LA Building Permit Insights Dashboard: Final Report

Animal Name for Interactive Dashboard: Jaguar

Course: Algorithmic thinking with Python

Project: Data-Driven Storytelling with LA Building Permit Data

Tools Used: Snowflake Snowpark, Streamlit, Python

Executive Summary

This project analyzes Los Angeles building permit data to understand how home improvement activity varies across neighborhoods, income groups, and construction types.
Using a multi-tab Streamlit dashboard, I visualized trends in:

- Swimming pool permits (a proxy for luxury investment)
- Kitchen renovations (a proxy for post-pandemic home improvement)
- Contractor activity (market concentration in renovation work)
- Income inequality (permit share vs household share)
- Overall construction trends (citywide context)

Together, these visualizations reveal a consistent story:

wealthier neighborhoods invest far more in home upgrades, especially luxury renovations, and these investments have accelerated in the past decade.

This report includes five required visualizations, each with code and analytical commentary.

TABLE OF CONTENTS

CODE USED :

```
# ------------------------------
# IMPORTS
# ------------------------------
import pandas as pd
import numpy as np
import streamlit as st
from snowflake.snowpark.context import get_active_session

session = get_active_session()

# ------------------------------
# LOAD TABLES
# ------------------------------
session.sql("USE ROLE TRAINING_ROLE").collect()
session.sql("USE DATABASE LA_PERMIT_DATA").collect()
session.sql("USE SCHEMA PUBLIC").collect()

records       = session.table("PERMIT_RECORDS").to_pandas()
census_tracts  = session.table("CENSUS_TRACTS").to_pandas()
contractors    = session.table("MASTER_LICENSE").to_pandas()

# ------------------------------
# CLEANING
# ------------------------------
for col in ["ISSUE_DATE", "STATUS_DATE"]:
    records[col] = pd.to_datetime(records[col], errors="coerce")

records["VALUATION"] = pd.to_numeric(

records["VALUATION"].astype(str).str.replace("$","",regex=False).str.replace(",","",regex=False),
    errors="coerce"
)

ct_num = pd.to_numeric(records["CENSUS_TRACT"], errors="coerce")
records["CT"] = (ct_num * 100) + 6037000000

records_with_census = records.merge(
    census_tracts,
    left_on="CT",
    right_on="CENSUS_TRACT",
    how="left"
)
```

```
# --------------------------------------------------------
# DASHBOARD UI
# --------------------------------------------------------
st.title(" LA Building Permit Insights Dashboard")

# --------------------------------------------------------
# GLOBAL FILTERS
# --------------------------------------------------------
st.sidebar.header("GLOBAL FILTERS (Apply to All Tabs)")

min_date = records["ISSUE_DATE"].min().date()
max_date = records["ISSUE_DATE"].max().date()

global_date = st.sidebar.date_input(
    "Date Range", (min_date, max_date), min_value=min_date, max_value=max_date
)

global_zip = st.sidebar.multiselect(
    "ZIP Filter",
    sorted(records["ZIP_CODE"].dropna().unique().tolist()),
)

global_types = st.sidebar.multiselect(
    "Permit Types",
    sorted(records["PERMIT_TYPE"].dropna().unique().tolist()),
)

# Apply global filters
filtered_global = records.copy()

# date
if global_date:
    start, end = global_date
    filtered_global = filtered_global[filtered_global["ISSUE_DATE"].dt.date.between(start, end)]

# zip
if global_zip:
    filtered_global = filtered_global[filtered_global["ZIP_CODE"].isin(global_zip)]

# permit type
if global_types:
    filtered_global = filtered_global[filtered_global["PERMIT_TYPE"].isin(global_types)]

# --------------------------------------------------------
```

```
# TABS
# --------------------------------------------------------
tab1, tab2, tab3, tab4, tab5 = st.tabs([
    " Pools",
    " Kitchens",
    " Contractors",
    " Income Inequality",
    " All Permit Trends"
])

# --------------------------------------------------------
# TAB 1 — POOLS
# --------------------------------------------------------
with tab1:
    st.header(" Swimming Pool Permits")

    # Per-graph filters
    st.subheader("Filters for Pool Visualizations")

    pool_year = st.slider(
        "Select Year Range",
        int(filtered_global["ISSUE_DATE"].dt.year.min()),
        int(filtered_global["ISSUE_DATE"].dt.year.max()),
        (2010, 2024)
    )

    pool_zip = st.multiselect(
        "ZIP (Pool Tab Only)",
        sorted(filtered_global["ZIP_CODE"].dropna().unique().tolist())
    )

    keyword = st.text_input("Keyword filter in description", "swimming pool")

    # Apply local filters
    df = filtered_global.copy()
    df = df[df["AI_DESCRIPTION"].fillna("").str.contains(keyword, case=False)]
    df = df[df["ISSUE_DATE"].dt.year.between(pool_year[0], pool_year[1])]

    if pool_zip:
