2023GFA_ANA_500_02 Foundations of Data Analytics

Announcements
Read Me First
Syllabus
Course Guide
Assignments

Modules
Discussions
Learning Journals
Exams

gretl

Library Resources
Writing Center
STEM Center
ODEI
Purchase Course Materials

My Grades
Help

Review Test Submission: Mod 04 Week 8 P&P Assignment

User	Kohei Nishitani
Course	2023GFA_ANA_500_02 Foundations of Data Analytics
Test	Mod 04 Week 8 P&P Assignment
Started	10/20/23 5:25 PM
Submitted	10/20/23 5:59 PM
Due Date	10/20/23 11:59 PM
Status	Completed
Attempt Score	290 out of 350 points
Time Elapsed	33 minutes
Instructions	This is the paper and pencil assignment for Module 04, Week 8. Hurray! You are very nearly done with ANA 500. In this assignment in addition to continuing to explore ordinary least squares regression, particularly multiple variable or multivariable regression including terms representing interaction between variables and second-order variables this assignment will help you prepare for the comprehensive final examination. I am uploading a complete Word doc for this P&P assignment below. As was the case last week, this document contains everything you need to complete this assignment as well as a discussion about some of the concepts covered and references to your textbook and additional material. The intention is to help you develop an intuitive understanding for what is going on with this type of regression. As always, if you have question please ask! ANA500 Week 8 gretl and PP assignments.docx

Question 1 10 out of 10 points

Add the variable age to your original simple linear model to develop a multivariable linear model. (Continue as before using the restricted dataset.) Use that model to answer the following questions.

Are the intercept and all coefficients statistically significant?

Question 2 10 out of 10 points

Add the variable age to your original simple linear model to develop a multivariable linear model. (Continue as before using the restricted dataset.) Use that model to answer the following questions.

How much does a traditional-style house increase in value for each additional square foot of living space?

Question 3 10 out of 10 points

Now return to the full, unrestricted dataset. Rerun the multivariable model including size (sqft) and age. Use your results to answer the following questions.

The intercept and all coefficients are NOT statistically significant.

Question 4 10 out of 10 points

Now return to the full, unrestricted dataset. Rerun the multivariable model including size (sqft) and age. Use your results to answer the following questions.

The difference in cost for an additional square foot of living space from the traditional-style house to all houses is _____.

Question 5 10 out of 10 points

Now return to the full, unrestricted dataset. Rerun the multivariable model including size (sqft) and age. Use your results to answer the following questions.

The mean of house prices increases by _____.

Question 6 10 out of 10 points

Now return to the full, unrestricted dataset. Rerun the multivariable model including size (sqft) and age. Use your results to answer the following questions.

An increase of one year in age decreases the house price (for all houses) by _____.

Question 7 10 out of 10 points

Now return to the full, unrestricted dataset. Rerun the multivariable model including size (sqft) and age. Use your results to answer the following questions.

The value of the intercept for the multivariable linear model is "practically" realistic.

Question 8 10 out of 10 points

Add higher (second) order terms for the variables sqft and age to your multivariable model. Use your results from running that model to answer the following questions. (NOTE!!! Do not round until you are ready to submit your final answer. It does make a big difference here.)

The intercept and all coefficients are not statistically significant.

Question 9 0 out of 10 points

Add higher (second) order terms for the variables sqft and age to your multivariable model. Use your results from running that model to answer the following questions. (NOTE!!! Do not round until you are ready to submit your final answer. It does make a big difference here.)

Estimate the marginal effects of price as a function of size, i.e. $\partial price / \partial sqft$, and age, i.e. $\partial price / \partial age$, for the smallest or oldest house in the data, the largest or newest house in the data, and a 2300 square foot or a 20 year old house. To help here are the relevant equations:

The complete model is:

(This equation is in the attached Word doc.)

The derivative with respect to size (sqft) to find the marginal effect of size is:

(This equation is in the attached Word doc.)

The derivative with respect to age to find the marginal effect of age is:

(This equation is in the attached Word doc.)

The coefficient of the sqft term is _____.

