

Real State Analysis

1. Data Preparation and cleaning:

Importing necessary libraries (Pandas, NumPy, Matplotlib, Seaborn).

Reading and preprocessing the customers' and properties' data.

Handling missing values, duplicates, and data types.

Creating additional columns for analysis.

2. Statistical Analysis:

Generating descriptive statistics for numeric data.

Analyzing totals and averages by building and country.

Examining frequency distributions by state and age intervals.

Investigating correlations between age and price.

3. Data Visualization:

Creating visualizations to illustrate insights

Deal satisfaction across countries (Bar chart)

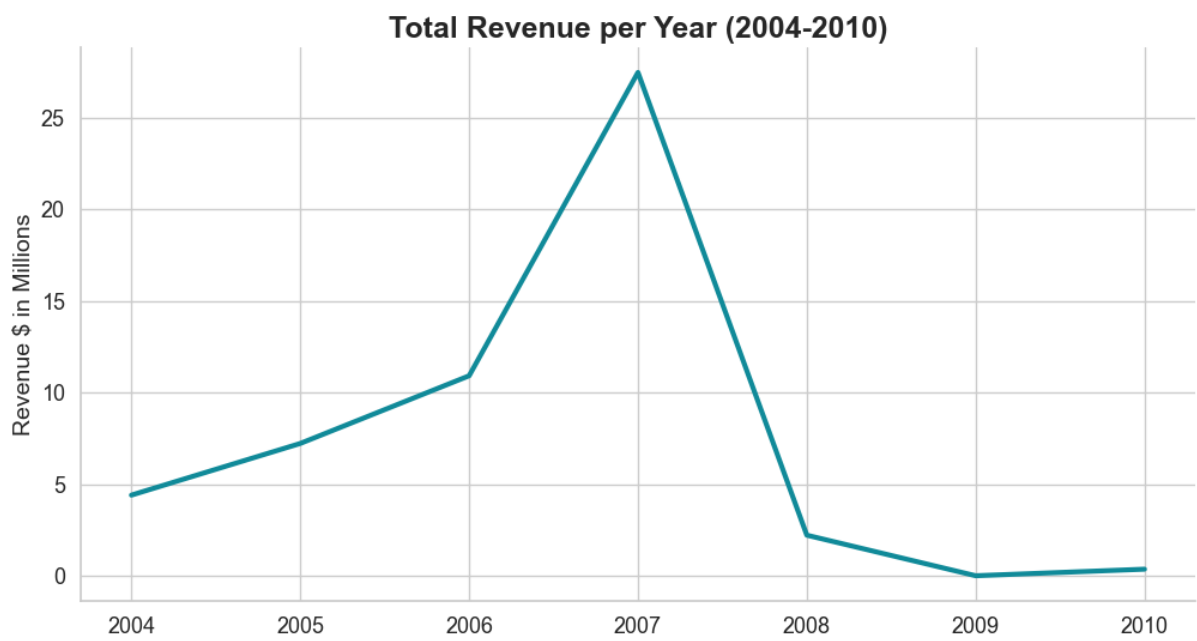
Customer age distribution (Histogram)

Segmentation by state (Pareto diagram).

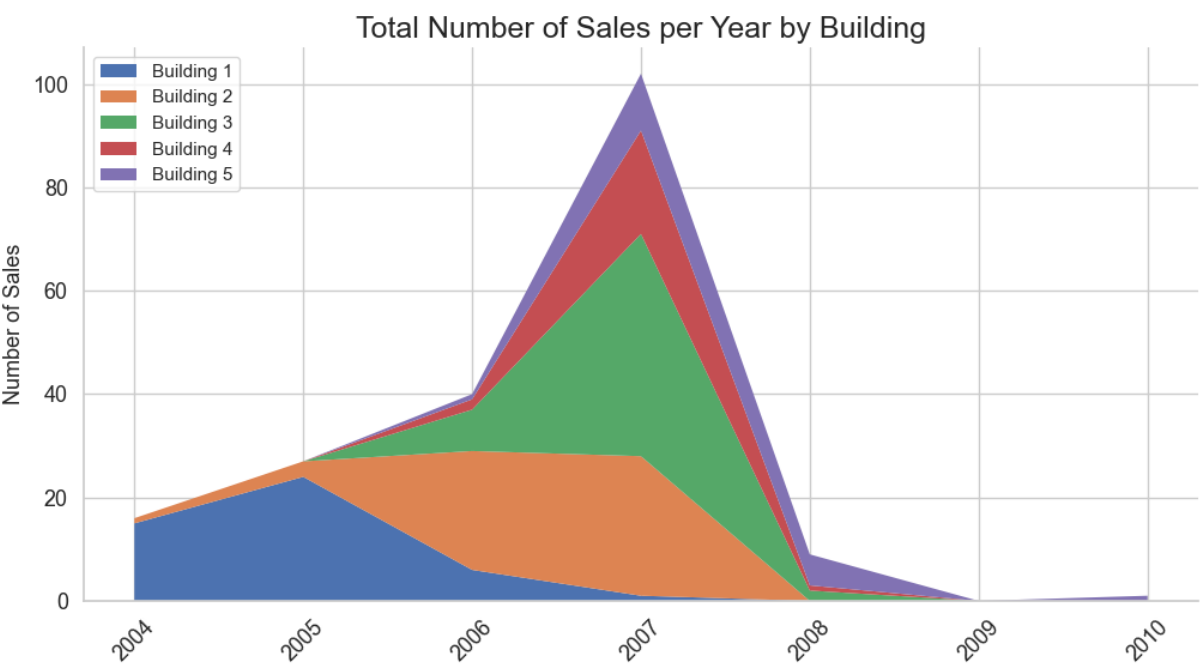
Total sales per year (Line chart).

