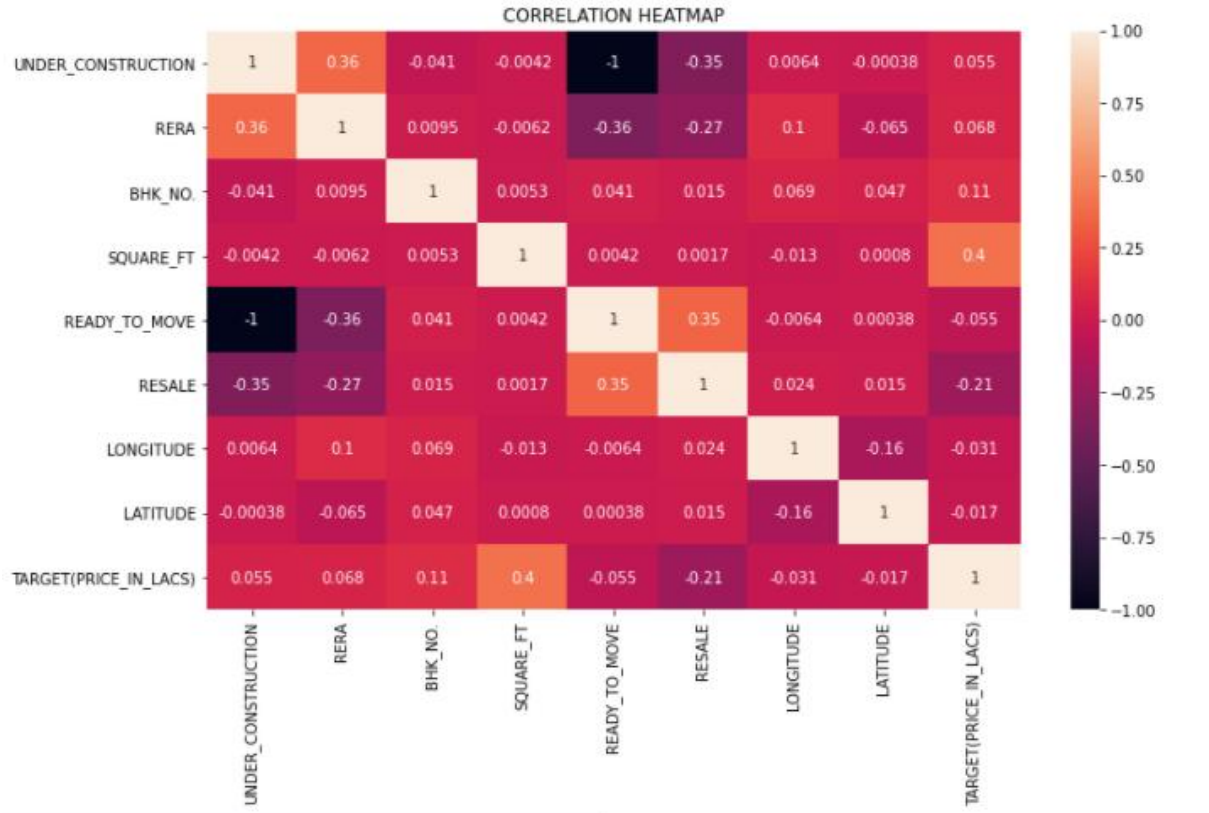
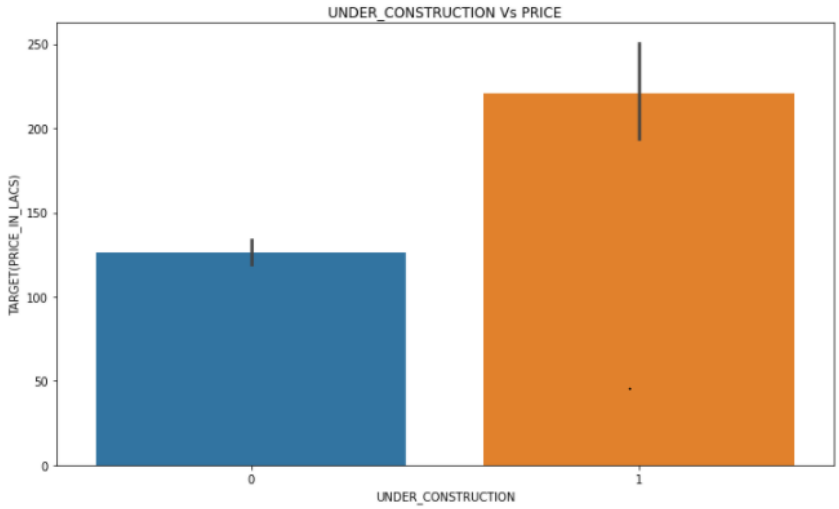
DATA ANALYSIS ON HOUSE PRICE PREDICTION

