Data Science Case Study

By Leo Karlo C. dela Cruz

United States Census Bureau
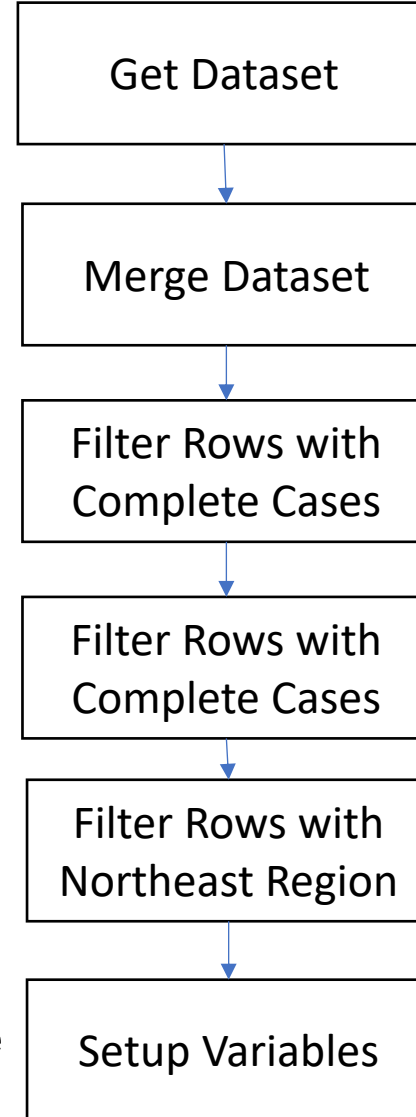
// Census.gov / **Manufactured Housing Survey Public Use File**

Tables

- Manufactured Housing Survey Public Use File: 2017   [<1.0 MB]
- Manufactured Housing Survey Public Use File: 2016   [<1.0 MB]
- Manufactured Housing Survey Public Use File: 2015   [<1.0 MB]
- Manufactured Housing Survey Public Use File: 2014   [<1.0 MB]

Current Dataset can estimate Housing Price

*need data on actual demand per housing price

Flowchart

- Get Dataset
- Merge Dataset
- Filter Rows with Complete Cases
- Filter Rows with Complete Cases
- Filter Rows with Northeast Region
- Setup Variables

Dependent Variable

NOMINAL VARIABLES:
- ControlId
- Region
- Status
- Titled
- Location
- Foundation
- Secured
- Jstatus
- Jprice
- Jbedroom
- Jtitle
- Jlocation
- Jfoundation
- Jsecured

NUMERIC VARIABLES:
- Weight
- Price (Housing Price)
- Sqft (Area)
- Weight Adjusted
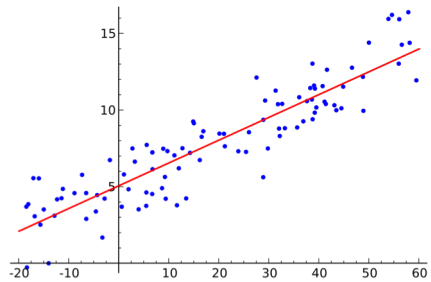
ORDINAL VARIABLES:
- ShipYear
- ShipMonth
- Sections
- Bedroom

Get Complete Northeast Region

Housing Demand Projections

Use Complete Northeast Region Data

Linear Regression

Assumptions:
- **Linearity**
- **Homoskedasticity**
- **Multivariate of Normality**
- **Independence of Errors**
  Lack of Multicollinearity

Why Linear Regression?

- **Interpretability**
- **Dependent variable is Numeric**
- **Demand projections is not time-dependent. There are other factors affecting price**
- **Clearly identify significant independent variables (e.g. sqft, bedroom, title) affecting the dependent variable (Housing Price)**

Housing Price Results
- **Applied backward elimination to identify significant factors** affecting the housing price

Housing Price increases if:
1. As **Year Increases**
2. As **Size of Home Increases**
3. As **Area (sqft.) Increases**
4. As **No. Bedroom Increases**
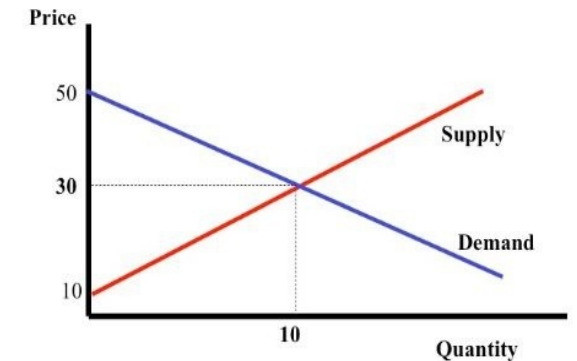5. If **Home is Placed outside manuf. Home communities**

Housing Price decreases if:
1. As **Weight Decreases**
2. If **Title is Personal Prop**
3. If **Home is Unsecured**
4. If **Price Indicator is Reported**
5. If **Location Indicator is Reported**
6. If **Foundation Indicator is Reported**
   (see graphs at the back)