Question 10 10 out of 10 points

Add higher (second) order terms for the variables sqft and age to your multivariable model. Use your results from running that model to answer the following questions. (NOTE!!! Do not round until you are ready to submit your final answer. It does make a big difference here.)

Estimate the marginal effects of price as a function of size, i.e. $\partial price / \partial sqft$, and age, i.e. $\partial price / \partial age$, for the smallest or oldest house in the data, the largest or newest house in the data, and a 2300 square foot or a 20 year old house. To help here are the relevant equations:

The complete model is:

(This equation is in the attached Word doc.)

The derivative with respect to size (sqft) to find the marginal effect of size is:

(This equation is in the attached Word doc.)

The derivative with respect to age to find the marginal effect of age is:

(This equation is in the attached Word doc.)

The coefficient of the term is _____.

Question 11 0 out of 10 points

Add higher (second) order terms for the variables sqft and age to your multivariable model. Use your results from running that model to answer the following questions. (NOTE!!! Do not round until you are ready to submit your final answer. It does make a big difference here.)

Estimate the marginal effects of price as a function of size, i.e. $\partial price / \partial sqft$, and age, i.e. $\partial price / \partial age$, for the smallest or oldest house in the data, the largest or newest house in the data, and a 2300 square foot or a 20 year old house. To help here are the relevant equations:

The complete model is:

(This equation is in the attached Word doc.)

The derivative with respect to size (sqft) to find the marginal effect of size is:

(This equation is in the attached Word doc.)

The derivative with respect to age to find the marginal effect of age is:

(This equation is in the attached Word doc.)

The coefficient of the age term is _____.

Question 12 10 out of 10 points

Add higher (second) order terms for the variables sqft and age to your multivariable model. Use your results from running that model to answer the following questions. (NOTE!!! Do not round until you are ready to submit your final answer. It does make a big difference here.)

Estimate the marginal effects of price as a function of size, i.e. $\partial price / \partial sqft$, and age, i.e. $\partial price / \partial age$, for the smallest or oldest house in the data, the largest or newest house in the data, and a 2300 square foot or a 20 year old house. To help here are the relevant equations:

The complete model is:

(This equation is in the attached Word doc.)

The derivative with respect to size (sqft) to find the marginal effect of size is:

(This equation is in the attached Word doc.)

The derivative with respect to age to find the marginal effect of age is:

(This equation is in the attached Word doc.)

The coefficient of the age term is _____.

Question 13 10 out of 10 points

Add higher (second) order terms for the variables sqft and age to your multivariable model. Use your results from running that model to answer the following questions. (NOTE!!! Do not round until you are ready to submit your final answer. It does make a big difference here.)

Estimate the marginal effects of price as a function of size, i.e. $\partial price / \partial sqft$, and age, i.e. $\partial price / \partial age$, for the smallest or oldest house in the data, the largest or newest house in the data, and a 2300 square foot or a 20 year old house. To help here are the relevant equations:

The complete model is:

(This equation is in the attached Word doc.)

The derivative with respect to size (sqft) to find the marginal effect of size is:

(This equation is in the attached Word doc.)

The derivative with respect to age to find the marginal effect of age is:

(This equation is in the attached Word doc.)

The estimated marginal effect of sqft on price for the smallest house, 662 square feet, is _____.

Question 14 10 out of 10 points

Add higher (second) order terms for the variables sqft and age to your multivariable model. Use your results from running that model to answer the following questions. (NOTE!!! Do not round until you are ready to submit your final answer. It does make a big difference here.)

Estimate the marginal effects of price as a function of size, i.e. $\partial price / \partial sqft$, and age, i.e. $\partial price / \partial age$, for the smallest or oldest house in the data, the largest or newest house in the data, and a 2300 square foot or a 20 year old house. To help here are the relevant equations:

The complete model is:

(This equation is in the attached Word doc.)

The derivative with respect to size (sqft) to find the marginal effect of size is:

(This equation is in the attached Word doc.)

The derivative with respect to age to find the marginal effect of age is:

(This equation is in the attached Word doc.)

