CSCI 3022
Final Exam
Spring 2019

Name:

Student ID:

Read the following:

- **RIGHT NOW**! Write your name on the top of your exam.

- You are allowed **one** 3×5in notecard of **handwritten** notes (both sides). No magnifying glasses!

- You may use a calculator provided that it cannot access the internet or store large amounts of data.

- You may **NOT** use a smartphone as a calculator.

- Clearly mark answers to multiple choice questions on the provided answer line.

- Mark only one answer for multiple choice questions. If you think two answers are correct, mark the answer that **best** answers the question. No justification is required for multiple choice questions.

- If you do not know the answer to a question, skip it and come back to it later.

- For free response questions you must clearly justify all conclusions to receive full credit. A correct answer with no supporting work will receive no credit.

- You have **150 minutes** for this exam.

Page	Points	Score
MC	32	
SA	8	
FR1	20	
FR2	20	
FR3	20	
Total	100	
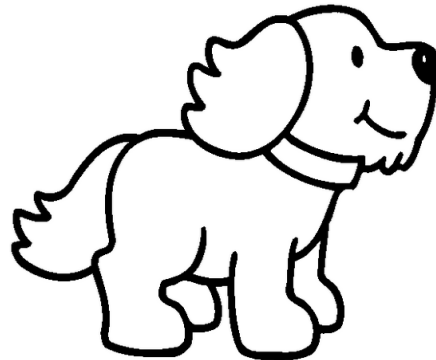
Potentially Useful Values and Formulas

Standard Normal Distribution: Here $\Phi(z)$ is the cumulative distribution function for the standard normal distribution evaluated at z. Its equivalent form in Python is $\Phi(z) = \text{stats.norm.cdf(z)}$,

$$\Phi(4.75) \approx 1.000 \quad \Phi(3.00) = 0.999 \quad \Phi(2.58) = 0.995 \quad \Phi(2.32) = 0.990 \quad \Phi(2.20) = 0.986$$

$$\Phi(2.00) = 0.977 \quad \Phi(1.96) = 0.975 \quad \Phi(1.80) = 0.964 \quad \Phi(1.75) = 0.960 \quad \Phi(1.64) = 0.950$$

$$\Phi(1.60) = 0.945 \quad \Phi(1.50) = 0.933 \quad \Phi(1.40) = 0.919 \quad \Phi(1.28) = 0.900 \quad \Phi(1.20) = 0.885$$

$$\Phi(1.04) = 0.850 \quad \Phi(1.00) = 0.841 \quad \Phi(0.93) = 0.825 \quad \Phi(0.84) = 0.800 \quad \Phi(0.76) = 0.775$$

$$\Phi(0.67) = 0.750 \quad \Phi(0.60) = 0.725 \quad \Phi(0.52) = 0.700 \quad \Phi(0.45) = 0.675 \quad \Phi(0.40) = 0.655$$

$$\Phi(0.32) = 0.625 \quad \Phi(0.25) = 0.600 \quad \Phi(0.19) = 0.575 \quad \Phi(0.13) = 0.550 \quad \Phi(0.00) = 0.500$$

Student's t-Distribution: The following values of the form $t_{\alpha,v}$ are the critical values of the t-distribution with v degrees of freedom, such that the area under the pdf and to the right of $t_{\alpha,v}$ is α. Its equivalent form in Python is $t_{\alpha,v} = \text{stats.t.ppf}(1 - \alpha, v)$.

$$t_{0.05,23} = 1.714 \qquad t_{0.025,23} = 2.069$$
$$t_{0.05,22} = 1.717 \qquad t_{0.025,22} = 2.074$$
$$t_{0.05,21} = 1.721 \qquad t_{0.025,21} = 2.080$$
$$t_{0.05,20} = 1.725 \qquad t_{0.025,20} = 2.086$$

