

Quiz 4 — 2/8/2024

Instructions

This take-home quiz is due on **Thursday, February 8 at 23:59**.

You may use your own course materials, as well as any materials directly linked from the course website. **No collaboration allowed.**

Type your answers **directly in this Jupyter notebook**, and submit this notebook (just the `ipynb` file) using the submission form on the [course website](#).

Problem 1

The data frame `RailsTrails` in the `Stat2Data` package contains data from a sample of 104 homes that were sold in 2007 in the city of Northampton, Massachusetts.

For this quiz, we want to see if a home's proximity to a biking trail is related to its selling price. Perhaps, for example, proximity to the trail could add value to the home.

Here, we focus on the estimated 2007 price of a home, in thousands of dollars. These values are in the variable `Price2007`. The variable `Distance` measures the distance in miles to the nearest entry point to the trail network.

a.

Load the data frame `RailsTrails` from the `Stat2Data` package. Display the first few rows.

In []:

b.

Use R to fit a simple linear regression model that predicts the estimated 2007 price of a home based on its distance to the nearest entry point to the the trail network.

Provide **only** the summary output for this part.

In []:

c.

Interpret the value of the slope of the fitted model, in the context of this setting.

Write your answer here. Double-click to edit.

Feedback. See the bottom of page 2 of Lesson 7, as well as Example 2b in Lesson 6 for guidance on this part.

Note that the estimated slope of the fitted simple linear regression model gives you information about how changes in the explanatory variable affect the response variable **on average** (or equivalently, **in expectation**).

d.

What estimated 2007 price does the fitted model predict for a house that is 1.5 miles from the nearest entry point to the trail network? **Report your answer in dollars.**

Use the code cell below as a calculator to show your work.

In []:

Feedback. See STAT2 Exercise 1.21a, assigned for homework, for a similar example.

e.

The first house in the dataset, at 406 Acrebrook Drive, has an estimated 2007 price of \$203,500, and is 2.4 miles from the nearest entry point to the trail network. Find the residual for this observation. **Report your answer in dollars.**

Use the code cell below as a calculator to show your work.

In []:

Feedback. See Example 2d in Lesson 7 and STAT2 Exercise 1.21b for similar examples.

f.

Make a residual vs. fitted values plot.

In []:

g.

Based on your residual vs. fitted values plot in part f, does the model satisfy the equal variance condition? Briefly explain.

Your explanation should be in complete sentences, with correct spelling and grammar.

Write your answer here. Double-click to edit.

Feedback. See Lesson 8 Part 2 for a similar example.

h.

Based on your residual vs. fitted values plot in part f, does the model satisfy the linearity condition? Briefly explain.

Your explanation should be in complete sentences, with correct spelling and grammar.

Write your answer here. Double-click to edit.

Feedback. See Lesson 8 Part 2 for a similar example.

Grading rubric

Problem	Weight
---------	--------

Problem	Weight
1a	1
1b	1
1c	1
1d	1
1e	1
1f	1
1g	1
1h	1
Max Score	80