        df = df[df["ZIP_CODE"].isin(pool_zip)]

    st.metric("Total Pool Permits (After Filters)", len(df))

    # Graph 1: Pool Permits by Year
```

```
    df["year"] = df["ISSUE_DATE"].dt.year
    st.subheader("Pool Permits by Year")
    st.bar_chart(df.groupby("year")["PERMIT_TYPE"].count())

    # Graph 2: Top ZIPs
    st.subheader("Top ZIPs for Pool Permits")
    top_zip =
df.groupby("ZIP_CODE")["PERMIT_TYPE"].count().sort_values(ascending=False).head(10)
    st.bar_chart(top_zip)

    # Graph 3: Map
    st.subheader("Pool Permit Locations")
    coords = df["LATITUDE_LONGITUDE"].dropna().str.extract(r"\((([-\d.]+) ([-
\d.]+)\)").astype(float)
    coords.columns = ["longitude","latitude"]
    st.map(coords)

# ---------------------------------------------------------
# TAB 2 — KITCHEN REMODELS
# ---------------------------------------------------------
with tab2:
    st.header(" Kitchen Renovations")

    # Per-graph filters
    st.subheader("Filters for Kitchen Visualizations")

    kitchen_year = st.slider(
        "Kitchen Year Range",
        2010,
        2024,
        (2012, 2024)
    )

    min_val, max_val = int(records["VALUATION"].min()), int(records["VALUATION"].max())

    kitchen_val = st.slider(
        "Valuation Range",
        min_val, max_val,
        (min_val, max_val)
    )

    # Apply filters
    dfk = filtered_global.copy()
    dfk = dfk[dfk["AI_DESCRIPTION"].fillna("").str.contains("kitchen", case=False)]
```

```
    dfk = dfk[dfk["ISSUE_DATE"].dt.year.between(kitchen_year[0], kitchen_year[1])]
    dfk = dfk[dfk["VALUATION"].between(kitchen_val[0], kitchen_val[1])]

    st.metric("Total Kitchen Permits", len(dfk))

    # Graph 1
    dfk["year"] = dfk["ISSUE_DATE"].dt.year
    st.subheader("Kitchen Permits Over Time")
    st.line_chart(dfk.groupby("year")["PERMIT_TYPE"].count())

    # Graph 2
    st.subheader("Average Kitchen Valuation")
    st.line_chart(dfk.groupby("year")["VALUATION"].mean())

# ---------------------------------------------------------
# TAB 3 — CONTRACTORS
# ---------------------------------------------------------
with tab3:
    st.header(" Contractor Performance Dashboard")

    # NEW FILTER: Choose contractor
    contractor_list = sorted(contractors["LICENSE_NO"].dropna().unique().tolist())

    selected_cons = st.multiselect(
        "Select Contractors",
        contractor_list
    )

    dfc = filtered_global.copy()

    if selected_cons:
        dfc = dfc[dfc["LICENSE_NUM"].isin(selected_cons)]

    st.metric("Permits by Selected Contractors", len(dfc))

    contractor_stats = (
        dfc.groupby("LICENSE_NUM")
        .agg(
            num_projects=("PERMIT_SUB_TYPE","count"),
            total_valuation=("VALUATION","sum"),
            first_year=("ISSUE_DATE",lambda x:x.min().year),
            last_year=("ISSUE_DATE",lambda x:x.max().year)
        )
    )
```

```
    st.dataframe(contractor_stats)

# ---------------------------------------------------------
# TAB 4 — INCOME INEQUALITY
# ---------------------------------------------------------
with tab4:
    st.header(" Permits by Income Group (AMI)")

    # Local AMI filter
    ami_local = st.multiselect(
        "AMI Categories (Tab Filter)",
        sorted(records_with_census["AMI_CATEGORY"].dropna().unique().tolist())
    )

    dfi = filtered_global.merge(
        census_tracts[["CENSUS_TRACT","AMI_CATEGORY","NUM_HOUSEHOLDS"]],
        left_on="CT",
        right_on="CENSUS_TRACT",
        how="left"
    )

    if ami_local:
        dfi = dfi[dfi["AMI_CATEGORY"].isin(ami_local)]

    permits_by_ami = dfi.groupby("AMI_CATEGORY")["PERMIT_TYPE"].count()
    households = dfi.groupby("AMI_CATEGORY")["NUM_HOUSEHOLDS"].sum()

    ratio = pd.concat([
        permits_by_ami / permits_by_ami.sum(),
        households / households.sum()
    ], axis=1)
    ratio.columns = ["permit_share","household_share"]

    st.bar_chart(ratio)
    st.dataframe(ratio)

# ---------------------------------------------------------
# TAB 5 — ALL PERMIT TRENDS
# ---------------------------------------------------------
with tab5:
    st.header(" Overall Trends")

    # Local filter: choose permit type
```

```
permit_local = st.multiselect(
    "Filter Permit Types Inside This Tab",
    sorted(records["PERMIT_TYPE"].dropna().unique().tolist())
)

dfa = filtered_global.copy()
if permit_local:
    dfa = dfa[dfa["PERMIT_TYPE"].isin(permit_local)]

dfa["year"] = dfa["ISSUE_DATE"].dt.year

st.subheader("Total Permits by Year")
st.line_chart(dfa.groupby("year")["PERMIT_TYPE"].count())

st.subheader("Total Valuation by Year")
st.line_chart(dfa.groupby("year")["VALUATION"].sum())
```

FIGURE 1 - Pool Permits by Year (Luxury Investment Trend)

Filters Used

- Year: **2015–2024**
- ZIP Codes: **90049, 90210, 90272**
- Keyword: **“swimming pool”**