The estimated marginal effect of sqft on price for the largest house, 7,897 square feet, is _____.

Question 15 10 out of 10 points

Add higher (second) order terms for the variables sqft and age to your multivariable model. Use your results from running that model to answer the following questions. (NOTE!!! Do not round until you are ready to submit your final answer. It does make a big difference here.)

Estimate the marginal effects of price as a function of size, i.e. $\partial price / \partial sqft$, and age, i.e. $\partial price / \partial age$, for the smallest or oldest house in the data, the largest or newest house in the data, and a 2300 square foot or a 20 year old house. To help here are the relevant equations:

The complete model is:

(This equation is in the attached Word doc.)

The derivative with respect to size (sqft) to find the marginal effect of size is:

(This equation is in the attached Word doc.)

The derivative with respect to age to find the marginal effect of age is:

(This equation is in the attached Word doc.)

The estimated marginal effect of sqft on price for the 2300 square foot house is _____.

Question 16 10 out of 10 points

Add higher (second) order terms for the variables sqft and age to your multivariable model. Use your results from running that model to answer the following questions. (NOTE!!! Do not round until you are ready to submit your final answer. It does make a big difference here.)

Estimate the marginal effects of price as a function of size, i.e. $\partial price / \partial sqft$, and age, i.e. $\partial price / \partial age$, for the smallest or oldest house in the data, the largest or newest house in the data, and a 2300 square foot or a 20 year old house. To help here are the relevant equations:

The complete model is:

(This equation is in the attached Word doc.)

The derivative with respect to size (sqft) to find the marginal effect of size is:

(This equation is in the attached Word doc.)

The derivative with respect to age to find the marginal effect of age is:

(This equation is in the attached Word doc.)

The estimated marginal effect of age on price for the oldest house, 662 square feet, is _____.

Question 17 0 out of 10 points

Add higher (second) order terms for the variables sqft and age to your multivariable model. Use your results from running that model to answer the following questions. (NOTE!!! Do not round until you are ready to submit your final answer. It does make a big difference here.)

Estimate the marginal effects of price as a function of size, i.e. $\partial price / \partial sqft$, and age, i.e. $\partial price / \partial age$, for the smallest or oldest house in the data, the largest or newest house in the data, and a 2300 square foot or a 20 year old house. To help here

are the relevant equations:

The complete model is:

(This equation is in the attached Word doc.)

The derivative with respect to size (sqft) to find the marginal effect of size is:

(This equation is in the attached Word doc.)

The derivative with respect to age to find the marginal effect of age is:

(This equation is in the attached Word doc.)

The estimated marginal effect of age on price for the newest house, 7,897 square feet, is _____.

Question 18 10 out of 10 points

Add higher (second) order terms for the variables sqft and age to your multivariable model. Use your results from running that model to answer the following questions. (NOTE!!! Do not round until you are ready to submit your final answer. It does make a big difference here.)

Estimate the marginal effects of price as a function of size, i.e. $\partial price / \partial sqft$, and age, i.e. $\partial price / \partial age$, for the smallest or oldest house in the data, the largest or newest house in the data, and a 2300 square foot or a 20 year old house. To help here are the relevant equations:

The complete model is:

(This equation is in the attached Word doc.)

The derivative with respect to size (sqft) to find the marginal effect of size is:

(This equation is in the attached Word doc.)

The derivative with respect to age to find the marginal effect of age is:

(This equation is in the attached Word doc.)

The estimated marginal effect of age on price for the 20 year old house is _____.

Question 19 10 out of 10 points

Add higher (second) order terms for the variables sqft and age to your multivariable model. Use your results from running that model to answer the following questions. (NOTE!!! Do not round until you are ready to submit your final answer. It does make a big difference here.)

Estimate the marginal effects of price as a function of size, i.e. $\partial price / \partial sqft$, and age, i.e. $\partial price / \partial age$, for the smallest or oldest house in the data, the largest or newest house in the data, and a 2300 square foot or a 20 year old house. To help here are the relevant equations:

The complete model is:

(This equation is in the attached Word doc.)