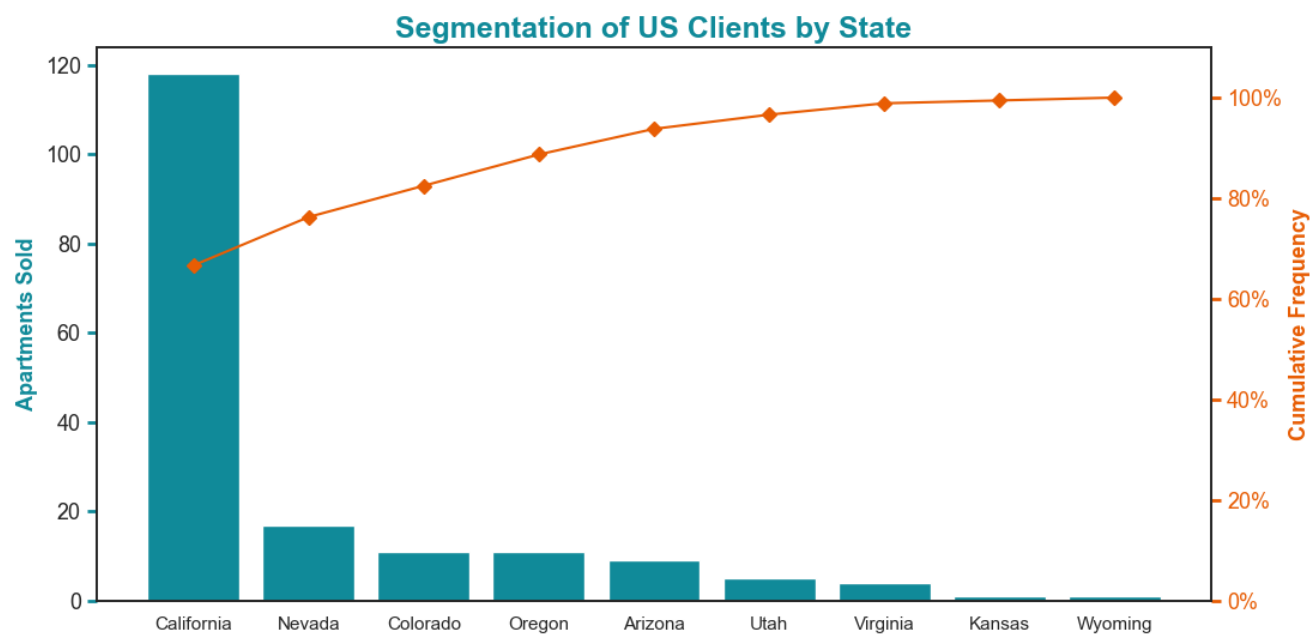
Yearly sales distribution across buildings (Stacked Area Chart).