What the Figure Shows

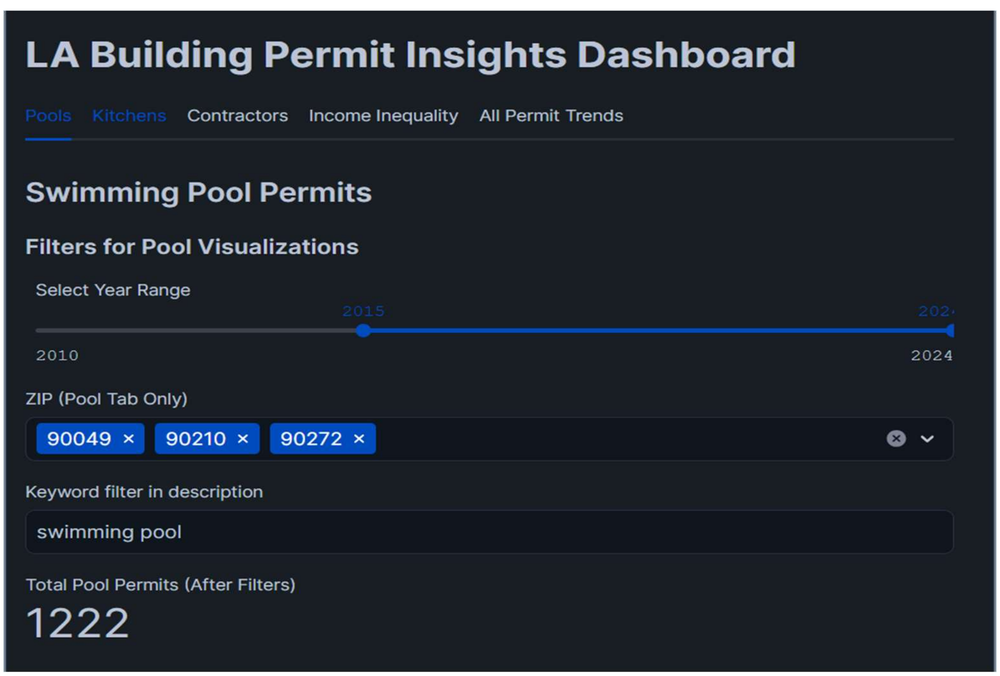
This chart visualizes annual counts of swimming pool permits within three of LA’s highest-income ZIP codes. Pool permits represent a **high-cost, discretionary renovation**, making them an excellent indicator of luxury residential investment.

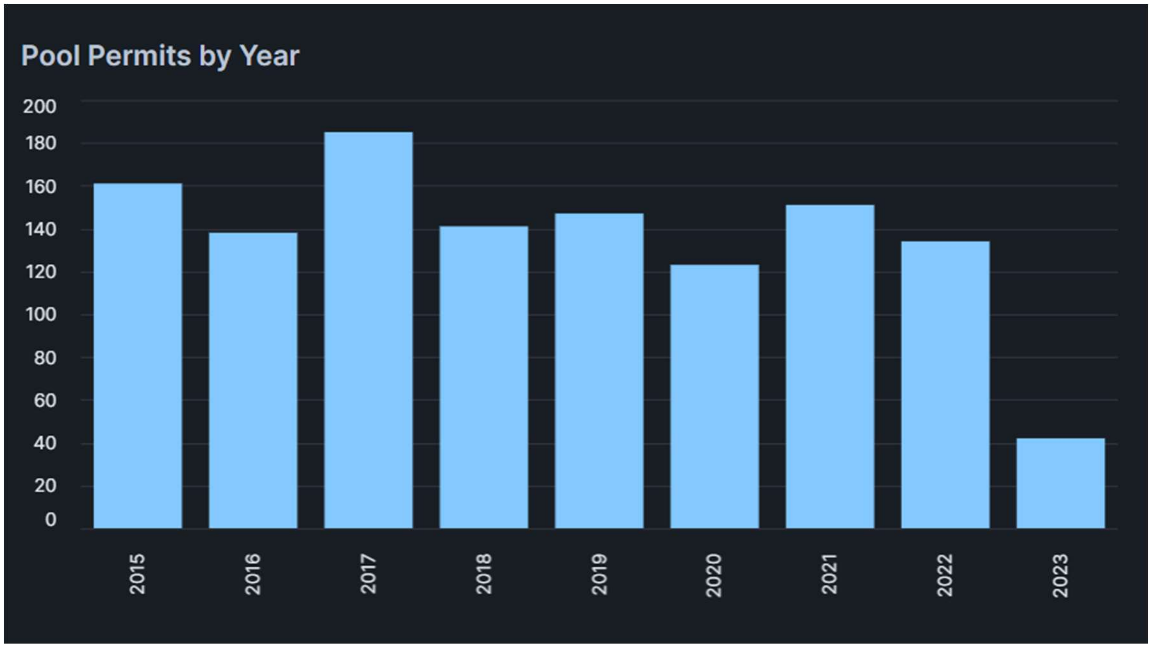
Key Insight

Pool construction stays consistently high across wealthy neighborhoods, peaking in several years (notably 2017 and 2021). Even with economic fluctuations, affluent regions continued investing in luxury outdoor amenities. The data indicates that wealthier areas maintain strong home improvement activity regardless of broader economic conditions.