The derivative with respect to size (sqft) to find the marginal effect of size is:

(This equation is in the attached Word doc.)

The derivative with respect to age to find the marginal effect of age is:

(This equation is in the attached Word doc.)

The lower bound on a 95% confidence interval around the marginal effect of size on price is _____.

Question 20 10 out of 10 points

Add higher (second) order terms for the variables sqft and age to your multivariable model. Use your results from running that model to answer the following questions. (NOTE!!! Do not round until you are ready to submit your final answer. It does make a big difference here.)

Estimate the marginal effects of price as a function of size, i.e. $\partial price / \partial sqft$, and age, i.e. $\partial price / \partial age$, for the smallest or oldest house in the data, the largest or newest house in the data, and a 2300 square foot or a 20 year old house. To help here are the relevant equations:

The complete model is:

(This equation is in the attached Word doc.)

The derivative with respect to size (sqft) to find the marginal effect of size is:

(This equation is in the attached Word doc.)

The derivative with respect to age to find the marginal effect of age is:

(This equation is in the attached Word doc.)

The upper bound on a 95% confidence interval around the marginal effect of size on price is _____.

Question 21 10 out of 10 points

We know that house prices are affected by the size of the house. It also seems that house prices are greater for owner-occupied houses. Let's quantify this. There are a couple different ways we can do this. First, consider how the size of a house that is owner-occupied affects house prices. Doing this, we can estimate the marginal effect of this.

Make an OLS model for the housing data you have. Include the variables; sqft, (square feet size), bedrooms (number of bedrooms), baths (number of bathrooms), age (years), owner (=1 if owner occupied, 0 if vacant or a rental), pool (=1 is yes, 0 if no). Add an interaction term between sqft and owner, i.e. price x owner.

(Insert equation from Word doc here)

Generate a "log-linear" model where you use the natural log of the price, i.e. ln(price) and the normal values of all other variables. In addition, this time to check normality generate a Q-Q Plot from your results. Remember that we determined earlier that we could make the distribution for price more normal by taking the natural log of it. Use your results to answer the following questions.

The intercept and all the coefficients are statistically significant.

Question 22 10 out of 10 points

We know that house prices are affected by the size of the house. It also seems that house prices are greater for owner-occupied houses. Let's quantify this. There are a couple different ways we can do this. First, consider how the size of a house that is owner-occupied affects house prices. Doing this, we can estimate the marginal effect of this.

Make an OLS model for the housing data you have. Include the variables; sqft, (square feet size), bedrooms (number of bedrooms), baths (number of bathrooms), age (years), owner (=1 if owner occupied, 0 if vacant or a rental), pool (=1 is yes, 0 if no). Add an interaction term between sqft and owner, i.e. price x owner.

(Insert equation from Word doc here)

Generate a "log-linear" model where you use the natural log of the price, i.e. ln(price) and the normal values of all other variables. In addition, this time to check normality generate a Q-Q Plot from your results. Remember that we determined earlier that we could make the distribution for price more normal by taking the natural log of it. Use your results to answer the following questions.

The coefficient of sqft_owner is significant at the 0.05 level of significance.

Question 23 10 out of 10 points

We know that house prices are affected by the size of the house. It also seems that house prices are greater for owner-occupied houses. Let's quantify this. There are a couple different ways we can do this. First, consider how the size of a house that is owner-occupied affects house prices. Doing this, we can estimate the marginal effect of this.

Make an OLS model for the housing data you have. Include the variables; sqft, (square feet size), bedrooms (number of bedrooms), baths (number of bathrooms), age (years), owner (=1 if owner occupied, 0 if vacant or a rental), pool (=1 is yes, 0 if no). Add an interaction term between sqft and owner, i.e. price x owner.

(Insert equation from Word doc here)

Generate a "log-linear" model where you use the natural log of the price, i.e. ln(price) and the normal values of all other variables. In addition, this time to check normality generate a Q-Q Plot from your results. Remember that we determined earlier that we could make the distribution for price more normal by taking the natural log of it. Use your results to answer the following questions.

The R-squared value is _____.