F-Distribution: The following values of the form F_{α,v_1,v_2} are the critical values of the F-distribution with v_1 and v_2 degrees of freedom, such that the area under the pdf and to the right of F_{α,v_1,v_2} is α. Its equivalent form in Python is $F_{\alpha,v_1,v_2} = \text{stats.f.ppf}(1 - \alpha, v_1, v_2)$.

$$F_{0.1,2,9} = 3.006 \qquad F_{0.1,3,12} = 2.606$$
$$F_{0.05,2,9} = 4.256 \qquad F_{0.05,3,12} = 3.490$$
$$F_{0.025,2,9} = 5.271 \qquad F_{0.025,3,12} = 5.715$$
$$F_{0.01,2,9} = 8.022 \qquad F_{0.01,3,12} = 5.953$$

Bayes' theorem	$p(A \mid B) = \dfrac{p(B \mid A)p(A)}{p(B)}$	Law of total probability	$p(E) = \sum\limits_{i=1}^{N} p(E \mid F_i)p(F_i)$
Union of sets	$p(A \cup B) = p(A) + p(B) - p(A \cap B)$	Conditional probability	$p(A \mid B) = \dfrac{p(A \cap B)}{p(B)}$
Sigmoid function	$\text{sigm}(z) = \dfrac{1}{1 + e^{-z}}$	Regression	$\hat{\sigma}^2 = \dfrac{SSE}{n - 2}, \quad SE(\hat{\beta}) = \dfrac{\hat{\sigma}}{\sqrt{\sum(x_i - \bar{x})^2}}$
Get that bread!		Binomial coefficients:	$C(n, k) = \binom{n}{k} = \dfrac{n!}{k! \, (n - k)!}$

Three types of F you might use:	$F = \dfrac{(SST - SSE)/p}{SSE/(n - p - 1)}$	$F = \dfrac{SSB/df_{SSB}}{SSW/df_{SSW}}$	$F = \dfrac{(SSE_{red} - SSE_{full})/(p - k)}{SSE_{full}/(n - p - 1)}$
Some confidence intervals:	$\bar{x} \pm z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$	$\bar{x} \pm t_{\alpha/2,n-1}\dfrac{\sigma}{\sqrt{n}}$	$\left[\dfrac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}}, \dfrac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}}\right]$

Multiple choice problems: Write your answers in the boxes, or they will not be graded!

Figure 1. Tasty or snuggly?

1. (2 points) You are writing a program to determine whether an object is a puppy or a blueberry muffin (see Figure 1), so you know which objects you can eat and which objects you can snuggle. The program tests the null hypothesis that the object is a puppy (i.e., that it is *not* edible) against the alternate hypothesis that the object is a blueberry muffin (i.e., that it *is* edible). If your program determines that an object is a blueberry muffin, you will eat it without question. What type of error is preferable?

    A. Type I error (false positive)

    B. Type II error (false negative)

    C. Type III error

    D. Type IV error

    E. Your program will classify an object perfectly every time, so there is no need to worry about errors

    B

2. (2 points) You are playing your favorite video game when an announcement appears on the screen: "ENTERING THE DATA SCIENCE LEVEL." Your character appears in front of a graveyard. Each headstone has a birth date, a death date, and an animal mascot of Data Science semesters past. An instruction appears on your screen: "Press F-statistic to pay respects."

    Curious as to whether or not the features of (i) animal type, (ii) weight, (iii) hair majesty, and (iv) number of appendages affects the life span of each animal in a statistically significant manner, you perform a multiple linear regression F test on the data. You compute an F-statistic of $F = 2.5$, and you calculate that the F-critical value (F_{crit}) for your significance level is at 3.2. Assuming your F test is valid, which **one** of the following corresponds to an appropriate conclusion?