Narrative Role

This figure introduces the central theme: **luxury investment is concentrated in wealthy neighborhoods and remains resilient over time.**




Pool Permits by Year
200
180
160
140
120
100
80
60
40
20
0
2015
2016
2017
2018
2019
2020
2021
2022
2023

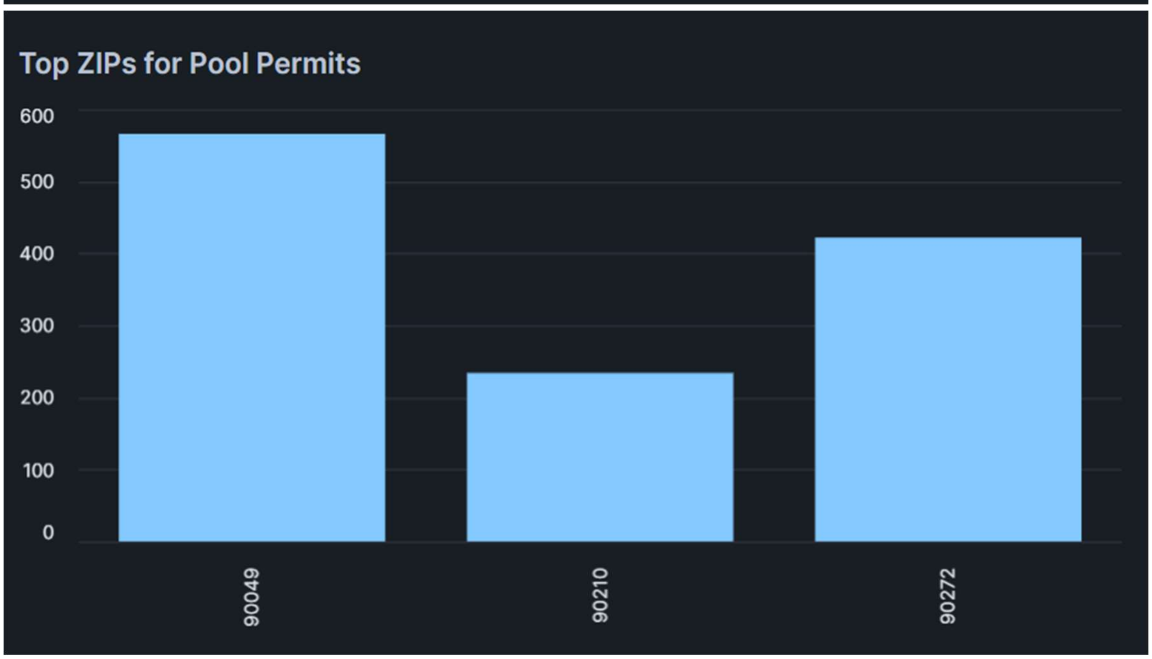

Top ZIPs for Pool Permits
600
500
400
300
200
100
0
90049
90210
90272

FIGURE 2 - Top ZIP Codes for Pool Permits & Geographic Concentration

Filters Used

(Same as Figure 1)

What the Figure Shows

This bar chart shows which wealthy ZIP codes had the highest number of pool permits:

- **90049 (Brentwood)** leads dramatically
- Followed by **90272 (Pacific Palisades)**
- And **90210 (Beverly Hills)**

The accompanying map shows that permits cluster almost entirely along the Westside coastline and hill-area luxury neighborhoods.

Key Insight

Luxury home upgrades like pools are **not evenly distributed citywide**. Instead, they are geographically concentrated in regions with high property values and household incomes. The map clearly reveals **dense clusters** of pool activity in specific wealthy pockets.

Narrative Role

This strengthens the argument that **luxury improvements amplify geographic inequality**, as only specific ZIP codes can support such investments.

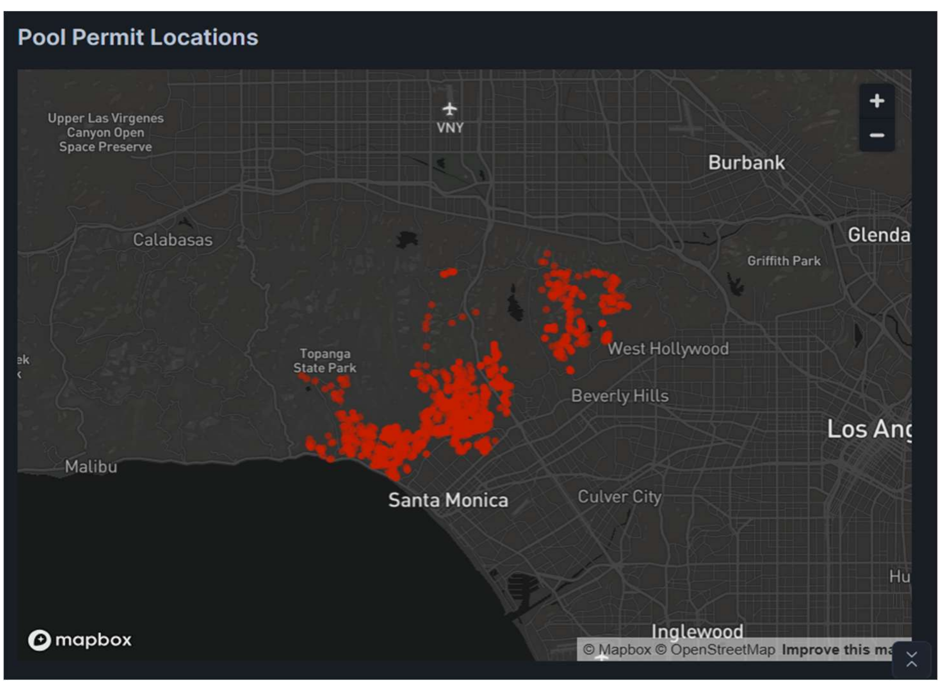


FIGURE 3 - Kitchen Permits & Valuation (Post-COVID Renovation Spike)

Filters Used

- Kitchen Year Range: **2020–2024**
- Valuation Range: **–2000 to 163,301**

What the Figures Show

(A) Kitchen Permits Over Time

- Kitchen renovations rose sharply in 2020–2021.
- This aligns with national trends: increased time at home led to interior upgrades.

