

# Smiley Real Estate

## Insights by Ruth Magana

# Index

[Best Markets](#)

[Market Trends](#)

[Best Agents](#)

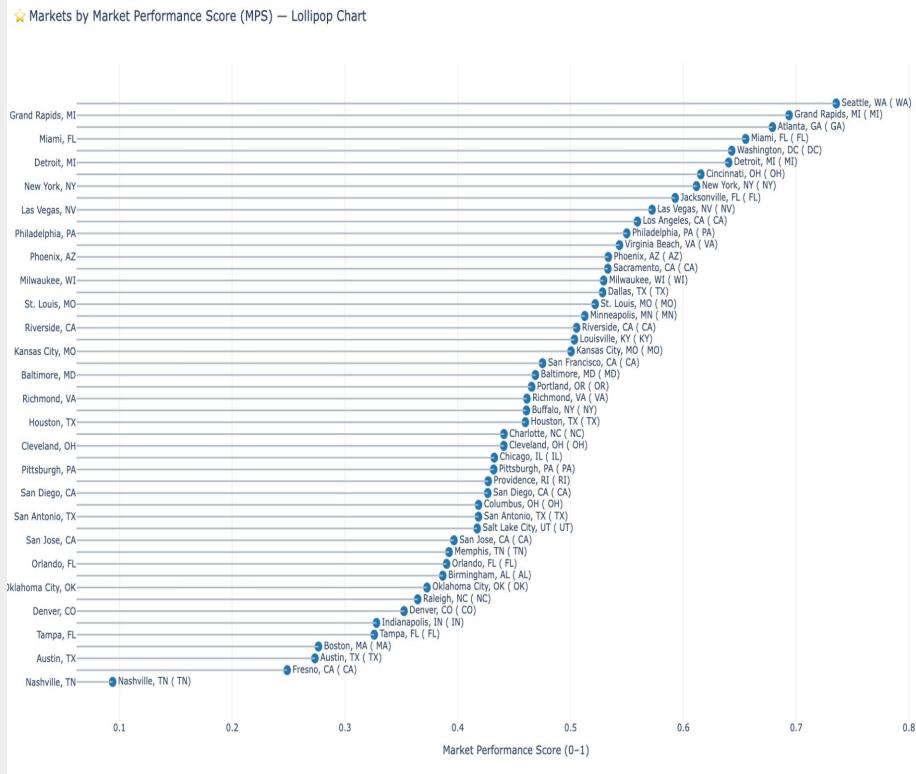
[Growth Areas](#)

[What if I had more time and data sources?](#)

[My Process](#)

# Best Markets

## Colab Section



# What are currently the 5 best markets for Smiley?

🏆 Top 5 Markets — Podium Ranking by Market Performance Score (MPS)

Market Performance Score: The Explanation is on the Next Slide! :)



# ★ Introducing the Market Performance Score (MPS)

To get a **holistic** view of market quality, the **MPS was developed**, which blends multiple performance factors into a single comparable metric across all markets.

Markets operate on very different scales (population, sales count, revenue, time).

To avoid larger markets automatically scoring higher, each metric is **normalized to a 0–1 scale**, which ensures:

- balanced weighting
- no single metric dominates due to scale differences

Some metrics are *inverted* (like days on market or time to sale), so **lower values score higher**

[Revenue] for Smiley was calculated as 1% of the home's value (represented by the **sales amount**).

The final score rewards a mix of **internal performance, customer success, market value, and long-term potential**:

**MPS =**

0.30 • Revenue (profitability)  
+ 0.15 • Sales Volume (throughput)  
+ 0.20 • Conversion Rate (customer efficiency)  
+ 0.10 • Speed to Sale (operational performance)  
+ 0.10 • Median Sale Price (value profile)  
+ 0.10 • Population (market potential)  
+ 0.05 • Days on Market (market temperature)

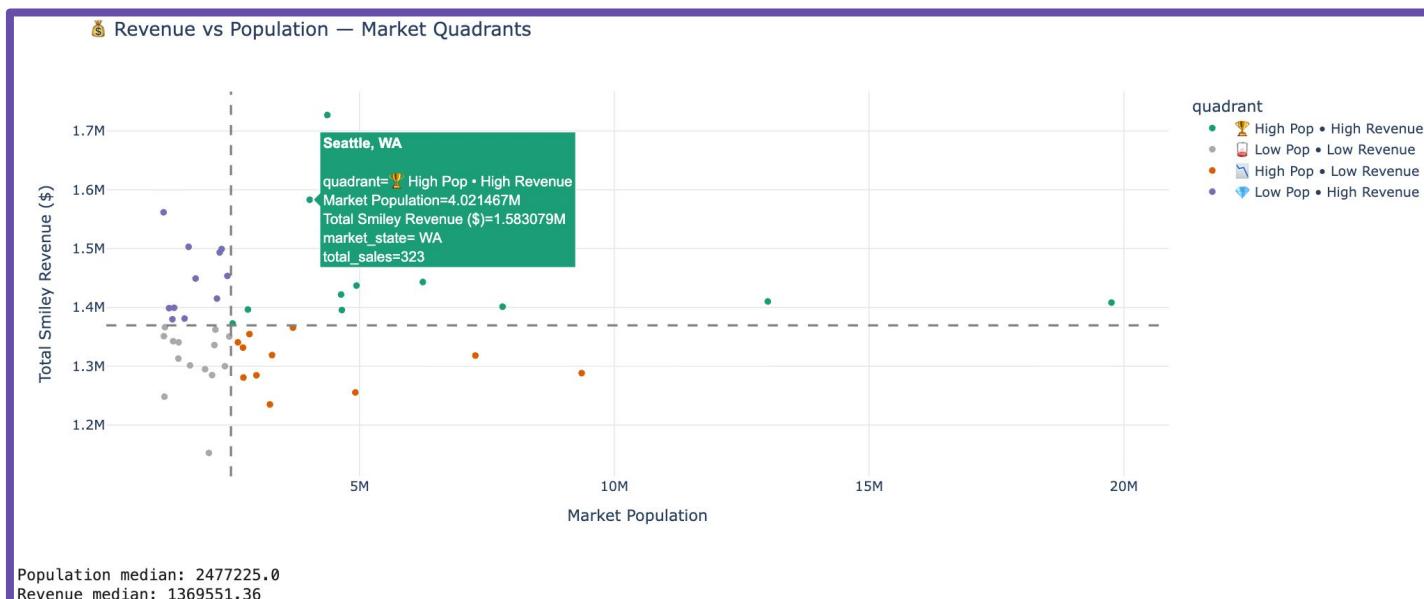
[Speed to Sale] = Avg Days between customer creation and sale date

External metrics like price increases/decreases were intentionally excluded to reduce bias from market-level volatility the company doesn't control.