    A. All of the features are statistically significantly different from 0

    B. At least one of the features is statistically significantly different from 0

    C. None of the features is statistically significantly different from 0

    D. At least one of the features is **not** statistically significantly different from 0

    E. All of the features' slope coefficients are statistically significantly different from 0

    F. At least one of the features' slope coefficients is statistically significantly different from 0

    G. None of the features' slope coefficients is statistically significantly different from 0

    H. At least one of the features' slope coefficients is **not** statistically significantly different from 0

    G

3. (2 points) You are at a psychic getting a reading done. You look into his crystal ball and see a multiple linear regression model. The psychic holds out two cards face down. One has an R^2 value on the other side, and the other has an adjusted R^2 value (R^2_{adj}). Both values correspond to the same multiple linear regression model, but one of the values is much greater than the other one. The psychic is not very good at his job, however, and asks you to tell him which card has the greater value. Which do you pick?

A. The R^2 card because the model does not fit the data well

B. The R^2 card because the model might not have enough features

C. The R^2 card because the model might have extraneous (too many) features

D. The R^2_{adj} card because the model does not fit the data well

E. The R^2_{adj} card because the model might not have extraneous features

F. Neither - they are always the same value

G. You go home because you just wanted to see your future data science grade and this guy is terrible at his job

```
C
```

4. (2 points) You are playing a card game with a hedgehog named Alpaca. Alpaca is known to be a bit of a cheater. The probability that Alpaca is cheating, without knowing anything else, is 50%. Alpaca is a very good cheater, and when Alpaca cheats, they have an 80% chance of winning, as opposed to a 50% chance of winning when not cheating. Alpaca just won. What is the probability that Alpaca was cheating (to 3 decimal places)?

A. 0.1

B. 0.25

C. 0.315

D. 0.4

E. 0.5

F. 0.615

G. 0.815

H. 0.9

```
F
```

5. (2 points) Samuel, your hedgehog companion, is working in a film class, and has recently posted a trailer for a movie he's planning. You show this to 100 of your internet friends and they each rate how much they like the trailer on a scale from 0 to 100. Then, you compute a 90% confidence interval for the mean rating of the trailer. The confidence interval you obtain is [40, 60]. You want to check this result using a larger sample so you share the video with 300 **more** people, for a new total sample size of 400. The mean rating from your new, larger sample is lower than the smaller sample by 2, but the standard deviation is the same. Using the sample size of 400, what is the 90% confidence interval for the mean rating of Samuel's movie trailer?

A. $[39, 61]$

B. $[43, 53]$

C. $[40, 60]$

D. $[45, 55]$

E. $[38, 58]$

```
B
```

6. (2 points) You are walking down the street living your best life when a man dressed entirely in yellow and green attire yells at you from across the street, clutching Joseph the hedgehog. He says,

"I have a challenge for you, Data Scientist. If you pass this challenge, I shall release Joseph. If not, then I will take Joseph to a farm upstate."

You bravely accept the challenge and begin walking toward this bizarre man. He sneers and asks:

"Oh, so you are approaching me? Well then, suppose you take an average of 4 steps per second. Which of these distributions is most appropriate if you wanted to find the probability that you take 7 steps in some amount of time?"

A. Geometric

B. Exponential

C. Normal

D. Negative binomial

E. Uniform

F. Poisson

G. Binomial

H. Bernoulli

> F

7. (2 points) Consider the following function. The function output constitutes a sample from which one of the following distributions?

```
def neat_o_complete_o(p):
    x = 0
    y = 5
    while y > 0:
        draw = np.random.choice([0,1], p=[1-p, p])
        x += 1
        if draw == 0:
            y -= 1
    return x
```

A. Geometric

B. Exponential

C. Normal

D. Negative binomial

E. Uniform

F. Poisson

G. Binomial

H. Bernoulli

> D

8. (2 points) Consider the following function. The function output constitutes a sample from which one of the following distributions?

```
def what_the_function():
    x = 0
    y = 5
    draw = stats.uniform.rvs(loc=-5, scale=20) # uniform random draw between -5 and 15
    if (draw >= x) and (draw <= y):
        return 1
    else:
        return 0
```

