**Real Estate Price Prediction**

In this use case, we would like to create a machine-learning model that can predict the price of a house/apartment based on multiple criteria.

The exercise will focus on the Dubai real estate market and multiple data has been sourced to help you build this model.

### Data

The data you are given is the following:

* *Gdp\_Quarterly:* GDP per sector until Q3 2023
* *Gross\_Domestic\_Product\_At\_Constant\_Prices*: Annual GDP per sector until 2022
* *Consumer\_Price\_Index*: monthly CPI per sector from 2015 till July 2024
* *EIBOR\_2015\_2024* : EIBOR rates from 2015 till 2024
* Transactions\_training\_1: Real estate Transactions from 2023 till 2024. To be used for training. Target variable: “Actual\_Worth”
* Transactions\_training\_2: the training set has been split into 2 as it is voluminous.
* Transactions\_Training\_1 and Transactions\_Training\_2 need to be concatenated to have the complete training set.
* Transactions\_test : Real estate Transactions in 2024. This is the test data you will have to score once you build the model

### Guidance:

### Can you describe the type of data you have in the different data sets? (Distribution, missing value, extreme values, correlation, etc.)

### What features you will include/build in your training data set?

### How many years of data you will use for your training?

### Is this a classification model or a regression model? What models/algorithms will be suitable for such task?

### What KPIs will you use to select the best model?

### Can you explain the models you will be using and highlight the main difference?

### How useful was having a GPU in building your model?

### Scoring:

### By 2:30pm (on the day of the Hackathon), a new file “Transactions\_test\_with\_score” will be released and it will have the prices of the units you have scored.

### You will need to use this to calculate the Absolute Average Error and provide it to the judges.