**Tunisia House Prices**

**Dataset Report**

**Motivation**

Due to the lack of datasets that contain information about Tunisia, I have decided to start a series of datasets covering various constructive Tunisian data; starting with this house pricing data set.

The creation of this dataset is an attempt to provide helpful information on house pricing in Tunisia, covering thousands of online listings provided with a considerably fair amount of information on each listing. A rather interesting aspect of the data is its geographic accuracy, this enables a lot of creative analysis opportunities.

**Context**

This dataset is the result of scraping "various" house sales websites. It contains more than 8000 observations with 25 attributes (excluding the id attribute). All records correspond to a house, an apartment, or a villa within the country of Tunisia.

All data was scraped in October 2022 and was processed to provide accurate data. It serves as an excellent introduction to implementing machine learning since it requires little to no data cleaning, a considerable size of data, and an understandable list of attributes.

You will find bellow two versions of this dataset one containing rows missing the prices and one with only the rows where the price is available.

**Content**

The dataset’s columns are as follows, their names are pretty much self-explanatory:

* **id**: 128-bit universally unique identifier.
* **price\_tnd**: the price of a given listing in Tunisian Dinars.
* **price\_eur**: the price of a given listing in Tunisian Euros. This was converted on 29 October 2022 with a conversion rate of 1 TND equals 0.31 euros.
* **location**: the neighborhood of a given listing.
* **city**: the city of a given listing.
* **governorate**: the governorate of a given listing out of 24 Tunisian governorates.
* **area**: the area of a giving listing in square meters.
* **room**: the number of rooms.
* **bathroom**: the number of bathrooms.
* **age**: the age of a listing provided in intervals with the value 0 indicating it's new.
* **state**: the state of a listing provided in 3 different values; 1 indicating it's in a normal state, 2 indicating it requires renovation and 0 indicating it's brand new.
* **latt**: latitude.
* **long**: longitude.
* **distance\_to\_capital**: the distance between a listing and the capital Tunis.
* **garage**: boolean, indicating the existence of a garage.
* **garden**: boolean, indicating the existence of a garden.
* **concierge**: boolean, indicating the existence of a concierge.
* **beach\_view**: boolean, indicating whether the listing has a beach view or not.
* **mountain\_view**: boolean, indicating whether the listing has a mountain view or not.
* **pool**: boolean, indicating the existence of a pool.
* **elevator**: boolean, indicating the existence of an elevator.
* **furnished**: boolean, indicating whether the listing is furnished or not.
* **equipped\_kitech**: boolean, indicating whether the listing kitchen is equipped or not.
* **central\_heating**: boolean, indicating whether the listing has central heating or not.
* **air\_conditionting**: boolean, indicating whether the listing has air conditioning or not.