A. Geometric

B. Exponential

C. Normal

D. Negative binomial

E. Uniform

F. Poisson

G. Binomial

H. Bernoulli

> H

9. (2 points) Suppose that you are performing a binary logistic regression classification to assign a class label $y \in \{0, 1\}$ to each data point and you model the probability that data point x belongs to Class 1 by

$$p(y = 1 \mid x) = \text{sigm}(\hat{\beta}_0 + \hat{\beta}_1 x)$$

where $\hat{\beta}_0 = 3$ and $\hat{\beta}_1 = 2$. If $p(y = 1 \mid x) \geq 0.5$, you classify as Class 1, otherwise, you classify as Class 0. How would your model classify a data point with $x = 2$?

$$\text{Recall:} \quad \text{sigm}(z) = \frac{1}{1 + e^{-z}}$$

A. inconclusive

B. class 0

C. class 1

D. $\hat{y} = 0.5$

E. $\hat{y} = 7$

F. the limit does not exist

C

10. (2 points) For the same logistic regression model given in the previous problem, what is the decision boundary?

A. $x = -3/2$

B. $x = -2/3$

C. $x = -1/2$

D. $x = 0$

E. $x = 1/2$

F. $x = 2/3$

G. $x = 3/2$

A

11. (2 points) Let $f(x)$ be the PDF of a normal distribution with mean 1 and variance 1. Compute

$$\int_{-0.4}^{2.2} f(x)dx$$

A. 0

B. 0.081

C. 0.115

D. 0.196

E. 0.331

F. 0.642

G. 0.804

H. 0.885

I. 0.919

J. 1

G

12. (2 points) Suppose you generate 1,000 confidence intervals for the mean of a population, using fixed significance level α. You discover that 798 of them in fact do contain the true mean. Which of the following is the most appropriate estimate of the significance level α?

    A. 0.01

    B. 0.02

    C. 0.05

    D. 0.1

    E. 0.2

    F. 0.8

$$\boxed{E}$$

13. (2 points) Suppose you compute a sample mean, drawn from a population that is normally distributed with known variance σ^2. Which combination of significance level α and sample size n produces the **narrowest** confidence interval for the mean?

    A. $\alpha = 0.1$ and $n = 25$

    B. $\alpha = 0.1$ and $n = 36$

    C. $\alpha = 0.05$ and $n = 25$

    D. $\alpha = 0.05$ and $n = 36$

    E. $\alpha = 0.01$ and $n = 25$

    F. $\alpha = 0.01$ and $n = 36$

$$\boxed{B}$$

14. (2 points) The average weight of a Daurian hedgehog is about 590 g. The weight of the these hedgehogs is known to be normally distributed, but an enterprising young data scientist/hedgehog enthusiast is concerned that the variation in the hedgehog weights is larger than acceptable. In an attempt to estimate the variance, she selects $n = 9$ hedgehogs at random and weighs them. The sample yields a sample variance of 30 g^2. Which of the following gives a 95% CI for the variance?

Recall that $\chi^2_{\alpha,\nu}$ and $t_{\alpha,\nu}$ are the χ^2 and t statistics (respectively) with α probability above it and ν degrees of freedom.

    A. $\dfrac{8 \cdot 30}{\chi^2_{0.025,8}} \leq \sigma^2 \leq \dfrac{8 \cdot 30}{\chi^2_{0.975,8}}$

    B. $\dfrac{8 \cdot 30}{\chi^2_{0.05,8}} \leq \sigma^2 \leq \dfrac{8 \cdot 30}{\chi^2_{0.95,8}}$

    C. $\dfrac{8 \cdot 30}{-\chi^2_{0.025,8}} \leq \sigma^2 \leq \dfrac{8 \cdot 30}{\chi^2_{0.025,8}}$

    D. $30 - \chi^2_{0.025,8} \cdot \sqrt{\dfrac{30}{9}} \leq \sigma^2 \leq 30 + \chi^2_{0.025,8} \cdot \sqrt{\dfrac{30}{9}}$

