Business Statistics and Analysis

VALUE: ***Current Market Value of the Housing Unit***

Business Statistics and Analysis

VALUE: ***Current Market Value of the Housing Unit***

Business Statistics and Analysis

VALUE: ***Current Market Value of the Housing Unit***

Business Statistics and Analysis

VALUE: ***Current Market Value of the Housing Unit***

$$VALUE = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k$$

Business Statistics and Analysis

VALUE: ***Current Market Value of the Housing Unit***

$$VALUE = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k$$

Business Statistics and Analysis

VALUE: ***Current Market Value of the Housing Unit***

$$VALUE = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k$$

- Use only the 2013 data.
- Consider only ‘Single Family Housing’.

  TYPE = 1 and STRUCTURETYPE = 1

Business Statistics and Analysis

VALUE: ***Current Market Value of the Housing Unit***

$$VALUE = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k$$

- Use only the 2013 data.
- Consider only ‘Single Family Housing’.

  TYPE = 1 and STRUCTURETYPE = 1

Business Statistics and Analysis

VALUE: ***Current Market Value of the Housing Unit***

$$VALUE = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k$$

- Use only the 2013 data.
- Consider only ‘Single Family Housing’.

  TYPE = 1 and STRUCTURETYPE = 1

Business Statistics and Analysis

Some more cleansing of data

- The VALUE variable may have a negative or very low value.
- For our analysis we will delete all housing units which have a market value of less than 1,000$.
- This would also delete all data on housing units which are rental units, OWNRENT = ‘2’

Business Statistics and Analysis

Some more cleansing of data

- The VALUE variable may have a negative or very low value.
- For our analysis we will delete all housing units which have a market value of less than 1,000$.
- This would also delete all data on housing units which are rental units, OWNRENT = ‘2’

Business Statistics and Analysis

Some more cleansing of data

- The VALUE variable may have a negative or very low value.
- For our analysis we will delete all housing units which have a market value of less than 1,000$.
- This would also delete all data on housing units which are rental units, OWNRENT = ‘2’

Business Statistics and Analysis

Some more cleansing of data

- The VALUE variable may have a negative or very low value.
- For our analysis we will delete all housing units which have a market value of less than 1,000$.
- This would also delete all data on housing units which are rental units, OWNRENT = ‘2’

Business Statistics and Analysis

- Choose only those housing units that have a market value which is 1000 dollars or greater, that is, VALUE ≥ $1000.

- Choose only the ‘Single Family Units’, that is, STRUCTURETYPE = 1 and TYPE = 1.

Business Statistics and Analysis

- Choose only those housing units that have a market value which is 1000 dollars or greater, that is, VALUE ≥ $1000.

- Choose only the ‘Single Family Units’, that is, STRUCTURETYPE = 1 and TYPE = 1.

Business Statistics and Analysis

- Choose only those housing units that have a market value which is 1000 dollars or greater, that is, VALUE ≥ $1000.

- Choose only the ‘Single Family Units’, that is, STRUCTURETYPE = 1 and TYPE = 1.

Business Statistics and Analysis

- Choose only those housing units that have a market value which is 1000 dollars or greater, that is, VALUE ≥ $1000.

- Choose only the ‘Single Family Units’, that is, STRUCTURETYPE = 1 and TYPE = 1.

Business Statistics and Analysis

- Choose only those housing units that have a market value which is 1000 dollars or greater, that is, VALUE ≥ $1000.

- Choose only the ‘Single Family Units’, that is, STRUCTURETYPE = 1 and TYPE = 1.

Business Statistics and Analysis

After data cleaning...

- Calculate various descriptive statistics for the VALUE variable.
- Visualize the empirical distribution of VALUE.
  - Plot a histogram of VALUE.
  - Plot a histogram of Ln(VALUE).

Which of these two histograms more closely resemble a Bell curve?

- Select your set of independent variables or the X variables.

Business Statistics and Analysis

After data cleaning...

- Calculate various descriptive statistics for the VALUE variable.
- Visualize the empirical distribution of VALUE.
  - Plot a histogram of VALUE.
  - Plot a histogram of Ln(VALUE).

Which of these two histograms more closely resemble a Bell curve?

- Select your set of independent variables or the X variables.

Business Statistics and Analysis

After data cleaning...

- Calculate various descriptive statistics for the VALUE variable.
- Visualize the empirical distribution of VALUE.
  - Plot a histogram of VALUE.
  - Plot a histogram of Ln(VALUE).

Which of these two histograms more closely resemble a Bell curve?

- Select your set of independent variables or the X variables.

Business Statistics and Analysis

After data cleaning...

- Calculate various descriptive statistics for the VALUE variable.
- Visualize the empirical distribution of VALUE.
  - Plot a histogram of VALUE.
  - Plot a histogram of Ln(VALUE).

Which of these two histograms more closely resemble a Bell curve?

- Select your set of independent variables or the X variables.

Business Statistics and Analysis

After data cleaning...

- Calculate various descriptive statistics for the VALUE variable.
- Visualize the empirical distribution of VALUE.
  - Plot a histogram of VALUE.
  - Plot a histogram of Ln(VALUE).

Which of these two histograms more closely resemble a Bell curve?

- Select your set of independent variables or the X variables.

Business Statistics and Analysis

After data cleaning...

