEENS
2	9 ROAD	40.7887349	-73.84758675	COLLEGE POINT	14191	QUEENS
3	BOWERY	40.72319488	-73.99293972	NEW YORK	11723	MANHATTAN
4	BOWERY	40.71771097	-73.99519132	NEW YORK	11723	MANHATTAN
5	BOWERY	40.71765882	-73.99521658	NEW YORK	11723	MANHATTAN
6	BOWERY	40.71688481	-73.99565312	NEW YORK	11723	MANHATTAN
7	BOWERY	40.72650772	-73.99174155	NEW YORK	11724	MANHATTAN
8	BOWERY	40.72352699	-73.99284227	NEW YORK	12075	MANHATTAN

- This Location Dimension consists for street name, latitude, longitude, city, zip code and Borough

COMPLAINTS

```
WITH complaintdim AS (
    SELECT
        row_number() OVER () AS Complaints_DIM_id, *
    FROM (
        SELECT DISTINCT Descriptor,
        Complaint_Type
        FROM `cis-4400-group-
project-403921.permit2.building_use` )
    )

    SELECT *
    FROM complaintdim
```

Complaints_DIM_id	Descriptor	Complaint_Type
1	Illegal Conversion Of Residential Building/Space	Building/Use
2	No Certificate Of Occupancy/Illegal/Contrary To CO	Building/Use
3	Zoning - Non-Conforming/Illegal Vehicle Storage	Building/Use
4	Illegal Commercial Use In Resident Zone	Building/Use
5	SRO - Illegal Work/No Permit/Change In Occupancy/Use	Building/Use

- This complaints dimension consists of complaint type and the descriptions of the complaints.

DATE

```

WITH DateDim AS (
    SELECT
        ROW_NUMBER() OVER() as date_dim_id,
        FORMAT_DATE("%Y%m%d", d) as date_integer,
        d AS full_date,
        EXTRACT(YEAR FROM d) AS year,
        EXTRACT(WEEK FROM d) AS year_week,
        EXTRACT(DAY FROM d) AS year_day,
        EXTRACT(MONTH FROM d) AS month,
        FORMAT_DATE('%B', d) as month_name,
        FORMAT_DATE('%w', d) AS week_day,
        FORMAT_DATE('%A', d) AS day_name
    FROM (
        SELECT DISTINCT Created_Date AS d
        FROM `cis-4400-group-
project-403921.permit2.building_use` )
)

SELECT *
FROM DateDim

```

date_dim_id	date_integer	full_date	year	year_week	year_day	month
1	20230921	2023-09-21T22:00:00.000Z	2023	38	21	9
2	20230829	2023-08-29T19:00:00.000Z	2023	35	29	8
3	20220716	2022-07-16T13:00:00.000Z	2022	28	16	7
4	20230522	2023-05-22T14:00:00.000Z	2023	21	22	5
5	20210719	2021-07-19T19:00:00.000Z	2021	29	19	7
6	20210719	2021-07-19T18:00:00.000Z	2021	29	19	7
7	20210719	2021-07-19T18:00:00.000Z	2021	29	19	7
8	20200224	2020-02-24T16:00:00.000Z	2020	8	24	2

- The date dimension consists of the full date, year, day and month

STATUS

```
WITH StatusDim AS (
    SELECT
        ROW_NUMBER() OVER () AS Status_DIM_id,
        Status
    FROM (
        SELECT DISTINCT Status
        FROM `cis-4400-group-
project-403921.permit2.building_use` )
)
```

```
SELECT * FROM StatusDim
```

Status_DIM_id	Status
1	Assigned
2	Closed
3	Open

- The Status dimension consists of the status of the complaints.

CREATED DATE VIEW

```
WITH date AS (
    SELECT *
    FROM {{ ref('DateDim') }} )
SELECT date_dim_id as CreatedDateViewDim, full_date,
date_integer, year, year_week, year_day, month, month_name, week_
day, day_name
FROM date
```

CreatedDateVie...	full_date	date_integer	year	year_week	year_day	month
1	2023-09-21T22:... >	20230921	2023	38	21	9
2	2023-08-29T19:... >	20230829	2023	35	29	8
3	2022-07-16T13:... >	20220716	2022	28	16	7
4	2023-05-22T14:... >	20230522	2023	21	22	5
5	2021-07-19T19:... >	20210719	2021	29	19	7
6	2021-07-19T18:... >	20210719	2021	29	19	7
7	2021-07-19T18:... >	20210719	2021	29	19	7
8	2020-02-24T16:... >	20200224	2020	8	24	2

- This is the created date view dimension, consisting day, month, year and full date.