    E. $590 - \chi^2_{0.025,8} \cdot \sqrt{\dfrac{30}{9}} \leq \sigma^2 \leq 590 + \chi^2_{0.025,8} \cdot \sqrt{\dfrac{30}{9}}$

    F. $590 - t_{0.025,8} \cdot \sqrt{\dfrac{30}{9}} \leq \sigma^2 \leq 590 + t_{0.025,8} \cdot \sqrt{\dfrac{30}{9}}$

$$\boxed{A}$$

15. (2 points) Suppose that the random variable X has mean 5 and standard deviation 3. Let Y be the random variable given by $Y = X^2 - 2$. What is the expected value of Y?

    A. 5

    B. 17

    C. 23

    D. 25

    E. 32

    F. 36

<div style="border:1px solid black; display:inline-block; padding:10px;">E</div>

16. (2 points) A game of **Plinko** is to be played on the board shown below. The pegs are unbiased, meaning that the disc has equal probability of moving left or right at each peg. Furthermore, the disc can only be dropped from directly above the top-most peg. What is the expected value of your winnings with a single disc?

$40	$10	$0	$30
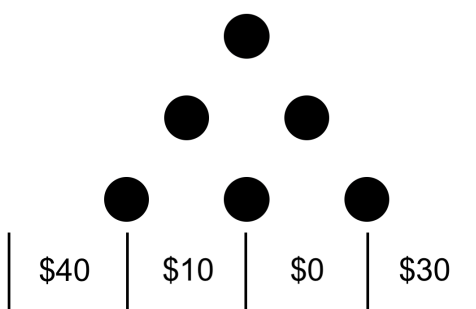
    A. $8

    B. $10

    C. $12.5

    D. $20

    E. $26.67

    F. $80

<div style="border:1px solid black; display:inline-block; padding:10px;">C</div>

The rest of this page may be used for scratch work.

17. (4 points) What is a p-value? Describe, in words, what a p-value is meant to quantify. Then, provide a definition in terms of a probability (you may use words in your probability terms). What do we use it for?

    Solution:

    In words: A p-value is a quantification of how probable (or how likely) it is that we would observe the data we saw, or anything more extreme, in a world where the null hypothesis is true. (Could also replace "data" with "test statistic")

    In math: p-value $= Pr(\text{data or more extreme} \mid H_0 \text{ true})$

    We use p-values to characterize how unlikely what we observed is, assuming the null hypothesis was true. If what we saw is super unlikely (low enough p-value), then we will **reject the null hypothesis**.

18. (4 points) Amy, the famous hedgehog data scientist, likes to perform hypothesis tests to soothe her weary bones at the end of a long day of science. On this particular day, Amy draws a sample of size 100 from the set of all CU students, and computes the sample mean height. Amy decides that if her sample mean exceeds 70 inches (5 ft 10 in), then that is sufficient evidence to convince her that CU students are significantly taller than 67 inches (5 ft 7 in), which is about average for adults in the United States. From her previous work, Amy knows that the standard deviation of the distribution of all CU students' heights is 20 inches.

    What significance level is Amy using in this hypothesis testing experiment? Show all work.

    Solution:
    $$Z_{crit} = \frac{70 - 67}{20/\sqrt{100}} = 1.5$$
    The amount of probability above $Z = 1.5$ gives the significance level:
    $$\alpha = 1 - \Phi(1.5) = 0.067 \quad \longrightarrow \quad \boxed{\alpha = 0.067}$$

Free response problems: If your answers do not fit on the page and the one right after it, <u>make a note</u> of where the work is continued!

19. (20 points) Suppose you have collected a sample from a population whose probability density function is governed by some unknown parameter θ. You compute a test statistic, X, which has the following probability density function, where C is a constant.

$$f(x) = \begin{cases} \frac{1}{2} & \theta - C \leq x \leq \theta + C \\ 0 & \text{otherwise} \end{cases}$$

Please answer the following questions, and be sure to show your work—for sufficient space, a blank page follows this one.

