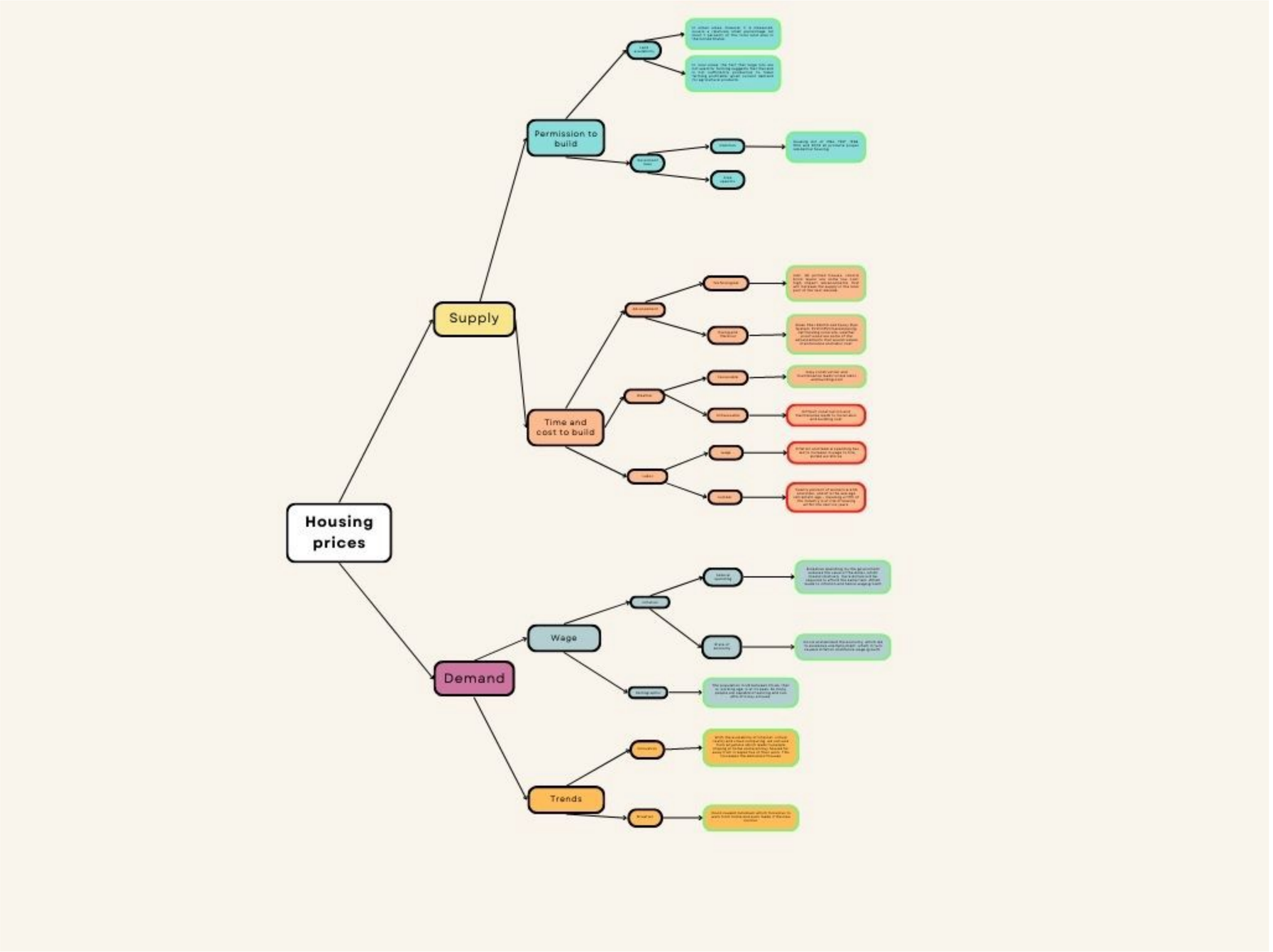
The housing market in the United States for the next decade

Introduction

- House prices increase every year, so there is a need for a system to predict house prices in the future. House price prediction can help the supplier (developer) determine the selling price of a house and can help the buyer (customer) to arrange the right time to purchase a house.

- Real estate has the image of being a tangible asset that comprises the property and the land on which it is built. Due to this reason, it is a significant portion of most people’s wealth.

- As unmovable as the land and the property seem, it is much more similar to the highly volatile stocks and bonds we dread to invest in. The similarity lies in the variability of its price. Which is highly dependent on supply and demand.



The above diagram is an issue tree that leads to the hypothesis. And it follows the MECE framework.

