Math 247: Homework Set 1 (Spring 2020) **Name:_____**

Non-StatCrunch_____/12 points Due date: _____

StatCrunch _____/8 points

Copying homework from the back of the book and giving yourself credit for it is not okay! This is considered academically dishonest and, if you do so, your homework score will be a ZERO.

Requirements: In order to receive credit on homework problems, you must

- Briefly **summarize** the question asked by the problem. *Writing just a few words (e.g. "find mean and standard deviation") is fine; you don't have to write down the entire problem. You can also include the question in the answer by writing a sentence "The mean is _____, and the standard deviation is _____"*
- **Show your work**
  - If the question is conceptual, write the answer and BE SURE to include an explanation in your own answer if the problem asks for one.
  - If the problem involves computation by hand then clearly show your work.
- **Check your answer**s in the back of the book. If your answer is incorrect, then you have to go back to *find and fix the error*.
- If the problem involves computation by **StatCrunch then copy your Statcrunch work in a Google Docs document and print it out**. Write the answers to the follow-up questions for that problem by typing or by writing them later by hand on the printout.

Self-Assessment: Determine total number of correctly done problems (they don't have to be done correctly the first time…just make sure you find and fix any errors in your work!) and put that score in the last column.

If you give credit for a problem that hasn't been completed, you will lose 1 point on the overall grade (for example, if you put down 9/9 in Assignment 1 but have completed only 6 problems, your entire score on the homework packet will lose 3 points). **Please be honest and accurate in your assessment.** ***Homework is due on the day of the exam***.

Odd Answers are in the back of the textbook. Even Answers are on the last page of this assignment sheet.

	Read this section:	**Do these problems:**	**Completed/ Total (you fill in)**
1	**Section 1.1:** What Are Data? (1.1 is a "Read only" section) **Section 1.2:** Classifying and Storing Data	**1.2 Exercises,** page 27: 1, 2, 3, 4, 5, 6, 7, 9, 11	☐ /9
2	**Section 1.4:** Collecting Data to Understand Causality	**1.4 Exercises,** page 30: 31, 32, 33, 34, 35, 36, 39, 41, 43, 44, 49	☐ /11
3	**Section 2.1:** Visualizing Variation in Numerical Data **Section 2.2:** Summarizing Important Features of a Numerical Distribution	**2.1, 2.2 Exercises,** page 63: 1, 3, 5, 7, 8, 11, 13, 15, 17, 25 (for #25 make a dotplot of each data set by hand) **StatCrunch**: 29, 33 (see instructions for 29, 33 below) #29 (use StatCrunch to make a dotplot and a histogram) #33 (use StatCrunch)	☐ /12
4	**Section 2.3:** Visualizing Variation in Categorical Variables **Section 2.4**: Summarizing Categorical Distributions **Section 2.5:** Interpreting Graphs	**Note: Please read these sections on your own and do the homework problems. This material will not be covered in lecture**. **2.3, 2.4, 2.5 Exercises**, page 70: 39, 41, 47, 48, 49	☐ /5

5	**Section 3.1:** Summaries for Symmetric Distributions	**3.1 Exercises**, page 120: 1, 2, 3, 4, 5 (do #5 by hand!), 11, 19 (do #19 by hand!) **StatCrunch exercises:** 7, 13, 25 (on #25 you will have to unstack the data)	☐ /10
6	**Section 3.3:** Summaries for Skewed Distributions	**3.3 Exercises,** page 124: 39, 41, 43,(do #41, 43 by hand!) **StatCrunch exercises:** 45 (data is from #3.24)	☐ /4
7	**Section 3.4:** Comparing Measures of Center	**3.4 Exercises**, page 126: 47, 48, 49, 50, 51, 53	☐ /6
8	**Section 3.5:** Using Boxplots for Displaying Summaries	**3.5 Exercises**, page 127: 55, 57, 59, 61, 64, 63 (do #64 first) **IMPORTANT:** Additional instructions for #63 and #64: Find the upper and lower outlier limits, then identify the outliers in the data set, then sketch the box plot, using an asterisk (*) for the outlier.	☐ /6

Even Answers, Chapter 1:
1.2: 11 people
1.4: (a) Shoe size: Numerical (b) Eye color: Categorical
1.6: Answers will vary (e.g., Relationship status)
1.32: Observational study
1.34: Controlled experiment
1.36: Observational study
1.44: (a) Confounding factor: Answers will vary but for example, people who choose a vegetarian diet might be more willing to adopt other healthy habits, such as exercise or eating more fruits and vegetables, which would lead to greater weight loss.
(b) Don't allow people to choose their diet. Randomly assign people to a vegetarian group or a non-vegetarian group.

Even answers, Chapter 2:
1.48: Because the data is categorical, not numerical, a bar chart would have been better. The use of a histogram is inappropriate here since histograms should be used only with numerical data.

Even answers, Chapter 3
3.4: Some value between 3 and 4 (say 3.5?) televisions per home is typical. Since this is a symmetric distribution we are just identifying the center of the data.
3.48: Symmetric: Use mean and standard deviation; Skewed: Use median and IQR
3.50: Owners should use the mean salary since it's much higher than the median. If you were a player, you should use the median salary since it's lower and more representative of what a typical player actually makes. The data is clearly strongly skewed to the right which makes sense since the "stars" will have super high salaries but the bulk of the players will make much less.
3.64: IQR = $\text{IQR} = Q_3 - Q_1$ = 90 – 80 = 10

Left Outlier Limit = $Q_1 - 1.5 \cdot \text{IQR}$ = 80 – 1.5(10) = 65

Since 26 is less than 65, we know this value is an outlier.

Right Outlier Limit = $Q_3 + 1.5 \cdot \text{IQR}$ = 90 + 1.5(10) = 105

Since no percentage is higher than 105, there are no potential outliers.

(b) The relationship is most likely due to per capita wealth in each country. This is a **confounder** between number of TV's and lifespan, and is a more meaningful predictor of lifespan.
(c) Clearly buying TV's will not extend anyone's life. Correlation is not causation.