(a) (4 points) For what value of the constant C is $f(x)$ a valid probability density function? Remember to show all work!

   Solution:
   $f(x) \geq 0$ for all x, so we just need to check that f integrates to 1.

   $$\int_{-\infty}^{\infty} f(x)dx = \int_{\theta-C}^{\theta+C} \frac{1}{2}dx$$
   $$= \frac{1}{2}(\theta + C - (\theta - C))$$
   $$= \frac{1}{2}(2C) = C \overset{\heartsuit}{=} 1 \quad \longrightarrow \quad \boxed{C = 1}$$

(b) (4 points) You want to see if there is sufficient evidence to conclude that the true parameter value is *greater than* $\theta = 1$. State the relevant null and alternative hypotheses.

   Solution:

 $$H_0 : \theta = 1$$
   $$H_1 : \theta > 1$$

(c) (6 points) Sketch the probability density function $f(x)$, assuming that the null hypothesis from part (b) is true. Label your axes, the density function f, a few important $x-$ and $y-$tick marks along your axes.

   Suppose the test statistic that you compute is $X = 3/2$. Depict this in your sketch as a vertical line and label with "TS" or "Test Statistic". Clearly mark and label in your sketch what the area/value/point that gives the p-value associated with the hypothesis test from part (b).

   Solution:
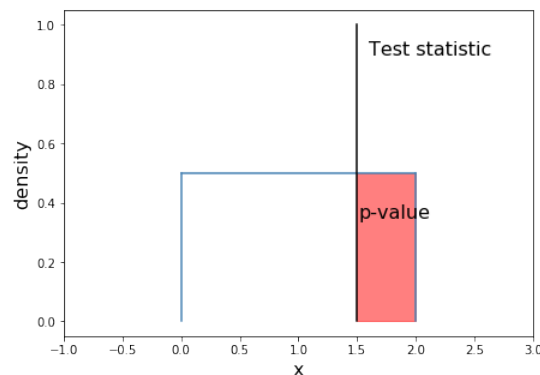
(d) (4 points) Compute the p-value associated with the hypothesis test from part (b), assuming you have the test statistic $X = 3/2$. You must set this up **as an integral** first, then simplify and perform the calculation by hand to receive credit.

Solution:

$$\text{p-value} = P(x \text{ as extreme or more extreme than TS } | \text{ H0 true})$$

$$= \int_{TS}^{\infty} f(x)dx = \int_{3/2}^{2} \frac{1}{2}dx \quad \longleftarrow f(x) \text{ assuming H0 is true}$$

$$= 0.5 \times 0.5 \quad \longleftarrow \text{ it's a rectangle, yo!}$$

$$= \boxed{0.25}$$

(e) (2 points) What can we conclude, at the 10% significance level?

Solution: p-value $= 0.25 > \alpha = 0.1$, so we **fail to reject the null hypothesis**

20. (20 points) Suppose you use statsmodels OLS to perform a simple linear regression of the form $y = \beta_0 + \beta_1 x$ on data consisting of $n = 23$ observations and obtain the following results:

```
                        OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.331
Model:                            OLS   Adj. R-squared:                  0.299
Method:                 Least Squares   F-statistic:                     10.39
Date:                Sun, 39 Dec 2025   Prob (F-statistic):            0.00407
Time:                         3:45:67   Log-Likelihood:                -89.784
No. Observations:                  23
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         -1.6176      5.562     -0.291      0.774     -13.184      9.948
x              2.9037      0.901      ?????      ?????      ?????      ?????
==============================================================================
```

Please answer the following questions, and be sure to show your work—for sufficient space, a blank page follows this one.

(a) (6 points) Compute the missing 95% confidence interval for the slope parameter.

Solution:
The confidence interval is a 95% t CI, given by: $\hat{\beta}_1 \pm t_{\alpha/2, n-2} \cdot SE(\hat{\beta}_1)$

The table gives $\hat{\beta}_1 = 2.9037$ and $SE(\hat{\beta}_1) = 0.901$, there are $n = 23$ data points, and we have $\alpha = 0.05$ for a 95% CI.