- **CORRELATION HEATMAP**



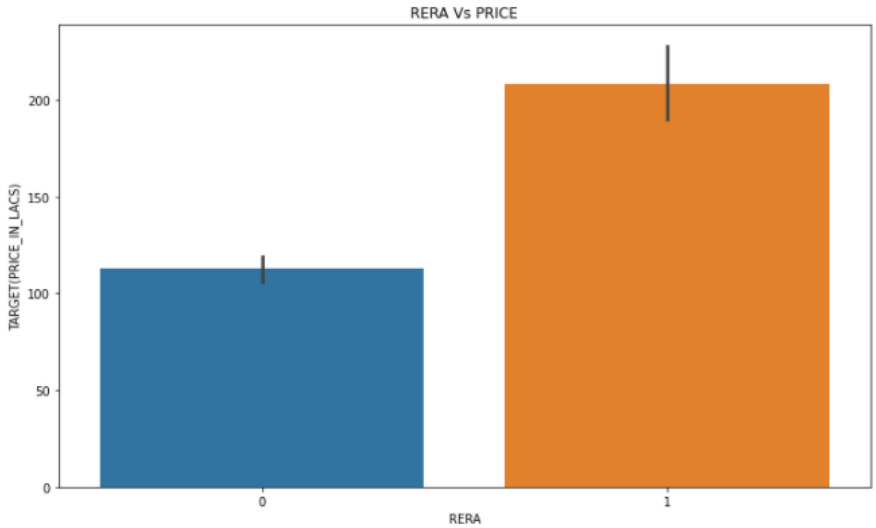
- *INFERENCE:*
  1. From the above Heatmap we can figure out the correlation between the features.
  2. If two features are highly correlated, we can figure out the value of one from the another
  3. In this case, we can drop the feature which lower correlation with the target column i.e the target column.
  4. In order to find which features are highly correlated, we have to choose a threshold (say 0.75) and then declare every feature with correlation above that threshold to be highly correlated.
  5. In this dataset no features are that highly correlated.

- **UNDER_CONSTRUCTION Vs TARGET PRICE BARPLOT**



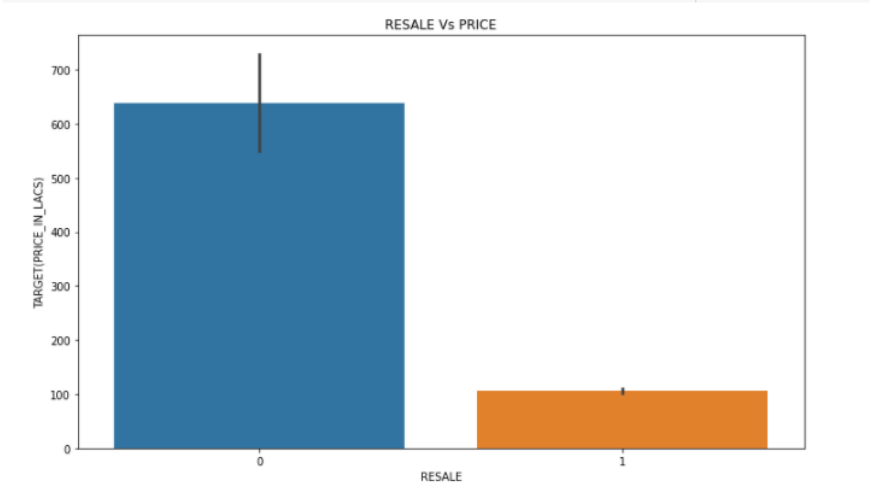
- *INFERENCE*:
    1. From this above barplot it is clear that the price of the house which is under construction sells for a higher price compared to the house whose construction is complete.

- **RERA Vs TARGET PRICE BARPLOT**



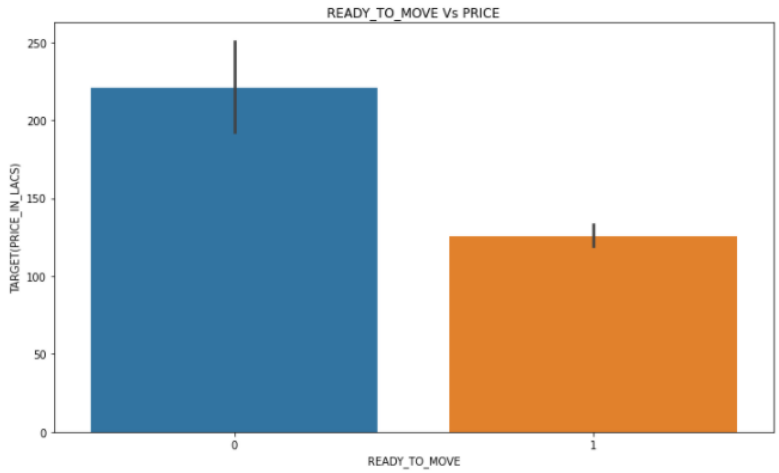
- *INFERENCE*:
    1. From the above Barplot we can clearly understand that the price of the house which has RERA approval sells for a higher price compared to the house which isn't RERA approved.

