STAT 1
Fall 2020
Midterm Exam 1

Name (Print): _____

Last 4 digits of SID: _____

Instructions:

- The exam is open book and open notes. You may not receive assistance from other people, including through forums or question and answer sites.
- There are 90 possible points. Each question part shows how many points it is worth.
- You are encouraged to use StatCrunch to complete this exam. See the StatCrunch page in Canvas for tutorial videos. If you prefer, you are also welcome to use Excel or another statistical software tool. You may also work by hand.
- **You must show your work on this exam!** If you are working in StatCrunch, this includes taking a screenshot of any output and pasting it into your exam document. For incorrect solutions, partial credit will be awarded for work shown.
- Your exam must be typed. Any graphs or tables should be pasted in from StatCrunch. You may submit a Word document or PDF.
- **Since we are online, your submission of this exam serves as your signed agreement to the following exam honesty policy:**

Honesty Statement and Pledge:

I have not given or received any aid or assistance from other students or online question and answer sites for the full duration of the exam. Everything I have written on this exam represents my own work and knowledge. I sign this knowing that infringements on the University's Academic Integrity policy may result in failure or expulsion.

Signature: _____ **Date:** _____

Part I

1. (4 points) When is is appropriate to use Chebyshev's Rule?

2. (4 points) When is is appropriate to use the Empirical Rule?

3. The whitefly, which causes defoliation of shrubs and trees and a reduction in salable crop yields, has emerged as a pest in Southern California. In a study to determine factors that affect the life cycle of the whitefly, an experiment will be conducted in which whiteflies are placed on two different types of plants at three different temperatures. The researchers are interested in the total number of eggs laid by caged females under each treatment.

   (a) (2 points) Is this an observational study or an experiment? Explain.

   (b) (8 points) Identify the:
   - Experimental units.
   - Response variable.
   - Factor(s).
   - Levels.
   - Treatments.

4. (5 points) Write a few sentences discussing something interesting you've learning in this class.

Part II

For this part of the exam, you will work with data on roller coasters. The StatCrunch link and link to download this data for use with other software are available in Canvas. This data gives information on a sample of 61 roller coasters from a number of different theme parks. The variables are:

- `Park`: the theme park where the roller coaster can be found
- `State / Country`: the state and country where the theme park can be found
- `Duration`: length of time the ride lasts (seconds)
- `Speed`: the roller coaster's top speed (miles per hour)
- `Height`: the roller coaster's top height (feet)
- `Drop`: longest single drop on the roller coaster (feet)
- `Length`: total track length (feet)
- `Yr Opened`: year the roller coaster opened
- `Inversion`: whether or not riders on the roller coaster go upside down

5. Suppose you wanted to collect your own data on how fun it is (or isn't, if you don't enjoy that sort of thing) to ride different roller coasters. For each coaster you ride, you will give it a rating from 1-10.
   (a) (8 points) Describe how you might get a random sample of roller coasters to ride. Defend your approach by considering the pros and cons.

   (b) (3 points) What type of sampling did you recommend?

6. Choose one qualitative variable from the roller coaster data. State clearly which variable you chose.
   (a) (3 points) Obtain a relative frequency distribution of this variable.

   (b) (4 points) Obtain a graph of this variable. Explain why you chose this type of graph.

   (c) (3 points) Find an appropriate measure of center for this variable. Explain why you chose this measure of center.

7. Choose one quantitative variable from the roller coaster data. State clearly which variable you chose.

   (a) (3 points) Create a histogram for this variable.

   (b) (3 points) Find an appropriate measure of center for this variable. Explain why you chose this measure of center.

   (c) (3 points) Find an appropriate measure of variation for this variable. Explain why you chose this measure of variation.

   (d) (4 points) Interpret the measure of center and the measure of variation you found in the previous two problems.

   (e) (5 points) Find the five-number summary for this variable. What is the 50th percentile?

   (f) (4 points) Create a box plot for this variable.

8. Choose two variables for a regression. State clearly which variables you chose. (Hint: this will be easier if you don't choose `Yr Opened` as one of your variables.)

   (a) (4 points) Why did you choose these two variables? (Why might their relationship be interesting?)

   (b) (3 points) Which variable will be the predictor and which will be the response? Why?

   (c) (6 points) Create a scatter plot for these two variables. Is it appropriate to perform a regression? Explain.

   (d) (7 points) Find the regression equation for these variables. Interpret the slope and intercept.

   (e) (4 points) What is the change in your response variable for a 100 unit change in your predictor variable? Be sure to write your answer in the context of the variables you chose.