This yields:

$$CI = 2.9037 \pm t_{0.025, 21} \cdot 0.901$$
$$= 2.9037 \pm 2.08 \cdot 0.901$$
$$= \boxed{[1.03, \quad 4.78]}$$

(b) (2 points) Based on your confidence interval, do we have reason to believe that β_1 is different from zero? Fully justify your response.

Solution:
Yes, because the CI does not contain 0. (at the 95% confidence level)

(c) (6 points) Suppose we want to test the hypothesis that the slope coefficient β_1 is *greater than* 1. State the null and alternate hypotheses and perform a relevant hypothesis test at the 5% significance level. Clearly state your conclusions.

Solution:
$$H_0 : \beta_1 = 1$$
$$H_1 : \beta_1 > 1$$

Test statistic is $t = \dfrac{\hat{\beta}_1 - 1}{SE(\hat{\beta}_1)} = \dfrac{1.9037}{0.901} = 2.113$ and it follows a t distribution with $\nu = n - 2 = 21$ degrees of freedom.

12

The upper tail rejection region is bounded below by $t_{0.05,\nu=21} = 1.721$, so any test statistic more extreme that this will lead us to reject the null hypothesis.

Since our test statistic (2.113) is greater than $t_{crit} = 1.721$, we **reject the null hypothesis** and conclude that there is evidence at the 5% significance level that the slope coefficient $\beta_1 > 1$

(d) (2 points) What fraction of the total variation in the response is explained by the simple linear regression model? Fully justify your response.

Solution:

The coefficient of determination is defined to be the fraction of variation that is explained by the model. This is $\boxed{R^2 = 0.331}$

(e) (4 points) Suppose this linear regression model relates the feature $X =$ "amount of time spent handling hedgehogs" (in minutes) and response $Y =$ "exam scores". Use your previous answers and the output above to evaluate the **strength** and **significance** of the relationship between hedgehog handling and exam scores. Fully justify your response.

Solution:

Parts (a) and (b) suggest that there **is a significant relationship** between hedgehog handling and exam scores.

Buuuuuuut Part (d) indicates that **this relationship is quite weak** because the SLR model using hedgehog handling as a feature does not explain much of the variation in exam scores.

You can also note that the slope coefficient $\hat{\beta}_1$ is quite small, but an argument based on R^2 is much better.

21. (20 points) You are in Tony's CSCI 3022 Introduction to Data Science course, and you are talking to your favorite CA about your grade. She tells you about her hypothesis that students who laugh at Tony's bad jokes in class get better grades than those who do not. Being an aspiring data scientist yourself, you decide to test her hypothesis. You collect a data set consisting of 12 students' grades on a particular quiz. You separate the students into three groups: (1) students who have *No Reaction* to Tony's bad jokes, (2) students who give a *Pity Laugh* at the bad jokes, and (3) students who *Laugh Out Loud* at the bad jokes. You obtain the following data for the students' grades, and decide to do a **one-way ANOVA** test to compare the three groups of students.

No Reaction (N)	Pity Laugh (P)	Laugh Out Loud (L)
4	7	8
5	8	9
7	8	11
8	9	12

(a) (4 points) Clearly state the null and alternative hypotheses for the one-way ANOVA test to compare the three groups of students and determine whether or not there is evidence that there is some difference in performance on this quiz.

Solution:

Let μ_N, μ_P and μ_L correspond to the mean grades in the No reaction, Pity laugh, and Laugh out loud groups.

$$H_0 : \mu_N = \mu_P = \mu_L$$
$$H_1 : \mu_i \neq \mu_j \text{ for some } i \neq j$$

(b) (8 points) Compute the relevant **test statistic** to test the hypotheses from part (a). Put a $\boxed{\textbf{box}}$ around your answer for the test statistic. Show all work!

