

HW 07

David Gerard

2020-03-05

Instructions

- Write your solutions in the app.R starter file.
- Only commit plain text files (like .R files).
- Do not modify the paths of any files.
- **Make sure to commit regularly.** Lack of informative and frequent commits will result in point deductions.
- Only include the necessary code, not any extraneous code, to answer the questions.
- Learning objectives:
 - Use your shiny skills to build an App that interactively analyzes the housing sales dataset.

Housing Sales App

Researchers were interested in predicting residential home sales prices in a Midwestern city as a function of various characteristics of the home and surrounding property. Data on 522 transactions were obtained for home sales during the year 2002. The 13 variables are

- **Price:** Sales price of residence (in dollars)
- **Area:** Finished area of residence (in square feet)
- **Bed:** Total number of bedrooms in residence
- **Bath:** Total number of bathrooms in residence
- **AC:** 1 = presence of air conditioning, 0 = absence of air conditioning
- **Garage:** Number of cars that a garage will hold
- **Pool:** 1 = presence of a pool, 0 = absence of a pool
- **Year:** Year property was originally constructed
- **Quality:** Index for quality of construction. **High**, **Medium**, or **Low**.
- **Style:** Categorical variable indicating architectural style
- **Lot:** Lot size (in square feet)
- **Highway:** 1 = highway adjacent, 0 = highway not adjacent.

We've seen these data a few times before.

Build a Shiny App that has the following attributes:

1. Three tabs. The first tab is for univariate analysis. The second tab is for bivariate analysis. The third tab is for the a spreadsheet of the *numeric* variables in the data.
2. The inputs/outputs for the univariate analysis should be:
 - The variable.

- Should we do the analysis on the log scale?
 - The number of bins in the histogram.
 - The null value for a one-sample t -test.
 - The results of the one-sample t -test.
3. For the univariate analysis, you should make a histogram if the variable is numeric and make a barplot otherwise.
 4. For the univariate analysis, the results of the test should be done on the log or non-log scale according to the user options.
 5. The inputs for the bivariate analysis should be:
 - The two variables.
 - Whether we should log each variable.
 - Whether we should add an OLS line.
 6. You should have a scatter plot if both variables are numeric, a boxplot if one is numeric and one is categorical, and a jitter plot if both are categorical. Only numeric variables should be logged.
 7. The spreadsheet tab should contain a Data Table with **only the numeric variables**. Use a `map*()` function to select these.
 8. Try to make your Shiny app as visually similar to my app as you can.

Hint: You can make this a lot easier by taking advantage of the modularity built into ggplot2:

```
p1 <- ggplot(mtcars, aes(x = disp, y = mpg))
p1

p1 <- p1 + geom_point()
p1

p1 <- p1 + scale_x_log10()
p1
```

Hint: I found it easier to use `geom_boxploth()` from the ggstance library instead of using `coord_flip()`.

Hint: Think carefully about which variables should be treated as categorical and which should be treated as quantitative.

Below are screenshots of my app under different user inputs:

