Dalton Murray

Data Science for Business

INT 7623 - 3404-202420\_INT7623\_OL

Dr. Kamel Rushaidat

April 15, 2024

**Assignment #4 – Regression & K-Means**

**Part 1: Regression**

1. Reposition the points so that the correlation coefficient (r) to 1. What does it mean to have r =1?

2. Reposition the points so that the correlation coefficient (r) to -1. What does it mean to have r =-1?

3. Reposition the points so that the correlation coefficient (r) to 0 or very close to zero. What does it mean to have r =0?

Include screenshots for every part and make a comparison between the three different scenarios in terms of the correlation between the two variables. Discuss your results

**Part 2: K-Means Clustering**

Choose the first strategy to initial the centroids by “choosing them by yourself”. Include screen shots for the steps. How many iterations the algorithm did till it finds the best clusters? (5 points)

2. Choose the second strategy to randomly choose the centroids. How many iterations the algorithm did till it finds the best clusters? (5 points)

3. Choose the third strategy by using the Farthest point as the centroids. How many iterations the algorithm did till it finds the best clusters? (5 points)

Discuss your conclusion about using the three different strategies. Add any interesting facts/notes that you found when tried this visualization. (6 points)

I have neither given nor received unauthorized aid in completing this work, nor have I presented someone else's work as my own.

*Dalton Murray*