Project Proposal

Meet Oswal
Kinjalk Srivastva

ESTIMAND/ QUESTION

This project seeks to estimate the causal effect of house size on housing price. Specifically, we aim to determine how variations in house size impact the selling price while accounting for key confounding variables.

DATA DESCRIPTION

The dataset includes a range of variables relevant to housing market analysis, with datapoints on property characteristics, housing status, pricing, and location details. The primary variables of interest are:

1. **House Size (Treatment):** The square footage of living space in a house
2. **Price (Outcome):** The price at which a house is listed or recently sold.
3. **Acre Lot (Confounder):** Property or land size, as larger lots may enable the construction of larger houses and may also directly affect the property's market value.
4. **City (Optional Confounder):** City as a variable captures the geographic and market-specific differences that may influence both house size and price.

CASUAL MODEL

For the above-mentioned dataset, the DAG which can model our Estimand, and the Treatment-Outcome-Confounder variables is illustrated bellow:
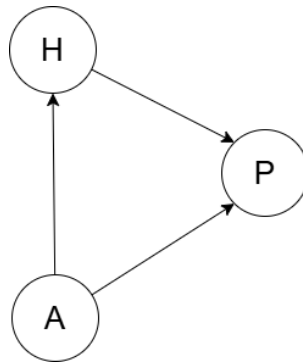
House size (H) is assumed to have a direct causal effect on Price (P).
Acre Lot (A) is assumed to may affect both House size (H) and Price(P) directly.

Note: Additional covariates such as City may also serve as potential confounders, influencing both the house size and market value. This DAG will be refined as our understanding of the relationship between variables develops.

PROPOSED STATISTICAL MODEL

To estimate the casual effect of house size on house price, we propose a linear regression model. This model will account for Acre Lot as confounder. We will use Bayesian Cross Validation to improve predictive accuracy and ensure model robustness.

This model setup will allow for flexibility in specifying prior distributions, which can adjust based on data characteristics or subject matter knowledge.