CLOSED DATE VIEW

```
WITH date AS (
    SELECT *
    FROM {{ ref('DateDim') }} )
```

```
)
SELECT date_dim_id as ClosedDateViewDim, full_date,
date_integer, year, year_week, year_day, month, month_name, week
_day, day_name
FROM date
```

ClosedDateVie...	full_date	date_integer	year	year_week	year_day	month
1	2023-09-21T22:... >	20230921	2023	38	21	9
2	2023-08-29T19:... >	20230829	2023	35	29	8
3	2022-07-16T13:... >	20220716	2022	28	16	7
4	2023-05-22T14:... >	20230522	2023	21	22	5
5	2021-07-19T19:... >	20210719	2021	29	19	7
6	2021-07-19T18:... >	20210719	2021	29	19	7
7	2021-07-19T18:... >	20210719	2021	29	19	7
8	2020-02-24T16:... >	20200224	2020	8	24	2

- This is the Closed Date View dimension, which pertains to the date in which the complaint was closed

DOB Permit Issuance

Permit Fact Table

```
WITH permittable AS (
    SELECT
        LONGITUDE, LATITUDE, House__, Permit_Type, Permit_Sequence__, Job__,
        Issuance_Date,
        COALESCE(Permit_Subtype, 'NA') AS Permit_Subtype,
        COALESCE(Permit_Status, 'NA') AS Permit_Status,
        COALESCE(Permittee_s_Business_Name, 'NA') AS Permittee_s_Business_Name,
        COALESCE(Permittee_s_License_Type, 'NA') AS Permittee_s_License_Type,
        COALESCE(Permittee_s_License__, 'NA') AS Permittee_s_License__,
        COALESCE(Borough, 'NA') AS Borough,
        COALESCE(CAST(Zip_Code AS STRING), 'NA') AS Zip_Code,
```

```

        COALESCE(CAST(COUNCIL_DISTRICT AS STRING), 'NA') AS
COUNCIL_DISTRICT
    FROM `cis-4400-group-project-403921.permit2.permit
issue`
    WHERE /*FORMAT_DATE("%Y", Issuance_Date) = '2023'
AND*/ Issuance_Date is not NULL
    OR (LATITUDE is not null and LONGITUDE is not null)
),

created_date_dimension AS (
    SELECT *
    FROM {{ ref('created_date_view') }})
),

closed_date_dimension AS (
    SELECT *
    FROM {{ ref('closed_date_view') }})
),

LocationDim AS (
    SELECT *
    FROM {{ ref('Location') }})
),

permitDim AS (
    SELECT *
    FROM {{ ref('permit') }})
),

JobDim AS (
    SELECT *
    FROM {{ ref('job') }})
)

SELECT
    pd.Permit_DIM_id,
    l.Location_DIM_id,
    j.JobDimID,
    dcl.ClosedDateViewDim,
    dcr.CreatedDateViewDim
FROM
    permittable AS p
LEFT JOIN

```

```

    closed_date_dimension AS dcl ON dcl.full_date =
p.Issuance_Date
LEFT JOIN
    created_date_dimension AS dcr ON dcr.full_date =
p.Issuance_Date
LEFT JOIN
    LocationDim AS l ON (l.LATITUDE=p.LATITUDE AND
l.LONGITUDE=p.LONGITUDE) /*OR (l.Borough=p.Borough AND
l.Zip_Code=p.Zip_Code AND p.LATITUDE is NULL AND
p.LONGITUDE is NULL) */
LEFT JOIN
    permitDim AS pd ON pd.Permit_Type = p.Permit_Type
    AND pd.Permit_Subtype=p.Permit_Subtype
    AND pd.Permittee_s_License__ = p.Permittee_s_License__
    AND pd.Permittee_s_License_Type =
p.Permittee_s_License_Type
    AND pd.Permit_Sequence__=p.Permit_Sequence__
    AND pd.Permit_Status = p.Permit_Status
    AND pd.Permittee_s_Business_Name =
p.Permittee_s_Business_Name
LEFT JOIN
    JobDim AS j ON j.Job__ = p.Job__
WHERE l.Location_DIM_id is not NULL

```

Permit_DIM_id	Location_DIM_id	JobDimID	ClosedDateViewDim	CreatedDateViewDim
377276	225975	892768	6118	6118
737995	225975	892768	6118	6118
381871	283	339	648	648
818487	283	339	648	648
1033650	283	339	648	648
797743	283	339	648	648
144135	101069	462	1818	1818
961492	363079	607313	1404	1404
322351	423	912524	211	211

- This is our Permit fact table consisting for location_Dim_id, JobDimID, CloseddateviewDim, CreatedDateViewDim and the surrogate key Permit_DIM_id.