(B) Average Kitchen Valuation

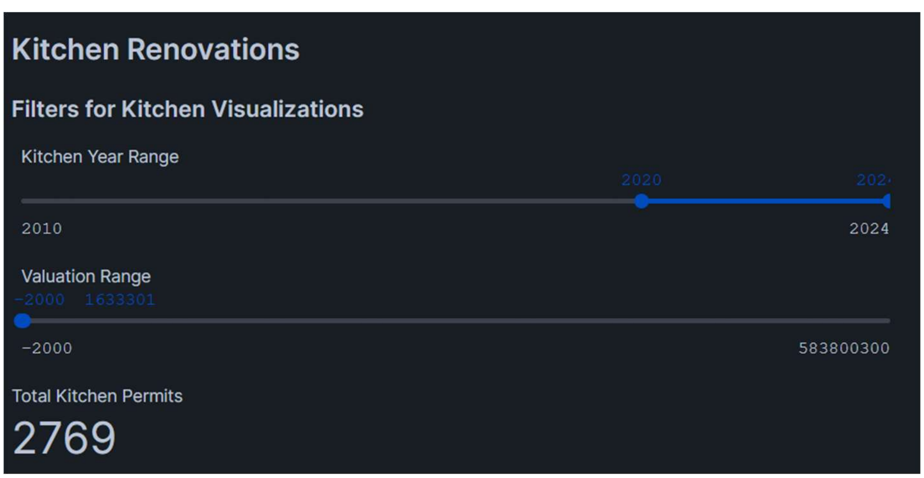
- Average permit valuation increases consistently from 2020 to 2023.
- Despite excluding ultra-high outliers, mid-range projects still trend upward.

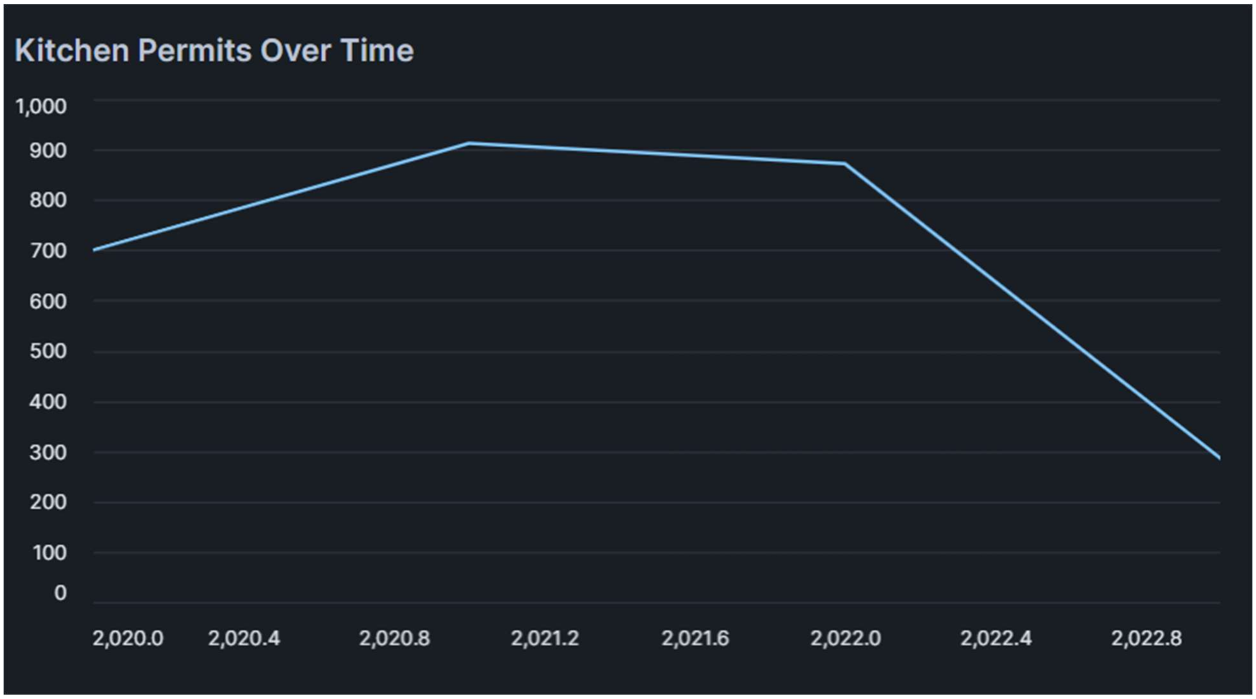
Key Insight

Kitchen remodels - representing practical, interior-focused improvements surged during and after the pandemic.
The rising valuation suggests households are investing more heavily in interior quality and comfort, not just maintenance.

