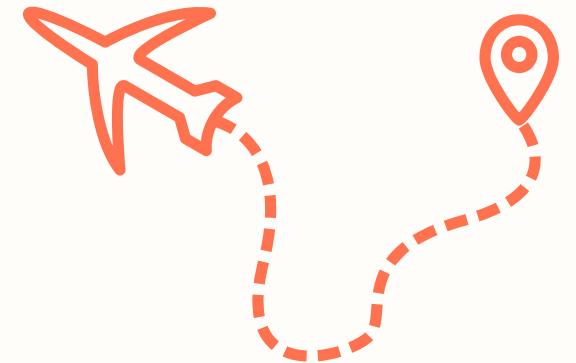




Reimagining a
Destination



Orlando Economic Development

CS 689 Term Project
Alissa Crist
acrist@bu.edu



Project Scope

Many theme parks investing billions into construction projects to expand or improve resorts

This analysis seeks to explore the impact this surge in post-pandemic tourism investment has had on the Orlando community and economy

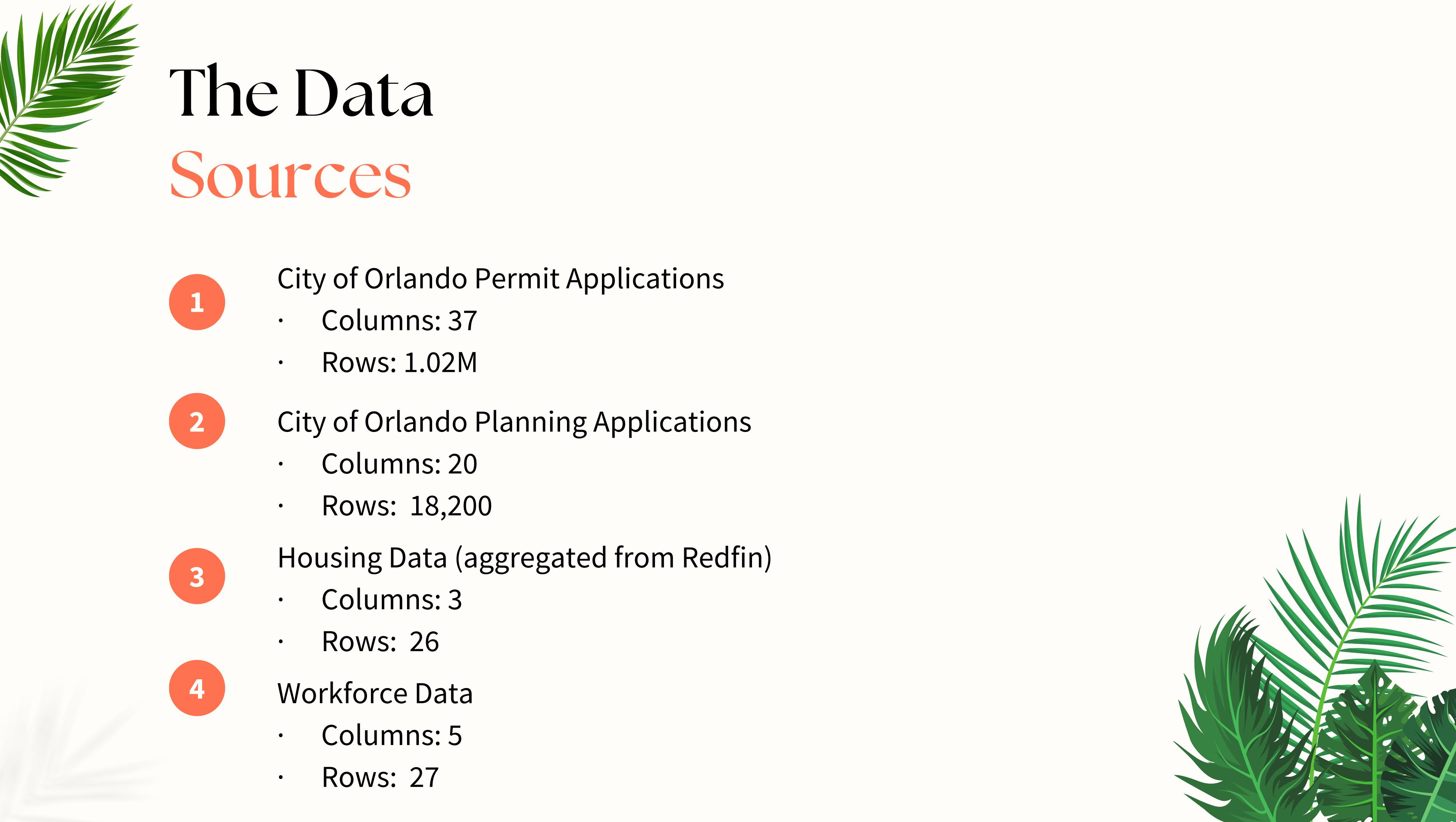




Project Scope

Questions

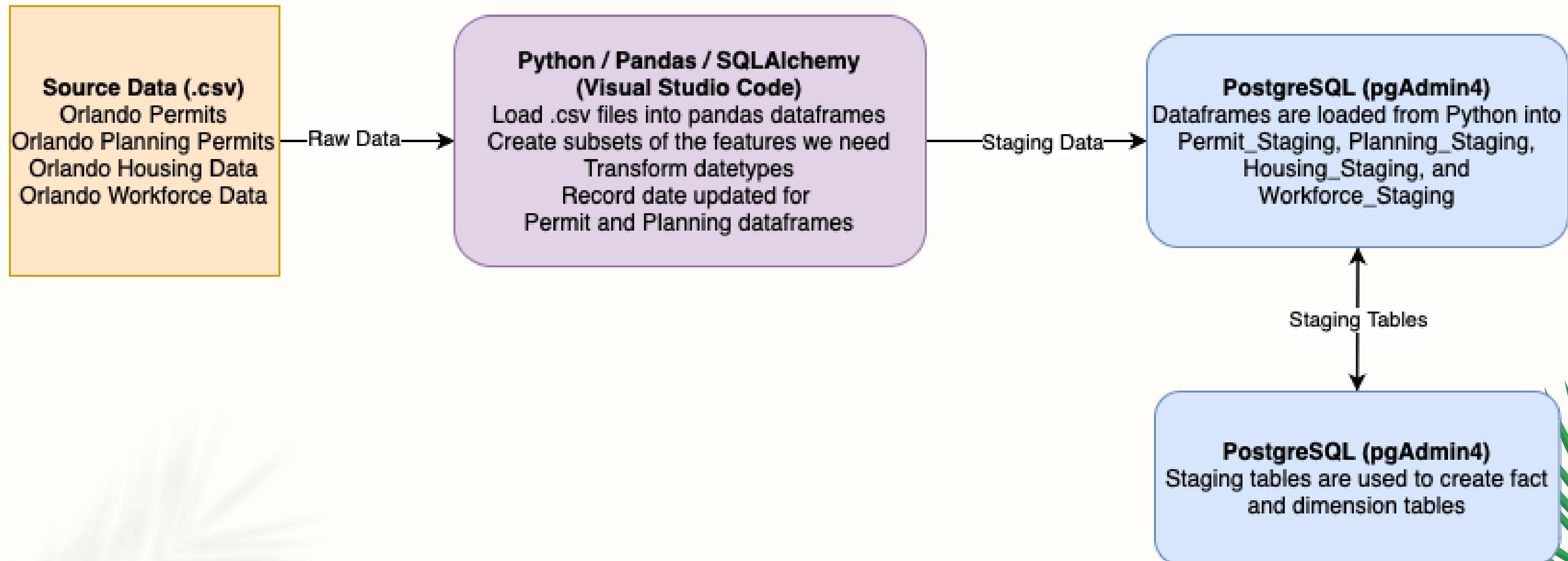
- 1 Which neighborhoods in Orlando saw the most applications for COMMERCIAL development in the past year?
- 2 Which neighborhoods in Orlando saw the most permit applications for RESIDENTIAL development in the past year?
- 3 Which contractor types were most hired in Orlando in the last year?
- 4 How has both residential and commercial development in the Orlando area impacted home prices, sales, and inventory?
- 5 How has COMMERCIAL development in the Orlando area impacted workforce numbers?



