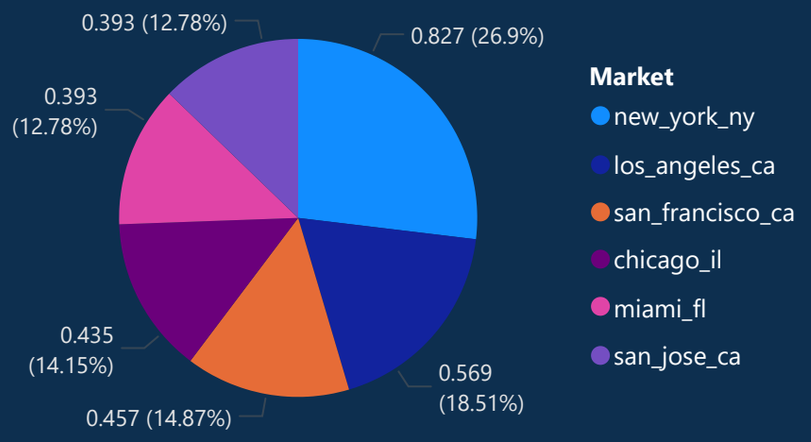
Best Markets Smiley Real Estate

Rank	Market ID	Total Listings Closed	Median Closed Price	Total Population	Composite Score
1	new_york_ny	13542	695000	8550405	0.83
2	los_angeles_ca	6606	725000	3898747	0.57
3	san_francisco_ca	2345	1200000	873965	0.46
4	chicago_il	8251	450000	2746388	0.44
5	san_jose_ca	1205	1050000	1021786	0.39
6	miami_fl	2876	600000	463347	0.39
7	washington_dc	3210	620000	705749	0.37
8	dallas_tx	6148	350000	1304379	0.37
9	atlanta_ga	3987	400000	498044	0.35
10	boston_ma	2876	700000	692600	0.34

Top 5 Real Estate Markets by Composite Score



Methodology

1) I looked at the market data to ensure all numbers were accurate and consistent. Looked for any errors. What i found: made sure the values in columns *price_new_list_median* and *price_closed_median* were consistent with the other column values.

2) My goal was to determine what was a fair way to identify top markets. I did not want to rely on one factor since we had many metrics to go based off. I decided the composite method was the best way to do that. I used python to calculate it for accuracy and speed.

- Calculation : *Listing_norm = Listing -Min(Listings)/ Max(Listings) - Min(Listings)* we apply this to our other factors as well (Population and Median Price) => *Composite Score = (Listing_norm+Population_norm+MedianPrice_norm)/3*

3) Lasty, the data was cleaned and analyzed and ready to be uploaded into Power Bi for visualization and insights.

Key Insights

Top Performing Market:

- ***New_York_NY*** is our top performing market with a composite score of (0.83). It has the highest amount of listings closed (13,542) and the biggest population (8.55M). Even though the median price is high ($695K), the large amount of transaction is what makes it our top performing market.

High Median Price Doesn't Guarantee Top Score:

- **San_Francisco_CA** and ***San_Jose_CA*** have the highest median closed prices ($1.2M and $1.05M), but their Composite Scores are lower (0.457 and 0.393).This shows us that a high median price by itself does not mean the market performs the best.

Emerging Markets with High Volume but Lower Prices:

- ***Dallas_TX*** and ***Los_Angeles_CA*** have significant transaction volumes (6,148 and 6,606 listings) with moderate closed prices ($350K and $725K).
- These markets perform well because of the mix of volume and population, not only looking at high prices.

Composite Score Highlights Overall Opportunity:

- Our composite scoring looks at closed listings, population and median price. Markets like ***Chicago_IL*** (0.44) and ***Miami_FL*** (0.39) are our steady performers across all areas meaning they do not lead in any category.

Best Agents Smiley Real Estate

Market: All

Agent: All

Agent	Market	Buyer Closes 2024	Seller Closes 2024	Average Rating 2024	Composite Score
Hassan White	las_vegas_nv	30	50	3.29	32.66
Robert Adams	fresno_ca	29	49	2.65	31.73
Latoya Lewis	denver_co	29	48	3.60	31.52
Mary Baker	boston_ma	28	48	4.99	31.40
Jamal Thompson	kansas_city_mo	28	49	2.87	31.37
Kenji Parker	buffalo_ny	30	47	2.61	31.32
Jessica Hall	phoenix_az	28	49	2.57	31.31
John Clark	washington_dc	29	46	4.98	31.00
Priya Tran	memphis_tn	29	47	2.61	30.92
Michael Hall	portland_or	29	47	2.30	30.86
Emily Patel	st_louis_mo	30	45	3.81	30.76
Sofia Mohamed	san_diego_ca	30	45	3.25	30.65
Malik Mitchell	detroit_mi	27	48	3.24	30.65
Daniel Smith	houston_tx	25	50	3.22	30.64
Mei Gonzalez	providence_ri	27	48	2.74	30.55
Isabella Johnson	indianapolis_in	29	45	4.65	30.53
Luis Green	richmond_va	26	49	2.60	30.52

Methodology

1) I looked at the agent data to ensure all numbers were accurate and consistent. Looked for any errors. What I found: I wanted to merge first and last names for a cleaner view on PowerBi, I used python for that.