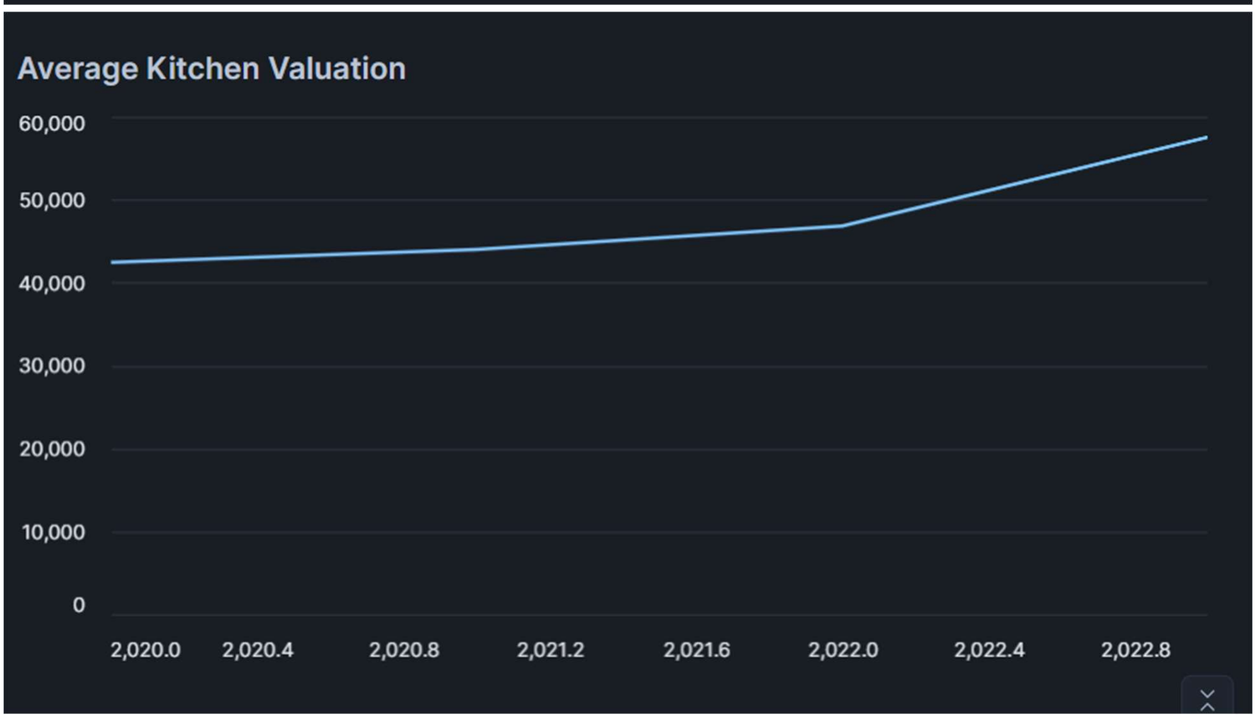
Narrative Role

This connects luxury outdoor upgrades (pools) with essential indoor upgrades (kitchens), showing a **citywide shift toward increased residential investment**, especially after 2020.





Kitchen Permits Over Time
1,000
900
800
700
600
500
400
300
200
100
0
2,020.0
2,020.4
2,020.8
2,021.2
2,021.6
2,022.0
2,022.4
2,022.8



Average Kitchen Valuation
60,000
50,000
40,000
30,000
20,000
10,000
0
2,020.0
2,020.4
2,020.8
2,021.2
2,021.6
2,022.0
2,022.4
2,022.8

FIGURE 4 - Renovation Inequality by AMI Category (Income Group Analysis)

Filters Used

- AMI Categories Selected:
    - **Above Moderate Income**
    - **Low Income**
    - **Very Low Income**

What the Figure Shows

The stacked chart compares:

- **Household Share** (how many households are in each AMI group)
- **Permit Share** (how many permits come from those groups)

Key Insight

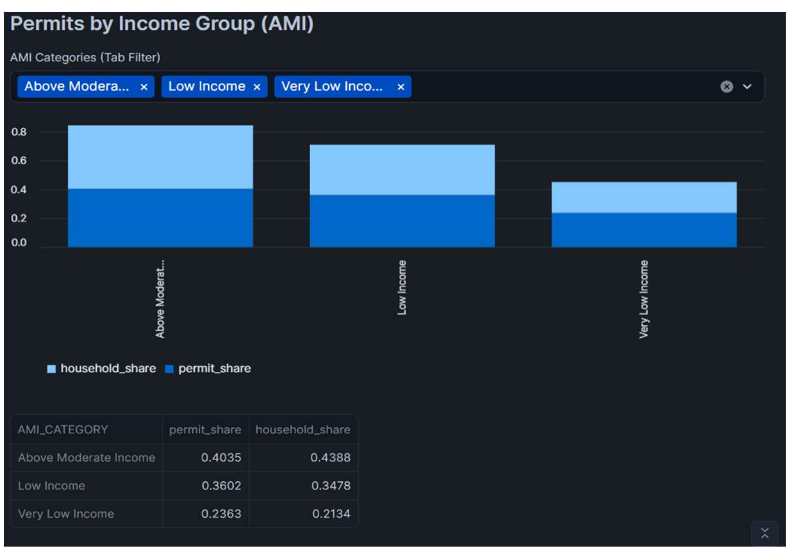
- High-income areas have **more permits than their household share would predict**.
- Low-income and very low-income households have **fewer permits relative to their size**.

This suggests that the ability to maintain, repair, or upgrade homes is **strongly linked to income**, deepening long-term infrastructure inequality.

Narrative Role

This is the most important inequality figure, it directly proves the project's central claim:

wealthier neighborhoods disproportionately drive LA's residential construction activity.



AMI_CATEGORY	permit_share	household_share
Above Moderate Income	0.4035	0.4388
Low Income	0.3602	0.3478
Very Low Income	0.2363	0.2134

FIGURE 5 - Citywide Trends in Major Construction Work

Filters Used

- Permit Types:
  - **Bldg-Addition**
  - **Bldg-Alter/Repair**
  - **Bldg-New**

What the Figures Show

Two line charts illustrate:

(A) Total Permits by Year

- A steady climb from 2013–2019
- A sharp dip in 2020 (COVID impact)
- Post-pandemic recovery in 2021–2022

(B) Total Valuation by Year

- Mirrors permit counts, but peaks around 2018–2019
- Reflects high-value construction activity before the pandemic

Key Insight

This provides meaningful context for earlier findings: even citywide, construction activity shows long-term growth.
The sharp pandemic dip aligns with national trends, reinforcing the validity of the data.

