DS-1 Proposal

Proposal for Research Funding:

Understanding Indian Housing Preferences through Advanced Feature Engineering and Visualization

Respected, Prof. Jonathan Lee

I am writing to seek funding for a research project that aims to gain a deeper understanding of Indian housing preferences, with a focus on how individuals choose and utilize their living spaces in both busy urban apartments and other types of residences.

The central objective of this research initiative is to explore the intricate dynamics that govern housing choices in India. We aim to decipher the nuanced needs and preferences of Indian residents by employing advanced feature engineering techniques. This will enable us to distill complex data into comprehensible insights, allowing for a comprehensive view of the factors influencing housing choices.

This research endeavor holds immense potential to shed light on the intricate fabric of Indian housing preferences. With your support, we can embark on this journey towards a more nuanced understanding of how individuals make decisions about their living spaces, and how this knowledge can be leveraged for the betterment of urban environments in India.

Your contribution to this endeavor will not only advance the field of housing research but also have a tangible impact on the lives of countless individuals and communities.

Thank you for considering this proposal.

Team Members:

- Yash Jivani
- Pratik Mane
- Shashank Mysore
- Saurabh Khatri
- Keerthi Jayram

Dataset Link -

https://www.kaggle.com/datasets/aemyjutt/indianhousesdataanalysis

Dataset Summary:

No. Of Columns = 11
No. Of Rows = 1260

(3b) The dataset must have at least 5 columns, 30 rows, with at least 3 numeric columns, and 1 categorical column – The dataset follows the given conditions.

Column Names and Description -

1. A. **Area** - the size or total living space of the house
2. B. **BHK** - A configuration comprising the number of bedrooms (BHK), living area (hall or living room), and kitchen.
3. C. **Bathroom** - Number of bathrooms within the residence
4. D. **Furnishing** -Apartment condition, including whether it is semi-furnished, unfurnished, etc.
5. E. **Locality** - Property Location
6. F. **Parking** - Is parking available at the house, and if so, how many parking spaces are provided?
7. G. **Price** - The cost of residence.
8. H. **Status** - The term "house status" refers to the house's condition, indicating whether it is ready for immediate occupancy or still being prepared, among other possibilities.
9. I. **Transaction** - The "type of property" refers to characteristics that classify a property, such as whether it is a new house or a resale house
10. J. **Type** - The type of housing, including options like single-family houses, multi-floor buildings, and apartments.
11. K. **Per_Sqft** - Price per square foot or price for one square foot of the house.

Outline of Questions to ask –

1. Does the area of the house affect the price of a house?
2. Does the number of rooms affect the price of a house?
3. Does the furnishing type such as semi- furnished, furnished, unfinished affect the price of the house?
4. Does the locality affect the price? If yes which are the top 5 localities where a house price is higher?
5. Does the locality affect the price? If yes which are the top 5 localities where a house price is higher?
6. Does the locality affect the price? If yes which are the top 5 localities where a house price is higher?

Columns Classification –

Dependent Variable (DV) / Explained Variable

- A. Price

Independent Variable (IV) /Explaining Variable

- Type

- Transaction
- Status
- Parking
- Locality
- Furnishing
- Bathroom
- BHK
- Area