2) My goal was to determine what was a fair way to identify top agents. I wanted to base the scoring for top agent by using the amount of deals closed with buyers and deals closed with sellers and the most important, average rating. I believe thats what makes them the best, having the best numbers across the board. I also wanted to make sure it was for each market, so i interpreted as we dont want two agents from New York, we want the best performer for that market. I determined using python was the most accurate way as well.

- Calculation :
- *Composite Score = (Buyer_closes_norm+Seller_closes_norm+Avg_rating_norm)/3*

3) Lasty, the data was cleaned and analyzed and ready to be uploaded into Power Bi for visualization and insights.

Key Insights

Top Performing Agents:
Hassan White is our strongest agent with a composite score (32.66) this mean he is closing with the largest number of both buyer and seller deals.

High Review Don't Equal Top Performance:
Agents with excellent rating (4.8-5.0) are not our top rankers because when looking at their total closing the are lower. In conclusion, this shows sales volume has a greater impact on overall success than just looking at customer ratings.

Balanced Closings:
Here we see that agents who maintained a balanced performance across both buyer and seller transaction are our top agents. This balance mean higher composite scores.

Composite Score Overall Value:
Our composite score is looking at buyer closes, seller closes, and review rating. Agents who are leaning towards the top are strong among most categories not just one. This ensures us we are looking at the bigger picture - reliable and well-rounded data.

Flow Chart Idea: Customer Call Activity and Sales Strategy

Start
We analyze all customers , look for any patterns that might stand out.

Identify Customers
During this step we are looking at call and sales data. We want to group customers by how recent they've made contact with our call handlers.

Call History
We want to focus on the frequency of the calls.

Multiple calls by call handler = High Engagement
Few Calls by call handler = Low Engagement

Sales Outcome
Here we want to focus whether or not they made sale when the calls were made?

Yes = the selling agent earns 2% of the home's value, and Smiley Real Estate earns 1% of the home's value

No= We consider an alternative strategy.

Next Strategy
We would want to focus on potential customers based on past engagement history, especially high potential customers.

Outcome
Work with marketing to come up with a better strategy for targeted calls to increase sales.

More Time?

If I had more time I would've wanted more focus on the area for growth. There is so much opportunity for different types of insights depending what the company wants focus on and the priorities of the team.

1) Merge Data Sources
I would merge calls , sales, and customer datasets using the customer_id. I would be able to track the customers journey. For example, how many times did it take to call customer_id 15? and if there was any sale for the days of those calls? Additionally having more detailed customer data would give us a better understanding of the customer's profile and behavior.

2) Analyze with more detail
With this detailed data set we are able to perform deeper analysis. We are able to see trends in regions, our customer type and customer history. This also includes sales amounts, we would be able to understand how much agents are earning and Smiley Real Estate earning. This would help many teams identify where the company focus needs to go.

3) Visualize patterns
With more time I would have liked to created visualizations clearly showing any patterns found. I believe these visualizations would bring a clearer picture to what needs focus especially dealing with such large data sets.

Process Overview

My process started with exploring the data to understand what i had and how i could use it to answer the questions about top agents and best markets. I focused on highlighting key insights that different teams could easily understand.

Tools Used
I primarily used Python for data cleaning, calculations and creating composite scores, VS Code for code editing Power BI for my visualization tool. I did use AI in order to guide me through my coding but always maintaining the integrity of my ideas.

Methodology
My main question here was how can I be fair when looking at all my data to determine who were the best agents and markets? how can I look at the "bigger picture"? i wanted to make sure I used multiple factor rather than just one metric. I used the composite method which uses different measures into a single score and from there rank them.

Challenges Faced
The toughest challenge was not knowing my audience and not having prior knowledge of the data I was working with it but at the same time it was fun figuring out what I could analyze from what I had. What helped throughout the process was thinking about past projects, continuously checking in with myself, is this too much information? is it not enough? I felt like my analysis and visualizations were straight to the point and recommendations that could be acted on tomorrow.