

# STATISTICAL DATA ANALYSIS FOR OPTIMIZED REAL ESTATE MARKETING STRATEGIES IN CALIFORNIA: DESCRIPTIVE STATISTICAL ANALYSIS

Leveraging Statistical Insights to Enhance  
Targeted Marketing Campaigns and  
Maximize Sales Efficiency

---



Rutvik Trivedi

# PROJECT OVERVIEW

---

- Problem Statement
- Understanding the Data
- Statistical Data Analysis and Visualization
- Statistical Calculations
- Insights
- Recommendations

# PROBLEM STATEMENT

---

A real estate company operating in California aims to **improve its marketing results** by targeting the right audience based on **data-driven decisions**.

The goal is to **enhance marketing efficiency** without increasing the budget through a **descriptive statistical analysis** of the customer and sales data.

# UNDERSTANDING THE DATA

---

## Dataset

The data for this analysis comes from the company's California real estate transactions database, containing various attributes related to sales, properties, and customer demographics.



## Statistical Frequency Distribution

Generation of frequency distribution tables for categorical variables. It helps visualize the data more efficiently and accurately.



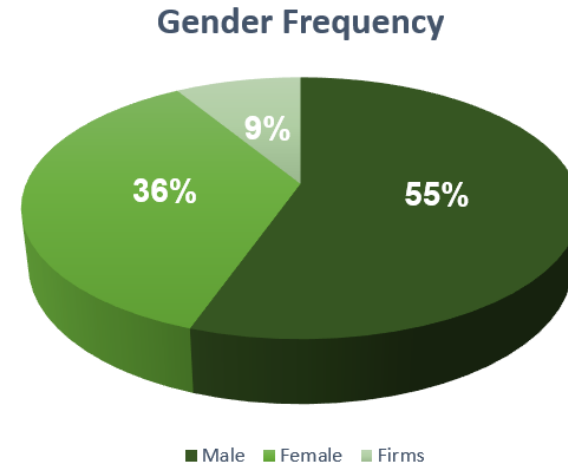
# UNDERSTANDING THE DATA

	A	B	C	D
1	Column Name	Category	Level of Measurement	
2	ID	Numerical	Quantitative	
3	Building	Numerical	Quantitative	
4	Year of sale	Numerical	Quantitative	
5	Month of sale	Numerical	Quantitative	
6	Type of property	Categorical	Qualitative	
7	Property #	Numerical	Quantitative	
8	Area (ft.)	Numerical	Quantitative	
9	Price	Numerical	Quantitative	
10	Status	Categorical	Qualitative	
11	Unnamed: 10	Numerical	Quantitative	
12	Customer ID	Categorical	Qualitative	
13	Entity	Categorical	Qualitative	
14	Name	Categorical	Qualitative	
15	Surname	Categorical	Qualitative	
16	Age at time of purchase	Categorical	Qualitative	
17	Interval	Categorical	Qualitative	
18	Y	Numerical	Quantitative	
19	M	Numerical	Quantitative	
20	D	Numerical	Quantitative	
21	Gender	Categorical	Qualitative	
22	Country	Categorical	Qualitative	
23	State	Categorical	Qualitative	
24	Purpose	Categorical	Qualitative	
25	Deal satisfaction	Numerical	Quantitative	
26	Mortgage	Categorical	Qualitative	
27	Source	Categorical	Qualitative	

# GENDER STATISTICS AND INSIGHTS

E7    =D7/\$D\$10

	A	B	D	E	F
1	<b>Real Estate California Database</b>				
2	<b>Gender</b>				
3					
4	<b>Frequency distribution table</b>				
5					
6		<b>Frequency</b>		<b>Relative frequency</b>	
7	Male	108		55%	
8	Female	70		36%	
9	Firms	17		9%	
10	<b>Total</b>	<b>195</b>		<b>100%</b>	
11					