The Data Sources

- 1 City of Orlando Permit Applications
 - Columns: 37
 - Rows: 1.02M
- 2 City of Orlando Planning Applications
 - Columns: 20
 - Rows: 18,200
- 3 Housing Data (aggregated from Redfin)
 - Columns: 3
 - Rows: 26
- 4 Workforce Data
 - Columns: 5
 - Rows: 27

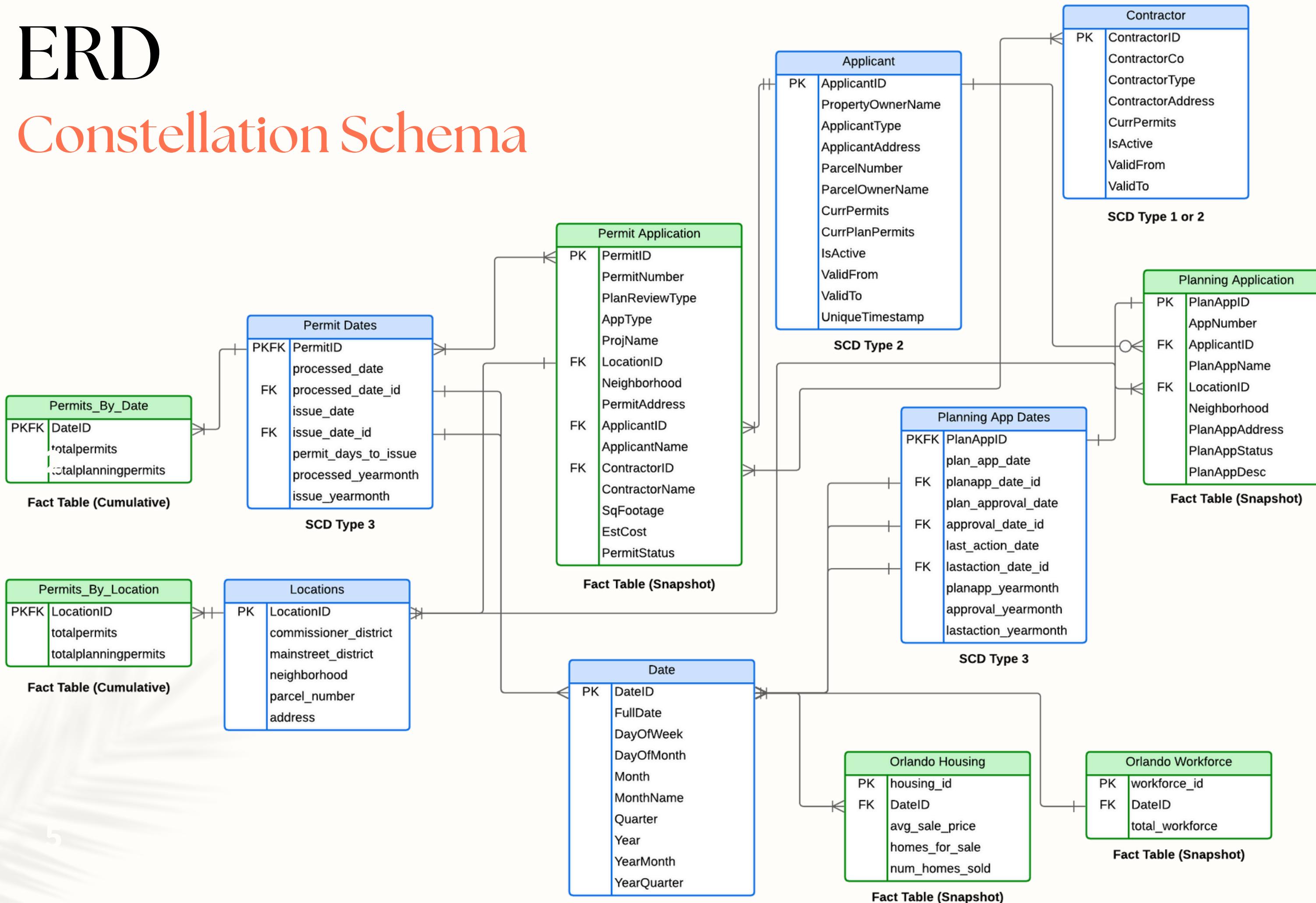
Data Flow Diagram





ERD

Constellation Schema





ETL Load - Python

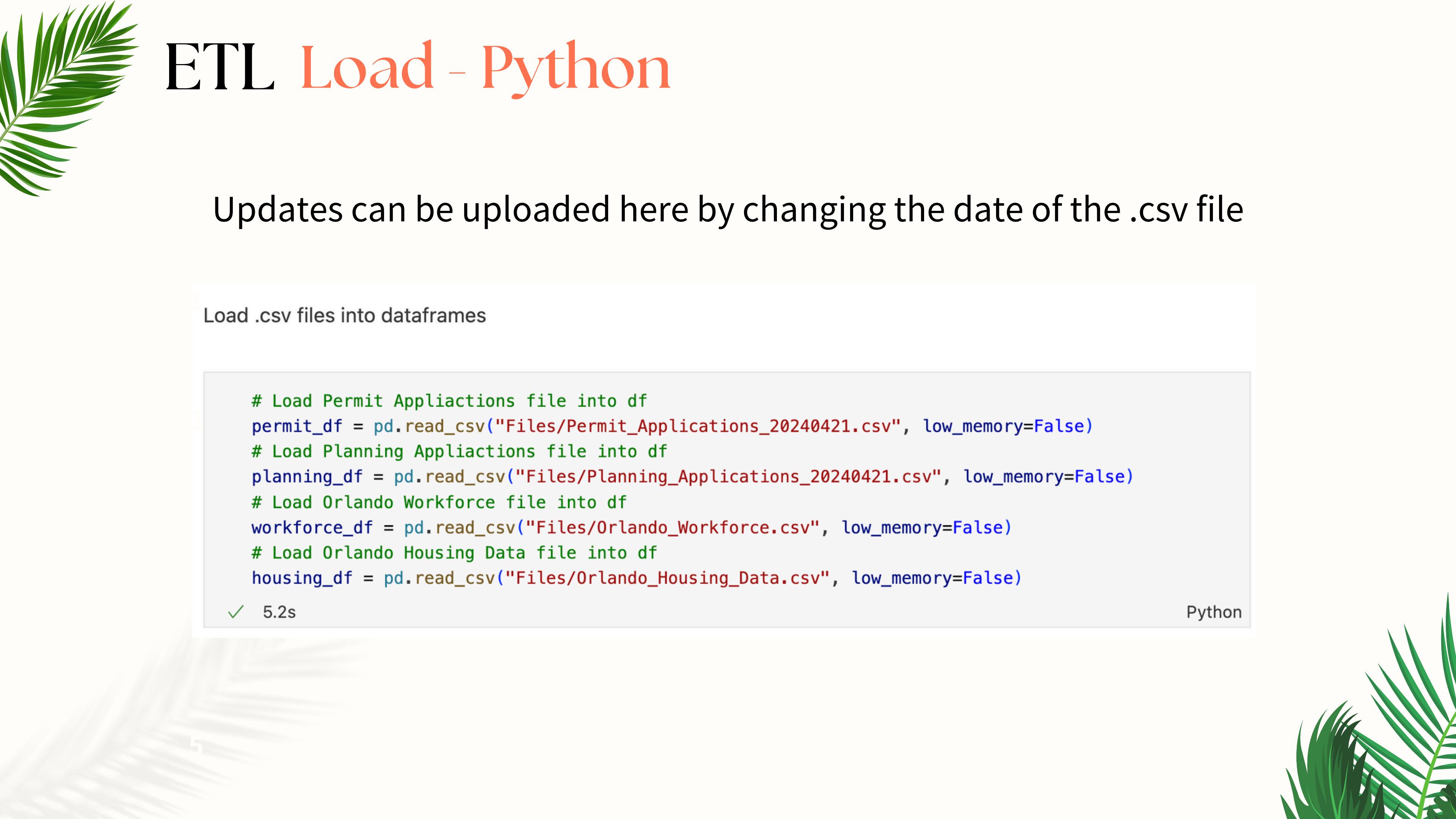
```
from sqlalchemy import create_engine

# PostgreSQL connection details
dbname = 'acrist_cs689_termproj'
user = 'postgres'
password = 'postgres1092'
host = 'localhost'

# Create SQLAlchemy engine
engine = create_engine(f'postgresql://{user}:{password}@{host}/{dbname}')

# Create staging tables in the orlando_development schema
permit_df_2022_sub.to_sql('Permit_Staging', engine, schema='orlando_development', if_exists='replace', index=False)
planning_df_2022_sub.to_sql('Planning_Staging', engine, schema='orlando_development', if_exists='replace', index=False)
workforce_df.to_sql('Workforce_Staging', engine, schema='orlando_development', if_exists='replace', index=False)
housing_df.to_sql('Housing_Staging', engine, schema='orlando_development', if_exists='replace', index=False)

# Commit changes
engine.dispose()
```