What is MECE?

- MECE is a method of grouping information into elements that are mutually exclusive (ME) and collectively exhaustive (CE).
- In other words, it is a process by which information—ideas, topics, issues, solutions—is arranged or, put in “MECE buckets with no overlapping between buckets and with each item having a place in one bucket only (ME), and with the buckets including all possible items relevant to the context.
- There are three ways to break down the problem in a MECE:
- Issue tree: Deals with the issue of the clients, where we break down the problem into smaller buckets and lead to the solution for each bucket.
- Decision tree: Decides on the correct decision to be made, where we break down the choice into probable more fundamental choices and reach a potential answer.
- Hypothesis tree: It is an answer-first approach, where we define the problem around the hypothesis. This is mainly done when we have prior knowledge about the field.

Applying MECE framework to our problem

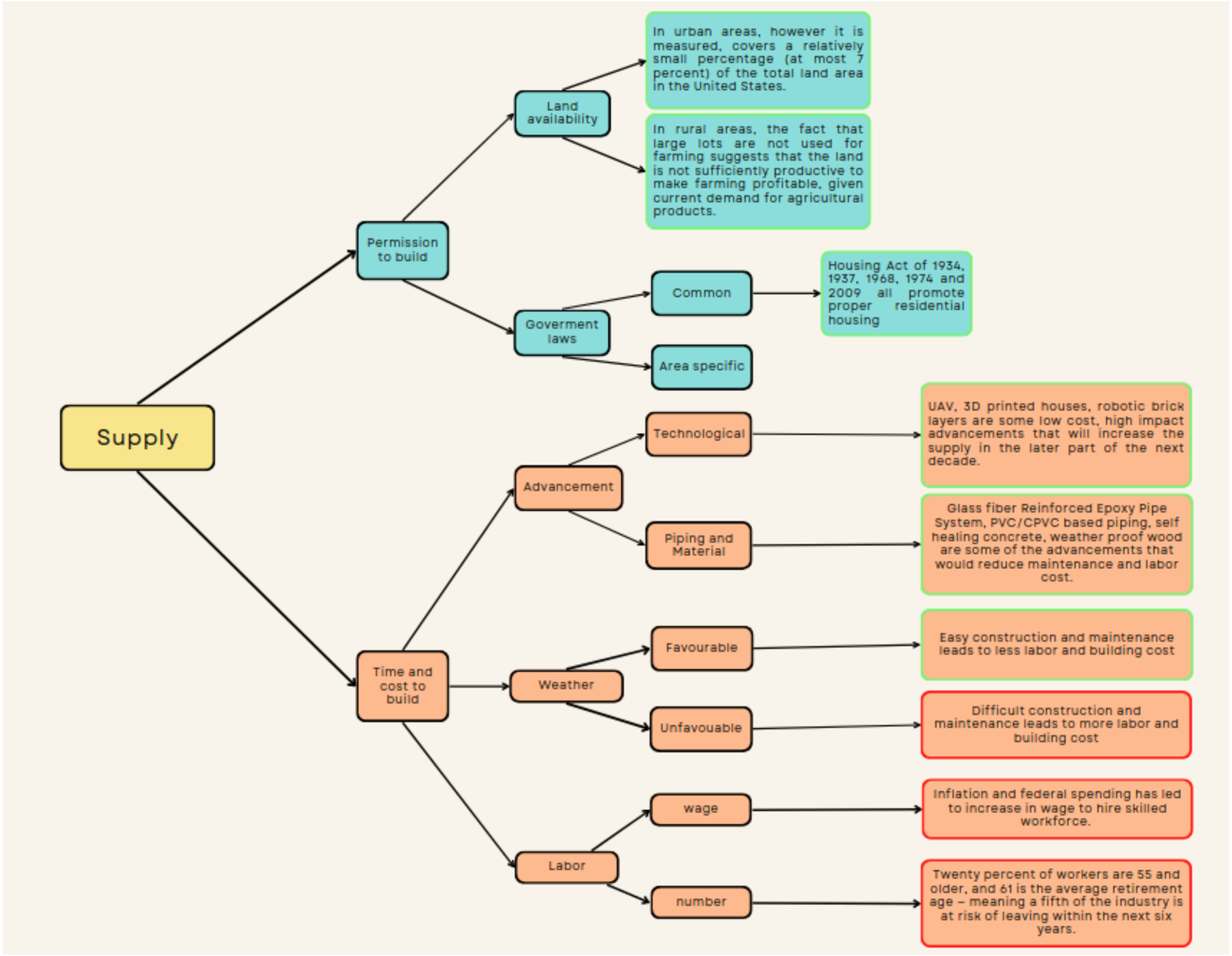
- In our case, we have to find out the factors that affect the housing market in the United States, it's not an issue or a decision to be made. So I decided to go with a **structure that breaks down the elephant in the room into smaller factors like the issue tree, which leads to the hypothesis on how those factors affect housing prices.**

- **Supply and Demand.** These are the root cause on which the housing prices depend(which makes the first level MECE).