# Why was the Market Performance Score (MPS) created?

**Smiley Real Estate operates across diverse markets** that differ in size, competitiveness, home value, speed, and customer behavior.

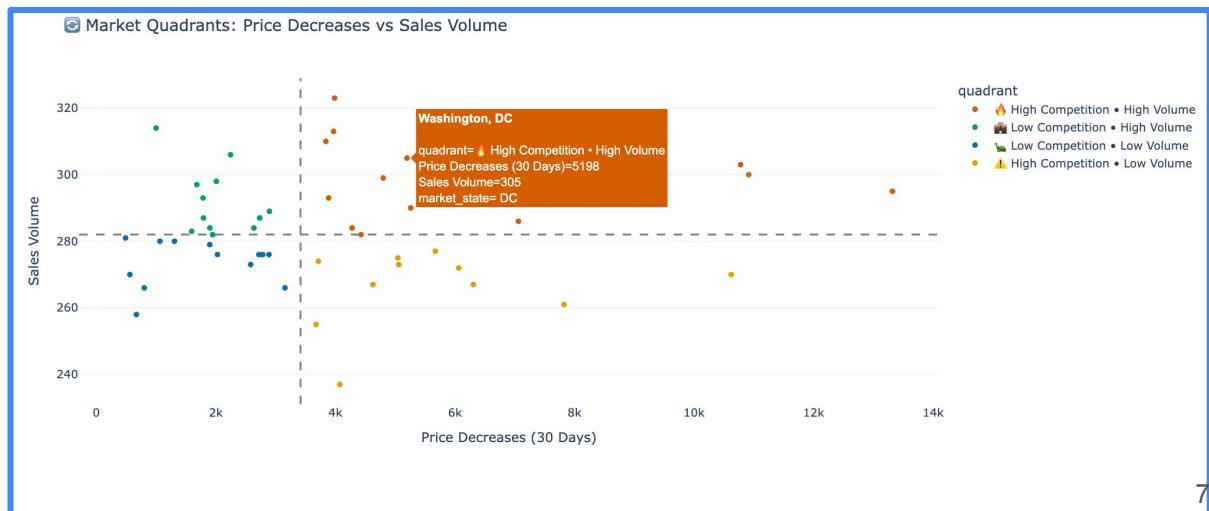
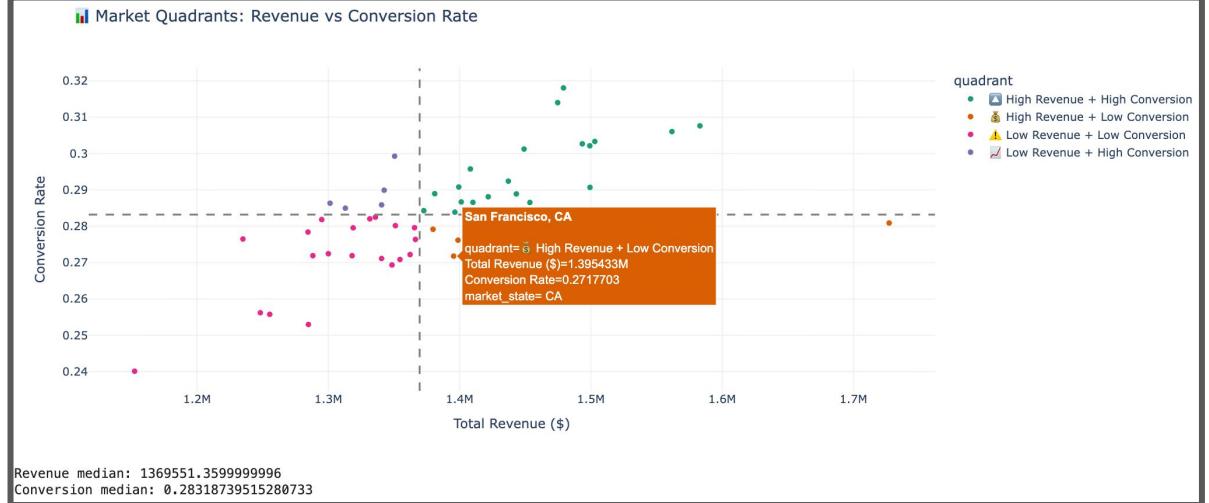
**Markets can be segmented using several two-metric quadrant frameworks**, such as **Revenue vs Population** (That identifies Profitable, Niche, Underperforming and Low-priority markets)



# Quadrants Categorization

Other useful quadrant views when analyzing a single dimension of market behavior:

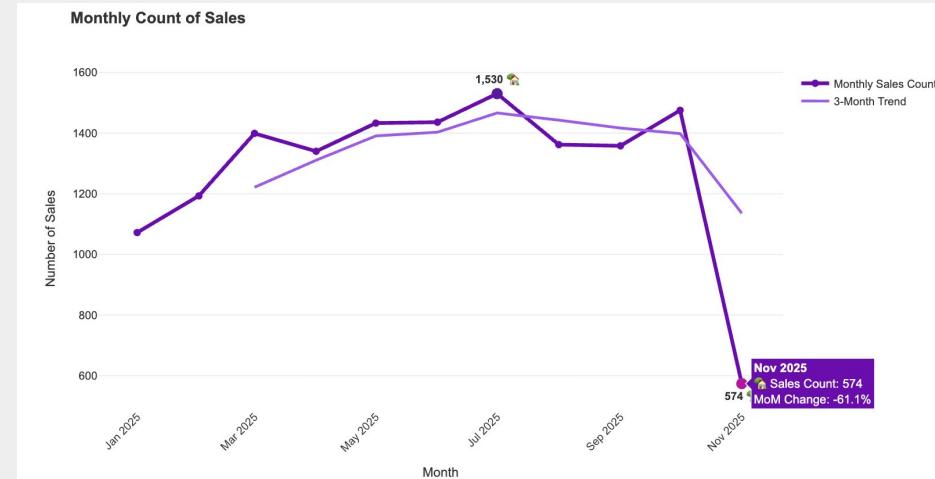
- **Revenue vs Conversion** (Value × Efficiency)
- **Price vs Volume** (Value × Throughput)
- **Speed vs Efficiency** (DOM × Conversion)
- **Competition vs Throughput** (Price Decreases × Sales)