Price Forecast is More Applicable for the Case
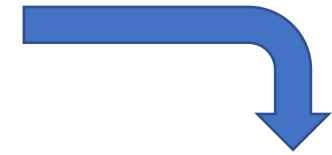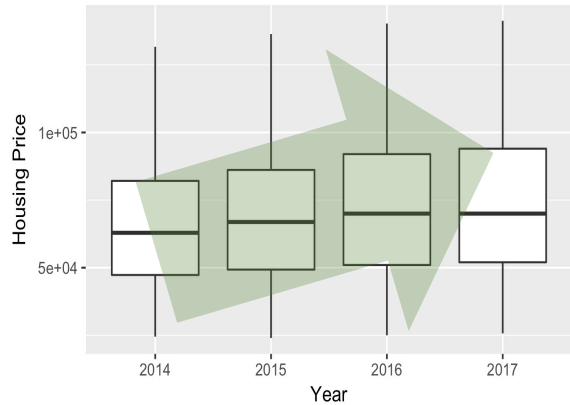
Supply Demand Curve

Further Studies:
- **Use Price Forecast to estimate Demand**
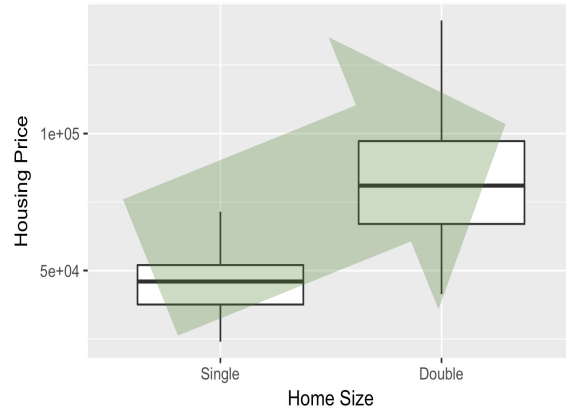- **Current Dataset don't exhibit the demand**

Yearly Housing Price
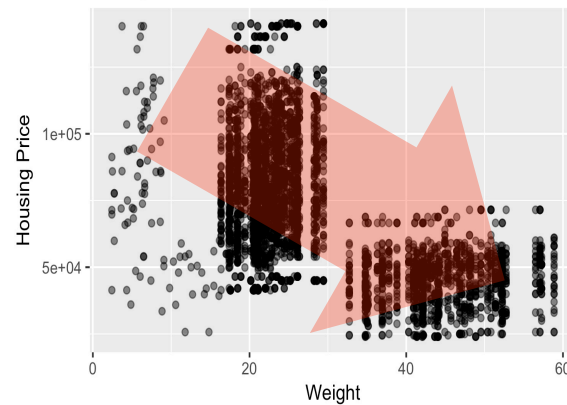Housing Prices at Year 2014 were less expensive

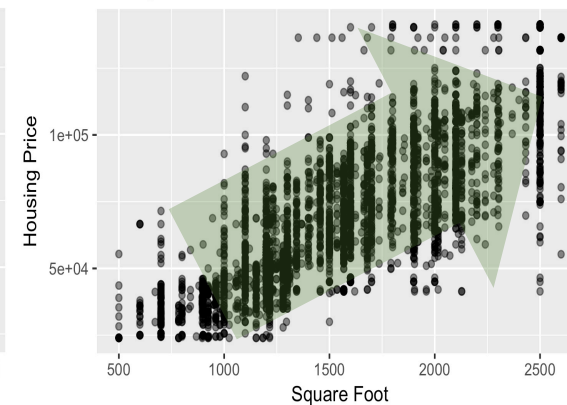**Size of Home**
Single Section Houses are less expensive

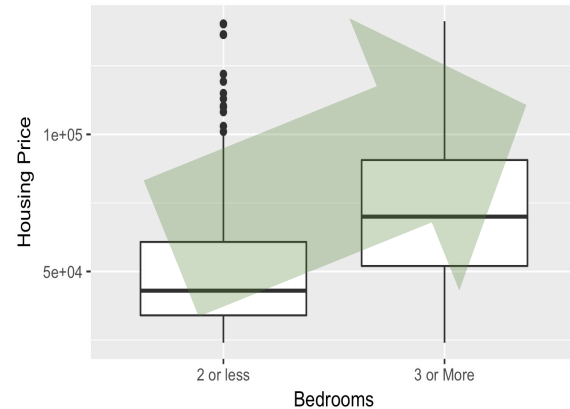**Sample Weight**
Housing Price decreases as Weight Increases

