An analysis of a real estate company based in the United States, operating in California, as it prepares to launch a marketing campaign aimed at a targeted audience.

Introduction

- In this project, I conducted an in-depth examination of a real estate company's database located in the United States. I performed statistical analysis on the data using Excel. For confidentiality purposes,customer names have been altered.

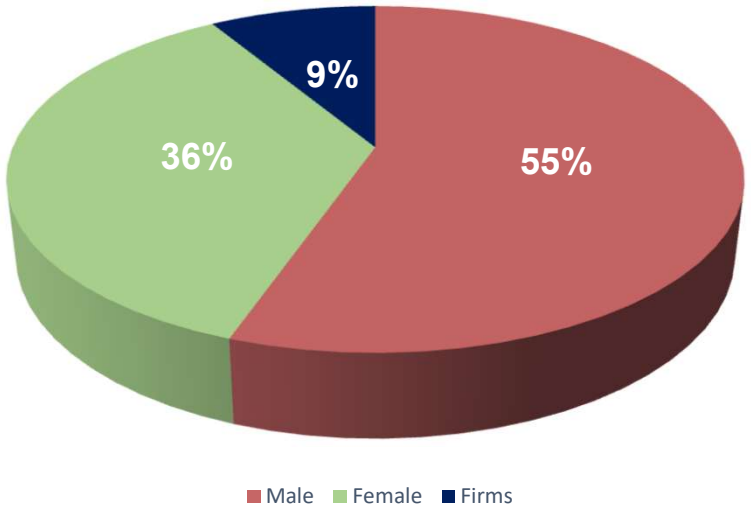
- The company is preparing to launch a marketing campaign with the goal of effectively reaching its target audience. The management believes that, through a brief analysis, improvements can be made without the necessity of investing additional resources.

- As a Data Analyst, my role is to analyze the data and determine which demographic groups are more likely to purchase our product. Once this analysis is completed, I will guide the marketing team in directing their efforts toward these specific groups

Gender

Frequency distribution table

	Frequency	Relative frequency
Male	108	55%
Female	70	36%
Firms	17	9%
Total	195	100%



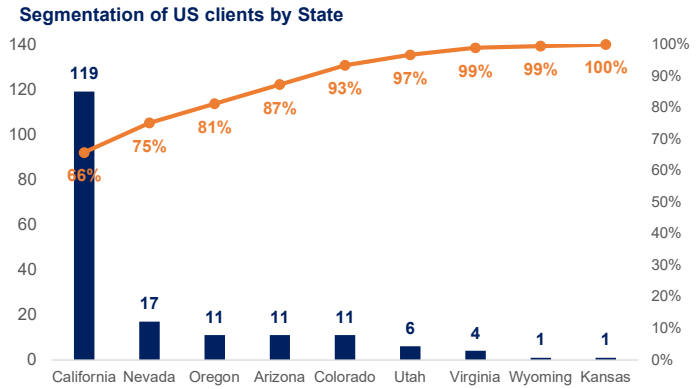
Conclusion

In the pie chart, it's evident that the majority of customers are males; however, this data is biased as it only represents the individuals who signed the contracts. It's highly probable that families purchase the apartments, but the data solely accounts for the contract signatories.

Location

Frequency distribution table

	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%		



Conclusion

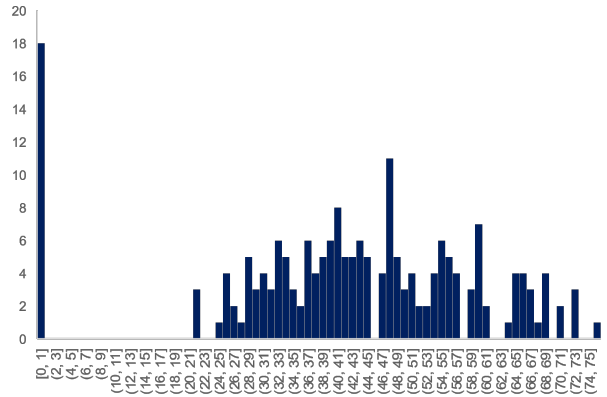
The Pareto chart illustrates that the majority, or 81%, of our customers originate from California, Nevada, and Oregon. As a result, these regions are our primary target locations.

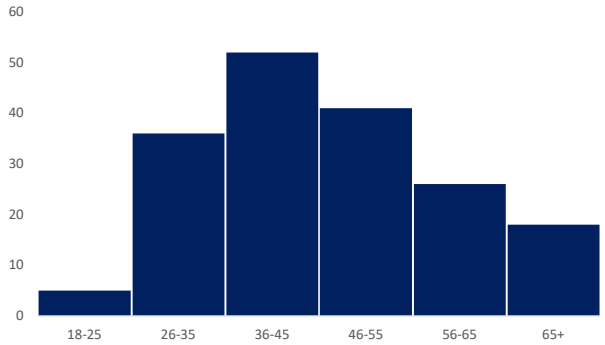
AGE (Central variable in our analysis since people rearly buy Real Estate more than once in their life.)

Frequency distribution table

	Frequency	Relative frequency
18-25	5	3%
26-35	36	20%
36-45	52	29%
46-55	41	23%
56-65	26	15%
65+	18	10%
Total	178	100%

Mean	46.15
Median	45.00
Mode	48.00
Skew	0.24
Variance	164.91
St. dev.	12.84





Conclusion

- The first graph represents the frequency distribution of the variable age. We have a high bar at 0 as that is the null values associated with corporate clients.
- The histogram illustrates that the majority of customers fall within the age range of 26 to 65, with an average age of 46 years. Specifically, a significant portion of customers are in the 36 to 45 years age group. It does not include firms data.

Relationship between age and price

Age and price	
Covariance	-176361.87
Correlation coefficient	-0.17

Conclusion

The scatter plot demonstrates that Age and Price are not correlated, as there is no discernible pattern or relationship between them

Result

- Men have a higher likelihood of signing contracts and may represent a more promising audience for our advertisements (though this conclusion is not entirely clear).
- California accounted for 68% of our sales, while the combined customer bases of Nevada, Oregon, Arizona, and Colorado made up 93% of our customers in the United States
- 71% percent of the sales were conducted with customers aged between 26 and 55, with an average age of 46 years and a standard deviation of 13 years. It is notable that younger individuals tend to be more active in purchasing real estate than their older counterparts.
- There is no clear relation between the age of the provided customers and the price they are willing to offer.