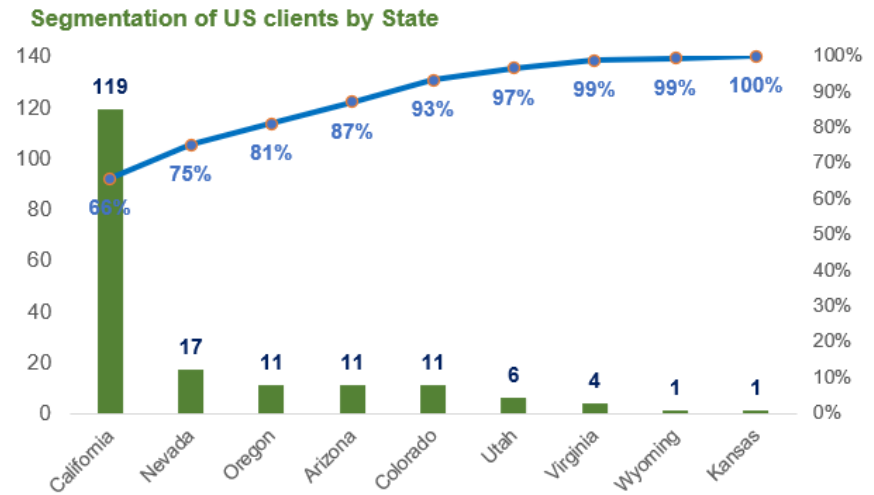


- **Males** are more likely to sign the contract and are potentially a better audience to target for our ads.
- Approximately **10%** of the sales came from the firms or the **corporate clients**

# LOCATION STATISTICS AND INSIGHTS

F7     $=C7/(\$C\$17-\$C\$16)$

	Frequency	Relative frequency	Cumulative frequency	Cumulative US only
California	119	45%	45%	66%
Nevada	17	6%	51%	75%
Oregon	11	4%	55%	81%
Arizona	11	4%	59%	87%
Colorado	11	4%	63%	93%
Utah	6	2%	66%	97%
Virginia	4	1%	67%	99%
Wyoming	1	0%	67%	99%
Kansas	1	0%	68%	100%
None (abroad)	86	32%	100%	
Total	267	100%		

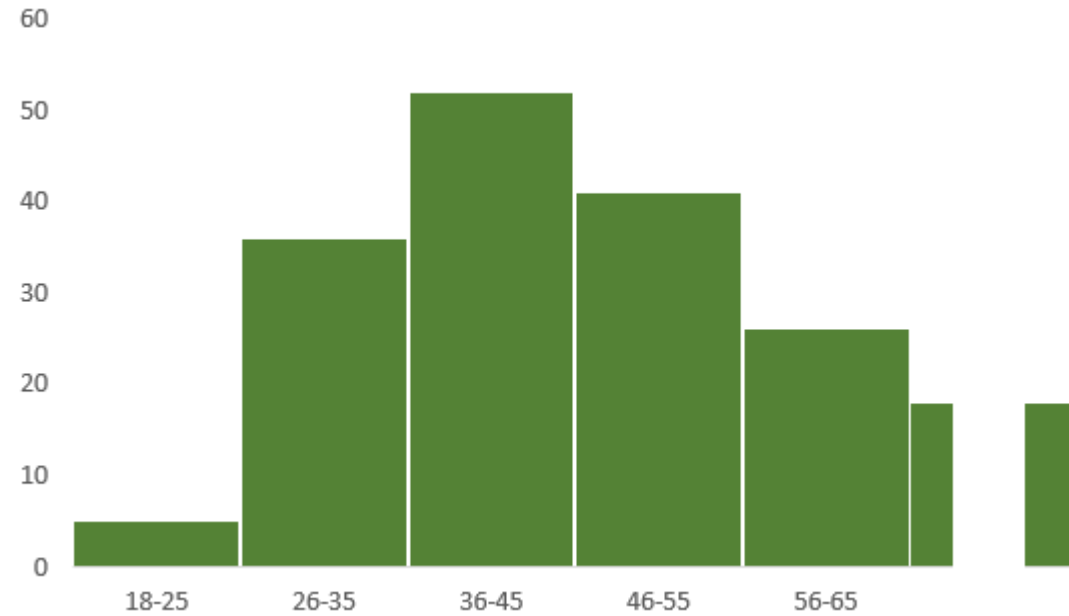


- **68%** of the sales came from California with Nevada, Oregon, Arizona, and Colorado forming **93%** of the US customer base.