EDA of Estate Data

Univariate Analyses Bivariate Analyses Spreadsheet

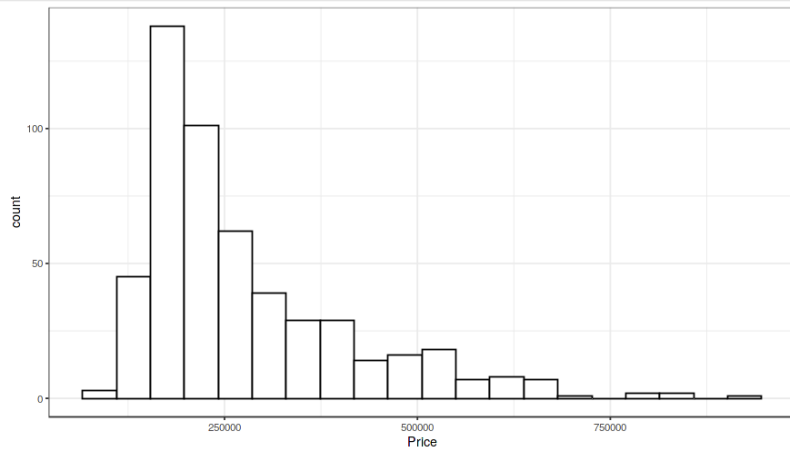
Variable
 Price

☐ Log

Bins
 1 20 100

Null Value
 0

P-value	Lower	Upper
0.00	266034.80	289753.50



EDA of Estate Data

Univariate Analyses Bivariate Analyses Spreadsheet

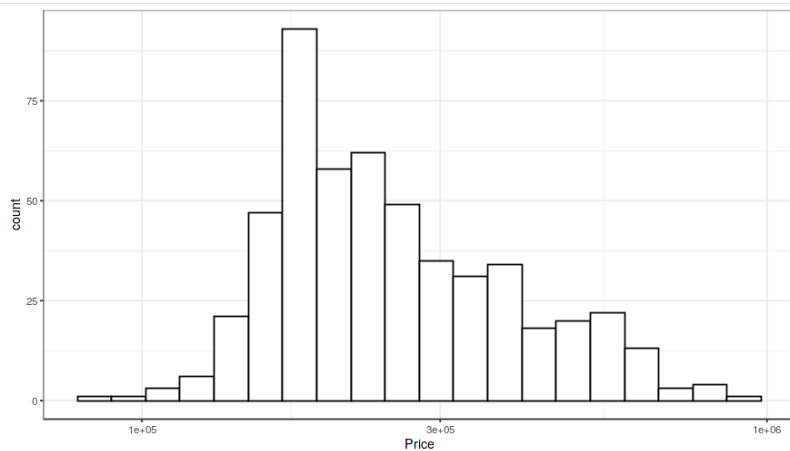
Variable
 Price

☒ Log

Bins
 1 20 100

Null Value
 0

P-value	Lower	Upper
0.00	17.89	17.99



EDA of Estate Data

Univariate Analyses Bivariate Analyses Spreadsheet

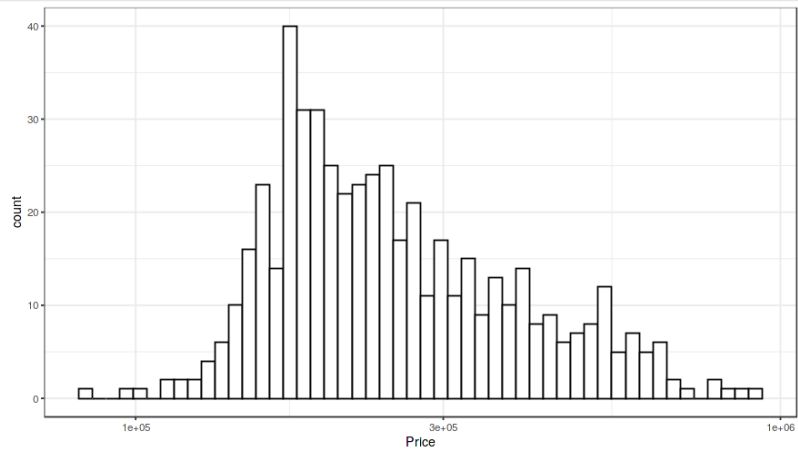
Variable
 Price

☒ Log

Bins
 1 50 100

Null Value
 0

P-value	Lower	Upper
0.00	17.89	17.99



EDA of Estate Data