Question 24 10 out of 10 points

We know that house prices are affected by the size of the house. It also seems that house prices are greater for owner-occupied houses. Let's quantify this. There are a couple different ways we can do this. First, consider how the size of a house that is owner-occupied affects house prices. Doing this, we can estimate the marginal effect of this.

Make an OLS model for the housing data you have. Include the variables; sqft, (square feet size), bedrooms (number of bedrooms), baths (number of bathrooms), age (years), owner (=1 if owner occupied, 0 if vacant or a rental), pool (=1 is yes, 0 if no). Add an interaction term between sqft and owner, i.e. price x owner.

(Insert equation from Word doc here)

Generate a "log-linear" model where you use the natural log of the price, i.e. ln(price) and the normal values of all other variables. In addition, this time to check normality generate a Q-Q Plot from your results. Remember that we determined earlier that we could make the distribution for price more normal by taking the natural log of it. Use your results to answer the following questions.

Most of the variance in the data is explained by the model.

Question 25 10 out of 10 points

We know that house prices are affected by the size of the house. It also seems that house prices are greater for owner-occupied houses. Let's quantify this. There are a couple different ways we can do this. First, consider how the size of a house that is owner-occupied affects house prices. Doing this, we can estimate the marginal effect of this.

Make an OLS model for the housing data you have. Include the variables; sqft, (square feet size), bedrooms (number of bedrooms), baths (number of bathrooms), age (years), owner (=1 if owner occupied, 0 if vacant or a rental), pool (=1 is yes, 0 if no). Add an interaction term between sqft and owner, i.e. price x owner.

(Insert equation from Word doc here)

Generate a "log-linear" model where you use the natural log of the price, i.e. ln(price) and the normal values of all other variables. In addition, this time to check normality generate a Q-Q Plot from your results. Remember that we determined earlier that we could make the distribution for price more normal by taking the natural log of it. Use your results to answer the following questions.

Most of the variance in the data is explained by the model.

The P-value from the F-test, i.e. P-value < 0.05, indicates the utility of the model is statistically sound.

Question 26 10 out of 10 points

We know that house prices are affected by the size of the house. It also seems that house prices are greater for owner-occupied houses. Let's quantify this. There are a couple different ways we can do this. First, consider how the size of a house that is owner-occupied affects house prices. Doing this, we can estimate the marginal effect of this.

Make an OLS model for the housing data you have. Include the variables; sqft, (square feet size), bedrooms (number of bedrooms), baths (number of bathrooms), age (years), owner (=1 if owner occupied, 0 if vacant or a rental), pool (=1 is yes, 0 if no). Add an interaction term between sqft and owner, i.e. price x owner.

(Insert equation from Word doc here)

Generate a "log-linear" model where you use the natural log of the price, i.e. ln(price) and the normal values of all other variables. In addition, this time to check normality generate a Q-Q Plot from your results. Remember that we determined earlier that we could make the distribution for price more normal by taking the natural log of it. Use your results to answer the following questions.

Most of the variance in the data is explained by the model.

The coefficient of the variable sqft_owner is essentially zero but its very small value does take away a bit from house prices if a house is owner-occupied.

Question 27 0 out of 10 points

We know that house prices are affected by the size of the house. It also seems that house prices are greater for owner-occupied houses. Let's quantify this. There are a couple different ways we can do this. First, consider how the size of a house that is owner-occupied affects house prices. Doing this, we can estimate the marginal effect of this.

Make an OLS model for the housing data you have. Include the variables; sqft, (square feet size), bedrooms (number of bedrooms), baths (number of bathrooms), age (years), owner (=1 if owner occupied, 0 if vacant or a rental), pool (=1 is yes, 0 if no). Add an interaction term between sqft and owner, i.e. price x owner.

(Insert equation from Word doc here)

Generate an OLS model using price as the dependent variable and the normal values of all other variables. In addition, this time to check normality generate a Q-Q Plot from your results. Remember that we determined earlier that we could make the distribution for price more normal by taking the natural log of it. Use your results to answer the following questions.