However, each quadrant highlights only one part of the story, the complete view is the MPS

# Market Trends

## From Smiley Real Estate Data





# Home Price Distribution — Key Percentile Insights

## Smiley Real Estate Sales Data (Year 2025)

Understanding Price Segments in the Market

### Key market segments:

- The Entry Level (up to \$404,313 - p25): The lowest quartile of sales, representing properties accessible to entry-level buyers.
- The Typical Home (\$484,184 - Median): This is the midpoint of the market; half of the homes sell for less, and half sell for more.
- The Upper Tier (above \$563,108 - p75): The top 25% of sales, indicating the boundary for higher-priced listings.
- The Luxury Segment (above \$658,989 - p95): Only 5% of sales exceed this price point, marking the exclusive, top-tier segment.

🔍 Outlier threshold (99th percentile): \$713,178

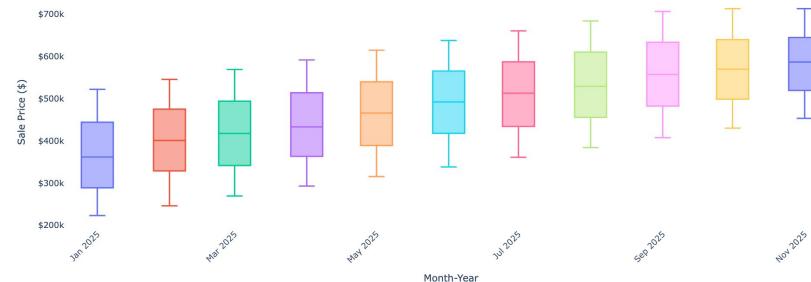
⚠️ Outliers removed from the following month-year groups:

- Mar 2025: 1 sale(s)
- Oct 2025: 74 sale(s)
- Nov 2025: 67 sale(s)

Distribution of Home Sale Prices (0–99th Percentile)



Distribution of Home Sale Prices per Month (0–99th Percentile)

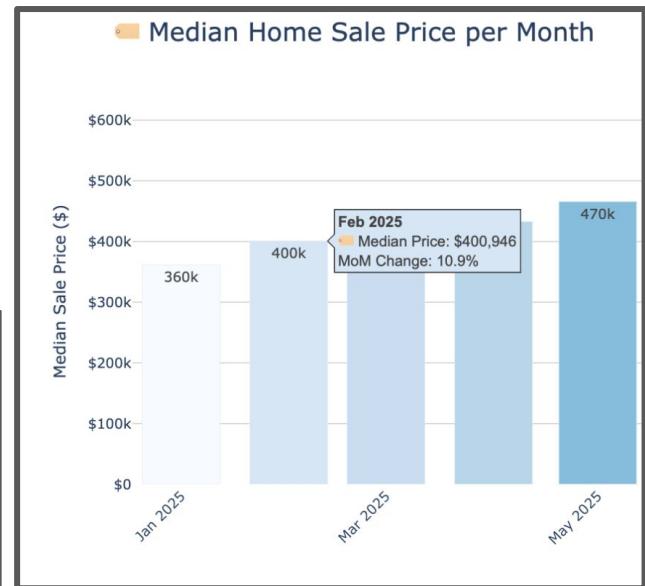


# Smiley Real Estate Median Home Sale Price per Month

The median home sale price for the period of January through November 2025 was \$484,184, which represents the typical price point in the market.

Over this period, the market saw consistent appreciation:

- The median monthly percentage change in sale price was 4.40%.
- The average monthly percentage change was 5.27%.
- Notably, the most significant surge occurred in February 2025, which recorded a 10.9% price increase compared to the previous month.



## How Long Does It Take to Close a Sale After First Call?

For Smiley Real Estate, the typical time from a customer's first call to a closed sale is approximately 5 days.

The process is highly efficient: three-quarters of all customers complete their purchase within 8 days, and 95% close within 12 days, showing that nearly all transactions are resolved within just under two weeks.



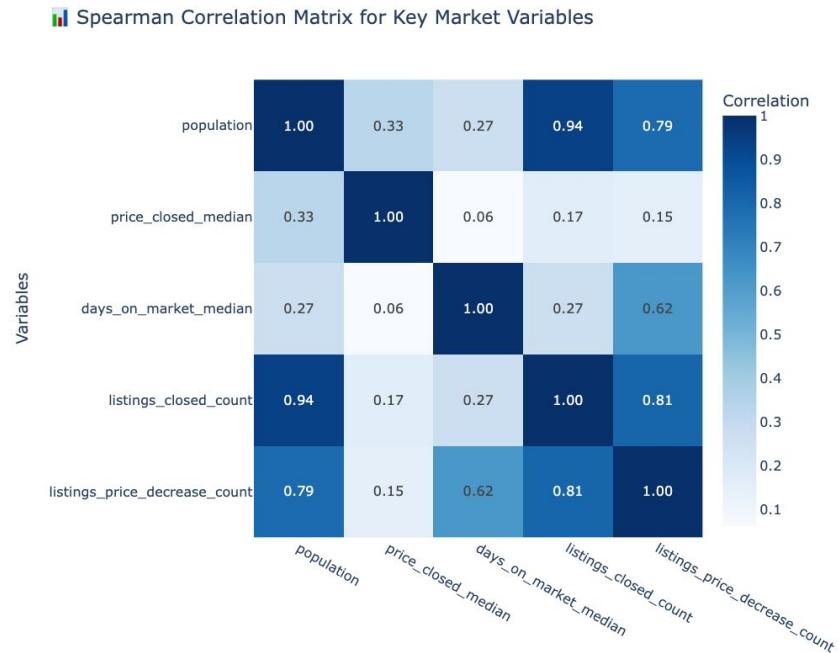
25th percentile: 3.0 days  
Median: 5.0 days  
75th percentile: 8.0 days  
95th percentile: 12.0 days

# Market Variables Correlations

In summary, our key insights are:

- 1) **Larger markets exhibit greater pricing competition** (Population and price decreases correlate strongly at  $p = 0.791$ )
- 2) **Price reductions are often a response to longer listing times** (Median days on market correlates with price cuts at  $p= 0.625$ ).