Univariate Analyses Bivariate Analyses Spreadsheet

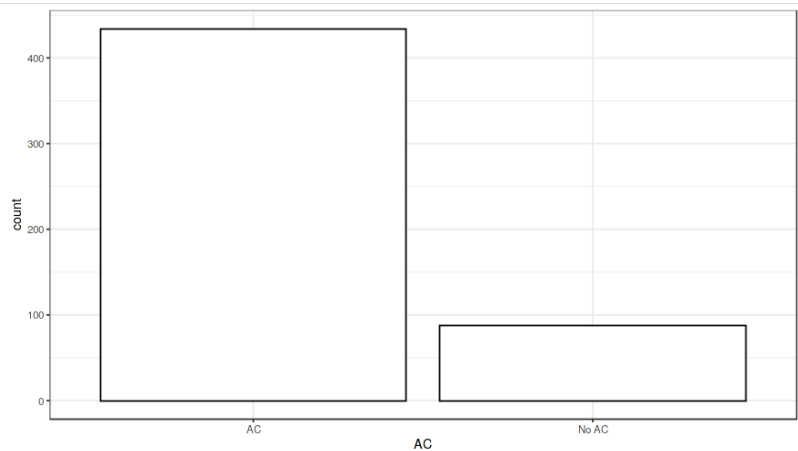
Variable
 AC

☒ Log

Bins
 1 50 100

Null Value
 0

data
 Not a numeric



EDA of Estate Data

Univariate Analyses Bivariate Analyses Spreadsheet

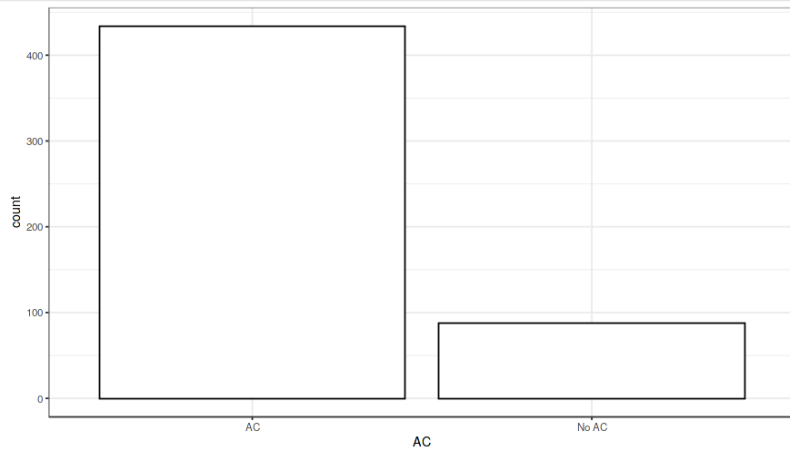
Variable
 AC

☐ Log

Bins
 1 50 100

Null Value
 0

data
 Not a numeric



EDA of Estate Data

Univariate Analyses Bivariate Analyses Spreadsheet

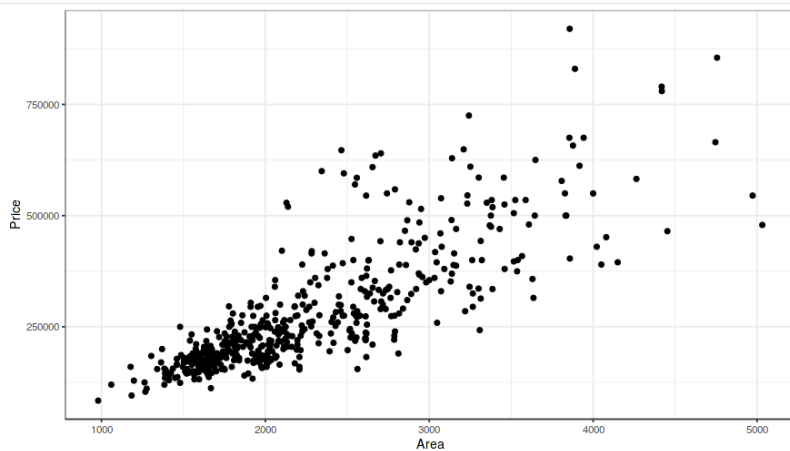
Variable 1
 Area

☐ Log

Variable 2
 Price

☐ Log

☐ OLS!



EDA of Estate Data

Univariate Analyses Bivariate Analyses Spreadsheet

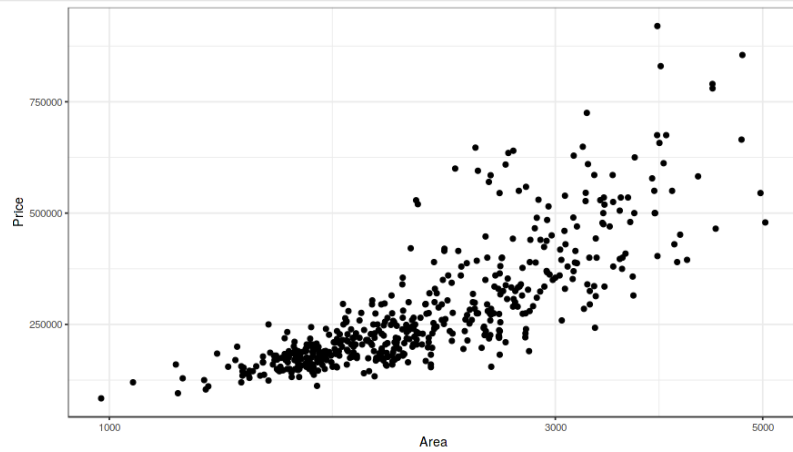
Variable 1
Area

☒ Log

Variable 2
Price

☐ Log

☐ OLS!



