EDA Quiz

Continue Course

Back to Week 2

14/14 points earned (100%)

Quiz passed!

---

1 / 1 points

1.
Which of the following are the three variables with the highest number of missing observations?

- *Misc.Feature, Fence, Pool.QC*
- *Misc.Feature, Alley, Pool.QC*

Correct
This question refers to the following learning objective(s):

Use frequency tables and bar plots to describe the distribution of a variable.

- *Pool.QC, Alley, Fence*
- *Fireplace.Qu, Pool.QC, Lot.Frontage*

---

1 / 1 points

2.
How many categorical variables are coded in R as having type *int*? Change them to factors when conducting your analysis.

- 0

1

2

Correct
This question refers to the following learning objective(s):

Identify variables as numerical and categorical.

3

---

1 / 1 points

3.
In terms of price, which neighborhood has the highest standard deviation?

StoneBr

Correct
This question refers to the following learning objective(s):

When describing the distribution of a numerical variable, mention its shape, center, and spread, as well as any unusual observations.

Timber

Veenker

NridgHt

---

1 / 1 points

4.
Using scatter plots or other graphical displays, which of the following variables appears to be the best single predictor of *price*?

Lot.Area

- *Bedroom.AbvGr*
- *Overall.Qual*

Correct
This question refers to the following learning objective(s):

Use scatterplots for describing the relationship between two numerical variables making sure to note the direction (positive or negative), form (linear or non-linear) and the strength of the relationship as well as any unusual observations that stand out.

- *Year.Built*

---

1 / 1 points

5.
Suppose you are examining the relationship between *price* and *area*. Which of the following variable transformations makes the relationship appear to be the most linear?

- Log-tranform neither *price* nor *area*
- Log-transform *price* but not *area*
- Log-transform *area* but not *price*
- Log-tranform both *price* and *area*

Correct
This question refers to the following learning objective(s):

- Recognize when transformations (e.g. log) can make the distribution of data more symmetric, and hence easier to model.
- Use scatterplots for describing the relationship between two numerical variables making sure to note the direction (positive or negative), form (linear or non-linear) and the strength of the relationship as well as any unusual observations that stand out.

---

1 / 1 points

6.

Suppose that your prior for the proportion of houses that have at least one garage is Beta(9, 1). What is your posterior? Assume a beta-binomial model for this proportion.

- Beta(954, 46)
- Beta(963, 46)
- Beta(954, 47)
- Beta(963, 47)

Correct

This question refers to the following learning objective(s):

Make inferences about a proportion using a conjugate Beta prior.

---

1 / 1 points

7.

Which of the following statements is true about the dataset?

- Over 30 percent of houses were built after the year 1999.
- The median housing price is greater than the mean housing price.
- 21 houses do not have a basement.

Correct

This question refers to the following learning objective(s):

Describe the distribution of a single variable.

- 4 houses are located on gravel streets.

---

1 / 1 points

8.

Test, at the $\alpha = 0.05$ level, whether homes with a garage have larger square footage than those without a garage.

- [x] With a p-value near 0.000, we reject the null hypothesis of no difference.

Correct

This question refers to the following learning objective(s):

Use the t-distribution for inference on a single mean, difference of paired (dependent) means, and difference of independent means.

- [ ] With a p-value of approximately 0.032, we reject the null hypothesis of no difference.
- [ ] With a p-value of approximately 0.135, we fail to reject the null hypothesis of no difference.
- [ ] With a p-value of approximately 0.343, we fail to reject the null hypothesis of no difference.

---

1 / 1 points

9.

For homes with square footage greater than 2000, assume that the number of bedrooms above ground follows a Poisson distribution with rate λ. Your prior on λ follows a Gamma distribution with mean 3 and standard deviation 1. What is your posterior mean and standard deviation for the average number of bedrooms in houses with square footage greater than 2000 square feet?

- [ ] Mean: 3.61, SD: 0.11
- [x] Mean: 3.62, SD: 0.16

Correct

This question refers to the following learning objective(s):

- Make inferences about data coming from a Poisson likelihood using a conjugate Gamma prior.

- Elicit prior beliefs about a parameter in terms of a Beta, Gamma, or Normal distribution.

- Mean: 3.63, SD: 0.09
- Mean: 3.63, SD: 0.91

---

1 / 1 points

10.

When regressing $\log(price)$ on $\log(area)$, there are some outliers. Which of the following do the three most outlying points have in common?

- They had abnormal sale conditions.
- They have only two bedrooms.
- They have an overall quality of less than 3.
- **They were built before 1930.**

Correct

This question refers to the following learning objective(s):

Identify outliers and high leverage points in a linear model.

---

1 / 1 points

11.

Which of the following are reasons to log-transform *price* if used as a dependent variable in a linear regression?

- **a. *price* is right-skewed.**

Correct

This question refers to the following learning objective(s):

Identify the assumptions of linear regression and assess when a model may need to be improved.

- b. *price* cannot take on negative values.
- c. *price* can only take on integer values.
- d. Both a and b

---

1 / 1 points

12.

How many neighborhoods consist of only single-family homes? (*Bldg.Type = 1Fam*)

- 0
- 1
- 2
- 3 (selected)

Correct

This question refers to the following learning objective(s):

Use contingency tables and segmented bar plots or mosaic plots to assess the relationship between two categorical variables.

---

1 / 1 points

13.

Using color, different plotting symbols, conditioning plots, etc., does there appear to be an association between $\log(area)$ and the number of bedrooms above ground (*Bedroom.AbvGr*)?

- Yes (selected)

Correct

This question refers to the following learning objective(s):

Use scatterplots and other graphical devices to describe the relationship between two numerical variables.

No

1 / 1 points

14.
Of the people who have unfinished basements, what is the average square footage of the unfinished basement?

590.36

595.25

Correct
This question refers to the following learning objective(s):

Describe the distribution of a single variable.

614.37

681.94