These findings are detailed in the Spearman correlation matrix.



Some strong correlations, like population with listings\_closed\_count ( $p = 0.938$ ) and closed listings with price decreases ( $p = 0.811$ ), are mostly driven by market size. Larger markets naturally have more activity, so these patterns offer limited strategic insight.

July  
17

# Monthly Sales Trends — Summary Insights

## Overall Trend

- Monthly home sales show **moderate variability**, with an **average MoM change of  $-2.57\%$** , indicating slight downward pressure on volume.
- However, the **median MoM increase of  $3.38\%$**  suggests that most months still experienced *mild growth*, with a few sharp declines pulling the average downward.

## Volatility

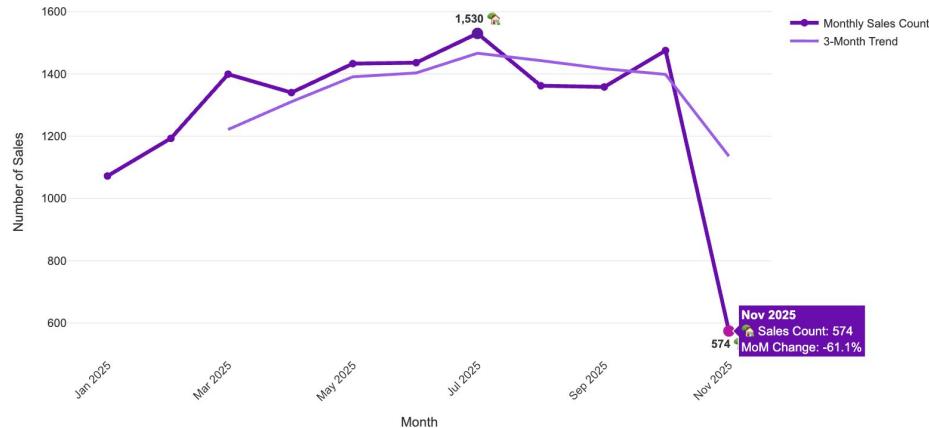
- The average **absolute month-to-month change of  $12.74\%$**  highlights meaningful fluctuations in market activity, signaling a dynamic sales environment rather than a stable, predictable trend.

## Highs and Lows

- Peak Month:**  
**July 2025 — 1,530 sales**, representing a healthy **+6.5% increase** over June. This mid-year surge aligns with typical seasonal strength in real estate activity.
- Lowest Month:**  
**November 2025 — 574 sales**, a significant **61.1% drop** from October. This sharp decline suggests an exceptional slowdown, potentially driven by incomplete data, seasonality, supply drops, or external market factors.

The market exhibits **seasonal peaks, occasional sharp contractions**, and overall **moderate volatility**. Monitoring these swings helps identify when markets accelerate or soften, informing better forecasting and resource planning.

## Monthly Count of Sales



**Note:** *November 2025 should be interpreted with caution, as its unusually low values may reflect incomplete or delayed reporting rather than true market performance.*

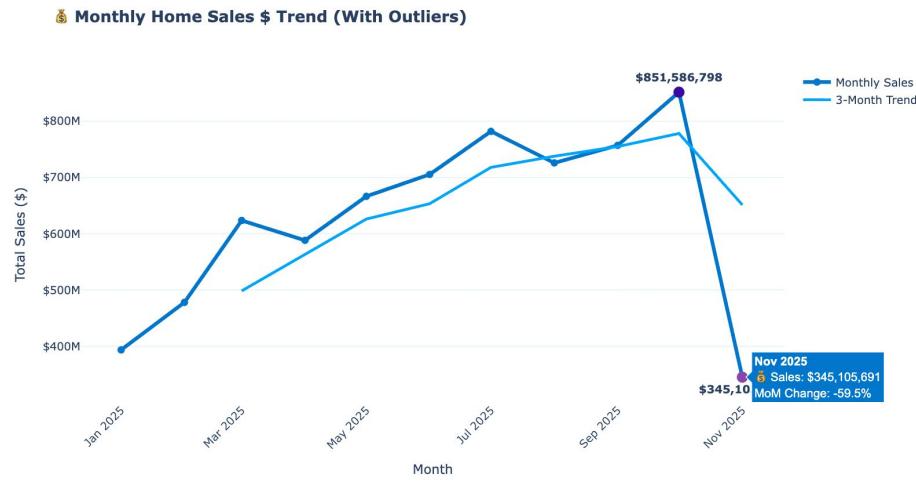
*For the purposes of this exercise, the month was retained to preserve the full dataset, but in a real-world scenario this datapoint would warrant further investigation and validation with the Smiley team before drawing conclusions.*

# Monthly home Sales Totals (\$)

Meaningful swings throughout the year. With all sales included, the market experiences an average month-over-month growth of 2.64%, a median of 8.35%, and a relatively high volatility of 17.10%, ranging from the lowest month in Nov 2025 (\$345M) to the peak in Oct 2025 (\$851M)

After removing extreme outliers above the 99th percentile (\$713K), the trend becomes more stable: average MoM growth softens to 1.40%, the median to 5.65%, and volatility drops to 15.41%.

The adjusted series still shows seasonal strength in Oct 2025 (\$798M) and weakness in Nov 2025 (\$296M).



