

Build a database to streamline Dream Homes NYC's operations and elevate the client experience

Final Project Presentation

Yuxuan Chen (Iris)

Yishan Liu

Xianghui Meng

Haoxiang Huo(Steven)

yc4637@columbia.edu

y15897@columbia.edu

xm2346@columbia.edu

hh3109@columbia.edu



Introduction-

“Our team will build Dream Homes NYC a centralized database that transforms unstructured records into our organized schema via an ETL process. The application on analytics and creation of Metabase dashboards enables faster, smarter decision-making, and gives agents and customers instant insights into listings, client behavior, and market trends.”

Task Division

Category	Task	Assigned Member(s)	Planned DDL
Business Understanding	Define detailed business requirements and client goals	Yuxuan Chen	Week 8
Database Schema	Design a normalized ERD (20+ tables in 3NF), including entities and relationships	Yishan Liu	Week 8
Data Simulation	Generate and clean sample data to simulate messy or semi-structured input	Steven Huo	Week 8
Table Creation	Write SQL DDL scripts to define all tables and constraints in PostgreSQL	Xianghui Meng	Week 8
Data Loading	Create Python/SQL scripts to load data into the database and write code for analytical procedure	Xianghui Meng	Week 9
Complex Queries	Write 8 complex SQL queries to support business decisions	Yuxuan Chen	Week 10
Dashboards	Design and build interactive dashboards using Metabase or Power BI	Steven Huo	Week 11
Customer Interaction Plan	Write Customer Interaction Plan	Yishan Liu	Weak 11
Final Report	Co-author project report, including intro, schema design, queries, and business insights	All Members	Week 12
Final Presentation	Prepare slides and live demo walkthrough; each member presents part of the work	All Members	Week 12

01.

Business Insights:



1. Centralized Office Profiles
2. Comprehensive Employee Directory
3. Detailed Client Records
4. Full Property Listings
5. Appointment Management
6. End-to-End Transaction Tracking
7. Normalized, Relational Schema
8. Automated ETL Pipeline
9. Data-Driven Insights
10. Live Dashboards & Analytics



Database Design Summary

Schema follows third normal form (3NF) with enforced referential integrity.

Key relationships:

- Office (1) → Employee (M)
- Client (1) → ClientRole (M)
- Property (1) → Feature (M), Media (M)
- Transaction (1) → Commission (1) → Lease (1) → PaymentRecord (M)
- Appointment connects agent, client, and property in time-stamped interactions.
- Documents and MarketingCampaigns link to properties or transactions.
- Natural keys (e.g., email, mls_number, transaction_code) support safe, deduplicated upserts.
- Design supports complex workflows across both sales and rentals

ETL PROCESS SUMMARY

- EnhancedDreamHomesETL loads dream_homes_nyc_dataset_v8.csv into dream_homes_db on PostgreSQL.
- Semi-structured CSV records are converted to typed rows using stable natural keys for safe re-runs.

Extraction:

- CSV read into pandas with UTF-8; empty cells → NaN; no line dropped.
- Original tokens preserved for typed converters in later stages.

Cleaning & Validation:

- Safe type converters (safe_int, safe_decimal, etc.) return NULL and log warnings on failure.
- Missing monetary fields skip Transaction insert but still load related Property.

Transformation:

- Normalize composite fields (e.g., addresses, bed/bath, client blurbs).
- Map categories via enum mappers to prevent spelling drift.
- Expand lists into child rows (e.g., features, appointments, documents).
- Fallback logic: missing offer_date → use listing_date or today; similar for offer_amount.