- Calculate various descriptive statistics for the VALUE variable.
- Visualize the empirical distribution of VALUE.
  - Plot a histogram of VALUE.
  - Plot a histogram of Ln(VALUE).

Which of these two histograms more closely resemble a Bell curve?

- Select your set of independent variables or the X variables.

Business Statistics and Analysis

After data cleaning...

- Calculate various descriptive statistics for the VALUE variable.
- Visualize the empirical distribution of VALUE.
  - Plot a histogram of VALUE.
  - Plot a histogram of Ln(VALUE).

Which of these two histograms more closely resemble a Bell curve?

- Select your set of independent variables or the X variables.

Business Statistics and Analysis

Some tips for selecting "X" variables...

1. Number of "X" variables need to be less than 16.
2. You need to appropriately code the categorical variables.
   - Collapsing categories.

   METRO3: '1' → Central city area

   '2', '3', '4' or '5' → *Not a* Central city area
3. You may try out the natural logarithmic or other transformation of variables.
4. Think about appropriateness of using the various "X" variables in the model.

Business Statistics and Analysis

Some tips for selecting "X" variables...

1. Number of "X" variables need to be less than 16.
2. You need to appropriately code the categorical variables.
   - Collapsing categories.

     METRO3: '1' → Central city area

     '2', '3', '4' or '5' → *Not a* Central city area
3. You may try out the natural logarithmic or other transformation of variables.
4. Think about appropriateness of using the various "X" variables in the model.

Business Statistics and Analysis

Some tips for selecting "X" variables...

1. Number of "X" variables need to be less than 16.
2. You need to appropriately code the categorical variables.
   - Collapsing categories.

     METRO3: '1' → Central city area

     '2', '3', '4' or '5' → *Not a* Central city area
3. You may try out the natural logarithmic or other transformation of variables.
4. Think about appropriateness of using the various "X" variables in the model.

Business Statistics and Analysis

Some tips for selecting "X" variables...

1. Number of "X" variables need to be less than 16.
2. You need to appropriately code the categorical variables.
   - Collapsing categories.

     METRO3: '1' → Central city area

     '2', '3', '4' or '5' → *Not a* Central city area
3. You may try out the natural logarithmic or other transformation of variables.
4. Think about appropriateness of using the various "X" variables in the model.

Business Statistics and Analysis

Some tips for selecting "X" variables...

1. Number of "X" variables need to be less than 16.
2. You need to appropriately code the categorical variables.
   - Collapsing categories.

     METRO3: '1' → Central city area

     '2', '3', '4' or '5' → *Not a* Central city area
3. You may try out the natural logarithmic or other transformation of variables.
4. Think about appropriateness of using the various "X" variables in the model.

Business Statistics and Analysis

Some tips for selecting "X" variables...

1. Number of "X" variables need to be less than 16.
2. You need to appropriately code the categorical variables.
   - Collapsing categories.

     METRO3: '1' → Central city area

     '2', '3', '4' or '5' → *Not a* Central city area
3. You may try out the natural logarithmic or other transformation of variables.
4. Think about appropriateness of using the various "X" variables in the model.

Business Statistics and Analysis

Some tips for selecting "X" variables...

1. Number of "X" variables need to be less than 16.
2. You need to appropriately code the categorical variables.
   - Collapsing categories.

     METRO3: '1' → Central city area

     '2', '3', '4' or '5' → *Not a* Central city area
3. You may try out the natural logarithmic or other transformation of variables.
4. Think about appropriateness of using the various "X" variables in the model.

Business Statistics and Analysis

Some tips for selecting "X" variables...

1. Number of "X" variables need to be less than 16.
2. You need to appropriately code the categorical variables.
   - Collapsing categories.

     METRO3: '1' → Central city area

     '2', '3', '4' or '5' → *Not a* Central city area
3. You may try out the natural logarithmic or other transformation of variables.
4. Think about appropriateness of using the various "X" variables in the model.

Business Statistics and Analysis

Some tips for selecting "X" variables...

1. Number of "X" variables need to be less than 16.
2. You need to appropriately code the categorical variables.
   - Collapsing categories.

     METRO3: '1' → Central city area

     '2', '3', '4' or '5' → *Not a* Central city area
3. You may try out the natural logarithmic or other transformation of variables.
4. Think about appropriateness of using the various "X" variables in the model.

Business Statistics and Analysis

Your report should cover the following...

1. The set of variables used in the regression model and a brief justification for their use.

2. The estimated regression model along with explanation of any variable transformations you have done.

3. Interpretation of the impact of various variables included in your model.

Business Statistics and Analysis

Your report should cover the following...

1. **The set of variables used in the regression model and a brief justification for their use.**

2. **The estimated regression model along with explanation of any variable transformations you have done.**

3. **Interpretation of the impact of various variables included in your model.**

Business Statistics and Analysis

Your report should cover the following...

1. **The set of variables used in the regression model and a brief justification for their use.**

2. **The estimated regression model along with explanation of any variable transformations you have done.**

3. **Interpretation of the impact of various variables included in your model.**

Business Statistics and Analysis

Your report should cover the following...

1. **The set of variables used in the regression model and a brief justification for their use.**

2. **The estimated regression model along with explanation of any variable transformations you have done.**

3. **Interpretation of the impact of various variables included in your model.**