# AGE STATISTICS AND INSIGHTS

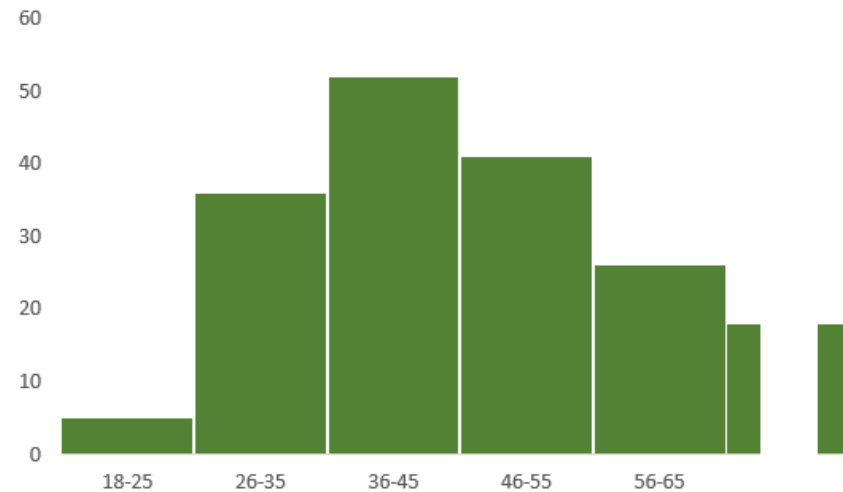
	A	B	C	D	E
1	<b>Real Estate California Database</b>				
2	<b>Age</b>				
3					
4	<b>Frequency distribution table</b>				
5					
6	<u>Frequency</u> <u>Relative frequency</u>				
7	18-25		5		3%
8	26-35		36		20%
9	36-45		52		29%
10	46-55		41		23%
11	56-65		26		15%
12	65+		18		10%
13	<b>Total</b>		<b>178</b>		<b>100%</b>
14					
15	<b>Mean</b>		46.15		
16	<b>Median</b>		45.00		
17	<b>Mode</b>		48.00		
18	<b>Skew</b>		0.24		
19	<b>Variance</b>		164.91		
20	<b>St. dev.</b>		12.84		
21					
22					



- **68%** of the sales came from California with Nevada, Oregon, Arizona, and Colorado forming **93%** of the US customer base.

# AGE STATISTICS AND INSIGHTS

---



- **71%** of sales were made with customers aged between **26 and 55 years** old, with a mean of **46 years** and a standard deviation of **13 years**.
- **Younger people** buying more properties than older people

# AGE AND PRICE STATISTICS

The screenshot shows an Excel spreadsheet with the following content:

	A	B	C	D	E	F
1	<b>Real Estate California Database</b>					
2	Relationship between age and price					
3						
4	<u>Age and price</u>					
5						
6	Covariance		-176361.87			
7	Correlation coefficient		-0.17			
8						
9						
10						
11						
12						

The formula bar at the top shows the formula for cell C6: `=COVARIANCE.S(Data!I6:I272,Data!P6:P272)`.

# AGE AND PRICE STATISTICS



- There is **no relationship** between the customer's age and the price they are willing to pay.

# RECOMMENDATIONS

---

- **Demographics:** Focus on males and age groups that have shown higher purchasing rates (26 to 55 years).
- **Corporate Clients:** Consider targeting firms or corporate clients as they make up approximately 10% of the sales.
- **Geographical Focus:** Concentrate marketing efforts in California, Nevada, Oregon, Arizona, and Colorado to cover 93% of the US customer base.
- **Property Type:** Highlight apartments in the campaign, as they are the most sold property type.
- **Timing:** Increase marketing efforts towards the end of the year to leverage seasonal buying trends.
- **Price Range:** Emphasize mid-range properties, which align with the majority of sales.