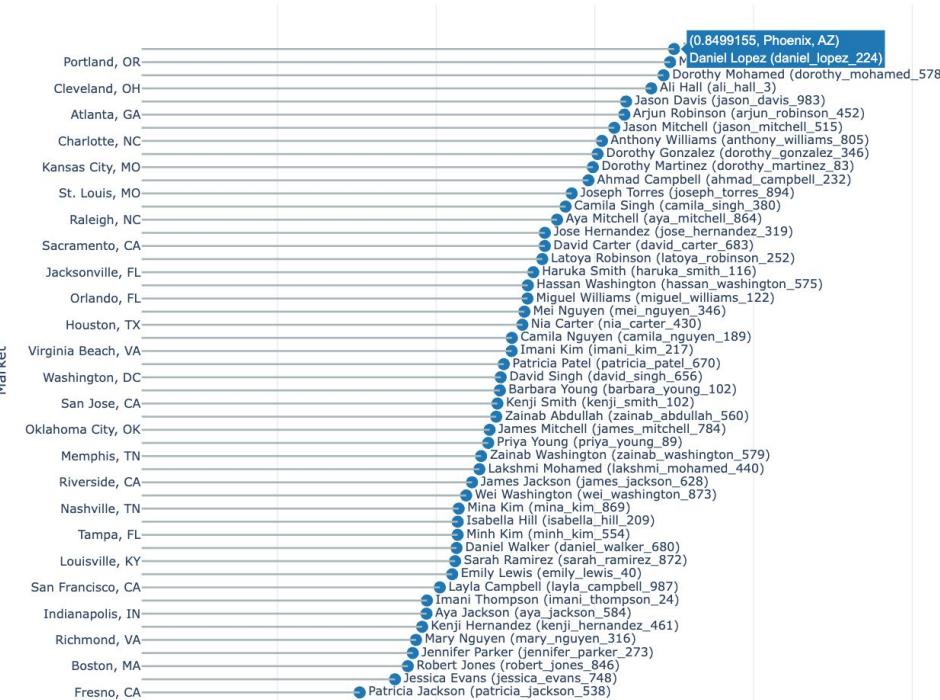
# Best Agents

## Colab Section



# Who is the best agent we have in each market?

★ Best Agent Per Market (APS Score)



Best Agent per Market — Table Version

Show code

...	market_name	agent_name	agent_id	median_sale	num_sales	total_revenue	conversion_rate	avg_time_to_sale	APS
0	Phoenix, AZ	Daniel Lopez	daniel_lopez_224	521605.0	29	148294.19	0.349398	6.551724	0.850
1	Portland, OR	Michael Hall	michael_hall_387	491495.5	28	136705.72	0.388889	6.000000	0.847
2	Milwaukee, WI	Dorothy Mohamed	dorothy_mohamed_578	523837.0	24	121035.73	0.461538	5.500000	0.843
3	Cleveland, OH	Ali Hall	ali_hall_3	511245.5	26	129712.20	0.426230	6.115385	0.836
4	Minneapolis, MN	Jason Davis	jason_davis_983	523532.0	24	122703.01	0.406780	6.875000	0.820
5	Atlanta, GA	Arjun Robinson	arjun_robinson_452	462604.5	24	111109.84	0.545455	6.166667	0.819
6	San Antonio, TX	Jason Mitchell	jason_mitchell_515	574107.0	22	124044.97	0.407407	5.909091	0.812
7	Charlotte, NC	Anthony Williams	anthony_williams_805	538084.0	27	140756.99	0.313953	7.000000	0.804
8	Cincinnati, OH	Dorothy Gonzalez	dorothy_gonzalez_346	440902.0	28	126905.19	0.424242	5.321429	0.802
9	Kansas City, MO	Dorothy Martinez	dorothy_martinez_83	516222.0	22	111035.48	0.400000	3.909091	0.799

Best Agents per Market by Agent Performance Score Ranking, [Table here](#)

# ★ How the Agent Performance Score (APS) Works

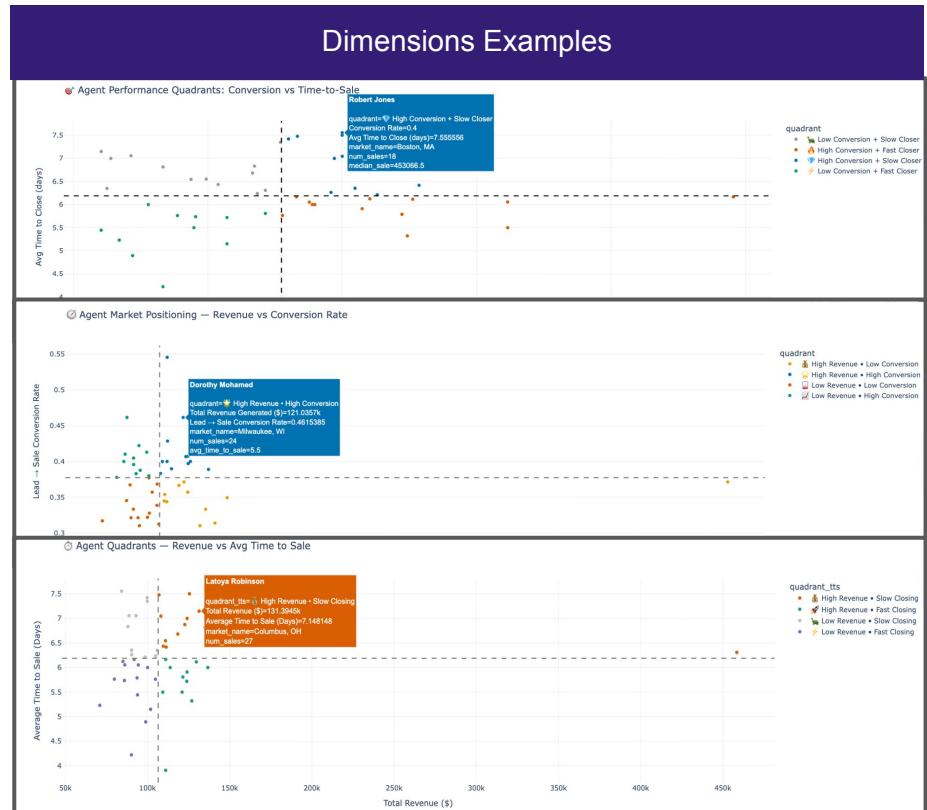
APS is a composite metric designed to fairly compare agents across all markets by **combining multiple dimensions** of performance into a single, standardized score.