DATA SIMULATION

A	B	C	D	E	F	G	H	I
transaction_id	mls_listing_number	property_address_full	listing_office_name	listing_office_address	listing_office_phone	property_type	status_current	transaction_
IN-2025-001	MLS8202793	536 Atlantic Ave Brooklyn NY 11211	Dream Homes Brooklyn	345 Court St Brooklyn NY 11231	(718) 555-0400	Co-op	SOLD	sale
IN-2025-002	MLS8408828	746 Washington St Jersey City NJ 07204	Dream Homes NJ	890 Bergen Ave Jersey City NJ 07306	(201) 555-0300	TH	SOLD	sale
IN-2025-003	MLS8569213	512 Washington St Jersey City NJ 07200	Dream Homes NJ	890 Bergen Ave Jersey City NJ 07306	(201) 555-0300	TH	RENTED	rental
IN-2025-004	MLS8102422	285 Smith St Brooklyn NY 11213	Dream Homes Brooklyn	345 Court St Brooklyn NY 11231	(718) 555-0400	Condo	SOLD	sale
IN-2025-005	MLS8561588	914 Court St Brooklyn NY 11230	Dream Homes Brooklyn	345 Court St Brooklyn NY 11231	(718) 555-0400	TH	RENTED	rental
IN-2025-006	MLS8532881	320 Greenwich Ave Fairfield CT 06802	Dream Homes CT	1200 Post Rd Fairfield CT 06824	(203) 555-0500	SFH	ACTIVE	sale
IN-2025-007	MLS8946341	700 River Rd Jersey City NJ 07034	Dream Homes NJ	890 Bergen Ave Jersey City NJ 07306	(201) 555-0300	TH	ACTIVE	sale
IN-2025-008	MLS8457049	951 Washington St Newark NJ 07257	Dream Homes NJ	890 Bergen Ave Jersey City NJ 07306	(201) 555-0300	Condo	ACTIVE	sale
IN-2025-009	MLS8915729	162 East 86th St Unit 7B New York NY 10032	Dream Homes Manhattan	120 Broadway Suite 2701 New York NY 10271	(212) 555-0100	TH	RENTED	rental
IN-2025-010	MLS8447640	773 Monroe St Newark NJ 07220	Dream Homes NJ	890 Bergen Ave Jersey City NJ 07306	(201) 555-0300	Condo	RENTED	rental
IN-2025-011	MLS8208952	687 River Rd Jersey City NJ 07276	Dream Homes NJ	890 Bergen Ave Jersey City NJ 07306	(201) 555-0300	TH	RENTED	rental
IN-2025-012	MLS8908181	301 Greenwich Ave Fairfield CT 06822	Dream Homes CT	1200 Post Rd Fairfield CT 06824	(203) 555-0500	TH	PENDING	sale
IN-2025-013	MLS8750840	416 West End Ave Unit 27B New York NY 10001	Dream Homes Manhattan	120 Broadway Suite 2701 New York NY 10271	(212) 555-0100	Co-op	PENDING	sale
IN-2025-014	MLS8399416	288 Atlantic Ave Brooklyn NY 11201	Dream Homes Brooklyn	345 Court St Brooklyn NY 11231	(718) 555-0400	TH	PENDING	sale
IN-2025-015	MLS8587052	520 Monroe St Edgewater NJ 07282	Dream Homes NJ	890 Bergen Ave Jersey City NJ 07306	(201) 555-0300	SFH	PENDING	sale
IN-2025-016	MLS8553789	446 Park Ave Unit 16A New York NY 10070	Dream Homes Manhattan	120 Broadway Suite 2701 New York NY 10271	(212) 555-0100	TH	PENDING	sale
IN-2025-017	MLS8297489	949 West End Ave Unit 10D New York NY 10110	Dream Homes Manhattan	120 Broadway Suite 2701 New York NY 10271	(212) 555-0100	Condo	ACTIVE	sale
IN-2025-018	MLS8483697	999 Post Rd Stamford CT 06864	Dream Homes CT	1200 Post Rd Fairfield CT 06824	(203) 555-0500	TH	ACTIVE	sale
IN-2025-019	MLS8025874	241 Fifth Ave Unit 10A New York NY 10027	Dream Homes Manhattan	120 Broadway Suite 2701 New York NY 10271	(212) 555-0100	Co-op	PENDING	sale
IN-2025-020	MLS8316323	835 Park Ave Unit 26B New York NY 10118	Dream Homes Manhattan	120 Broadway Suite 2701 New York NY 10271	(212) 555-0100	TH	SOLD	sale