- If supply is relatively more than the demand, houses are abundant for sale, but no one to buy them.

- And if the demand is more than the supply, more people buy but fewer houses to sell.

Supply


Supply
Permission to build
Land availability
In urban areas, however it is measured, covers a relatively small percentage (at most 7 percent) of the total land area in the United States.
In rural areas, the fact that large lots are not used for farming suggests that the land is not sufficiently productive to make farming profitable, given current demand for agricultural products.
Goverment laws
Common
Housing Act of 1934, 1937, 1968, 1974 and 2009 all promote proper residential housing
Area specific
Time and cost to build
Advancement
Technological
UAV, 3D printed houses, robotic brick layers are some low cost, high impact advancements that will increase the supply in the later part of the next decade.
Piping and Material
Glass fiber Reinforced Epoxy Pipe System, PVC/CPVC based piping, self healing concrete, weather proof wood are some of the advancements that would reduce maintenance and labor cost.
Weather
Favourable
Easy construction and maintenance leads to less labor and building cost
Unfavouable
Difficult construction and maintenance leads to more labor and building cost
Labor
wage
Inflation and federal spending has led to increase in wage to hire skilled workforce.
number
Twenty percent of workers are 55 and older, and 61 is the average retirement age – meaning a fifth of the industry is at risk of leaving within the next six years.

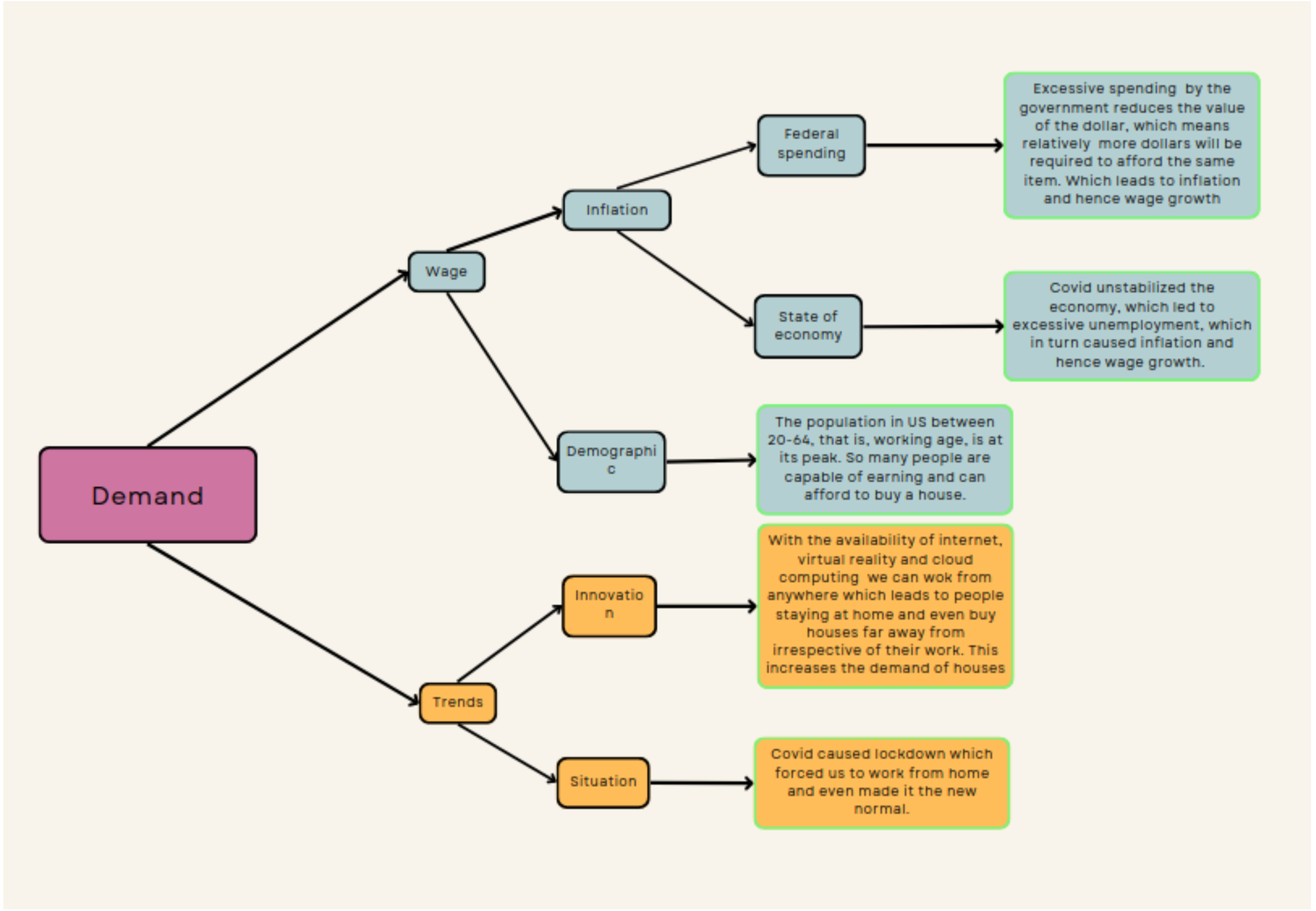
- The supply of the houses depends upon two factors, “permission to build” and “time and cost to build,” which can be broken down further into many other factors.