To build APS, we first calculate key KPIs for every agent:

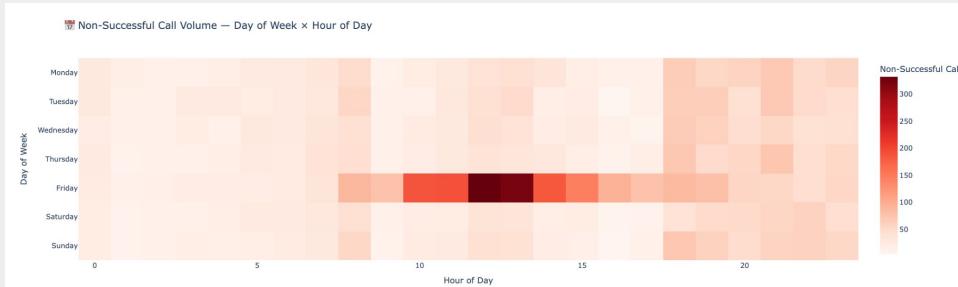
- **Total revenue generated**
- **Number of sales closed**
- **Median sale price**
- **Conversion rate** (sales ÷ assigned customers)
- **Average time to sale**

Finally, APS is computed as a weighted blend of these normalized KPIs.

The weighting prioritizes **revenue generation** and **sales volume**, while still rewarding **efficiency, deal quality, and conversion**.

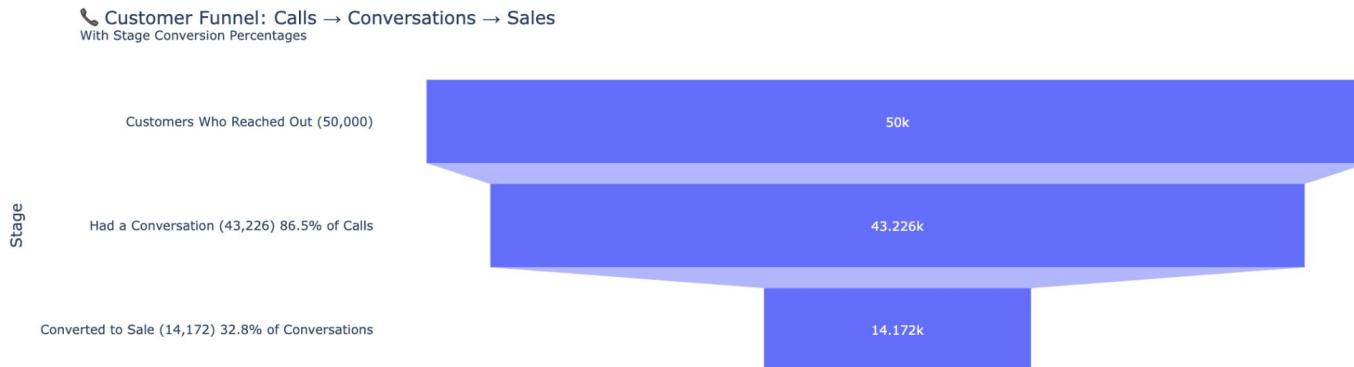


# Growth Areas



# Conversion Rates

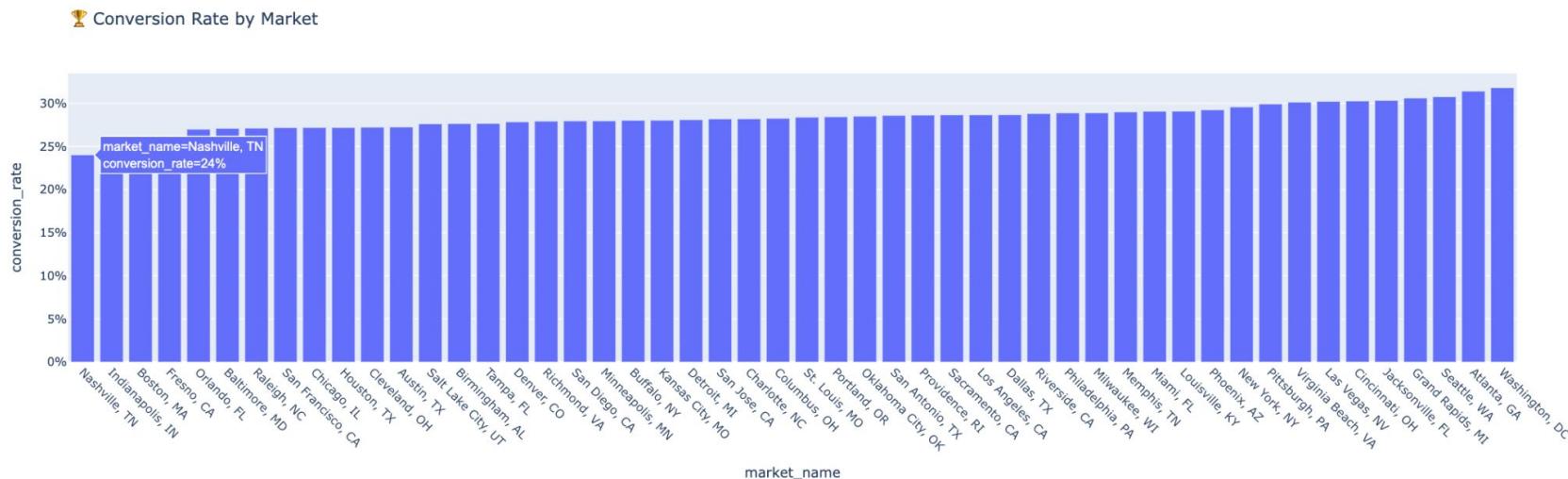
Only 28.3% of all customers that reached out to Smiley had an actual successful Sale, more analysis can be done to improve conversion to sale.



- Funnel Conversion Summary
- Calls → Conversations: 86.5%
- Conversations → Sales: 32.8%
- Calls → Sales: 28.3%

# Conversions by Market

The market with the lowest conversion rate is Nashville, TN we can create conversion strategies adjusted by Agent or States I suggest to use of quadrants shared previously to identify agents/markets.