IN-2025-021	MLS8058697	841 East 86th St Unit 28D New York NY 10097	Dream Homes Manhattan	120 Broadway Suite 2701 New York NY 10271	(212) 555-0100	SFH	ACTIVE	sale
IN-2025-022	MLS8353052	256 Post Rd Stamford CT 06902	Dream Homes CT	1200 Post Rd Fairfield CT 06824	(203) 555-0500	TH	SOLD	sale
IN-2025-023	MLS8379831	67 Broadway Unit 14C New York NY 10030	Dream Homes Manhattan	120 Broadway Suite 2701 New York NY 10271	(212) 555-0100	Co-op	ACTIVE	sale
IN-2025-024	MLS8239314	914 East 86th St Unit 20D New York NY 10032	Dream Homes Manhattan	120 Broadway Suite 2701 New York NY 10271	(212) 555-0100	Condo	SOLD	sale
IN-2025-025	MLS8320229	803 Bedford Ave Brooklyn NY 11205	Dream Homes Brooklyn	345 Court St Brooklyn NY 11231	(718) 555-0400	Co-op	ACTIVE	sale
IN-2025-026	MLS8847954	813 Monroe St Newark NJ 07192	Dream Homes NJ	890 Bergen Ave Jersey City NJ 07306	(201) 555-0300	Co-op	ACTIVE	sale
IN-2025-027	MLS8657500	571 Washington St Edgewater NJ 07254	Dream Homes NJ	890 Bergen Ave Jersey City NJ 07306	(201) 555-0300	Co-op	SOLD	sale
IN-2025-028	MLS8099770	912 Bergen Ave Jersey City NJ 07177	Dream Homes NJ	890 Bergen Ave Jersey City NJ 07306	(201) 555-0300	TH	PENDING	sale
IN-2025-029	MLS8383277	255 Newark Ave Edgewater NJ 07089	Dream Homes NJ	890 Bergen Ave Jersey City NJ 07306	(201) 555-0300	Co-op	RENTED	rental
IN-2025-030	MLS8684078	166 Bergen Ave Edgewater NJ 07271	Dream Homes NJ	890 Bergen Ave Jersey City NJ 07306	(201) 555-0300	Co-op	SOLD	sale
IN-2025-031	MLS8449539	642 Madison Ave Unit 2A New York NY 10004	Dream Homes Manhattan	120 Broadway Suite 2701 New York NY 10271	(212) 555-0100	Condo	RENTED	rental
IN-2025-032	MLS8505510	478 Bedford Ave Brooklyn NY 11201	Dream Homes Brooklyn	345 Court St Brooklyn NY 11231	(718) 555-0400	SFH	RENTED	rental
IN-2025-033	MLS8513172	311 Grand St Brooklyn NY 11217	Dream Homes Brooklyn	345 Court St Brooklyn NY 11231	(718) 555-0400	TH	SOLD	sale
IN-2025-034	MLS8878140	210 Park Ave Unit 12D New York NY 10118	Dream Homes Manhattan	120 Broadway Suite 2701 New York NY 10271	(212) 555-0100	Co-op	RENTED	rental
IN-2025-035	MLS8481295	250 Washington St Newark NJ 07144	Dream Homes NJ	890 Bergen Ave Jersey City NJ 07306	(201) 555-0300	TH	SOLD	sale
IN-2025-036	MLS8591990	838 Newark Ave Jersey City NJ 07105	Dream Homes NJ	890 Bergen Ave Jersey City NJ 07306	(201) 555-0300	Co-op	SOLD	sale
IN-2025-037	MLS8033667	138 Main St Greenwich CT 06870	Dream Homes CT	1200 Post Rd Fairfield CT 06824	(203) 555-0500	Co-op	SOLD	sale
IN-2025-038	MLS8079874	817 Smith St Brooklyn NY 11212	Dream Homes Brooklyn	345 Court St Brooklyn NY 11231	(718) 555-0400	TH	PENDING	sale
IN-2025-039	MLS8767776	755 Smith St Brooklyn NY 11211	Dream Homes Brooklyn	345 Court St Brooklyn NY 11231	(718) 555-0400	TH	SOLD	sale
IN-2025-040	MLS8650041	284 Washington St Edgewater NJ 07122	Dream Homes NJ	890 Bergen Ave Jersey City NJ 07306	(201) 555-0300	TH	SOLD	sale
IN-2025-041	MLS8670000	205 Rte 17 N Unit 200 New York NY 10115	Dream Homes NY	120 Broadway Suite 2701 New York NY 10271	(212) 555-0100	TH	PENDING	sale