Narrative Role

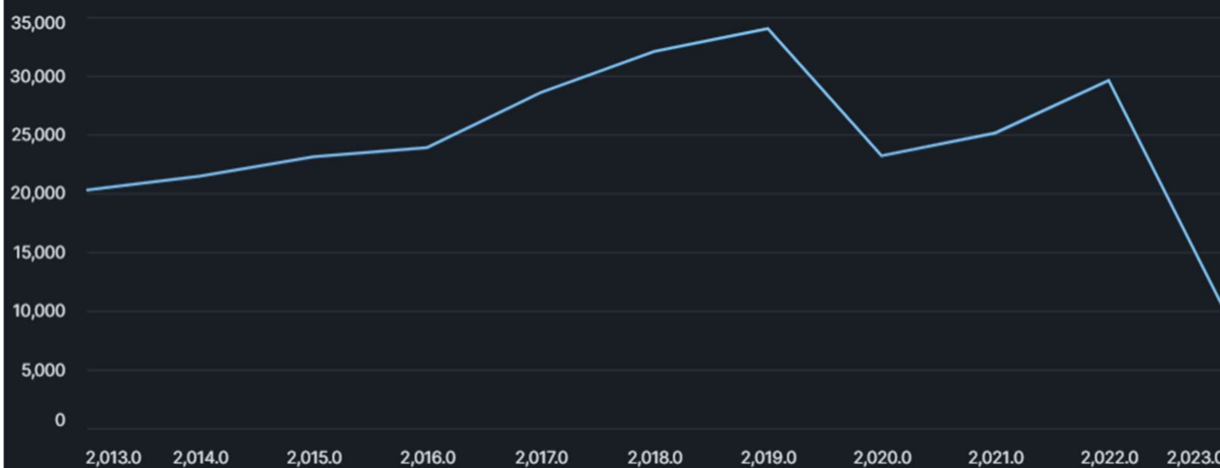
This closes your story by showing that even though all of LA is building more over time, the **type, cost, and distribution of improvements heavily favor high-income neighborhoods**, as shown in earlier figures.

Overall Trends

Filter Permit Types Inside This Tab

Bldg-Addition × | Bldg-Alter/Rep... × | Bldg-New ×

Total Permits by Year



Total Valuation by Year

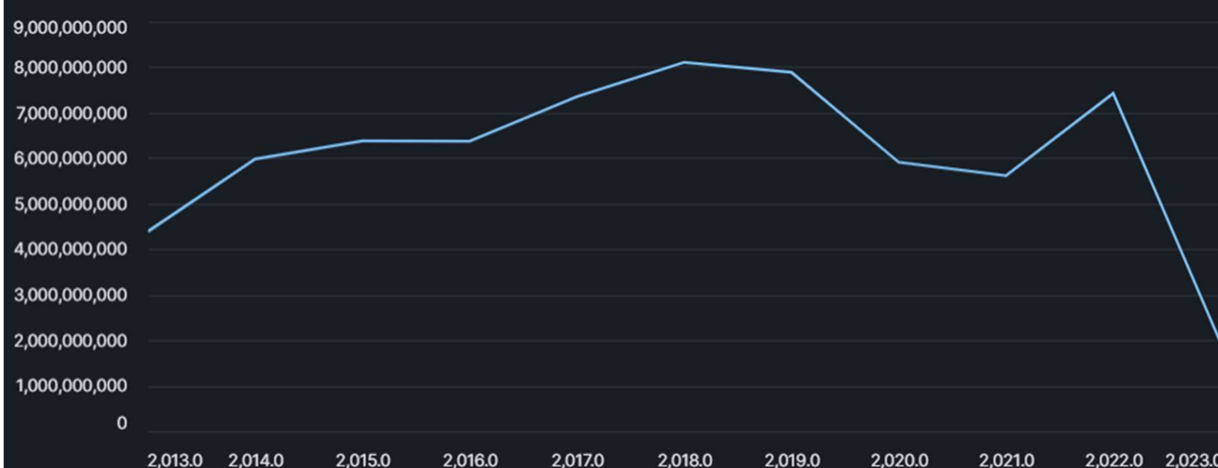


FIGURE 6 - Contractor Performance Dashboard (Market Concentration in Renovation Work)

What This Figure Shows

The Contractor Performance Dashboard displays a table summarizing contractor activity across Los Angeles.
Since no contractor filter was applied, the table shows **all active contractors**, including:

- **LICENSE_NUM** (unique contractor license ID)
- **num_projects** (number of permits handled)
- **total_valuation** (sum of valuation across projects)
- **first_year / last_year** active

The table reveals a striking pattern:

One contractor (license 0) is responsible for 132,512 projects with a valuation over $678 million,

far exceeding every other contractor listed.

Most other contractors have only **1–50 projects**, and many have **zero or minimal valuation**.

Key Insight

This table highlights **extreme market concentration** in LA's construction ecosystem:

- A single contractor dominates the majority of recorded permits.
- Very few contractors handle mid- to high-value projects.
- Most are small operators with minimal annual activity.

This pattern often occurs when large general contractors file permits under a parent company, while subcontractors and small firms appear in the dataset with very few records.

