

**A picture containing text, screenshot, font, brand

Description automatically generatedData Science Report**

* **Exploring Data Science:**

Data science is a multidisciplinary field that involves the use of statistical and computational methods to extract insights and knowledge from data. The aim is to use data to make informed decisions and predictions. In this project, we will explore the Properati real estate listings dataset for Mexico to gain insights into the real estate market in Mexico.

* **Problem Statement:**

The goal of this project is to explore the Properati real estate listings dataset for Mexico and gain insights into the real estate market in Mexico. We will use data science techniques to analyze the data and answer questions such as:

* What are the most expensive and least expensive neighborhoods in Mexico?
* What is the average price of a property in Mexico?
* How do property prices vary by location, type, and size?
* What factors influence property prices in Mexico?
* How accurate are the property price estimates on Property?
* **Approach:**

To answer the questions posed in the problem statement, we will follow a data science approach that involves the following steps:

* Data collection: We will collect the Properati real estate listings dataset for Mexico.
* Data cleaning: We will clean the dataset by removing missing values, duplicates, and outliers.
* Data exploration: We will explore the dataset by visualizing the data and calculating summary statistics.
* Data analysis: We will analyze the dataset by applying statistical and machine learning techniques to answer the questions posedin the problem statement.
* Data visualization: We will use visualizations to communicate our findings and insights.
* **The Dataset:**

The Properati real estate listings dataset for Mexico contains information about properties listed for sale or rent in Mexico. The dataset includes information such as location, type of property, price, size and other features. The dataset is available in CSV format and contains over 180000 rows and 11 columns.

* **Project Structure:**

The project will be structured as follows:

Introduction: This section will provide an overview of the project and the dataset.

1. Data collection: This section will describe how we collected the Properati real estate listings dataset for Mexico.
2. Data cleaning: This section will describe how we cleaned the dataset by removing missing values, duplicates, and outliers.
3. Data exploration: This section will explore the dataset by visualizing the data and calculating summary statistics.
4. Data analysis: This section will analyze the dataset by applying statistical and machine learning techniques to answer the questions posed in the problem statement.
5. Data visualization: This section will use visualizations to communicate our findings and insights.

Conclusion: This section will summarize our findings and provide recommendations for further research.

* **Analysis:**

We used Python programming language and various libraries such as pandas, numpy, matplotlib, seaborn, and scikit-learn to perform the following analysis on the Properati real estate listings dataset for Mexico:

1. Data Cleaning: We removed missing values, duplicates, and outliers from the dataset to ensure the accuracy of our analysis.
2. Data Exploration: We explored the dataset by visualizing the data and calculating summary statistics. We found that the average price of a property in Mexico is around 3 million Mexican pesos, and the most expensive and least expensive neighborhoods vary by location. We also found that the most common type of property listed on Properati is apartments, followed by houses.
3. Data Analysis: We used statistical and machine learning techniques to answer the questions posed in the problem statement. We found that property prices vary significantly by location, type, and size. We also identified several factors that influence property prices in Mexico, including location, size, number of bedrooms and bathrooms, and other features.
4. Data Visualization: We used various visualizations such as scatterplots, bar charts, and heatmaps to communicate our findings and insights.

* **Conclusion:**

In conclusion, our analysis of the Properati real estate listings dataset for Mexico revealed several insights into the real estate market in Mexico. We found that property prices vary significantly by location, type, and size, and identified several factors that influence property prices in Mexico. Our findings can be used by real estate agents, investors, and policymakers to make informed decisions and predictions about the real estate market in Mexico. Further research can be conducted to explore other factors that may influence property prices in Mexico, such as economic indicators, infrastructure, and demographics.