Stanford | Online High School

OM065 Quiz Instructions

Before the assessment:

1. Review all of these instructions and the Honor Code before taking the exam.

2. You have 2 hours, in one sitting to complete the exam.

3. Solve the problems on these pages, attaching additional pages if necessary. You may choose to create an R notebook for your work and add text fields to answer the questions in words and provide interpretations of the numbers you get when needed.

4. You may use R/ Python or another statistical package (in combination with R studio or another R interface, etc.) on your computer and the R commands reference manuals for this class (if you choose a statistical software other than R you must know it well enough to execute the code without consulting references). No notes, books, or other references, online or offline resources are allowed.

5. The OHS Honor Code applies, and a parent, guardian, teacher, or other adult must confirm (by providing a signature) that you have followed these instructions.

6. You must show all work to receive credit. Answers for which no work is shown will receive no credit (unless specifically stated otherwise).

8. For questions specific to this assessment, please contact:
   email preferred: Victoria D'Urso
   e-mail: vdurso@stanford.edu
   cell: 650-390-3828

During the assessment:

1. Supervise the student, ensuring he/she follows all of the above instructions.

2. Make sure the student's testing area is quiet and well lit, and that the student will not be disturbed.

3. Periodically update the student with the amount of time remaining.

After the assessment:

1. Have the student sign the Honor Code, then sign the Honor Code yourself.

2. The student should turn in the quiz solutions file to the appropriate dropbox on Canvas.

Good Luck!

Stanford | Online High School

OHS Honor Code Statement

Dr. V D'Urso OM065

For the student: I attest to the following (please check each box):

- ☑ I followed the instructions provided carefully and honestly, and
- ☑ The work returned to the Online High School is my own, produced within the stated time limits without assistance from books, notes, or other aids beyond those specifically permitted.

Student's Name (print): Neev Shaw

Student's Signature: Neev Shaw **Date:** 9/13/2025

For the parent / guardian: I attest to the following (please check each box):

- ☑ I have diligently supervised this student, ensuring that the instructions provided were followed carefully and honestly,
- ☑ The work returned to the Online High School is solely that of the student above, produced within the stated time limits without assistance from books, notes, or other aids beyond those specifically permitted, and
- ☑ I attest that the student's identity matches the name of the work submitted.
- ☑ I did not aid the student with the assessment except in the interpretation of the instructions.

Parent / Guardian's Name (print): Komal Shaw

Parent/ Guardian's Signature: Komal Shaw **Date: 9/13/2025**

Assessment date: 9/13/2025 **Start time:** 2:27 pm **Finish time:** 3:46 pm

You may use R or another statistical package as needed for any portion of the exam. Unless otherwise specified, assume a 5% significance level for any tests and 95% confidence level for any intervals. The point values for each part are in parentheses.

Some of the questions on this exam will refer to the dataset **HighPeaks**. Forty-six mountains in the Adirondacks of upstate New York are known as the High Peaks with elevations near or above 4000 feet. The file HighPeaks contains information on the *Elevation* (in feet) of each peak along with data on typical hikes including *Difficulty* rating (on a 1 to 7 scale with 7 being the most difficult), the *Ascent* (in feet), *Length* of a round trip distance (in miles), and expected trip *Time* (in hours).

1. Refer to the **HighPeaks** data.

(a) Fit the regression model of Y = *Time* on X = *Ascent*.

(i) What is the fitted model? (2)

(ii) Test the null hypothesis that *Ascent* is not related to *Time*. Provide the P-value and state the conclusion. (2)

(b) From your model in part (a), interpret the coefficient of *Ascent* in the context of this setting. (2)

(c) Construct a 95% confidence interval for β_1, the coefficient of *Ascent* in the model. (2)

(d) Construct a 95% prediction interval for trip *Time* when *Ascent* is 3000. (2)

(e) Are there any mountains that have unusually large residuals for this model? If so, what is the (approximate) magnitude of the largest standardized residual? (You can read this off of a graph; there is no need for a precise calculation.) (2)

2. In general (not referring to the **HighPeaks** data specifically) draw scatterplots that illustrate the *failure* of each of these two simple linear regression assumptions. Label your axes to indicate what is being plotted in each case.

(a) a plot that shows a *lack* of linearity. (2)

(b) a plot that shows a *lack* of homoscedasticity. (2)

3. Refer to the **HighPeaks** data.

(a) What is the correlation between Y = *Time* of a hike and X = *Elevation* of the mountain? (2)

(b) Does it look like *Elevation* should be very helpful in predicting *Time*? Conduct an appropriate test and explain (2)