ETL Load - Python

Updates can be uploaded here by changing the date of the .csv file

Load .csv files into dataframes

```
# Load Permit Applications file into df
permit_df = pd.read_csv("Files/Permit_Applications_20240421.csv", low_memory=False)
# Load Planning Applications file into df
planning_df = pd.read_csv("Files/Planning_Applications_20240421.csv", low_memory=False)
# Load Orlando Workforce file into df
workforce_df = pd.read_csv("Files/Orlando_Workforce.csv", low_memory=False)
# Load Orlando Housing Data file into df
housing_df = pd.read_csv("Files/Orlando_Housing_Data.csv", low_memory=False)
```

✓ 5.2s

Python

ETL Table Schema - SQL

	table_catalog name	table_schema name	table_name name	table_type character varying
1	acrist_cs689_termproj	orlando_development	Housing_Staging	BASE TABLE
2	acrist_cs689_termproj	orlando_development	Permit_Staging	BASE TABLE
3	acrist_cs689_termproj	orlando_development	Planning_Staging	BASE TABLE
4	acrist_cs689_termproj	orlando_development	Workforce_Staging	BASE TABLE
5	acrist_cs689_termproj	orlando_development	ApplicantDim	BASE TABLE
6	acrist_cs689_termproj	orlando_development	ContractorDim	BASE TABLE
7	acrist_cs689_termproj	orlando_development	LocationDim	BASE TABLE
8	acrist_cs689_termproj	orlando_development	DateDim	BASE TABLE
9	acrist_cs689_termproj	orlando_development	PermitAppFact	BASE TABLE
10	acrist_cs689_termproj	orlando_development	PlanningFact	BASE TABLE
11	acrist_cs689_termproj	orlando_development	PermitDateDim	BASE TABLE
12	acrist_cs689_termproj	orlando_development	PlanningDateDim	BASE TABLE
13	acrist_cs689_termproj	orlando_development	Permits_By_Date	BASE TABLE
14	acrist_cs689_termproj	orlando_development	Permits_By_Location	BASE TABLE
15	acrist_cs689_termproj	orlando_development	OrlandoHousingFact	BASE TABLE
16	acrist_cs689_termproj	orlando_development	OrlandoWorkForceFact	BASE TABLE

ETL SCD Type 2:ApplicantDim

Step 1: Load updated data from Python by changing the name of the .csv file to the latest we have

```
from sqlalchemy import create_engine

# PostgreSQL connection details
dbname = 'acrist_cs689_termproj'
user = 'postgres'
password = 'postgres1092'
host = 'localhost'

# Create SQLAlchemy engine
engine = create_engine(f'postgresql://{{user}}:{{password}}@{{host}}/{{dbname}}')

# Create staging tables in the orlando_development schema
permit_df_2022_sub.to_sql('Permit_Staging', engine, schema='orlando_development', if_exists='replace', index=False)
planning_df_2022_sub.to_sql('Planning_Staging', engine, schema='orlando_development', if_exists='replace', index=False)
workforce_df.to_sql('Workforce_Staging', engine, schema='orlando_development', if_exists='replace', index=False)
housing_df.to_sql('Housing_Staging', engine, schema='orlando_development', if_exists='replace', index=False)

# Commit changes
engine.dispose()
```

ETL SCD Type 2: ApplicantDim

Step 2: Create the ApplicantDim table and load using
update_applicant_dim_full() function in PostgreSQL

See .sql file for this function in full

ETL SCD Type 2: ApplicantDim

Step 3: Load original data (from 4/10/24) and show results

CurrPermits in descending order (original)

	applicantid [PK] integer	currpermits integer	currplanpermits integer	isactive boolean	validfrom timestamp without time zone	validto timestamp without time zone	updatetimestamp timestamp without time zone
1	34367	4	0	true	2024-04-23 13:50:16.010141	[null]	2024-04-23 13:50:16.010141
2	31196	4	0	true	2024-04-23 13:50:16.010141	[null]	2024-04-23 13:50:16.010141
3	31216	4	0	true	2024-04-23 13:50:16.010141	[null]	2024-04-23 13:50:16.010141
4	32806	4	0	true	2024-04-23 13:50:16.010141	[null]	2024-04-23 13:50:16.010141
5	30831	4	0	true	2024-04-23 13:50:16.010141	[null]	2024-04-23 13:50:16.010141
6	32604	4	0	true	2024-04-23 13:50:16.010141	[null]	2024-04-23 13:50:16.010141
7	30849	4	0	true	2024-04-23 13:50:16.010141	[null]	2024-04-23 13:50:16.010141
8	21421	4	0	true	2024-04-23 15:02:28.872106	[null]	2024-04-23 15:02:28.872106
9	3155	4	0	true	2024-04-23 15:02:28.872106	[null]	2024-04-23 15:02:28.872106
10	35475	4	0	true	2024-04-23 13:50:16.010141	[null]	2024-04-23 13:50:16.010141

ETL SCD Type 2: ApplicantDim

Step 4: Load **NEW** data (from 4/21/24) and show results

CurrPermits in descending order (NEW)

	applicantid [PK] integer	currpermits integer	currplanpermits integer	isactive boolean	validfrom timestamp without time zone	validto timestamp without time zone	updatetimestamp timestamp without time zone
1	32806	4	0	true	2024-04-23 13:50:16.010141	[null]	2024-04-23 13:50:16.010141
2	30831	4	0	true	2024-04-23 13:50:16.010141	[null]	2024-04-23 13:50:16.010141
3	21421	4	0	true	2024-04-23 15:04:01.062902	[null]	2024-04-23 15:04:01.062902
4	30849	4	0	true	2024-04-23 13:50:16.010141	[null]	2024-04-23 13:50:16.010141
5	37159	4	0	true	2024-04-23 13:50:16.010141	[null]	2024-04-23 13:50:16.010141
6	38094	4	0	true	2024-04-23 13:50:16.010141	[null]	2024-04-23 13:50:16.010141
7	18977	4	0	true	2024-04-23 15:04:01.062902	[null]	2024-04-23 15:04:01.062902
8	28646	4	0	true	2024-04-23 15:04:01.062902	[null]	2024-04-23 15:04:01.062902
9	31216	4	0	true	2024-04-23 13:50:16.010141	[null]	2024-04-23 13:50:16.010141
10	32604	4	0	true	2024-04-23 13:50:16.010141	[null]	2024-04-23 13:50:16.010141

