

Business Understanding

Made by Timofey Sedov [@moflotas](#) as a part of the **Data Mining course**

Determine business objectives

Background

Company under domain name [aqar.fm](#) is a real estate company that provides a platform for real estate agents to advertise their properties for sale or rent. The company is based in Saudi Arabia and has a website and mobile application.

Business objectives

The primary business objective is to predict the price of a property based on its features to automate the process of pricing properties for sale or rent. The related business questions are:

- What are the most important features that affect the price of a property?
- How does the location of a property affect its price?
- How does the size of a property affect its price?
- etc.

Business success criteria

The criteria for a successful outcome to the project from the business point of view is to have a model that can predict the price of a property with a high accuracy. The accuracy of the model will be measured by the mean absolute error (MAE) and the root mean squared error (RMSE).

Assess situation

Inventory of resources

Since we provide offline analytics without the help of the company, the only available resources are

- My computing resources (Laptop with GPU, Google Colab)
- [Dataset](#)
- [Website](#)

Requirements, assumptions and constraints

The requirements of the project are:

- The project should be completed by the end of the course.
- The results should be understandable by the company.
- The results should be of high quality.

The assumptions made by the project are:

- The dataset is representative of the company's data.

- The dataset is clean and does not contain any errors.

The constraints on the project are:

- The dataset is huge and might be hard to process

Risks and contingencies

The event that might occur to delay the project or cause it to fail is crisis in the country with consequences on the real estate market causing the prices to change drastically. The contingency plan is to use the latest data available.

Terminology

So far there is no terminology relevant to the project.

Costs and benefits

The cost of the project is the time spent on it and the benefit is the time saved by the company.

Determine data mining goals

Data mining goals

The data mining goals are:

- Predict the price of a property based on its features.
- Identify the most important features that affect the price of a property.

Data mining success criteria

The criteria for a successful outcome to the project in technical terms is to have a model that can predict the price of a property with a high accuracy. The accuracy of the model will be measured by the mean absolute error (MAE) and the root mean squared error (RMSE).

Produce project plan

Project plan

The project plan is divided into six phases:

- Business understanding (2 weeks)
- Data understanding (2 weeks)
- Data preparation (2 weeks)
- Modeling (2 weeks)
- Evaluation (2 weeks)
- Deployment (2 weeks)

Initial assessment of tools and techniques

The tools and techniques used in the project are:

- Python and it's libraries (Pandas, Numpy, Matplotlib, Seaborn, Scikit-learn, etc.) for data analysis and visualization
- Google Colab and it's GPU support for faster model training
- Git and Github for version control