EDA of Estate Data

Univariate Analyses Bivariate Analyses Spreadsheet

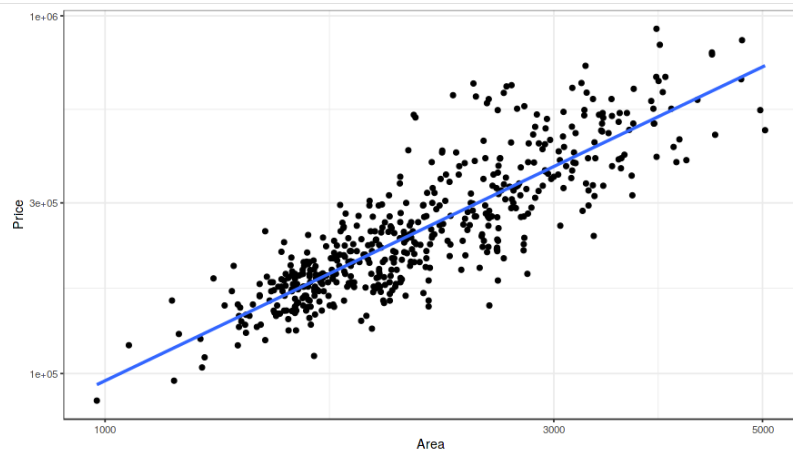
Variable 1
Area

☒ Log

Variable 2
Price

☒ Log

☒ OLS!



EDA of Estate Data

Univariate Analyses Bivariate Analyses Spreadsheet

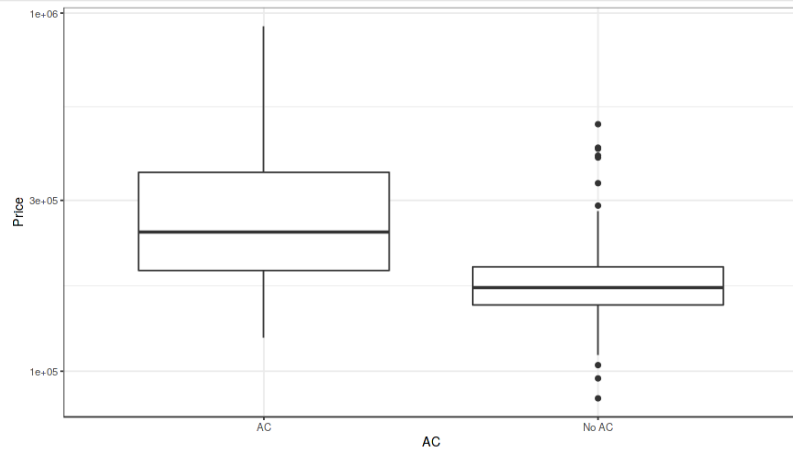
Variable 1
 AC

☒ Log

Variable 2
 Price

☒ Log

☒ OLS!



EDA of Estate Data

Univariate Analyses Bivariate Analyses Spreadsheet

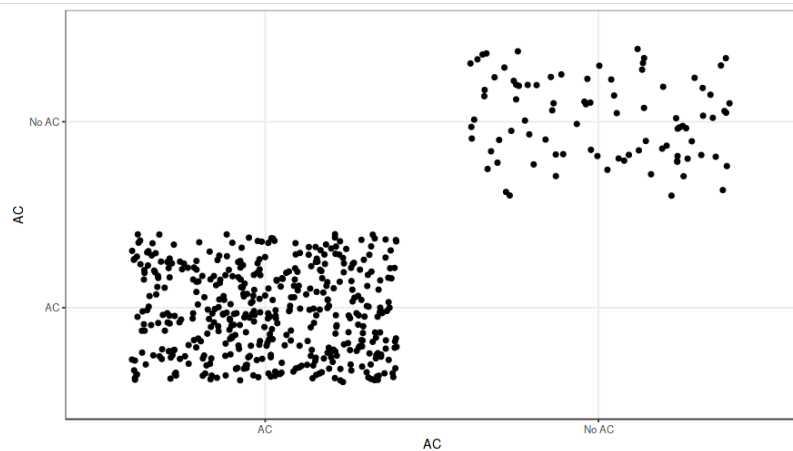
Variable 1
 AC

☒ Log

Variable 2
 AC

☒ Log

☒ OLS!