Date

```
WITH DateDim AS (
    SELECT
        ROW_NUMBER() OVER() as date_dim_id,
        FORMAT_DATE("%Y%m%d", d) as date_integer,
        d AS full_date,
        EXTRACT(YEAR FROM d) AS year,
        EXTRACT(WEEK FROM d) AS year_week,
        EXTRACT(DAY FROM d) AS year_day,
        EXTRACT(MONTH FROM d) AS month,
        FORMAT_DATE('%B', d) as month_name,
        FORMAT_DATE('%w', d) AS week_day,
        FORMAT_DATE('%A', d) AS day_name
    FROM (
        SELECT DISTINCT Issuance_Date AS d
        FROM `cis-4400-group-project-403921.permit2.permit
issue`
    )
)

SELECT *
FROM DateDim
```

date_dim_id	date_integer	full_date	year	year_week	year_day	month
1	19950915	1995-09-15	1995	37	15	9
2	19970129	1997-01-29	1997	4	29	1
3	19900612	1990-06-12	1990	23	12	6
4	19961127	1996-11-27	1996	47	27	11
5	19960206	1996-02-06	1996	5	6	2
6	19950111	1995-01-11	1995	2	11	1
7	19911202	1991-12-02	1991	48	2	12
8	19910215	1991-02-15	1991	6	15	2

- This is our date dimension that consists of years, week, month and day.

Location

```
With LocationDim AS (
SELECT
    row_number() OVER () AS Location_DIM_id, *
FROM (
    SELECT DISTINCT
        LONGITUDE,
        LATITUDE,
        House__,
        COALESCE(Borough, 'NA') AS Borough,
        COALESCE(CAST(Zip_Code AS STRING), 'NA') AS Zip_Code,
        COALESCE(CAST(COUNCIL_DISTRICT AS STRING), 'NA') AS
COUNCIL_DISTRICT
    FROM `cis-4400-group-project-403921.permit2.permit
issue`)
)
SELECT *
FROM LocationDim
```

Location_DIM_id	LONGITUDE	LATITUDE	House__	Borough	Zip_Code	COUNCIL_
149	-73.852425	40.741727	110-13	QUEENS	11368	21
150	-73.859082	40.745858	104-50	QUEENS	11368	21
151	-73.860047	40.739741	54-09	QUEENS	11368	21
152	-73.865487	40.740716	51-07	QUEENS	11368	21
153	-73.863943	40.742537	98-03	QUEENS	11368	21
154	-73.854313	40.749714	111-01	QUEENS	11368	21
155	-73.851559	40.740372	56-45	QUEENS	11368	21
156	-73.854219	40.749724	111-82	QUEENS	11368	21

- This is our location dimension for DOB permit issuance.

Job

```
WITH JobDim AS (
SELECT
    row_number () OVER () AS JobDimID, *
From (
```

```

    SELECT DISTINCT
        Job_Type,
        Job__,
        Job_doc__
    FROM `cis-4400-group-project-403921.permit2.permit
issue` )
)
SELECT *
FROM JobDim

```

JobDimID	Job_Type	Job__	Job_doc__
1	A2	400500490	1
2	NB	500193195	1
3	A3	400679896	1
4	DM	400674819	1
5	A2	101234273	2
6	A2	300404649	1
7	SG	200030331	1
8	A2	200028157	1
9	A2	2007324200	1

- This is our Job dimension that includes Jobtype, job number and job doc number

Permit

```

WITH PermitDIM AS (
    SELECT
        ROW_NUMBER() OVER () AS Permit_DIM_id,
        *
    FROM (
        SELECT DISTINCT
            Permit_Type,
            COALESCE (Permit_Subtype, 'NA') AS Permit_Subtype,
            COALESCE (Permit_Status, 'NA') AS Permit_Status,
            Permit_Sequence__,
            Permittee_s_First_Name,
            Permittee_s_Last_Name,
            COALESCE (Permittee_s_Business_Name, 'NA') AS
Permittee_s_Business_Name,

```

```

        COALESCE (Permittee_s_License_Type, 'NA') AS
Permittee_s_License_Type,
        COALESCE (Permittee_s_License__, 'NA') AS
Permittee_s_License__
    FROM `cis-4400-group-project-403921.permit2.permit
issue`
)
)

SELECT *
FROM PermitDIM

```

Permit_DIM_id	Permit_Type	Permit_Subtype	Permit_Status	Permit_Seque...	Permittee_s_Fi...	Permittee_...
817433	EW	OT	ISSUED	1	EMILY	MILLER
817434	EQ	FN	ISSUED	1	HOWARD	TOMPKIN
817435	AL	NA	ISSUED	1	TEDDY	KRAIN
817436	AL	NA	ISSUED	1	ANTHONY	PELEGREINO
817437	EW	OT	ISSUED	1	GEORGE	JANTOW
817438	SG	NA	ISSUED	1	JONATHON	BRAGOLI
817439	EQ	SF	ISSUED	3	MICHAEL	BRIONE
817440	PL	NA	RE-ISSUED	1	ALVIN	WITTLIN

