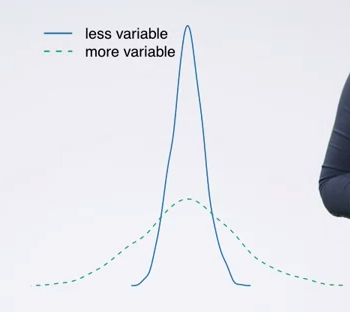
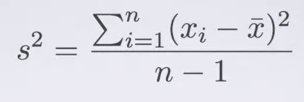
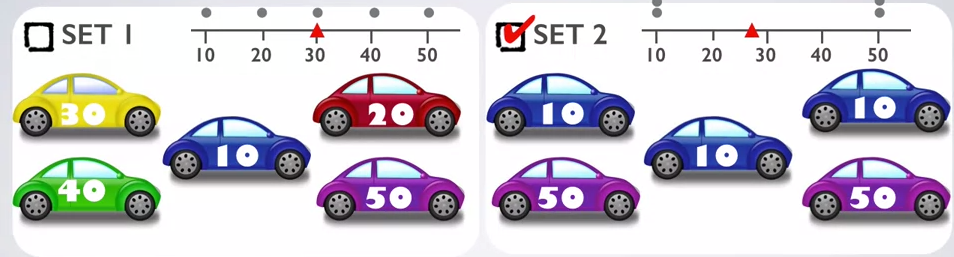
Coursera Stats Video Lectures – Measures of Spread

Week 1, Video 8

measures of spread, aka variability in the data



1. measure of spread: range
   1. difference between the min and the max values in the data
2. measure of spread: variance
   1. the average squared variation from the mean
   2. sample variance denoted by *s2*
   3. population variance denoted by *σ2*
   4. calculation 
   5. variance units are the square of the variable units. e.g., if we are measuring height, the variance would be expressed as *x* cm2
   6. Why square the variance
      1. gets rid of negatives
      2. squaring increases larger deviations more than smaller ones so they have a heavier weight
3. measure of spread: standard deviation
   1. roughly the average deviation around the mean
   2. same units as the data provided
   3. computed as the square root of the variance
   4. denoted as *s* for the sample and *σ* for the population
4. variability vs diversity
   1. diversity doesn’t inform variability.
   2. higher the variability, the less likely it is that data points cluster around the mean



1. measure of spread: inter-quartile range
   1. range of the middle 50% of the data,
   2. distance between the 1st quartile (25%) and the third quartile (75%)
   3. easiest to see in a box plot
   4. more reliable than the range because it doesn’t rely on the endpoints, which themselves could be outlyers