ETL SCD Type 2: ApplicantDim

Step 3: Load original data (from 4/10/24) and show results

CurrPlanPermits in descending order (original)

	applicantid [PK] integer	currpermits integer	currplanpermits integer	isactive boolean	validfrom timestamp without time zone	validto timestamp without time zone	updatetimestamp timestamp without time zone
1	193759	[null]	8	true	2024-04-23 15:02:28.872106	[null]	2024-04-23 15:02:28.872106
2	194072	[null]	7	true	2024-04-23 15:02:28.872106	[null]	2024-04-23 15:02:28.872106
3	194402	[null]	6	true	2024-04-23 15:02:28.872106	[null]	2024-04-23 15:02:28.872106
4	193782	[null]	6	true	2024-04-23 15:02:28.872106	[null]	2024-04-23 15:02:28.872106
5	193926	[null]	6	true	2024-04-23 15:02:28.872106	[null]	2024-04-23 15:02:28.872106
6	193906	[null]	6	true	2024-04-23 15:02:28.872106	[null]	2024-04-23 15:02:28.872106
7	193304	[null]	6	true	2024-04-23 15:02:28.872106	[null]	2024-04-23 15:02:28.872106
8	422501	[null]	5	true	2024-04-23 15:01:19.532113	[null]	2024-04-23 15:01:19.532113
9	293690	[null]	5	true	2023-03-01 00:00:00	[null]	2023-03-01 00:00:00
10	193675	[null]	5	true	2024-04-23 15:02:28.872106	[null]	2024-04-23 15:02:28.872106

ETL SCD Type 2: ApplicantDim

Step 4: Load **NEW** data (from 4/21/24) and show results

CurrPlanPermits in descending order (**NEW**)

	applicantid [PK] integer	currpermits	currplanpermits	isactive boolean	validfrom timestamp without time zone	validto timestamp without time zone	updatetimestamp timestamp without time zone
1	193759	[null]	8	true	2024-04-23 15:04:01.062902	[null]	2024-04-23 15:04:01.062902
2	194072	[null]	7	true	2024-04-23 15:04:01.062902	[null]	2024-04-23 15:04:01.062902
3	193782	[null]	6	true	2024-04-23 15:04:01.062902	[null]	2024-04-23 15:04:01.062902
4	193304	[null]	6	true	2024-04-23 15:04:01.062902	[null]	2024-04-23 15:04:01.062902
5	194402	[null]	6	true	2024-04-23 15:04:01.062902	[null]	2024-04-23 15:04:01.062902
6	193906	[null]	6	true	2024-04-23 15:04:01.062902	[null]	2024-04-23 15:04:01.062902
7	193926	[null]	6	true	2024-04-23 15:04:01.062902	[null]	2024-04-23 15:04:01.062902
8	425435	[null]	5	true	2024-04-23 15:01:19.532113	[null]	2024-04-23 15:01:19.532113
9	295466	[null]	5	true	2024-03-01 00:00:00	[null]	2024-03-01 00:00:00
10	193403	[null]	5	true	2024-04-23 15:04:01.062902	[null]	2024-04-23 15:04:01.062902

Queries | Question 1

Which neighborhoods in Orlando saw the most applications for commercial development in the past year?

```
SELECT
    Neighborhood,
    TotalCommercialApplications,
    RANK() OVER (ORDER BY TotalCommercialApplications DESC) AS NeighborhoodRank
FROM (
    SELECT
        l.Neighborhood,
        COUNT(p.PermitID) AS TotalCommercialApplications
    FROM
        orlando_development."PermitAppFact" p
    JOIN
        orlando_development."PermitDateDim" pd ON pd.PermitID = p.PermitID
    JOIN
        orlando_development."DateDim" d ON pd.processed_date_id = d.DateID
    JOIN
        orlando_development."LocationDim" l ON p.LocationID = l.LocationID
    WHERE
        d.FullDate BETWEEN DATE_TRUNC('year', CURRENT_DATE) -
        INTERVAL '1 year' AND DATE_TRUNC('year', CURRENT_DATE) - INTERVAL '1 day'
        AND p.PlanReviewType IN ('Commercial')
    GROUP BY
        l.Neighborhood
) AS Subquery
ORDER BY
    TotalCommercialApplications DESC;
```

Output | Question 1

Which neighborhoods in Orlando saw the most applications for commercial development in the past year?

	neighborhood character varying (255)	totalcommercialapplications bigint	neighborhoodrank bigint
1	Florida Center	3866	1
2	Central Business District	1920	2
3	Airport North	1182	3
4	Princeton/Silver Star	1052	4
5	South Orange	914	5
6	Lake Nona South	874	6
7	Millenia	706	7
8	Metro West	666	8
9	ORLANDO INTERNATIONAL AIRPORT	578	9
10	North Orange	478	10

Queries | Question 3

Which contractor types were most hired in Orlando in the last year?

```
SELECT
    ContractorType,
    PlanReviewType,
    TotalHires,
    Rank
FROM (
    SELECT
        c.ContractorType,
        p.PlanReviewType,
        COUNT(p.PermitID) AS TotalHires,
        RANK() OVER (PARTITION BY p.PlanReviewType ORDER BY COUNT(p.PermitID) DESC) AS Rank
    FROM
        orlando_development."PermitAppFact" p
    JOIN
        orlando_development."ContractorDim" c ON p.ContractorID = c.ContractorID
    JOIN
        orlando_development."PermitDateDim" pd ON p.PermitID = pd.PermitID
    JOIN
        orlando_development."DateDim" d ON pd.processed_date_id = d.DateID
    WHERE
        d.FullDate BETWEEN DATE_TRUNC('year', CURRENT_DATE) -
        INTERVAL '1 year' AND DATE_TRUNC('year', CURRENT_DATE) - INTERVAL '1 day'
        AND p.PlanReviewType IN ('Commercial', 'Residential 1/2', 'Residential 3 or more')
    GROUP BY
        c.ContractorType,
        p.PlanReviewType
) SubQuery
WHERE Rank <= 10
ORDER BY
    PlanReviewType
```

Output | Question 3

Which contractor types were most hired in Orlando in the last year?

	contractortype character varying (255) 	planreviewtype character varying (100) 	totalhires bigint 	rank bigint 
1	General	Commercial	6422	1
2	Electrical	Commercial	4880	2
3	Fire	Commercial	3712	3

11	General	Residential 1/2	17804	1
12	Electrical	Residential 1/2	8396	2
13	Plumbing	Residential 1/2	6678	3

20	General	Residential 3 or more	3440	1
21	Mechanical	Residential 3 or more	1410	2
22	Electrical	Residential 3 or more	1050	3

Queries | Question 4

How has both residential and commercial development in the Orlando area impacted home prices, sales, and inventory?

```
SELECT
    dd.YearMonth,
    ohf.Median_Sale_Price,
    ohf.Num_Homes_Sold,
    ohf.Homes_for_Sale,
    COALESCE(pbd.TotalPermits, 0) AS TotalPermits,
    COALESCE(pbd.TotalPlanningPermits, 0) AS TotalPlanningPermits
FROM
    orlando_development."OrlandoHousingFact" ohf
JOIN
    orlando_development."DateDim" dd ON ohf.DateID = dd.DateID
LEFT JOIN
    (SELECT DateID, SUM(TotalPermits) AS TotalPermits,
        SUM(TotalPlanningPermits) AS TotalPlanningPermits
     FROM orlando_development."Permits_By_Date"
     GROUP BY DateID) pbd ON dd.DateID = pbd.DateID
WHERE
    dd.FullDate BETWEEN DATE_TRUNC('year', CURRENT_DATE) - INTERVAL '2 year'
    AND DATE_TRUNC('year', CURRENT_DATE) - INTERVAL '1 day'
ORDER BY
    dd.YearMonth;
```

Output | Question 4

How has both residential and commercial development in the Orlando area impacted home prices, sales, and inventory?

