Le Thi Phuong Thao

Student: 413707007

HW0303

**Question 2.17**

1. *Plot house price against house size in a scatter diagram*

A graph of blue dots

AI-generated content may be incorrect.

1. *Estimate the linear regression model, draw a sketch of the fitted line*

A screenshot of a computer

AI-generated content may be incorrect.

We estimate that an additional 100 square feet of living area will increase the expected home price by $13,402.94 holding all else constant. The estimated intercept −115.4236 would imply that a house with zero square feet has an expected price of $−115,423.60.

A graph showing a red line and blue dots

AI-generated content may be incorrect.

1. *Estimate the quadratic regression model*

*A screenshot of a computer code

AI-generated content may be incorrect.*

We estimate that an additional 100 square feet of living area for a 2000 square foot home will increase the expected home price by $7,380.80 holding all else constant

1. *Graph the fitted curve for the model*

A graph with a line and a line

AI-generated content may be incorrect.

1. *Ans*
2. *Ans*
3. *Ans*

**Question 2.25**

1. *Construct a histogram of FOODAWAY and its summary statistics. What are the mean and median values? What are the 25th and 75th percentiles*

A graph with green bars and numbers

AI-generated content may be incorrect.A close up of numbers

AI-generated content may be incorrect.

P25 = 1st quarter = 12.04

P75 = 3rd quarter = 67.50

Mean = 49.27

Median = 32.55

1. *What are the mean and median values of FOODAWAY for households including a member with an advanced degree? With a college degree member? With no advanced or college degree member*

The answers are respectively

73.15 & 48.15

48.60 & 36.11

* 1. & 26.02

A screenshot of a computer

AI-generated content may be incorrect.

1. *Construct a histogram of ln(FOODAWAY) and its summary statistics. Explain why FOODAWAY and ln(FOODAWAY) have different numbers of observations*

A blue graph with black text

AI-generated content may be incorrect.

A close up of a number

AI-generated content may be incorrect.



There are 1200 observations in raw foodaway data and 1022 observations for its logarithm data since Foodway has 0 value, ln(0)=∞so log(foodway) is not computable

1. *Estimate the linear regression*

Slop b2



A unit increase in INCOME is associated with a 0.69% increase in FOODAWAY spending.

1. *Plot ln(FOODAWAY) against INCOME, and include the fitted line from part (d)*

A graph with blue dots and red lines

AI-generated content may be incorrect.

1. *Calculate the least squares residuals from the estimation in part (d). Plot them vs. INCOME. Do you find any unusual patterns, or do they seem completely random?*

A graph with blue dots

AI-generated content may be incorrect.

The residual is quite randomly distributed, there is more whitespaces of the residual plot when income is increased (there are a few high-income observations)

**Question 2.28**

1. *Summary statistics*

A group of numbers and a number of numbers

AI-generated content may be incorrect.

A graph of a graph

AI-generated content may be incorrect.

The wage histogram is right-skewed, meaning most people earn relatively low wages, while a small number earn very high wages.

A graph of a graph

AI-generated content may be incorrect.

The education histogram is more centered, with many people having around 10 to 16 years of education. It appears more balanced (less skewed) compared to wage.

Overall, wage is heavily skewed, while education is more normally distributed

1. *Estimate regression*

A screenshot of a computer program

AI-generated content may be incorrect.

Education has a positive and significant impact on wage. For each additional year of education, wage increases by about $2.15 on average. R squared means around 20.7% of the variation in wage is explained by education alone.

1. *Calculate the least squares residuals and plot them against EDUC. Are any patterns evident? If assumptions SR1–SR5 hold, should any patterns be evident in the least squares residuals*

A graph with blue dots

AI-generated content may be incorrect.

The magnitude of the residuals increases with the large of Educ, suggesting that the error variance is larger for larger values of EDUC.

1. *Estimate the quadratic regression by group and compare results*

A white rectangular object with black text

AI-generated content may be incorrect.

In all sub-sample, education has a positive and statistically significant effect on wages. Females show the highest estimated increase in education, while blacks have the lowest among these groups. Comparing gender and race, education has more influence on wages than race factor. Overall, education alone explains only part of wage variation, indicating other factors also play a role (R square is less than 30%)

1. Plot