- This is our Permit dimension consisting of permit type, permit subtype, permit status, sequence number, permitee first and last name

Created_Date_View

```

WITH date AS (
    SELECT *
    FROM {{ ref('DateModel') } }
)
```

```

SELECT date_dim_id as
CreatedDateViewDim,full_date,date_integer,year,year_week,year_day,month,month_name,week_day,day_name
FROM date

```

CreatedDateVie...	full_date	date_integer	year	year_week	year_day	month
1	1995-02-24	19950224	1995	8	24	2
2	1995-09-12	19950912	1995	37	12	9
3	1991-11-15	19911115	1991	45	15	11
4	1992-06-16	19920616	1992	24	16	6
5	1992-10-16	19921016	1992	41	16	10
6	1997-01-23	19970123	1997	3	23	1
7	1995-03-07	19950307	1995	10	7	3
8	1993-01-05	19930105	1993	1	5	1

- This dimension is for our created date view

Close Date View

```

WITH date AS(
    SELECT *
    FROM {{ ref('DateModel') }})
SELECT date_dim_id as
ClosedDateViewDim,full_date,date_integer,year,year_week,year_day,month,month_name,week_day,day_name
FROM date

```

ClosedDateVie...	full_date	date_integer	year	year_week	year_day	month
1	1994-11-03	19941103	1994	44	3	11
2	1997-02-26	19970226	1997	8	26	2
3	1997-03-26	19970326	1997	12	26	3
4	1992-04-07	19920407	1992	14	7	4
5	1995-06-02	19950602	1995	22	2	6
6	1993-02-26	19930226	1993	8	26	2
7	1995-10-10	19951010	1995	41	10	10
8	1995-02-07	19950207	1995	6	7	2

- this is our closed date view for our permit dataset

Big Query

```
SELECT *
FROM `dbt_jjiang.buildingFact`
LEFT JOIN `dbt_jjiang.LocationDim` USING
(Location_DIM_id)
LEFT JOIN `dbt_jjiang.Complaint` USING
(Complaints_DIM_id)
LEFT JOIN `dbt_jjiang.agencyDim` USING (agency_DIM_id)
LEFT JOIN `dbt_jjiang.status` USING (Status_DIM_id)
LEFT JOIN `dbt_jjiang.closed_dateView` USING
(ClosedDateViewDim)
LEFT JOIN `dbt_jjiang.created_dateView` USING
(CreatedDateViewDim);
```

Row	CreatedDateViewDir	ClosedDateViewDim	Status_DIM_id	agency_DIM_id	Complaints_DIM_id	Location_DIM_id	unique
1	10375	10375	2	1	1	7449	
2	23434	23434	2	1	1	16713	
3	60977	60977	2	1	2	44614	
4	8129	8129	2	1	1	5849	
5	23239	23239	1	1	1	16564	
6	31976	31976	2	1	1	22958	
7	28315	28315	2	1	4	20381	
8	29005	29005	2	1	1	41827	
9	37774	37774	2	1	1	44468	
10	60077	60077	2	1	1	44075	

Results per page: 50 ▾ 1 – 50 of 18300 | < < > >|

- This is our joined table for 311 (building/use)

```
SELECT *
FROM `dbt_jjiang.factforpermit`
LEFT JOIN `dbt_jjiang.permit` USING (Permit_DIM_id)
LEFT JOIN `dbt_jjiang.Location` USING (Location_DIM_id)
LEFT JOIN `dbt_jjiang.job` USING (JobDimID)
LEFT JOIN `dbt_jjiang.closed_date_view` USING
(ClosedDateViewDim)
```

```

LEFT JOIN `dbt_jjiang.created_date_view` as t USING
(CreatedDateViewDim)
/*where CreatedDateViewDim is not null*/
where t.year > 2022;

```

Row	CreatedDateViewDim	ClosedDateViewDim	JobDimID	Location_DIM_id	Permit_DIM_id	Permit_Type
1	8059	8059	1061873	364335	690047	NB
2	8059	8059	1061873	415943	690047	NB
3	7680	7680	1074639	86960	263325	EW
4	7680	7680	1074639	417471	263325	EW
5	3952	3952	450244	115929	690047	NB
6	3703	3703	1019329	91102	724467	EW
7	9233	9233	926865	241378	737011	EW
8	2090	2090	1558260	387204	592635	EW
9	3898	3898	747277	382339	592635	EW
10	7322	7322	291111	29889	449350	EW

Results per page: 50 ▾ 1 – 50 of 117693 |< < > >|

- This is our joined table for permit issuance used for our visuals and kpis.