- **RESALE Vs TARGET PRICE BARPLOT**



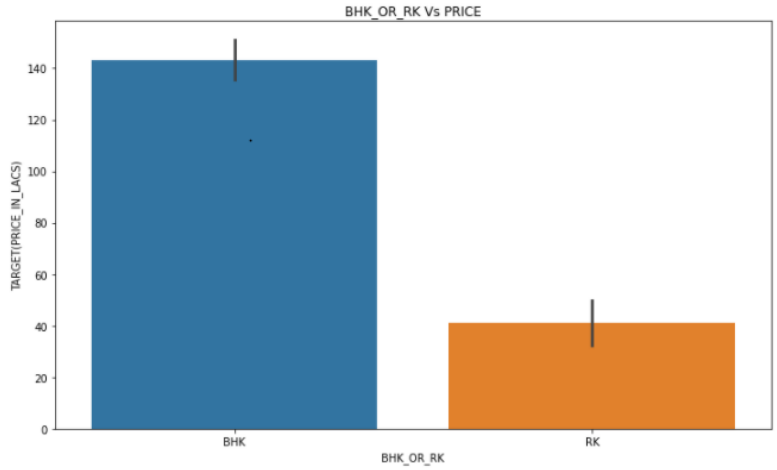
- *INFERENCE*:
    1. From the above Barplot it is clear that the price of the house which resells is very low compared to that of a new house.
    2. This result is expected as a house gets older its price reduces.

- **READY_TO_MOVE Vs TARGET PRICE BARPLOT**



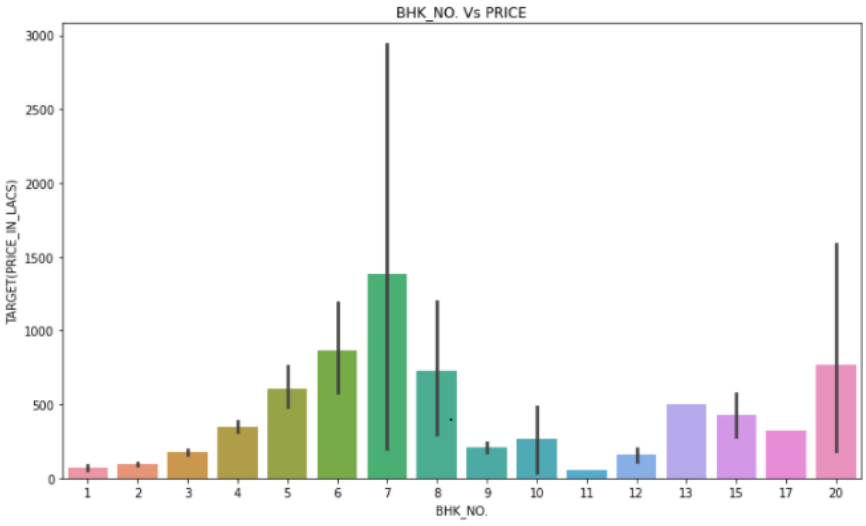
- *INFERENCE*:
    1. From the above Barplot we can see that ready to move house sells for a lower price.
    2. This may be due to the reason that most of the ready to move houses are resold and we know that when a house is resold its price decreases tremendously.

- **BHK_OR_RK Vs TARGET PRICE BARPLOT**



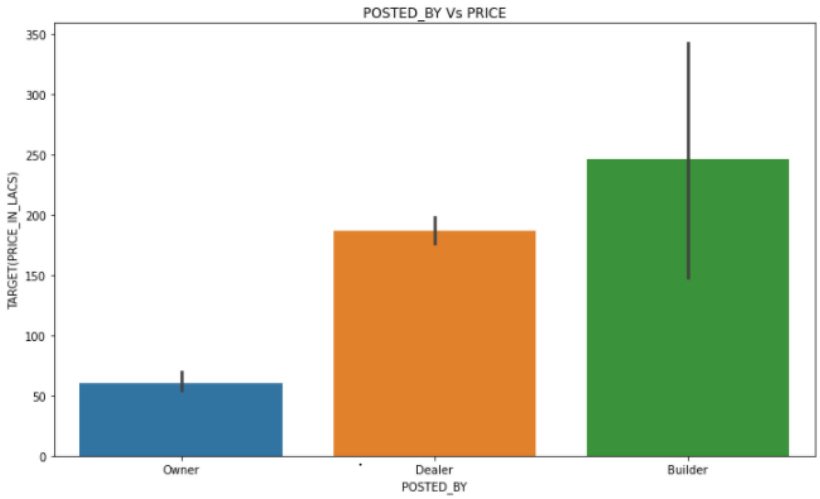
- *INFERENCE*:
    1. From the above Barplot it is clear that BHK(bedroom, hall, kitchen) sells for more price compared to that of RK (room, kitchen).
    2. It is natural that people pay high price for house with a Bedroom.