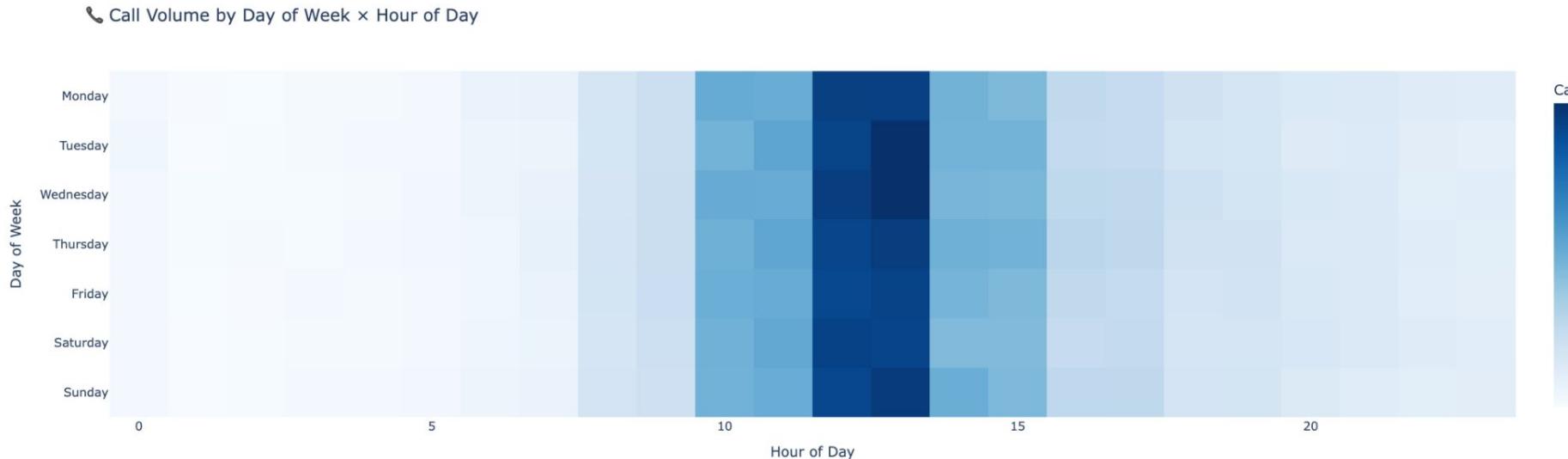
# Calls Volume Trend

The drop in the volume of calls in November, mirroring the trend observed in Sales, may be attributable to incomplete data capture. However, this requires further investigation to confirm if the issue lies with data logging or a genuine decrease in customer interaction.



# All Calls by Day and Hour

Most of the calls are around 10 to 15 hours, we can adjust schedule based on this demand hours

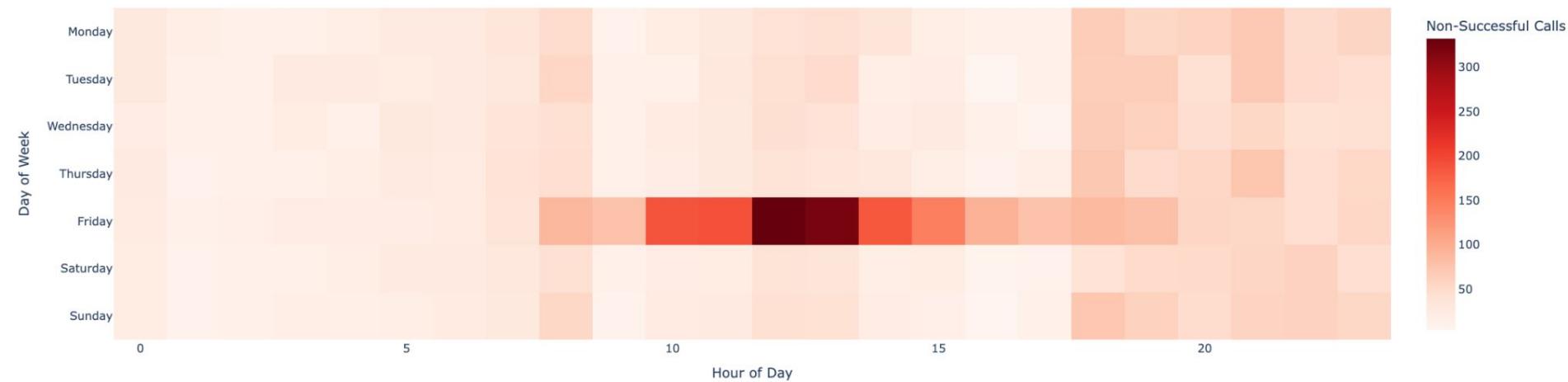


# Unsuccessful calls

Most unsuccessful calls, cases where customers were unable to speak with a Smiley call handler occurred on Fridays.

Schedules and call-handlers can be adjusted.

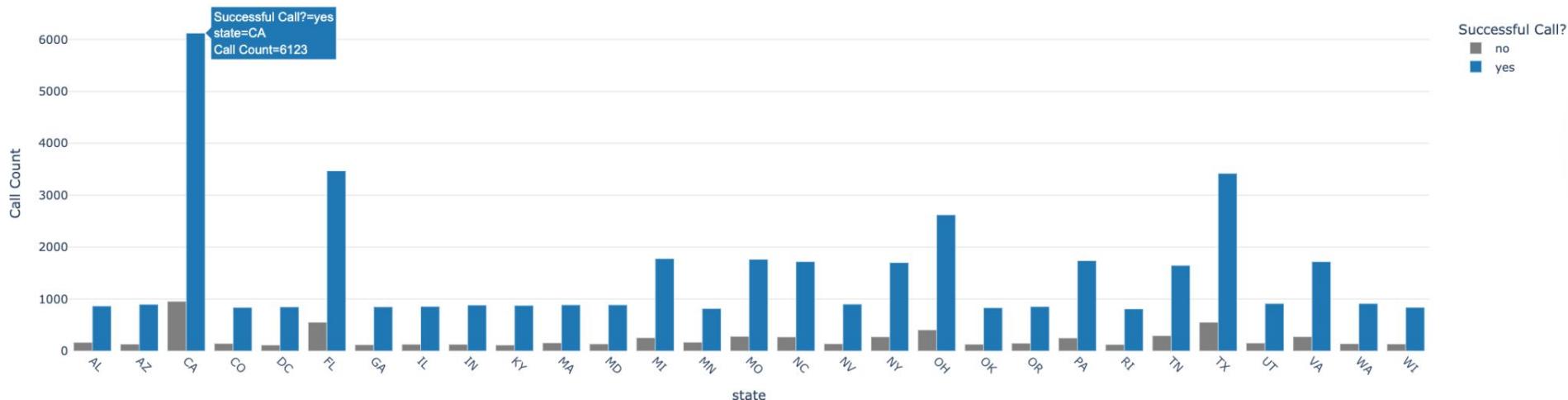
Non-Successful Call Volume — Day of Week × Hour of Day



# Calls Volume by State

Some states receive more calls than others, Marketing can create campaigns for low call volume states & Call-Team can increase the amount of call-handlers for high volume states.

