

Worksheet 1 - Pearson

Team: _____

Up to 35XP

Please hand in one finished copy of your work.
All data is available on our website

$$r = \frac{\frac{\sum XY}{N} - \left(\frac{\sum X}{N}\right)\left(\frac{\sum Y}{N}\right)}{\sqrt{\left(\frac{\sum X^2}{N} - \left(\frac{\sum X}{N}\right)^2\right) \left(\frac{\sum Y^2}{N} - \left(\frac{\sum Y}{N}\right)^2\right)}}$$

Save your work as I may ask you to demo it to the class.

1. Medicine Hat Tigers

Using Pearson's Correlation Coefficient calculate the coefficient between the weight and height of the Medicine Hat Tigers (the data is on the spreadsheet). You may need to convert the height to something more reasonable.

2. A School Music Program: A case study

We are interested in seeing whether a relationship exists between music performance grades and students' individual Iowa Test of Basic Skills (ITBS) scores. Using the provided test set, calculate the strength of the association between music and language, music and science, and music and math. Are any of these associations strong?

3. Arabic documents

We have 3 Arabic documents. One document is from the newspaper Akhbar, another from Almshaheer, and another we are not sure of. We have represented each document as a list of frequencies of the most frequent Arabic words. Using Pearson's Correlation Coefficient, how would you classify document A (Akhbar or Almshaheer)?

3. Who has similar movie tastes?

Using the movie database we constructed in class find out who (within your team) has the most similar movie tastes to one person in your team. (i.e., pick a person and find who in your team is most similar to that person).