Solution:

For this ANOVA we have $I = 3$ groups and $N = 12$ total observations.

The test statistic for the one-way ANOVA is

$$F = \frac{SSB/df_{SSB}}{SSW/df_{SSW}}$$

where we have $df_{SSB} = I - 1 = 2$ and $df_{SSW} = N - I = 12 - 3 = 9$.

We can pretty much just read off the group means: $\bar{y}_N = 6, \qquad \bar{y}_P = 8, \qquad \bar{y}_L = 10$

And the grand mean is: $\bar{\bar{y}} = \dfrac{1}{12}(4 \cdot 6 + 4 \cdot 8 + 4 \cdot 10) = 8$

We calculate the sums of squares:

$$\begin{aligned} SSB &= \sum_{i=1}^{3} n_i (\bar{y}_i - \bar{\bar{y}})^2 \\ &= 4(6-8)^2 + 4(8-8)^2 + 4(10-8)^2 \\ &= 4 \cdot 4 + 4 \cdot 0 + 4 \cdot 4 \\ &= 32 \end{aligned}$$

14

$$SSW = \sum_{i=1}^{3} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$$

$$= [(4-6)^2 + (5-6)^2 + (7-6)^2 + (8-6)^2] + [(7-8)^2 + (9-8)^2]+$$
$$[(8-10)^2 + (9-10)^2 + (11-10)^2 + (12-10)^2]$$
$$= [4+1+1+4] + [1+1] + [4+1+1+4]$$
$$= 22$$

So the test statistic is:
$$F = \frac{32/2}{22/9} = \frac{16 \cdot 9}{22} = \frac{8 \cdot 9}{11} = \frac{72}{11} \approx \boxed{6.545}$$

(c) (4 points) For a test at the $\alpha = 0.1$ significance level, perform a **rejection region** test for your test statistic from part (b). Be sure to clearly state (i) the distribution you are referencing (including any degrees of freedom), (ii) the critical value to which you are comparing your test statistic, and (iii) the conclusion of your test.

Solution:

The test statistic F follows an F distribution with numerator degrees of freedom $df_{SSB} = 2$ and denominator degrees of freedom $df_{SSW} = 9$. So:

$$F \sim F(2,9)$$

We'll be comparing to a critical value of $F_{crit} = F_{0.1,2,9} = 3.006$. For $\alpha = 0.1$, our test statistic is greater than the critical F ($F = 6.545 > 3.006 = F_{0.1,2,9}$), so we **reject the null hypothesis** and conclude that there is evidence that some pair of groups have a difference in means.

(d) (4 points) For this part of the problem only, suppose you collected more data from the No Reaction and Pity Laugh groups. You collected 32 samples from the No Reaction group, with a sample mean of 6 and sample variance of 20. From the Pity Laugh group, you also collected 32 samples, with a sample mean of 8 and sample variance of 30. Compute a p-value associated with a hypothesis test to determine if there is evidence that the Pity Laugh and No Reaction population means are different. **Do not pick a significance level**, but comment on the strength of this evidence for a difference in group means.

Solution:

Since we have more than 30 data points from each group, we can use the Central Limit Theorem. Our test statistic is:

$$Z = \frac{\bar{y}_N - \bar{y}_P - 0}{\sqrt{s_N^2/n_N + s_P^2/n_P}} = \frac{6-8}{\sqrt{20/32 + 30/32}} = -\frac{2}{\sqrt{50/32}} = -\frac{2}{\sqrt{25/16}} = -\frac{2}{5/4} = -\frac{8}{5} = -1.6$$

The p-value is:
p-value $= 2 \times \Phi(-1.6) = 2 \times 0.055 = \boxed{0.11}$

This p-value says there is about a 1-in-10 probability that we might observe the data that we saw, in a world where the null hypothesis (no difference in means) is true. This evidence is a mixed bag – it might be strong enough to convince some people to reject H_0 and conclude there is some difference in group means, but probably isn't strong enough to convince everyone.