Demo of Interactive Dashboard

Haoxiang Huo's Personal Collection

5310 group6 project NY Homes - Dashboard series

Property Features by Price Range (What features and amenities can I expect at different price points?)

feature_name	\$300K-500K	\$500K-750K	Over \$750K
renovated	2	1	4
parking	1	2	2
laundry	1	2	3
gym	0	1	8
storage	1	2	2
elevator	2	1	3
views	1	2	2
doorman	0	1	5
concierge	0	1	2
terrace	0	0	6
garden	0	0	1

Marketing Campaign ROI Analysis (Which marketing strategies work best ...)

campaign_type	success_rate	total_campaigns	properties_sold	avg_budget
online	70	10	10	1000
print	55	10	10	1000
open_house	25	10	10	1000
social_media	15	10	10	1000
email	10	10	10	1000
direct_mail	10	10	10	1000

Top Performing Agents Overview (Who are the best rental agents in the area?)

agent_name	properties_managed
Matthew Anderson	3
Karen Johnson	1
Jennifer Martin	1
Betty Thomas	1
Helen Martinez	1
Maria Jackson	1
Mark Moore	1
Robert Thomas	1
Daniel Harris	1
Daniel Garcia	1

Seasonal Sell Trends (When is the best time of year to sell client house?)

Property Value by Size Analysis (What house size gives me the best value for money?)

pgAdmin 4

Object Explorer Dashboard Dependencies ans_psq1.sql tri_psq1.sql dream_homes_db/... final_q.sql

Servers (1) Docker Postgres Databases (11) dream_homes_db

```
-- Dream Homes Real Estate Database - 12 Essential Client Queries
-- Question 1: What features and amenities can I expect at different price points?
WITH price_ranges AS (
  SELECT property_id, list_price,
  CASE
    WHEN list_price < 300000 THEN 'Under $300K'
    WHEN list_price < 500000 THEN '$300K-$500K'
    WHEN list_price < 750000 THEN '$500K-$750K'
    ELSE 'Over $750K'
  END as price_range
  FROM Property
  WHERE current_status IN ('active', 'sold') AND list_price > 0
)
SELECT pr.price_range,
  pf.feature_type,
  pf.feature_name,
  COUNT(*) as feature_count,
  ROUND(COUNT(*) * 100.0 / SUM(COUNT(*)) OVER (PARTITION BY pr.property_id), 2) as percentage_in_range
FROM price_ranges pr
JOIN PropertyFeature pf ON pr.property_id = pf.property_id
```

Data Output Messages Notifications

Showing rows: 1 to 9 Page No: 1 of 1

price_range	feature_type	feature_name	feature_count	percentage_in_range
\$500K-750K	amenity	parking	6	37.50
\$500K-750K	amenity	gym	5	31.25
\$500K-750K	amenity	laundry	5	31.25
Over \$750K	amenity	gym	9	24.32
Over \$750K	amenity	terrace	7	18.92
Over \$750K	amenity	doorman	6	16.67

Successfully run. Total query runtime: 134 msec. 9 rows affected.