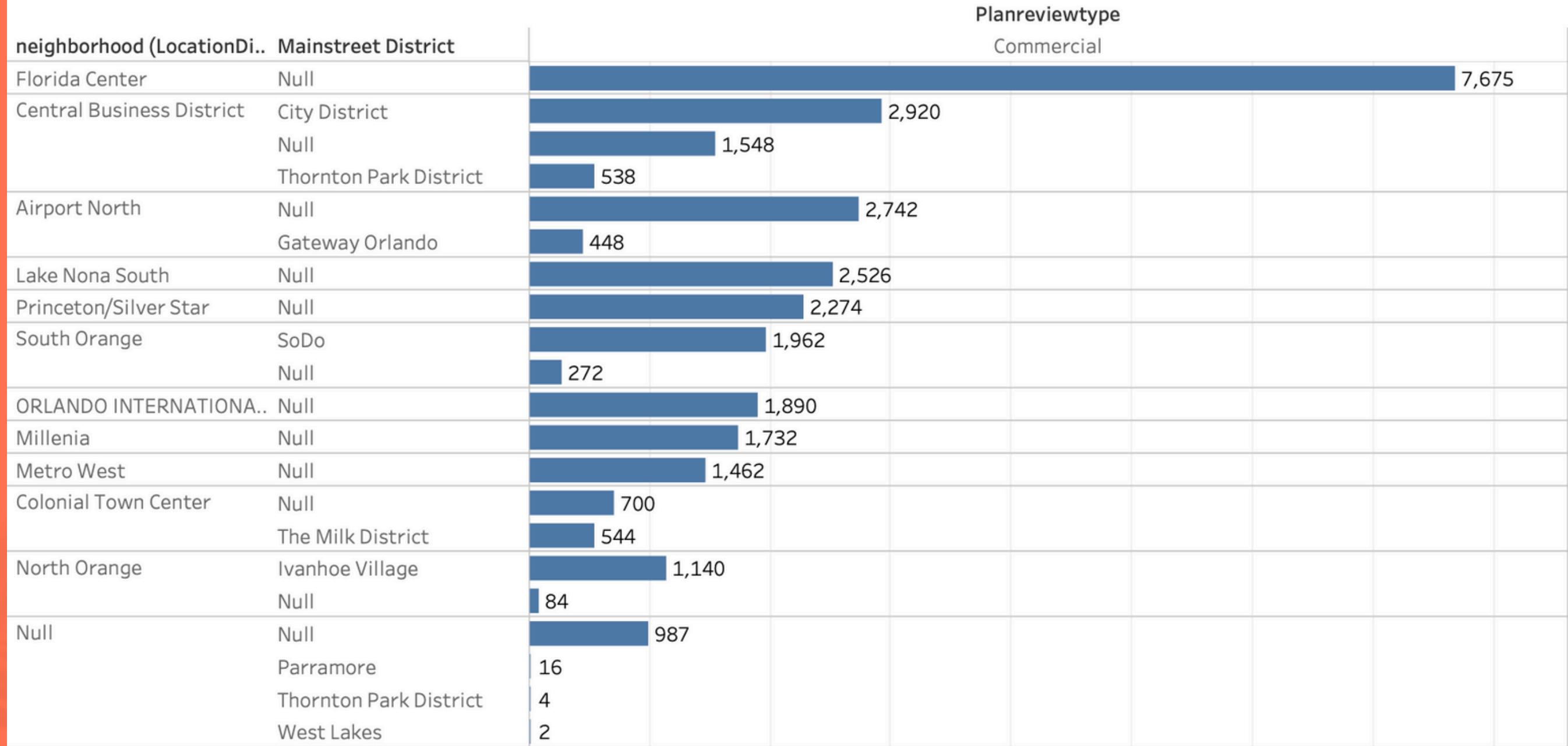
	yearmonth character varying (20) 	median_sale_price numeric 	num_homes_sold integer 	homes_for_sale integer 	totalpermits bigint 	totalplanningpermits bigint 
1	2022-01	397900	952	2818	14	0
2	2022-02	397500	1012	2536	324	8
3	2022-03	412500	1201	2513	318	5
4	2022-04	429900	1111	2765	364	7
5	2022-05	450000	1147	3684	0	0
6	2022-06	465000	1022	4879	318	7
7	2022-07	464000	901	6806	440	7
8	2022-08	453168	981	7351	320	9
9	2022-09	449411	770	7542	356	11
10	2022-10	445000	783	7392	0	1

Visualizations

Visualization | Question 1

Which neighborhoods in Orlando saw the most applications for commercial development in the past year?