Yearly Revenue Trends

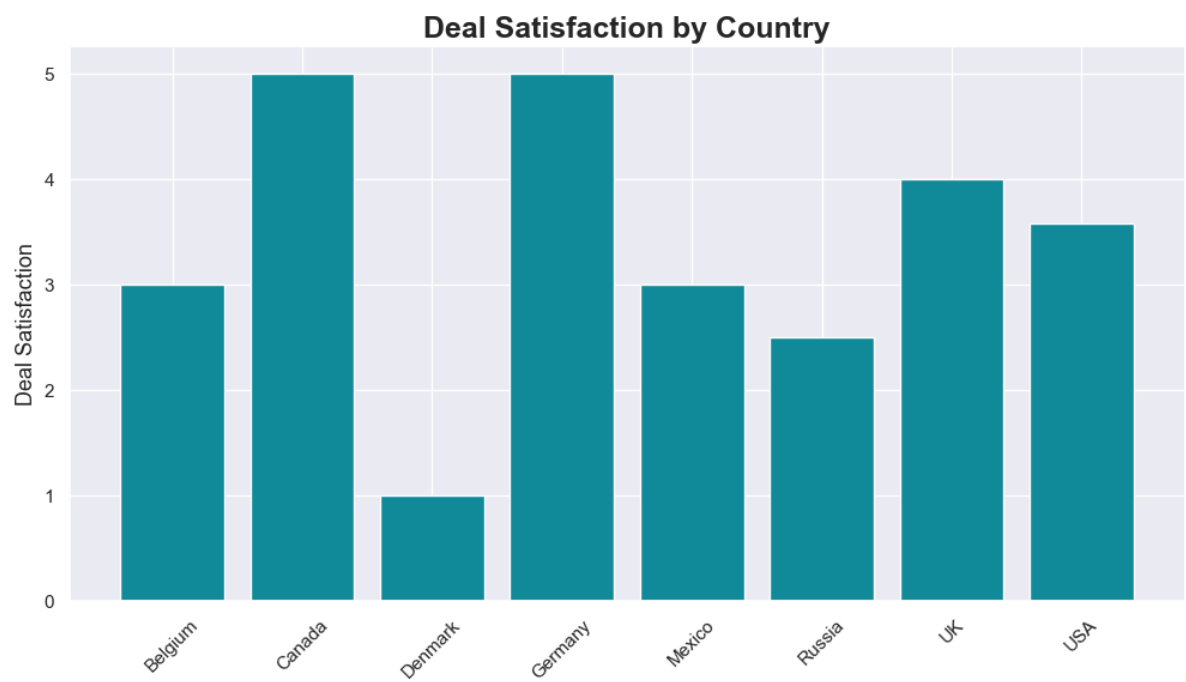


Sales per building

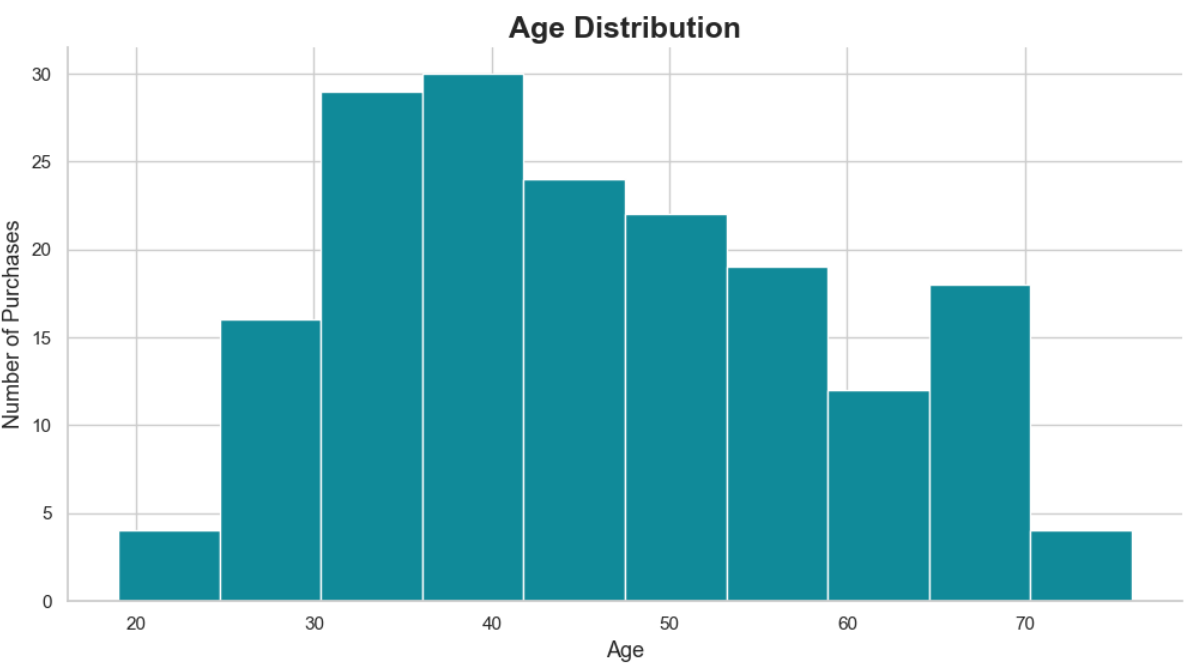




Average Deal satisfaction by country



Age Distribution



Key Insights:

The most sold building is Building 2, and it has the highest mortgage.

Customers aged between 36 and 42 tend to buy more buildings.

Buildings priced around \$201,705.60 and \$243,776.37 are sold the most.

There is a weak negative correlation between the price of the building and the age of purchase.

Canada and Germany have the highest deal satisfaction.

California has the highest number of sold properties in the USA.

The year 2007 generated the highest revenue.