4.2_Scores

Kyle Ramirez

12/26/2021

{r setup, include=FALSE} knitr::opts_chunk$set(echo = TRUE)

Assignment: ASSIGNMENT 4.2 SCORES

Name: Ramirez, Kyle

Date: 2021-12-26

What are the observational units in this study?

Answer: Counts of students, The total score, Regular or sports section

Identify the variables mentioned in the narrative paragraph and determine which are categorical and quantitative?

Answer: Section is categorical. Count and scores are quantitative.

Create one variable to hold a subset of your data set that contains only the Regular Section and one variable for the Sports Section.

Set the working directory to the root of your DSC 520 directory

setwd("/Users/Kyle/Documents/GitHub/KR/Ramirez_Kyle_DSC510/dsc520") scores_df <- read.csv("data/r4ds/scores.cs sports_score <- subset(scores_df, Section == "Sports") regular_score <- subset(scores_df, Section == "Regular") View(sports_score) View(regular_score)

Use the Plot function to plot each Sections scores and the number of students achieving that score. Use additional Plot Arguments to label the graph and give each axis an appropriate label. Once you have produced your Plots answer the following questions:

library(ggplot2) ggplot(sports_score, aes(x=Count, y=Score)) + geom_point() + ggtitle('Counts vs. Scores') + xlab('Counts') + ylab('Scores')

Comparing and contrasting the point distributions between the two section, looking at both tendency and consistency: Can you say that one section tended to score more points than the other? Justify and explain your answer.

Answer: The regular classes appeared to score higher than the sports section.

Did every student in one section score more points than every student in the other section? If not, explain what a statistical tendency means in this context.

Answer: It would seem like for mostly the sports section. Students scored higher in smaller numbers than they did in sections with higher counts. In the regular section, students scored better in the 20 counts amounts than in any other.

What could be one additional variable that was not mentioned in the narrative that could be influencing the point distributions between the two sections?

Answer: It seemed that the students that got a curriculum with a variety of applications instead of just the sports application to learning. They performed much better than the ones with just sports. Maybe there is a limitation as to what can be applied by only using sports analogies to learning?