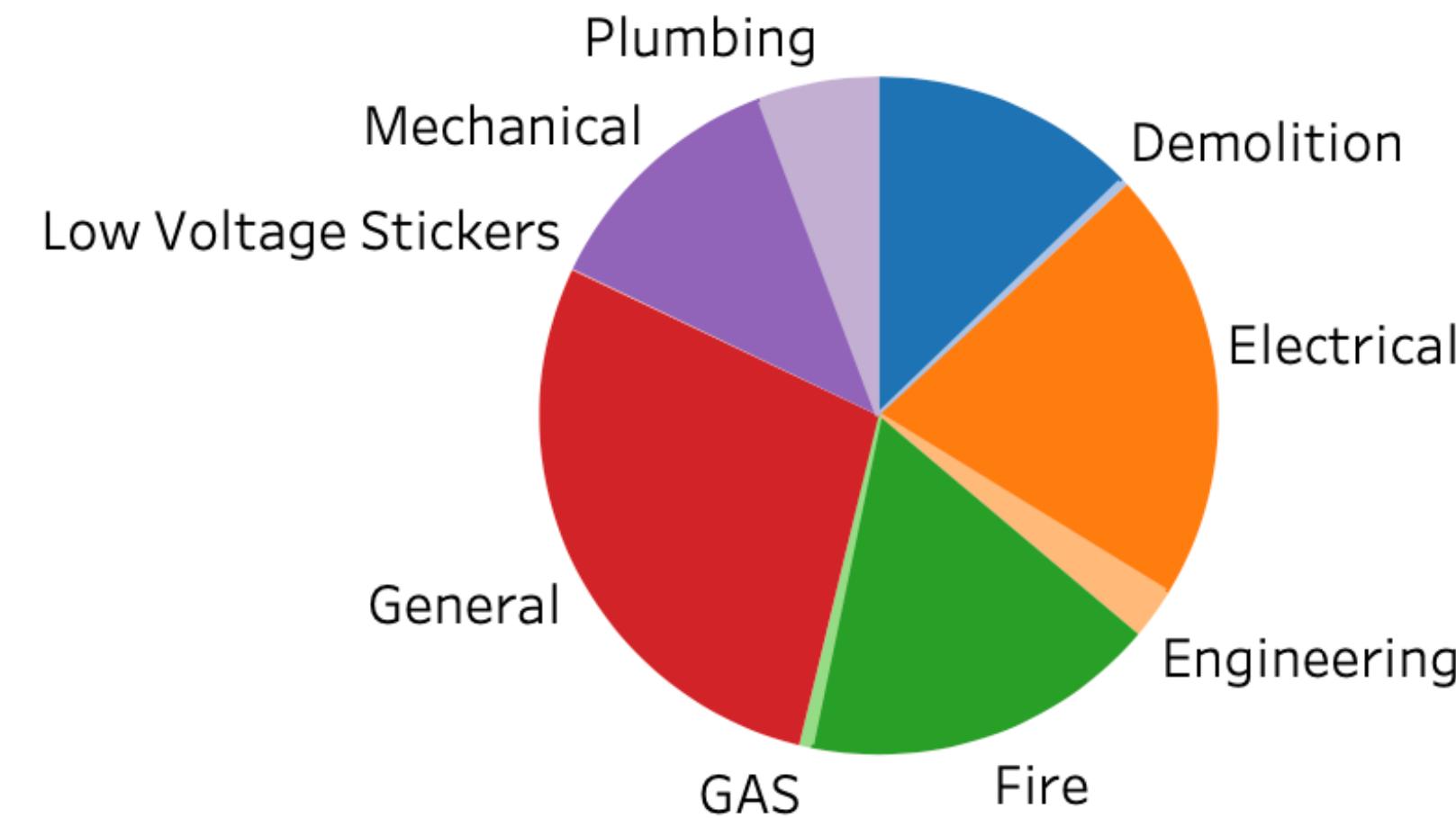
Most Commercial Permits by Neighborhood



Visualization | Question 3

Which contractor types were most hired in Orlando in the last year?

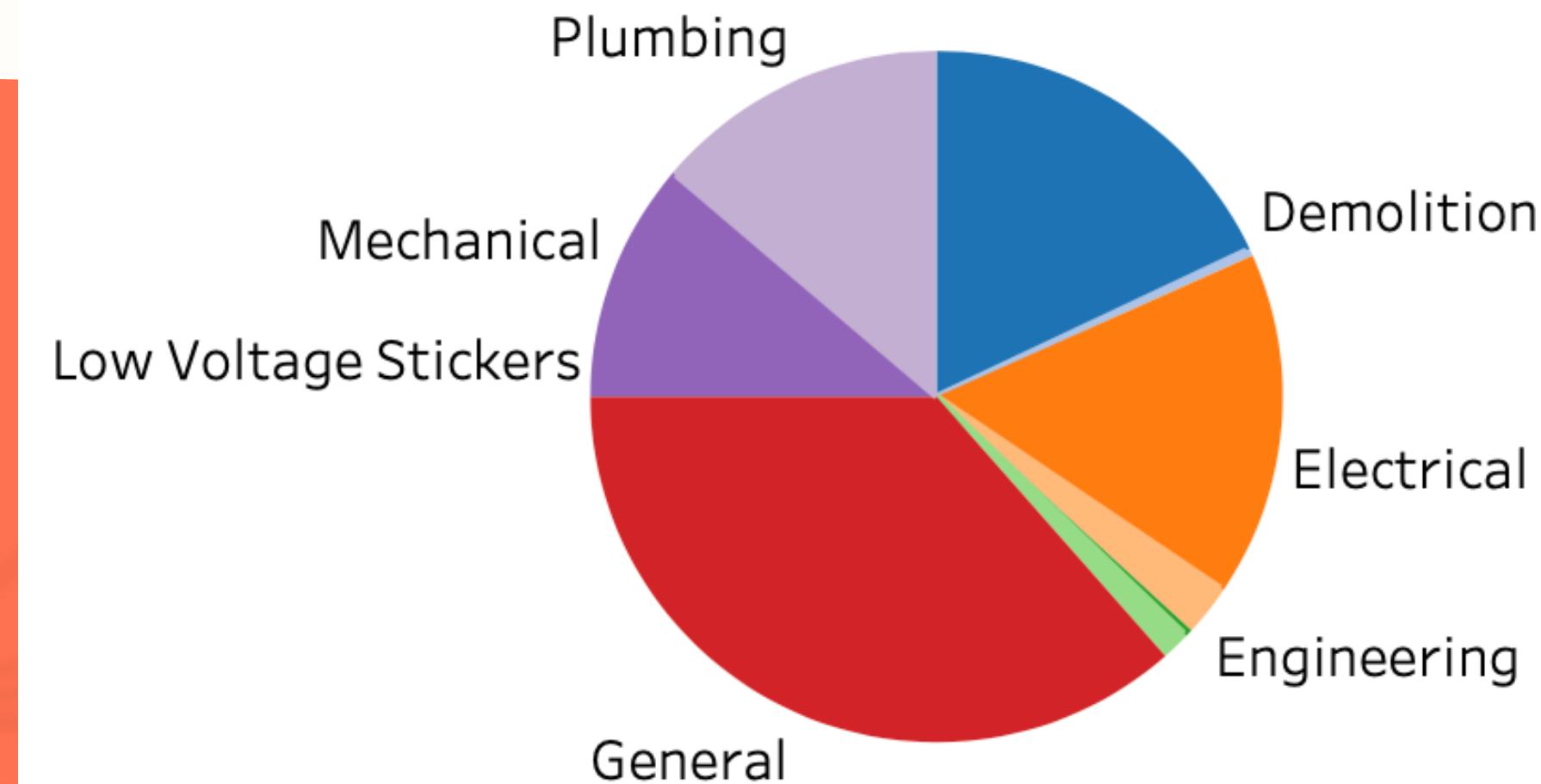
Commercial



Visualization | Question 3

Which contractor types were most hired in Orlando in the last year?