5. Schema

311 (Building / Use)

(Status Schema)

The screenshot shows the BigCommerce interface for managing the 'status' schema. On the left, the Explorer sidebar lists various resources, including 'closed_dateView', 'closed_date_view', 'created_dateView', 'created_date_view', 'factforpermit', 'job', 'my_first_dbt_model', 'my_second_dbt_mod...', 'permit', and 'status'. The 'status' resource is currently selected, indicated by a blue selection bar at the bottom of the sidebar.

The main panel displays the schema details for 'status'. At the top, there are tabs for SCHEMA, DETAILS, LINEAGE, DATA PROFILE, and D. The SCHEMA tab is active, showing a table with three columns: Field name, Type, and Mode. The table contains three rows:

Field name	Type	Mode	Key	Col
Status_DIM_id	INTEGER	NULLABLE	-	-
Status	STRING	NULLABLE	-	-

Below the table is a blue 'EDIT SCHEMA' button.

(Complaints Schema)

The screenshot shows the BigCommerce interface for managing the 'Complaint' schema. On the left, the Explorer sidebar lists resources under 'cis-4400-group-project-403921', including 'Saved queries (3)', 'External connections', 'dbt_asingh', 'dbt_jjiang', and 'DateDim'. The 'Complaint' resource is currently selected, indicated by a blue selection bar at the bottom of the sidebar.

The main panel displays the schema details for 'Complaint'. At the top, there are tabs for SCHEMA, DETAILS, LINEAGE, and DATA PROFILE. The SCHEMA tab is active, showing a table with four columns: Field name, Type, Mode, and Key. The table contains four rows:

Field name	Type	Mode	Key
Complaints_DIM_id	INTEGER	NULLABLE	-
Descriptor	STRING	NULLABLE	-
Complaint_Type	STRING	NULLABLE	-

(Date Dim Schema)

The screenshot shows the dbt UI interface. The top navigation bar includes 'Explorer', '+ ADD', and a search bar. The main title is 'DateDim'. Below the title, there are tabs for 'SCHEMA', 'DETAILS', 'LINEAGE', 'DATA PROFILE', and 'DATA'. The 'SCHEMA' tab is selected. A table lists the fields of the DateDim schema:

	Field name	Type	Mode	Key	Co
<input type="checkbox"/>	date_dim_id	INTEGER	NULLABLE	-	-
<input type="checkbox"/>	date_integer	STRING	NULLABLE	-	-
<input type="checkbox"/>	full_date	TIMESTAMP	NULLABLE	-	-
<input type="checkbox"/>	year	INTEGER	NULLABLE	-	-
<input type="checkbox"/>	year_week	INTEGER	NULLABLE	-	-
<input type="checkbox"/>	year_day	INTEGER	NULLABLE	-	-
<input type="checkbox"/>	month	INTEGER	NULLABLE	-	-
<input type="checkbox"/>	month_name	STRING	NULLABLE	-	-
<input type="checkbox"/>	week_day	STRING	NULLABLE	-	-
<input type="checkbox"/>	day_name	STRING	NULLABLE	-	-

The left sidebar shows a tree view of resources under 'cis-4400-group-project-403921', including 'Saved queries', 'External connections', 'dbt_asingh', 'dbt_jjiang', and various models like 'Complaint', 'DateDim', 'DateModel', etc.

(Location Schema)

The screenshot shows the dbt UI interface. The top navigation bar includes 'Explorer', '+ ADD', and a search bar. The main title is 'LocationDim'. Below the title, there are tabs for 'SCHEMA', 'DETAILS', 'LINEAGE', 'DATA PROFILE', and 'DATA'. The 'SCHEMA' tab is selected. A table lists the fields of the LocationDim schema:

	Field name	Type	Mode	Key	Co
<input type="checkbox"/>	Location_DIM_id	INTEGER	NULLABLE	-	-
<input type="checkbox"/>	Street_Name	STRING	NULLABLE	-	-
<input type="checkbox"/>	Latitude	FLOAT	NULLABLE	-	-
<input type="checkbox"/>	Longitude	FLOAT	NULLABLE	-	-
<input type="checkbox"/>	City	STRING	NULLABLE	-	-
<input type="checkbox"/>	Zip_Codes	STRING	NULLABLE	-	-
<input type="checkbox"/>	Borough	STRING	NULLABLE	-	-

The left sidebar shows a tree view of resources under 'dbt_jjiang', including 'Complaint', 'DateDim', 'DateModel', 'Location', and various models like 'LocationDim', 'agencyDim', etc.