Most of the variance in the data is explained by the model.

The marginal effect of size for owner-occupied 2000 square foot houses is _____.

NOTE: if you notice that the plot of residuals is expanding and indicates that there is a problem you are correct. Ordinarily we would transform the dependent variable price by taking the natural log of it and then building the model. However, doing this, in this case, produces the result that the price of houses does not vary with size which doesn't fit the real-world. If you were doing a real analysis, this is something you would (should) discuss in the results section of a report.

Question 28 0 out of 10 points

We know that house prices are affected by the size of the house. It also seems that house prices are greater for owner-occupied houses. Let's quantify this. There are a couple different ways we can do this. First, consider how the size of a house that is owner-occupied affects house prices. Doing this, we can estimate the marginal effect of this.

Make an OLS model for the housing data you have. Include the variables; sqft, (square feet size), bedrooms (number of bedrooms), baths (number of bathrooms), age (years), owner (=1 if owner occupied, 0 if vacant or a rental), pool (=1 is yes, 0 if no). Add an interaction term between sqft and owner, i.e. price x owner.

(Insert equation from Word doc here)

Generate an OLS model, for price as the dependent variable and the normal values of all other variables. In addition, this time to check normality generate a Q-Q Plot from your results. Remember that we determined earlier that we could make the distribution for price more normal by taking the natural log of it. Use your results to answer the following questions.

Most of the variance in the data is explained by the model.

The marginal effect of size for owner-occupied 2500 square foot houses is _____.

NOTE: if you notice that the plot of residuals is expanding and indicates that there is a problem you are correct. Ordinarily we would transform the dependent variable price by taking the natural log of it and then building the model. However, doing this, in this case, produces the result that the price of houses does not vary with size which doesn't fit the real-world. If you were doing a real analysis, this is something you would (should) discuss in the results section of a report.

Question 29 0 out of 10 points

We know that house prices are affected by the size of the house. It also seems that house prices are greater for owner-occupied houses. Let's quantify this. There are a couple different ways we can do this. First, consider how the size of a house that is owner-occupied affects house prices. Doing this, we can estimate the marginal effect of this.

Make an OLS model for the housing data you have. Include the variables; sqft, (square feet size), bedrooms (number of bedrooms), baths (number of bathrooms), age (years), owner (=1 if owner occupied, 0 if vacant or a rental), pool (=1 is yes, 0 if no). Add an interaction term between sqft and owner, i.e. price x owner.

(Insert equation from Word doc here)

Generate a "log-linear" model where you use the natural log of the price, i.e. ln(price) and the normal values of all other variables. In addition, this time to check normality generate a Q-Q Plot from your results. Remember that we determined earlier that we could make the distribution for price more normal by taking the natural log of it. Use your results to answer the following questions.

Most of the variance in the data is explained by the model.

The value of the coefficient for sqft_owner and its sign (plus or minus) makes sense.

Question 30 10 out of 10 points

We know that house prices are affected by the size of the house. It also seems that house prices are greater for owner-occupied houses. Let's quantify this. There are a couple different ways we can do this. First, consider how the size of a house that is owner-occupied affects house prices. Doing this, we can estimate the marginal effect of this.

Make an OLS model for the housing data you have. Include the variables; sqft, (square feet size), bedrooms (number of bedrooms), baths (number of bathrooms), age (years), owner (=1 if owner occupied, 0 if vacant or a rental), pool (=1 is yes, 0 if no). Add an interaction term between sqft and owner, i.e. price x owner.

(Insert equation from Word doc here)

Generate a "log-linear" model where you use the natural log of the price, i.e. ln(price) and the normal values of all other variables. In addition, this time to check normality generate a Q-Q Plot from your results. Remember that we determined earlier that we could make the distribution for price more normal by taking the natural log of it. Use your results to answer the following questions.

Most of the variance in the data is explained by the model.

The transformation by taking the natural log of the price appears to have made the price distribution more normal.