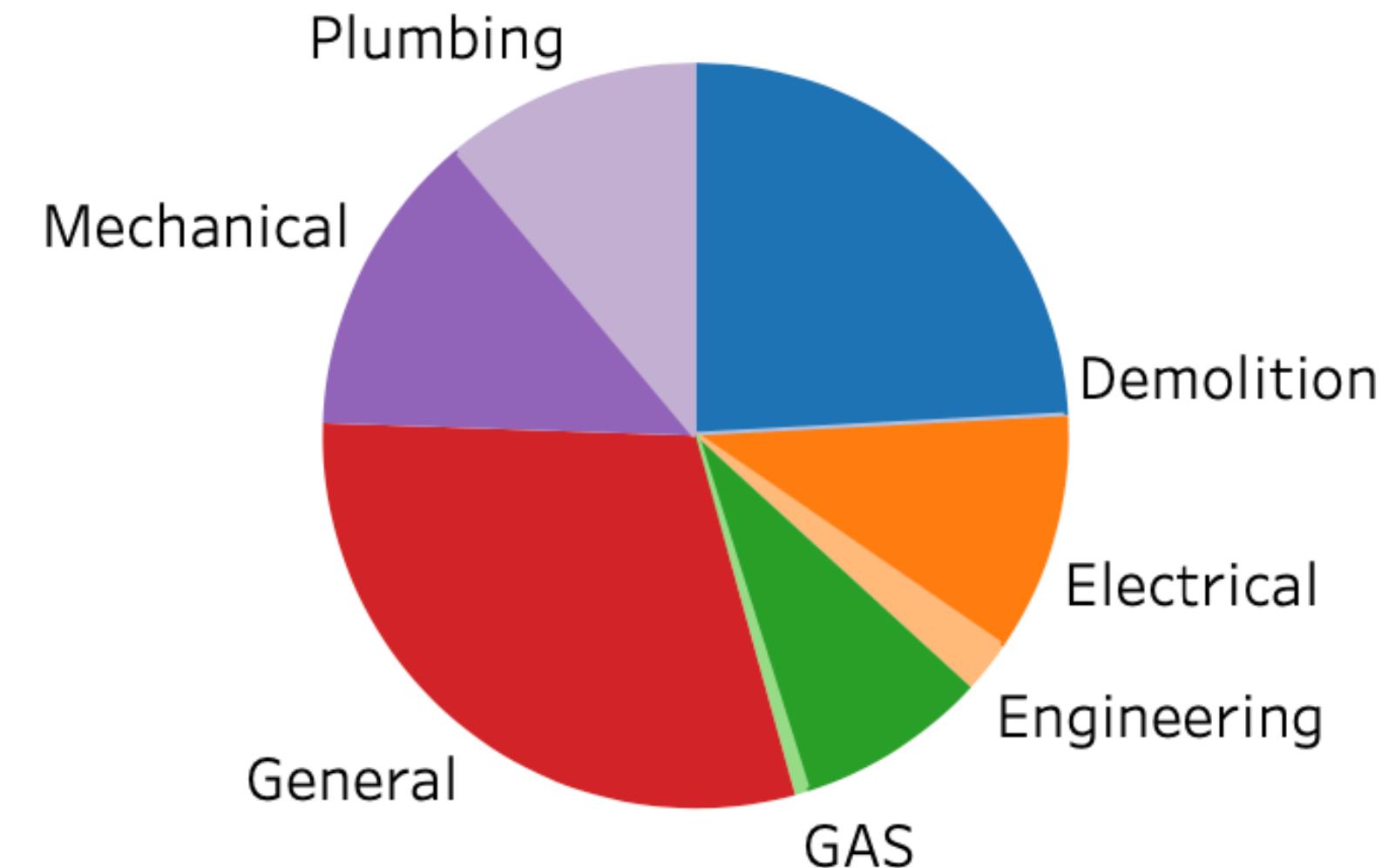
Residential 1/2



Visualization | Question 3

Which contractor types were most hired in Orlando in the last year?