(Agency Schema)

The screenshot shows the dbt UI interface. On the left, the Explorer sidebar lists resources under the project 'dbt_jjiang'. A specific schema named 'agencyDim' is selected and highlighted in blue. The main panel displays the schema details for 'agencyDim'. The 'SCHEMA' tab is active, showing a table of fields:

Field name	Type	Mode	Key
agency_DIM_id	INTEGER	NULABLE	-
agency	STRING	NULABLE	-
agency_name	STRING	NULABLE	-

Below the table is a blue 'EDIT SCHEMA' button.

(Created_Date_View Schema)

The screenshot shows the dbt UI interface. On the left, the Explorer sidebar lists resources under the project 'dbt_jjiang'. A specific schema named 'created_date_view' is selected and highlighted in blue. The main panel displays the schema details for 'created_date_view'. The 'SCHEMA' tab is active, showing a table of fields:

Field name	Type	Mode	Key
CreatedDateViewDim	INTEGER	NULABLE	-
full_date	DATE	NULABLE	-
date_integer	STRING	NULABLE	-
year	INTEGER	NULABLE	-
year_week	INTEGER	NULABLE	-
year_day	INTEGER	NULABLE	-
month	INTEGER	NULABLE	-
month_name	STRING	NULABLE	-
week_day	STRING	NULABLE	-
day_name	STRING	NULABLE	-

(Closed_Date_View Schema)

The screenshot shows the BigQuery UI with the following details:

Explorer tab is selected.

The current view is titled `closed_dateView`.

SCHEMA tab is selected.

Filter: Enter property name or value

Field name	Type	Mode	Key
ClosedDateViewDim	INTEGER	NULLABLE	-
full_date	TIMESTAMP	NULLABLE	-
date_integer	STRING	NULLABLE	-
year	INTEGER	NULLABLE	-
year_week	INTEGER	NULLABLE	-
year_day	INTEGER	NULLABLE	-
month	INTEGER	NULLABLE	-
month_name	STRING	NULLABLE	-
week_day	STRING	NULLABLE	-
day_name	STRING	NULLABLE	-

(Building / Use fact table Schema)

The screenshot shows the BigQuery UI with the following details:

Explorer tab is selected.

The current view is titled `buildingFact`.

SCHEMA tab is selected.

Filter: Enter property name or value

Field name	Type	Mode	Key	Collation	Default Value	F
ClosedDateViewDim	INTEGER	NULLABLE	-	-	-	-
CreatedDateViewDim	INTEGER	NULLABLE	-	-	-	-
Location_DIM_id	INTEGER	NULLABLE	-	-	-	-
agency_DIM_id	INTEGER	NULLABLE	-	-	-	-
Status_DIM_id	INTEGER	NULLABLE	-	-	-	-
Complaints_DIM_id	INTEGER	NULLABLE	-	-	-	-
unique_key	INTEGER	NULLABLE	-	-	-	-

DOB (Permit Issuance)

Fact table Schema

The screenshot shows the schema for the factforpermit fact table. The interface includes a sidebar for navigating resources and a main panel for viewing the schema. The schema table has columns for Field name, Type, Mode, Key, and Collation.

Field name	Type	Mode	Key	Collation
Permit_DIM_id	INTEGER	NULLABLE	-	-
Location_DIM_id	INTEGER	NULLABLE	-	-
JobDimID	INTEGER	NULLABLE	-	-
ClosedDateViewDim	INTEGER	NULLABLE	-	-
CreatedDateViewDim	INTEGER	NULLABLE	-	-

(Job Schema)

The screenshot shows the schema for the job dimension table. The interface includes a sidebar for navigating resources and a main panel for viewing the schema. The schema table has columns for Field name, Type, Mode, and Key.

Field name	Type	Mode	Key
JobDimID	INTEGER	NULLABLE	-
Job_Type	STRING	NULLABLE	-
Job__	INTEGER	NULLABLE	-
Job_doc__	INTEGER	NULLABLE	-

(Permit Schema)

The screenshot shows the dbt UI interface. On the left, there's an 'Explorer' sidebar with a search bar and a list of resources including 'closed_dateView', 'permit', 'status', and 'test3'. A specific resource, 'permit', is selected and highlighted with a blue background. On the right, the main panel has a title bar with tabs for 'permit', 'QUERY', 'SHARE', 'COPY', and 'DELETE'. Below the title bar, there are two tabs: 'SCHEMA' (which is active) and 'DETAILS'. Under the 'SCHEMA' tab, there's a 'Filter' input field and a table listing columns with their properties. The columns listed are:

Field name	Type	Mode	Key
Permit_DIM_id	INTEGER	NULLABLE	-
Permit_Type	STRING	NULLABLE	-
Permit_Subtype	STRING	NULLABLE	-
Permit_Status	STRING	NULLABLE	-
Permit_Sequence__	INTEGER	NULLABLE	-
Permittee_s_First_Name	STRING	NULLABLE	-
Permittee_s_Last_Name	STRING	NULLABLE	-
Permittee_s_Business_Name	STRING	NULLABLE	-
Permittee_s_License_Type	STRING	NULLABLE	-
Permittee_s_License__	STRING	NULLABLE	-

(Location Schema)

The screenshot shows the dbt UI interface. On the left, there's an 'Explorer' sidebar with a search bar and a list of resources including 'Complaint', 'DateDim', 'DateModel', 'Location', 'LocationDim', 'agencyDim', 'buildingFact', 'closed_dateView', 'closed_date_view', and 'created_dateView'. A specific resource, 'Location', is selected and highlighted with a blue background. On the right, the main panel has a title bar with tabs for 'Location', 'QUERY', 'SHARE', 'COPY', and 'DELETE'. Below the title bar, there are five tabs: 'SCHEMA' (active), 'DETAILS', 'LINEAGE', 'DATA PROFILE', and 'DAT'. Under the 'SCHEMA' tab, there's a 'Filter' input field and a table listing columns with their properties. The columns listed are:

Field name	Type	Mode	Key
Location_DIM_id	INTEGER	NULLABLE	-
LONGITUDE	FLOAT	NULLABLE	-
LATITUDE	FLOAT	NULLABLE	-
House__	STRING	NULLABLE	-
Borough	STRING	NULLABLE	-
Zip_Code	STRING	NULLABLE	-
COUNCIL_DISTRICT	STRING	NULLABLE	-

(Created_date_view Schema)

The screenshot shows the BigQuery UI with the following details:

- Left Panel (Explorer):** Shows a list of resources. The item "created_dateView" is selected and highlighted in blue.
- Top Bar:** The title bar shows "created_dateView".
- Right Panel (Schema View):**
 - Schema Tab:** Active tab.
 - Fields:**

Field name	Type	Mode	Key
CreatedDateViewDim	INTEGER	NULLABLE	-
full_date	TIMESTAMP	NULLABLE	-
date_integer	STRING	NULLABLE	-
year	INTEGER	NULLABLE	-
year_week	INTEGER	NULLABLE	-
year_day	INTEGER	NULLABLE	-
month	INTEGER	NULLABLE	-
month_name	STRING	NULLABLE	-
week_day	STRING	NULLABLE	-
day_name	STRING	NULLABLE	-

(Closed_date_view Schema)

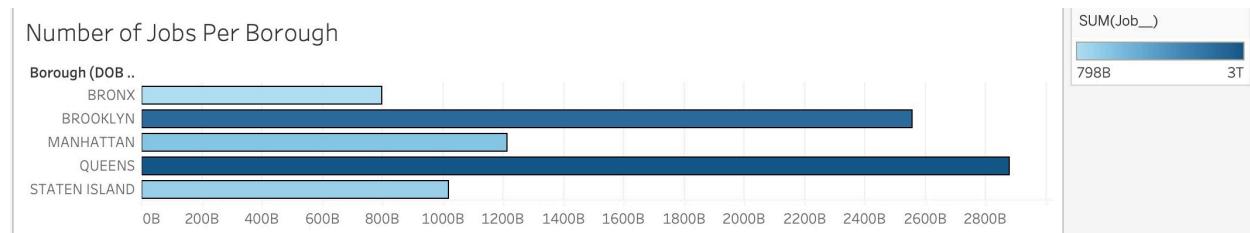
The screenshot shows the BigQuery UI with the following details:

- Left Panel (Explorer):** Shows a list of resources. The item "closed_dateView" is selected and highlighted in blue.
- Top Bar:** The title bar shows "closed_dateView".
- Right Panel (Schema View):**
 - Schema Tab:** Active tab.
 - Fields:**

Field name	Type	Mode	Key
ClosedDateViewDim	INTEGER	NULLABLE	-
full_date	TIMESTAMP	NULLABLE	-
date_integer	STRING	NULLABLE	-
year	INTEGER	NULLABLE	-
year_week	INTEGER	NULLABLE	-
year_day	INTEGER	NULLABLE	-
month	INTEGER	NULLABLE	-
month_name	STRING	NULLABLE	-
week_day	STRING	NULLABLE	-
day_name	STRING	NULLABLE	-

6. KPIs and Visuals

Number of Jobs Per Borough



-The bar chart illustrates the number of jobs conducted in 311 building/use and DOB permit issuance across New York City's boroughs reveals Queens as the dominant player, surpassing other boroughs. This observation demonstrates Queens' role in construction and development activities within the city. Queens' supremacy in job numbers implies a substantial economic impact, presenting opportunities for increased employment, heightened property values, and improved infrastructure. As the city evolves, continued monitoring of these trends will be crucial for adaptive urban planning that responds to the dynamic needs of New York City and its residents. This analysis provides valuable insights for stakeholders seeking to foster resilient and equitable development across the diverse landscape of the city.