Question 31 10 out of 10 points

Looking at the marginal effect when we simply added an interaction term didn't really work out. And, it seems that taking the natural log of price isn't very effective either. Much earlier we did see that using a quadratic equation resulted in the lowest values for the sum of squares error (SSE). So, let's consider the second way to look at this. Given that owner=1 if the house is owner-occupied and =0 if the house is vacant or a rental, then we can manipulate the equations to take that into account (see Section 5.10 of your textbook). We already expressed the formula for the estimated value of price with sqft_owner as an indicator variable. Let's look at the equations for the case when owner=1 and owner=0 and use a quadratic form.

(Insert the entire set of equations from the Word doc here!)

Now we build the model and get the coefficients. Use your results to answer the following questions.

When the house is vacant or a rental, i.e. when owner = 0, the marginal effect increases for increasing house size.

Question 32 10 out of 10 points

Looking at the marginal effect when we simply added an interaction term didn't really work out. And, it seems that taking the natural log of price isn't very effective either. Much earlier we did see that using a quadratic equation resulted in the lowest values for the sum of squares error (SSE). So, let's consider the second way to look at this. Given that owner=1 if the house is owner-occupied and =0 if the house is vacant or a rental, then we can manipulate the equations to take that into account (see Section 5.10 of your textbook). We already expressed the formula for the estimated value of price with sqft_owner as an indicator variable. Let's look at the equations for the case when owner=1 and owner=0 and use a quadratic form.

(Insert the entire set of equations from the Word doc here!)

Now we build the model and get the coefficients. Use your results to answer the following questions.

When the house is owner-occupied, i.e. when owner = 1, the marginal effect increases for increasing house size.

Question 33 10 out of 10 points

Looking at the marginal effect when we simply added an interaction term didn't really work out. And, it seems that taking the natural log of price isn't very effective either. Much earlier we did see that using a quadratic equation resulted in the lowest values for the sum of squares error (SSE). So, let's consider the second way to look at this. Given that owner=1 if the house is owner-occupied and =0 if the house is vacant or a rental, then we can manipulate the equations to take that into account (see Section 5.10 of your textbook). We already expressed the formula for the estimated value of price with sqft_owner as an indicator variable. Let's look at the equations for the case when owner=1 and owner=0 and use a quadratic form.

(Insert the entire set of equations from the Word doc here!)

Now we build the model and get the coefficients. Use your results to answer the following questions.

Now we build the model and get the coefficients. Use your results to answer the following questions.

The trend of increase price as house size increases makes "practical" sense.

Question 34

10 out of 10 points

Looking at the marginal effect when we simply added an interaction term didn't really work out. And, it seems that taking the natural log of price isn't very effective either. Much earlier we did see that using a quadratic equation resulted in the lowest values for the sum of squares error (SSE). So, let's consider the second way to look at this. Given that owner=1 if the house is owner-occupied and =0 if the house is vacant or a rental, then we can manipulate the equations to take that into account (see Section 5.10 of your textbook). We already expressed the formula for the estimated value of price with sqft_owner as an indicator variable. Let's look at the equations for the case when owner=1 and owner=0 and use a quadratic form.

(Insert the entire set of equations from the Word doc here!)

Now we build the model and get the coefficients. Use your results to answer the following questions.

When the house is owner occupied the marginal effect of increasing size is doubled.

Question 35

10 out of 10 points

Looking at the marginal effect when we simply added an interaction term didn't really work out. And, it seems that taking the natural log of price isn't very effective either. Much earlier we did see that using a quadratic equation resulted in the lowest values for the sum of squares error (SSE). So, let's consider the second way to look at this. Given that owner=1 if the house is owner-occupied and =0 if the house is vacant or a rental, then we can manipulate the equations to take that into account (see Section 5.10 of your textbook). We already expressed the formula for the estimated value of price with sqft_owner as an indicator variable. Let's look at the equations for the case when owner=1 and owner=0 and use a quadratic form.

(Insert the entire set of equations from the Word doc here!)

Now we build the model and get the coefficients. Use your results to answer the following questions.

Consider the plot of the residuals. Does it appear that they are evenly distributed about zero?

Friday, October 20, 2023 5:59:44 PM EDT

← OK