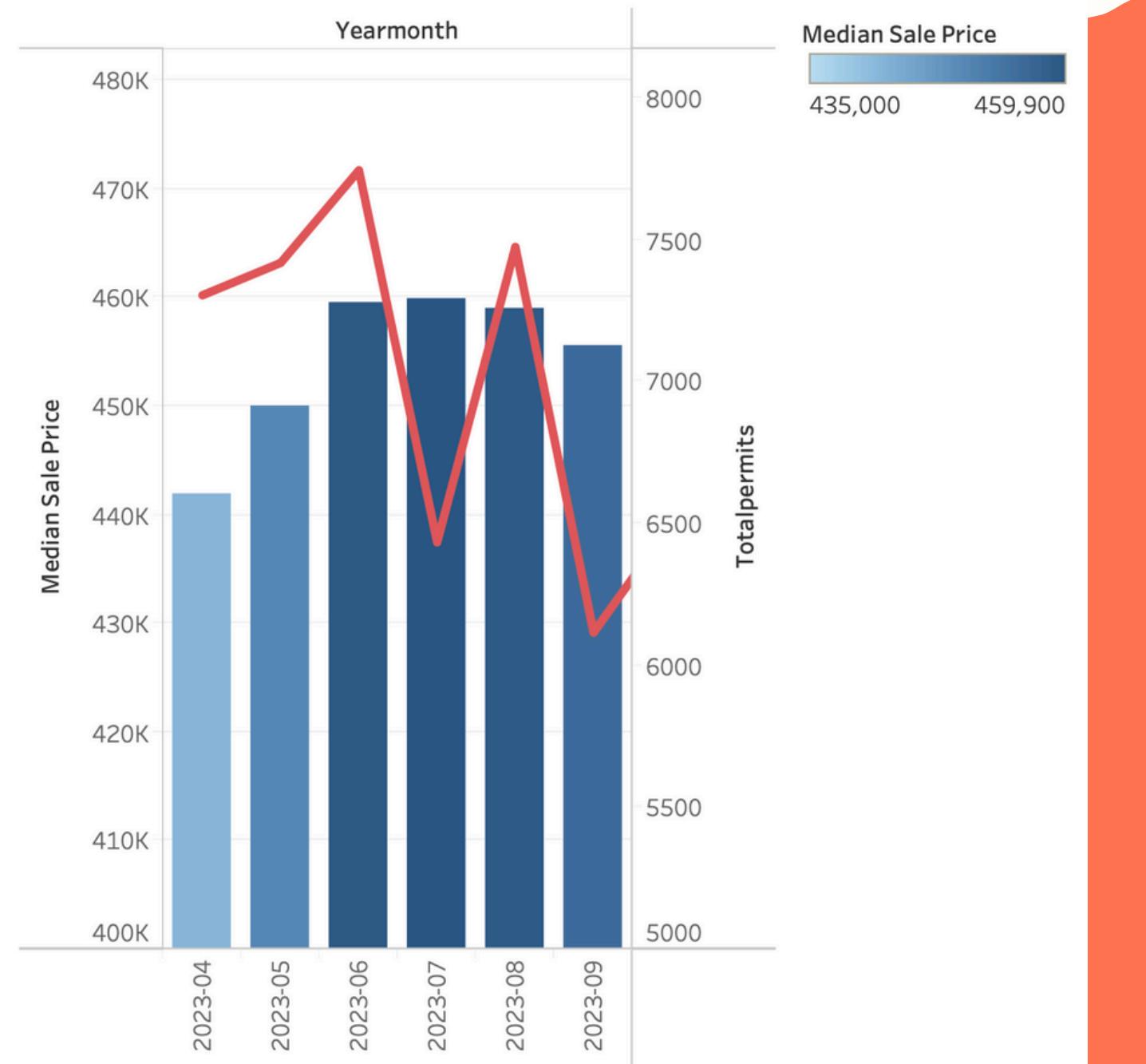
Residential 3+



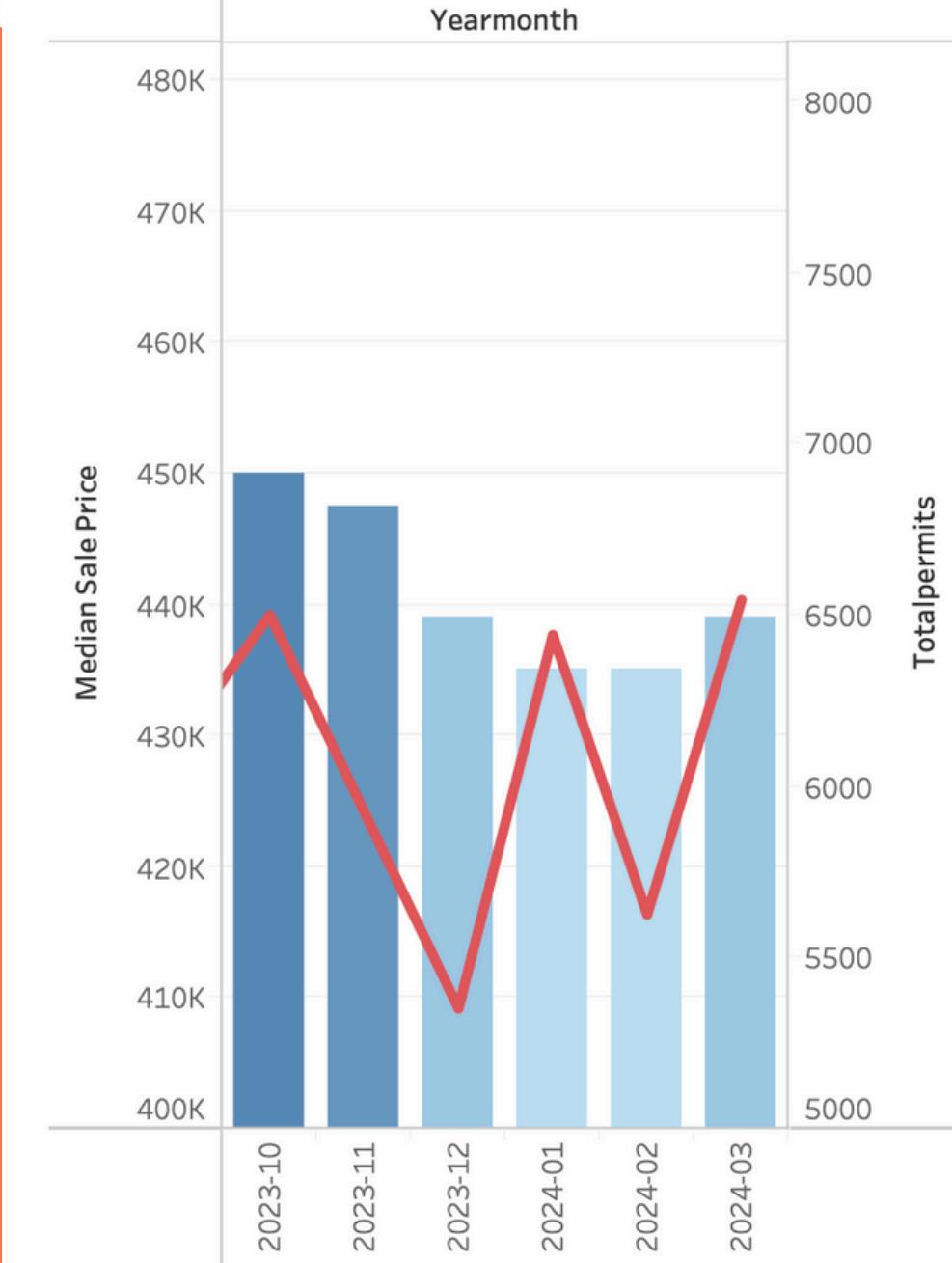
Visualization | Question 4

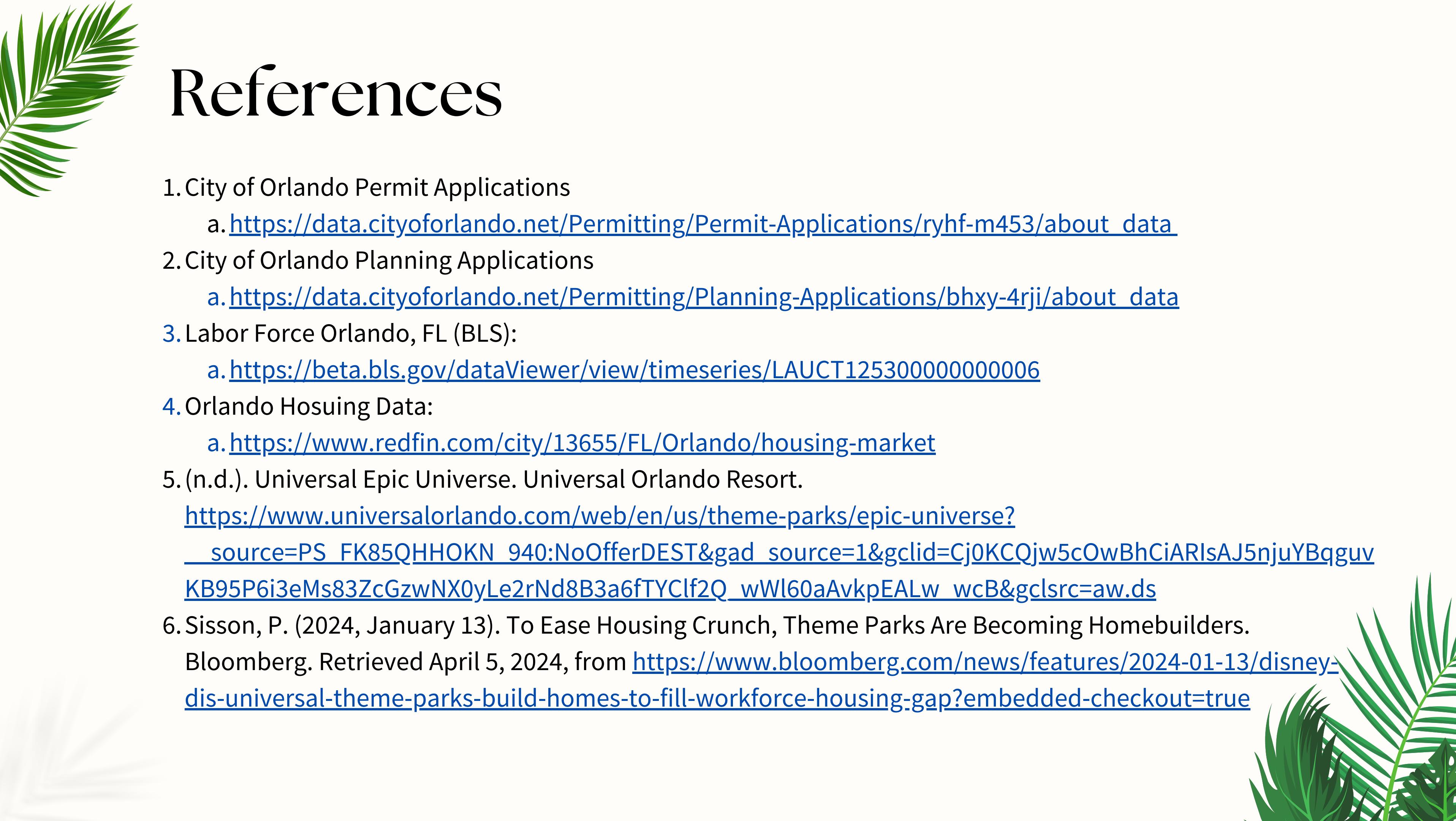
How has both residential and commercial development in the Orlando area impacted home prices, sales, and inventory?

Orlando Area Median Housing Prices / Permit Applications



Orlando Area Median Housing Prices / Permit Applications





References

1. City of Orlando Permit Applications
 - a. https://data.cityoforlando.net/Permitting/Permit-Applications/ryhf-m453/about_data
2. City of Orlando Planning Applications
 - a. https://data.cityoforlando.net/Permitting/Planning-Applications/bhxy-4rji/about_data
3. Labor Force Orlando, FL (BLS):
 - a. <https://beta.bls.gov/dataViewer/view/timeseries/LAUCT125300000000006>
4. Orlando Housing Data:
 - a. <https://www.redfin.com/city/13655/FL/Orlando/housing-market>
5. (n.d.). Universal Epic Universe. Universal Orlando Resort.
https://www.universalorlando.com/web/en/us/theme-parks/epic-universe?source=PS_FK85QHHOKN_940>NoOfferDEST&gad_source=1&gclid=Cj0KCQjw5cOwBhCiARIsAJ5njuYBqguvKB95P6i3eMs83ZcGzwNX0yLe2rNd8B3a6fTYClf2Q_wWl60aAvkpEALw_wcB&gclsrc=aw.ds
6. Sisson, P. (2024, January 13). To Ease Housing Crunch, Theme Parks Are Becoming Homebuilders. Bloomberg. Retrieved April 5, 2024, from <https://www.bloomberg.com/news/features/2024-01-13/disney-dis-universal-theme-parks-build-homes-to-fill-workforce-housing-gap?embedded-checkout=true>

Thank You!