Number of Permits permitted per borough



- Between all 5 boroughs just in 2023 there were 29,367 permits issued. 3,924 were in Bronx, 8,628 in Brooklyn, 8,732 in Queens, 5,445 in Manhattan, and 2,638 in Staten Island. These numbers are smaller from the last couple of years where in 2021 a total of 69,000 permits were issued.

Complaint Types per Borough



- The boroughs have their own colors and are color-coded in different pigments. The most common complaint was complaint number 1 which pertains to the illegal conversion of residential building spaces, followed by complaint 4 of illegal commercial use in the resident zone, number 3 zoning illegal vehicle storage and, 1, and 2 are nearly tied for construction and, lack of permits

7. Our project implementation involved (ETL) tools, and a Target Database Management System (DBMS). Specifically, we leveraged dbt for data transformation, and Google Cloud BigQuery served as our hosted project repository. SQL was employed for coding dimensions and facts within BigQuery and dbt. Additionally, Tableau was utilized for the creation of visualizations, contributing to interesting insights. This selection of tools played a fundamental role in achieving the project's objectives with efficiency and precision.

8.

a.

- The execution of our project was made possible through the strategic utilization of various software and database tools, each playing a pivotal role in coordinating and managing the project, as well as facilitating the execution of programming tasks. Our comprehensive toolset included:
 - Google BigQuery:
 - Utilized for high-performance, massively scalable analytics on large datasets. Google BigQuery allowed us to efficiently store and analyze vast amounts of data, ensuring seamless integration with our analytical processes.
 - dbt (data build tool):
 - Served as a critical component for transforming and modeling our data. With dbt, we were able to streamline the data transformation process, ensuring consistency and reliability in our analytical outputs.
 - Tableau:

- Played a central role in visualizing and interpreting our data. Tableau's intuitive interface and powerful visualization capabilities allowed us to create insightful visuals, enabling stakeholders to easily grasp complex insights derived from our analyses.

- Lucidchart:

- Employed for collaborative diagramming and visualization of our project workflows.

- SQL:

- Acted as the common language for interacting with our databases. SQL was fundamental in querying and manipulating data within our databases, providing a standardized and efficient means for the team to interact with the underlying data structures.

This integrated approach to utilizing these tools not only enhanced our project management capabilities but also streamlined our programming tasks. The synergy between Google BigQuery for data storage, dbt for efficient data transformation, Tableau for visualization, Lucidchart for collaborative workflow representation, and SQL as a universal language created a cohesive environment, fostering collaboration and ensuring the successful realization of our project objectives. The seamless integration of these tools empowered our team to navigate the complexities of our data-driven project with precision and agility.

b.

- Crafting the dimensional models took some time, and the ETL process, especially data extraction, posed its challenges. Learning to use dbt, a tool new to all of us, added to the complexity and slowed down our initial

data extraction efforts. Coding the fact table turned out to be quite a tough, with persistent errors and complexities in achieving the desired results. Handling null values also brought its set of difficulties, requiring careful decisions to maintain data integrity. Overall, SQL coding emerged as the trickiest part, involving both the nuances of fact table development and the thoughtful management of null values.

Looking back, the project taught us the importance of getting acquainted with tools like dbt beforehand. If we had the chance to do it again, we would've started working with dbt earlier in the semester. Additionally, focusing on data profiling for null value management from the get-go might have eased the challenges, we assumed it would be easy and did not put enough effort into that part. Nevertheless, these hurdles collectively deepened our understanding of the project and emphasized the need for a well-coordinated approach in dealing with the intricacies of data modeling and coding.

c.

- Some benefits that the data warehouse provides is the ability to see correlations and data between the 311 complaints (building /use) and the rate of permits being issued and the amount being issued around the five boroughs

d.

- Tackling the data warehouse project was a mix of satisfaction and challenge for us. It pushed our SQL skills to the limit and introduced us to the world of dbt and Google BigQuery. Despite the initial complexities, we're content with the practical skills gained. This experience has not only deepened our

understanding but also added practical value to our skill set. It's safe to say this project will make a solid addition to our resumes, marking a notable step forward in our professional journey.