Checkpoint I: Project Proposal

Group : G13
Date: 2021/10/02

Domain

Real Estate Market in Portugal, depending on several attributes like location, area, condition, number of rooms, bathrooms and other aspects in March 2021, as well as the criminality rates in their cities.

Dataset

- We are going to use two datasets, one about the Real-Estate Market in Portugal in March 2021 and the other about each City's criminal rates in 2020.
- The first dataset will be obtained from the website Kaggle. The dataset is named *Portugal properties – rent, buy and vacation*. This dataset has multiple attributes, which are Locations, Rooms, Price, Area, Bathrooms, Condition (if it's used, new or renovated), the Advertisement type (rent, buy or vacations) and the Property type (house or apartment).
- The second dataset will be obtained from the website Instituto Nacional de Estatística. We selected the Criminal rate (‰) as the indicator and the year 2020, all cities in Portugal and all types of Crimes (including the total) as dimensions. The attributes of this dataset are the City, the number of all types of Crimes, as well as the number of Crimes against physical integrity, Theft/robbery, Carjacking, Drunk driving, Driving without license and Crimes against national heritage.

Example Questions

- **Question 1:** How does the area of a property influence its price?
  - With this question we're hoping to be able to analyze if there's a notable relationship between the area of the property and the asking price of it.
  - To achieve this, we will use some attributes from the properties dataset, which are the Area of the property and the Price that's being asked for it.
- **Question 2:** With similar characteristics, how do the rent and selling price of properties compare?
  - This question refers mainly to Real Estate investors. It allows them to know an estimate of how much they can earn (monthly) from their investments.
  - To achieve this, we will use the properties dataset to analyze several of the attributes that directly relate with the topology of the property, as well as play with the type of ad it is (whether it is for sale or rent) and its asking price.
- **Question 3:** Given the same characteristics, how do the rent/selling prices of one property in a specific location compare to the prices of another property in a different location?
  - This question is particularly useful to help users compare the renting/selling cost of a property they wish in different locations.
  - To achieve this, we will use the different characteristics of a property and the place where it is located from the properties dataset.
- **Question 4:** Within a certain budget, what type of commodities can you expect from an apartment/house in a specific location?

- This is used to let users know what to expect when searching for a property in a certain location and with a specific budget.
- This is done by making an average of every condition in an apartment or house.

- **Question 5:** What is the neighbourhood within a city that has the cheapest average monthly rent?
  - This is a good feature to have if a user does not have a preference on the neighbourhood that he wants to live in, so it can help him decide where the cost is cheaper, or cheaper areas that may be associated with high-risk zones (high criminality, loud environment, etc.), for example.
  - To achieve this, we need to calculate the average cost of a monthly rent within all the neighbourhoods of a county and present them to the user.

- **Question 6:** How does the criminality rate in each city influence the prices of the properties?
  - With this question we're hoping to be able to analyze if the price of a property and the criminality in its city are related.
  - To achieve this, we will join the attributes of the location of the properties from the properties dataset with the crime-rate dataset and relate it to the price of the properties.

Data Sample

Question 1:

```
(from "portugal_ads_proprieties.csv")
Locations; Area; Price
Portimão, Faro; 85.0; 1150.0
Portimão, Faro; 30.0; 500.0
Portimão, Faro; 50.0; 600.0
```

Question 2:

```
(from "portugal_ads_proprieties.csv")
Locations; Rooms; Bathrooms; Condition; PropertyType; AdsType; ProprietyType; Price
Cascais e Estoril, Cascais, Lisboa; 1; 1; Used; Sell; Apartment; 155000.0
Cascais e Estoril, Cascais, Lisboa; 1; 1; Used; Rent; Apartment; 1200.0
```

Question 3:

```
(from "portugal_ads_proprieties.csv")
Rooms; Area; Bathrooms; Condition; AdsType; ProprietyType; Locations; Price
3; 117.8; 2; New; Sell; Apartament; Abragão, Penafiel, Porto; 130000.0
3; 140.0; 2; New; Sell; Apartament; Abraveses, Viseu; 200000.0
```

Question 4:

```
(from "portugal_ads_proprieties.csv")
Price; Locations; ProprietyType; Rooms; Bathrooms; Area; Condition; AdsType
700.0; Glória e Vera Cruz, Aveiro; Apartament; 1; 2; 59.0; New; Rent
```

Question 5:

```
(from "portugal_ads_proprieties.csv")
Locations; Price
Portimão, Faro; 220.0
```

Question 6:

```
(from "portugal_ads_proprieties.csv")
Locations; Price
Cacia, Aveiro; 650.0
Fernão Ferro, Seixal, Setúbal; 550.0
(from "portugal_criminality_rates.csv")
Localização geográfica (NUTS - 2013); Total
Aveiro; 28,8
Seixal; 23,3
```