- **BHK_NO. Vs TARGET PRICE BARPLOT**

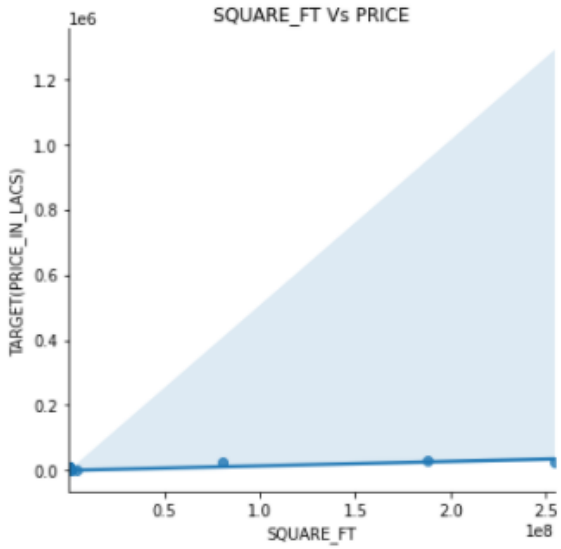


- *INFERENCE*:
    1. From the above barplot we can see that house price increases as the no.of.bedrooms increase until 7 BHK and then there is no clear pattern between BHK and house price.
    2. The reason behind this kind of pattern maybe because many people won't want to buy a house with that many bedrooms.

- **POSTED_BY Vs TARGET PRICE BARPLOT**

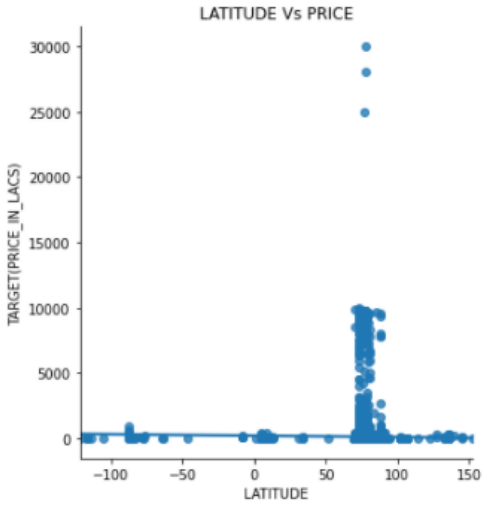


- *INFERENCE*:
    1. From the above graph we can see that houses which are sold by a Builder sells for the highest price, followed by that sold by a Dealer.
    2. This may be due to the fact that a Builder mostly sells new house and the house that aren't ready to occupy.
- **SQUARE_FT Vs TARGET PRICE BARPLOT**



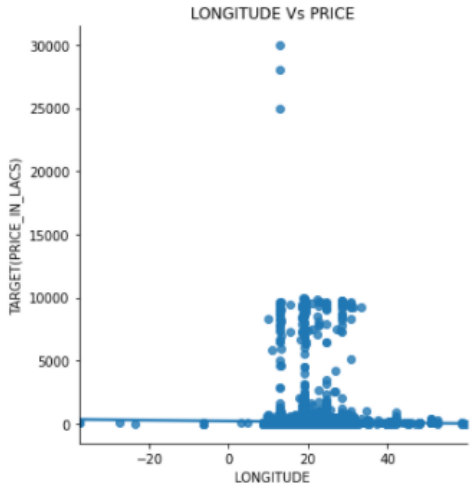
- *INFERENCE*:
    1. From the above plot we can see that the price of the house increases as the area of the house increases.
    2. It just proves the fact that bigger the house the price of the house increases.

- **LATITUDE Vs TARGET PRICE BARPLOT**



- *INFERENCE*:
    1. From the above plot we can infer that there is no concrete pattern between latitude and target price.
    2. So this feature isn't going to help us predict the price. So we can drop this feature.

- **LONGITUDE Vs TARGET PRICE BARPLOT**



- *INFERENCE*:
    1. From the above plot we can infer that there is no clear pattern between longitude and price of house.
    2. So we can drop this feature.