- Property Features by Price Range:** Parking, laundry, and gym dominate mid-tier homes, luxury tiers feature terraces and doormen.
- Marketing ROI Analysis:** Email campaigns show highest budget and sales volume, online/social media yield better success rates.
- Top Performing Agents:** Matthew Anderson leads in properties managed, followed by Karen Johnson and Jennifer Martin.
- Seasonal Sell Trends:** Market peaks in spring and early summer.

Analytics Applications

Data-Driven Real Estate Decision Solutions



Dream Homes NYC transforms data into productivity through intelligent coding: Clients drag sliders on interactive heatmaps while backend executes `'price_range_analysis()'` to instantly recommend budget-matched listings, landlords input amenities to trigger `'rental_premium_calculator(gym, doorman)'` for precise rental yield forecasts. Marketing teams optimize ad spend using auto-generated ROI dashboards, while sales meetings dynamically allocate leads via real-time conversion funnels. Data evolves from static reports to an intelligent operational engine - powering every business process from client engagement to transaction closure.

Conclusion

We transformed Dream Homes NYC's fragmented data into a scalable cloud-based warehouse with automated ETL, enabling real-time decision-making across all business functions while future-proofing for AI and IoT integrations.

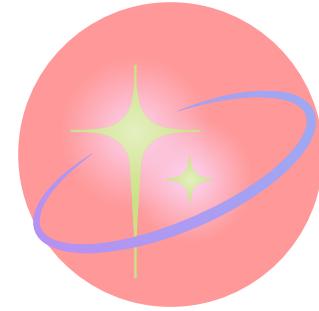
Centralized Data Foundation

1. Consolidated scattered documents into a normalized relational database
2. Implemented foreign keys/constraints for 100% data integrity



Self-Sustaining Automation

1. Python ETL pipelines with auto-rollback for nightly updates
2. Guaranteed report availability regardless of errors



Role-Specific Intelligence

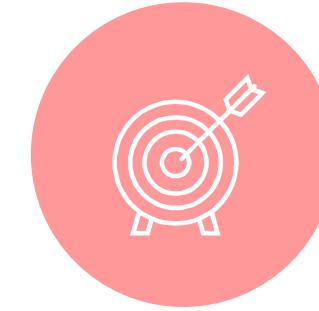
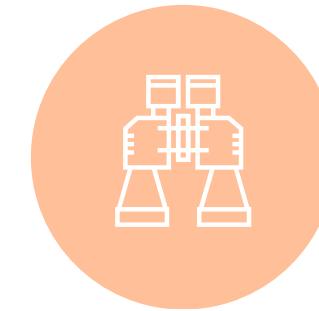
- 12 predefined queries answering:
1. Brokers' comps checks (2-sec mobile access)
 2. Marketing's channel ROI tracking
 3. Executive dashboards (revenue/ occupancy/ KPIs)



Conclusion

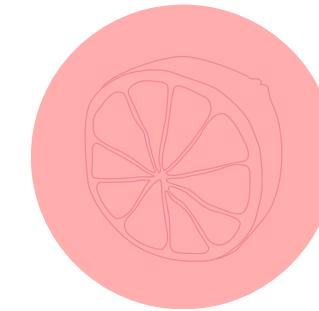
Decision Velocity

1. Reduced analysis cycles from weeks → hours
2. Live transaction pricing during property viewings.



Cloud-Native Scalability

1. Handles 10x query concurrency via horizontal scaling
2. Zero business code changes for infrastructure growth



Future-Proof Architecture

- Ready for:
 1. Smart home IoT ingestion
 2. ML valuation models
 3. Self-service customer portals.

Thanks for watching!