Square Footage of Home
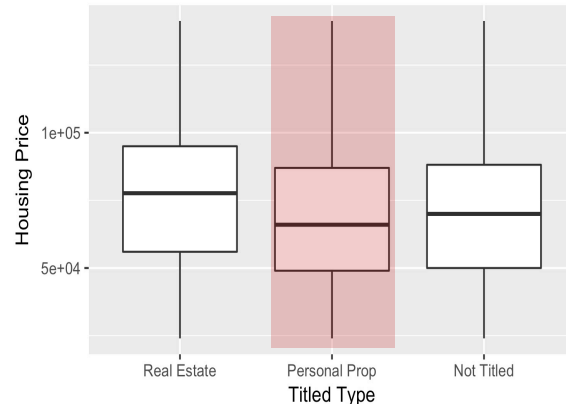Housing Price increase as Area increases

Number of Bedrooms
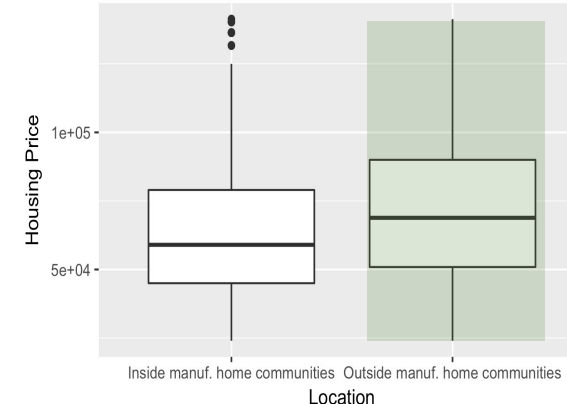Bedrooms with 2 or less are less expensive

How Home is Titled
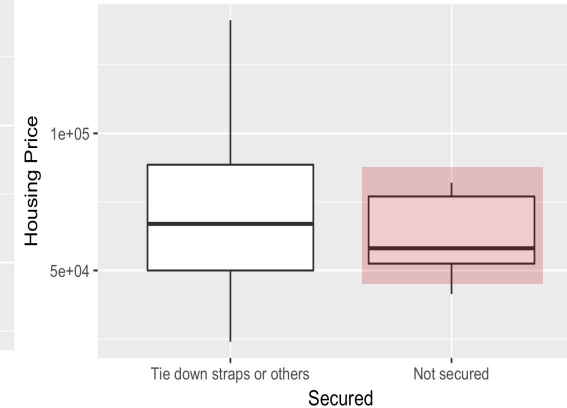Personal Prop Titles were less expensive

Where a Home is Placed
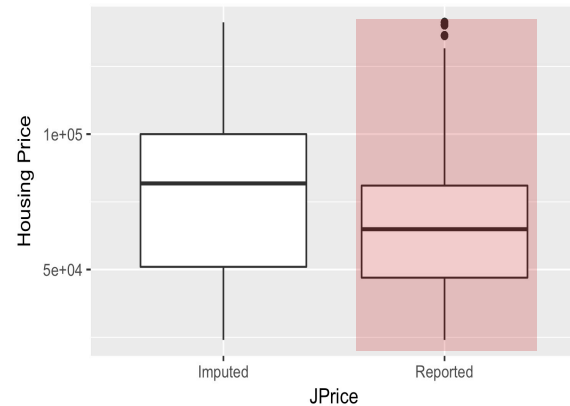Outside manuf. home communities are more expensive

How Home is Secured
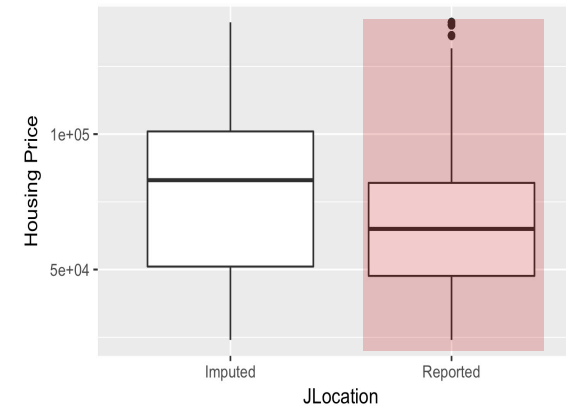Unsecured homes are less expensive

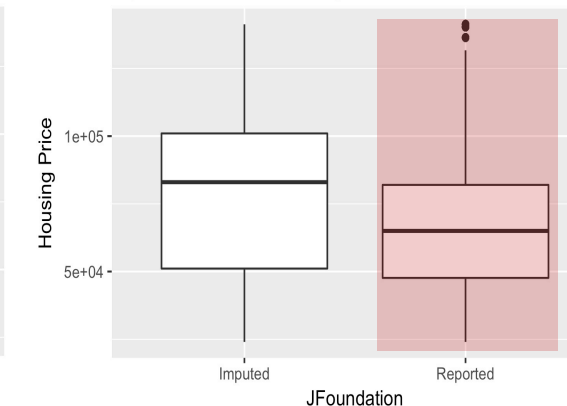**Imputed Indicator of Price**
Reported price are less expensive

Imputed Indicator of Location
Reported location are less expensive

Imputed Indicator of Foundation
Reported foundation are less expensive

Decrease in Housing Price

Increase on Housing Price