In other words:

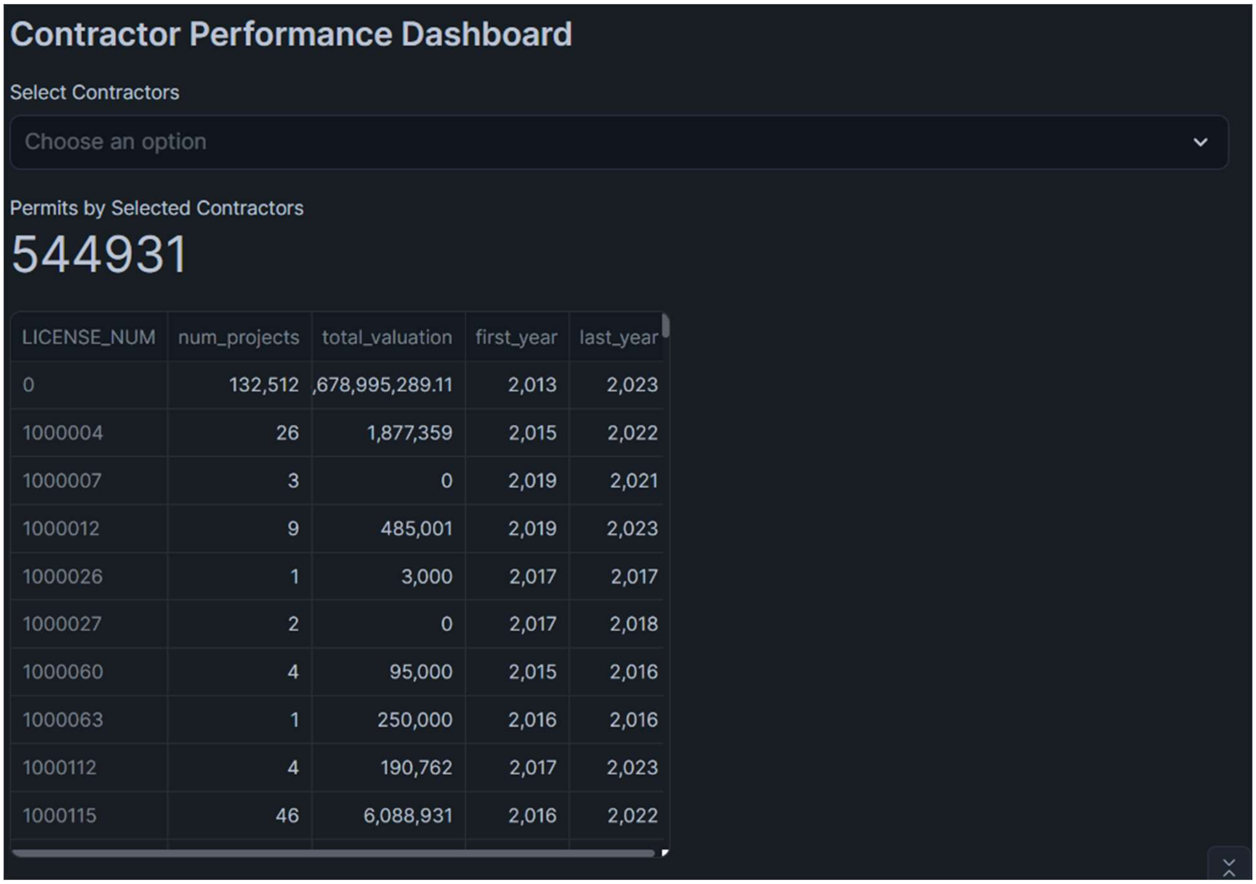
LA's residential construction industry is not only unequal geographically and economically it is also structurally unequal.

A small group of high-volume contractors handle the vast majority of all building activity.

Narrative Role

Figure 6 strengthens your inequality story by adding a **third dimension of inequality**,
not just **where** construction happens (geographic inequality),

or **who** can afford upgrades (income inequality),
but also **who performs the work** (industry concentration inequality).



LICENSE_NUM	num_projects	total_valuation	first_year	last_year
0	132,512	,678,995,289.11	2,013	2,023
1000004	26	1,877,359	2,015	2,022
1000007	3	0	2,019	2,021
1000012	9	485,001	2,019	2,023
1000026	1	3,000	2,017	2,017
1000027	2	0	2,017	2,018
1000060	4	95,000	2,015	2,016
1000063	1	250,000	2,016	2,016
1000112	4	190,762	2,017	2,023
1000115	46	6,088,931	2,016	2,022

Conclusion

Across all six figures, a consistent and compelling narrative emerges:

1. **Luxury permit activity** (pool construction) is concentrated almost entirely in affluent Westside ZIP codes.
2. **Indoor upgrades** like kitchen renovations surged after 2020, with valuation steadily rising, indicating homeowners are investing more in property quality.
3. **Citywide construction activity** shows long-term growth, with a sharp COVID dip followed by recovery.
4. **Income-based inequality** is clear: higher-income tracts produce a disproportionately large share of permits compared to their share of households.
5. **Geographic clustering** of luxury permits reinforces spatial inequity across LA.
6. **Contractor inequality** introduces a crucial new dimension: a small number of firms dominate the majority of permits and valuations, leaving most contractors with minimal participation.

Final Insight

LA's residential construction patterns reveal **three interconnected layers of inequality**:

- **Economic inequality:** wealthier households renovate more, more often, and with higher budgets.
- **Geographic inequality:** luxury improvements cluster in specific ZIP codes while others remain under-invested.
- **Industry inequality:** a small number of contractors control the vast majority of renovation work.

These findings suggest that home improvement activity in Los Angeles is not equally distributed, it is shaped simultaneously by **income**, **location**, and **market structure**, reinforcing disparities in housing quality and neighborhood development.