EDA of Estate Data

Univariate Analyses Bivariate Analyses Spreadsheet

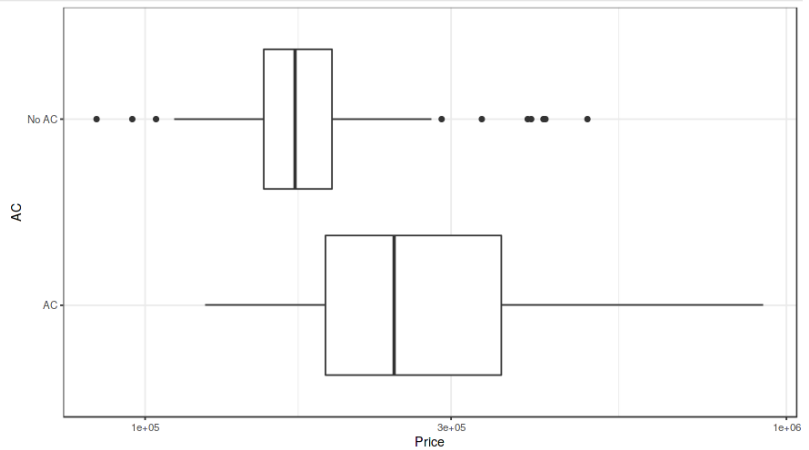
Variable 1
 Price

☒ Log

Variable 2
 AC

☒ Log

☒ OLS!



EDA of Estate Data

Univariate Analyses Bivariate Analyses Spreadsheet

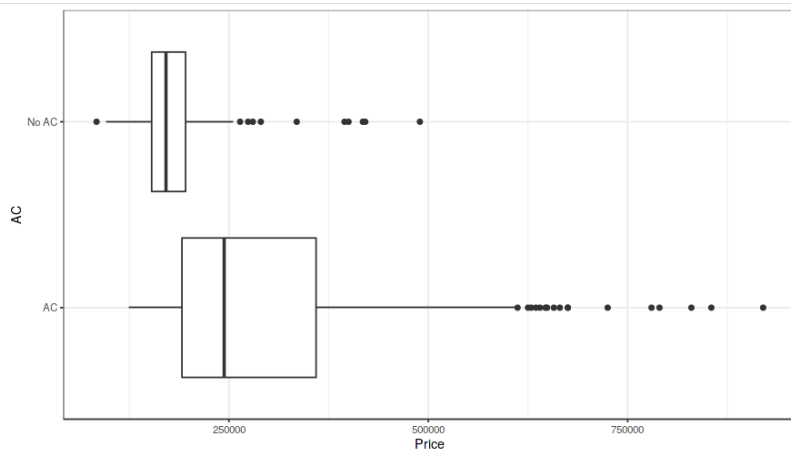
Variable 1
 Price

☐ Log

Variable 2
 AC

☒ Log

☒ OLS!



EDA of Estate Data

Univariate Analyses

Bivariate Analyses

Spreadsheet

Show 10 entries

Search:

Price	Area	Bed	Bath	Garage	Year	Lot
360000	3032	4	4	2	1972	22221
340000	2058	4	2	2	1976	22912
250000	1780	4	3	2	1980	21345
205500	1638	4	2	2	1963	17342
275500	2196	4	3	2	1968	21786
248000	1966	4	3	5	1972	18902
229900	2216	3	2	2	1972	18639
150000	1597	2	1	1	1955	22112
195000	1622	3	2	2	1975	14321
160000	1976	3	3	1	1918	32358

Price Area Bed Bath Garage Year Lot

Showing 1 to 10 of 522 entries

Previous 1 2 3 4 5 ... 53 Next