Successful vs Non-Successful Calls by State



What if I had more time  
and data sources?

# What if I had more time and data sources?



## Enhancing Data Granularity and Context

To move beyond basic descriptive metrics, focus on enriching both your internal and external data inputs:

### 1. External Economic and Demographic Context

Current analysis focuses heavily on internal performance. I need to benchmark this performance against the external reality of the 50 markets.

- **Market Contextualization:** Integrate and compare your sales volume and price data against **local inflation, job market health (unemployment rate), median household income, and demographic age** for each market. This helps to determine if a market's low performance is due to *agent failure* or *economic headwind*.
- **Property Type Analysis:** Ingest data on the **type of property sold** (e.g., apartment, townhouse, single-family house). This allows to normalize performance based on market specialty (e.g., an agent specializing in luxury houses should not be compared directly to one specializing in high-volume apartment sales).

## 2. Deepening Call and Customer Funnel Data

Current conversion rate is based on total customers/total sales. This is too broad.

- **Refining the Funnel:** Establish a clear **current customer journey/funnel** within your data warehouse. Identify key conversion stages (e.g., Lead ->Qualified ->Assigned Agent ->Viewing ->Closed).
- **Call Type Input:** Break down the **call types** from the **calls** table (e.g., initial inquiry, follow-up, closing negotiation, technical support). Assign different weights to conversion rates based on the call type's objective, providing a much more accurate measure of agent/handler effectiveness.
- **Clarifying Seller vs. Buyer Customers:** **Because of time I assumed sales table had all customer possible successful interactions in case that not true** Explicitly label customers as **Seller** (customer is selling a house, and Smiley assists) or **Buyer** to adjust performance metrics, as these processes have vastly different timelines and revenue profiles (Or if there is a buyers table that will improve the analysis)

## ⭐ Improving Agent and Handler Measurement

Focus on bringing **Qualitative Feedback** into quantitative models:

- **Net Promoter Score (NPS) Implementation:** Implement a system to regularly collect **Net Promoter Scores (NPS)** for both **agents and call handlers**. This moves beyond the outdated 2024 review data. NPS is a superior metric for gauging customer loyalty and satisfaction compared to simple review averages.
- **Integrating Qualitative Feedback:** Use the collected NPS and feedback data as a weighted factor in the **Agent Performance Score (APS)**. Agents with high sales volume but low NPS might be causing long-term brand damage.

## 🎯 Aligning Metrics with Company Strategy

The final and most crucial improvement is ensuring your scores reflect the company's true mission.

- **Mission-Driven Weighting:** Before setting the final **APS and MPS weights**, formally clarify the **company's mission and values** (e.g., *Is the primary value speed? Luxury? Volume?*). The weighting structure (e.g., 30% revenue, 20% conversion) should be a direct, measurable reflection of these organizational goals.

# My Process

Colab

The screenshot shows a Google Colab interface with the following structure:

- File Edit View Insert Runtime Tools Help**
- Commands + Code + Text ▶ Run all ▾**
- Table of contents**
- Data Processing**
  - Obtain and Scrub
  - ↳ 66 cells hidden
- Exploration of Data**
  - ↳ 7 cells hidden
- Interpret Stage**
  - Charts created for interpretation step, interpretation is available at slides
  - ↳ 32 cells hidden
- Modeling Stage**
  - ↳ 22 cells hidden

Sections and sub-sections listed in the table of contents include:

- Interpret Stage
  - Join Tables
  - Section: Agents
    - Top 10 Agents — R...
  - Section: Markets
    - Distribution of Ho...
    - Distribution of Ho...
    - Distribution of Clo...
    - Distribution of Clo...
  - Spearman Correlatio...
  - Spearman Correlation...
  - Analysis of Strong Sp...
  - Market Position: Po...
  - Market Position: Da...
  - Top 10 Markets ...
  - Top 10 Markets ...
  - Full Customer + Call...
  - Customer Funnel...
  - Conversion Rate ...

# My Process Part 1 & 2

My process followed the OSENM analytical framework (Obtain → Scrub → Explore → Interpret → Model).

**Obtain:** To begin, I imported the raw tables directly from Google Sheets into Google Colab, since the dataset was not stored in a database (which would have required a different process like an ETL). Colab's built-in Sheets integration made this step straightforward.

During the **Scrub** stage, I cleaned currency symbols, fixed text-based “NA” values, standardized datetimes, and engineered additional fields such as day-of-week and hour-of-day.

# My Process Part 3

## Exploratory Data Analysis (EDA) Findings

During the Explore phase, a thorough examination of the table structures and relationships yielded several important operational findings:

1. **Agent Identification:** It was discovered that multiple agents shared the same name. To ensure accurate and non-inflated analysis, the unique field **agent\_id** was prioritized as the true identifier for all subsequent calculations.
2. **Customer Onboarding Process:** An identical match was observed between customer creation dates and first call dates. This suggests that customers are integrated into the system at the exact time of their initial contact.

## Data Integrity Note (November Performance)

It was noted that all key performance metrics—including calls, sales, and related activity—experienced a distinct decrease in November. For the sake of preserving the complete dataset for this exercise, this month was retained. However, in a real-world analysis, this significant drop would necessitate immediate follow-up and validation with the Smiley team before drawing any definitive conclusions or recommendations.

I generated several initial charts: top 10 rankings, distributions, and volume plots to better understand the behaviors within each table.

## Part 4 & 5 final

**Interpret:** After building individual exploratory views, I constructed a unified analytical table using SQL joins directly within Colab. This consolidated dataset made it possible to produce the advanced visualizations and insights showcased in this slides.

In the **Model** phase of OSEMN, I developed two scoring frameworks: **APS (Agent Performance Score)** and **MPS (Market Performance Score)** which quantify performance using normalized metrics and weighted components. These models helped objectively rank markets and identify the top-performing agent within each market, enabling clearer strategic insights and more actionable recommendations for Smiley Real Estate.