

Project Introduction:

- 1. Analyzing data of all real Estate deals executed in Madinah in 2019 and the scope of this project will focus on this estate types (Land , Apartment, House, Villa, apartment building) .

Data Cleaning and Enringing:

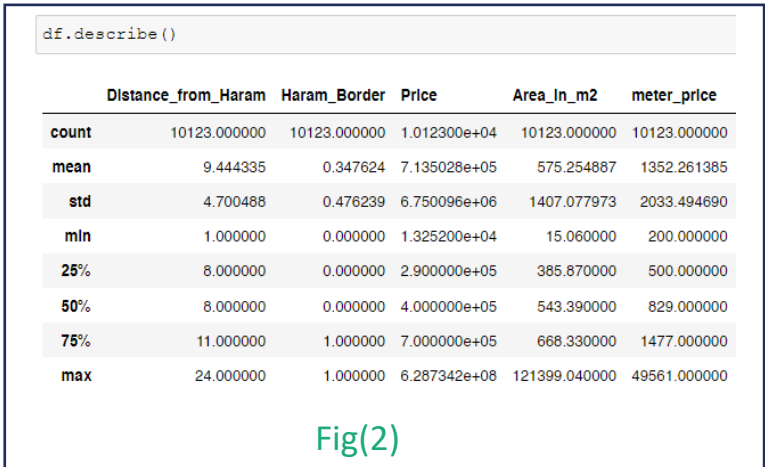
- 1. We check for null Value and rename columns data.
- 2. Reprocessing Neighborhood column and change name by deleting all the unwanted detail and unify it according to Madinah Municpilaty (from 749 unique value to 75 as shown in Fig(1))
- 3. After analyzing outliers in prices values, we find Significant Gap Between minimum and maximum price as describe in fig(2 , 3),so we delete all the rows where price above 10 million (15 rows).
- 4. We add 3 new feature, because it will give us a better understating for the data as the following:
 - 1. Distance from Haram (AL masjid Anabawi).
 - 2. Haram Border.
 - 3. Direction(North, South, East, West)

Model Selection:

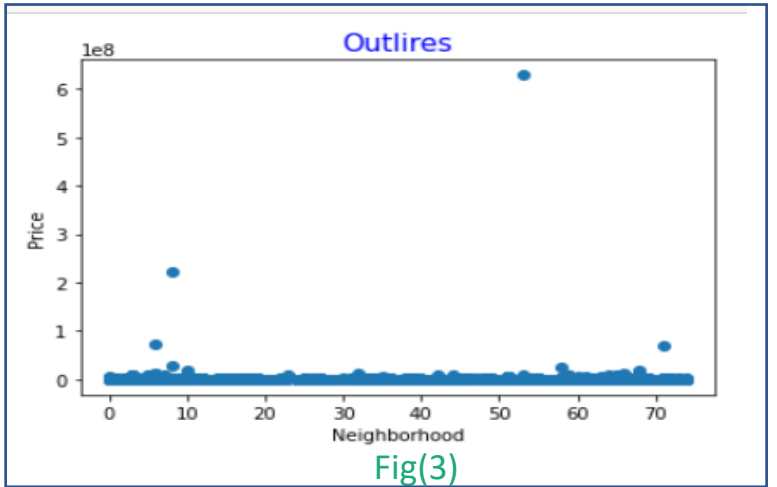
- 1. As a regression problem, we try with three model to evaluate (Leaner regression, Ridge, Polynomial).
- 2. The correlation between the Feature was low as the heat Map in fig(4)
- 3. The preliminary results of the scores wasn't good (around 30%), so will try to build XGBoost Regressor Model to evaluate more.



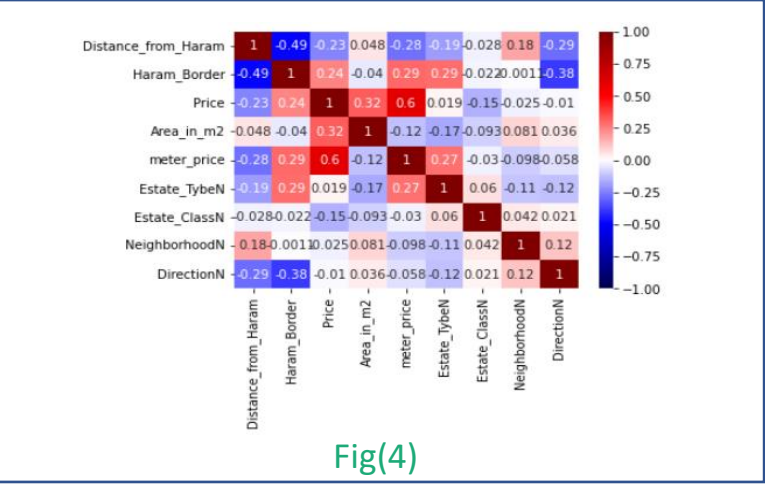
Fig(1)



Fig(2)



Fig(3)



Fig(4)