- And finally, the leaf nodes are the hypotheses that either support the supply or are against the supply.

- All the hypotheses are not equally weighted, which means some would matter more than others based on the time and situation.

- Supply and demand are relative, so we would have to look at the demand for housing to take an educated guess on which hypothesis should be given more importance to make a final call.

Demand


Demand
Wage
Trends
Inflation
Demographic
Federal spending
State of economy
Innovation
Situation
Excessive spending by the government reduces the value of the dollar, which means relatively more dollars will be required to afford the same item. Which leads to inflation and hence wage growth
Covid unstabilized the economy, which led to excessive unemployment, which in turn caused inflation and hence wage growth.
The population in US between 20-64, that is, working age, is at its peak. So many people are capable of earning and can afford to buy a house.
With the availability of internet, virtual reality and cloud computing we can wok from anywhere which leads to people staying at home and even buy houses far away from irrespective of their work. This increases the demand of houses
Covid caused lockdown which forced us to work from home and even made it the new normal.

- The demand for houses depends upon wage and Trends, which can be broken down into smaller factors.
- You might think that demand only depends on wages and trends?
- Many more factors include recession, interest rate, credit availability, etc.
- Due to inflation and which in turn is a result of excessive federal spending and a weak economy.
- Now that we have both supply and demand perspectives under our belt, we can take a better guess on which hypotheses should be weighted higher than others and then decide whether supply or demand would overpower each other or not.

Supply vs Demand

- One of the many and the most important factor in weighing the hypotheses is the urgency or, in other words, how fast will it come to fruition, is it affecting the market right now?

- In the case of demand, all the hypotheses have an immediate impact. For example, trends change fast, and due to covid, the whole world was pushed into the new normal of working remotely.

- Same for the wage growth, the demographic is affecting it right now, and inflation and recession are building up at this instant.

- But in the case of supply, there is only one factor with hypotheses that have an immediate impact, that too against the growth of supply, which is labor.

- All other factors supporting supply growth are slow growing and will have an impact in the near future.

Methodology and Models Approaches:

Variables: Year, land area, location of city, land price, building price, restaurants, health facilities, amusement parks, educational facilities, markets, worship places, and public transportations feasibility.

- **Regression Analysis and Particle Swarm Optimization**:

  Forecast the predicted value of house

  PSO is used for selection of affect variables and regression analysis is used to determine the optimal coefficient in prediction. The proved combination regression and PSO is suitable to get the minimum prediction error.

  Metrics: Error values are calculated based on Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE).

- **Time-series forecasting:**

  Price trend analysis

  Time series analysis is a technique which deals with time series data or trend analysis in which data points ordered in time. Time-series approach utilizes time-series data to look for the relationship between current prices and prevailing prices. In this model, the time is often the independent variable and the goal is to make a forecasting for the future.

  Components: Secular Trend, cyclical fluctuations, seasonal variations, and irregular component.

  Metrics: Gauge, forecast bias, mean average deviation (MAD), and mean average percentage error (MAPE)

- **Recommendation system**:

  Recommends location and model house

  This is a content based filtering approach that will do recommendation based on top-rated properties, most visited properties, google APIs recommend best location according to their budget in nearby area of interest.

  Metrics: Mean average Precision and Mean average Recall

Conclusion

- In easy words, building houses, however great the technology gets and smooth the bills and permit processes get, will take time. Whereas the fluctuations in demands are immediate. Hence, the house market prices will go to the moon for now!