

Unit 1 Assignment

1. An international conference on education is being held in Italy. At the beginning of the two-day conference, delegates are asked to fill out a survey, from which a sample of questions is shown below. For each of the questions, classify the variable of interest (shown in **bold**) as either quantitative, categorical and nominal, or categorical and ordinal.
- (a) What **country** are you from?
 - (b) What **time zone** do you live in?
 - (c) What **distance** did you travel to get to this conference?
 - (d) What is the highest **level of education** you have completed (high school, Bachelor's degree, Master's degree, Phd)
 - (e) Have you **visited Italy** before?
 - (f) How many **days** do you plan to stay in Italy?
 - (g) How would you rate the **quality of education** in your country's school systems?
(5 = Excellent, 4 = Very Good, 3 = Good, 2 = Fair, 1 = Poor)

2. The batting averages of 34 Major League Baseball players are ordered and shown below:

0.178	0.202	0.210	0.219	0.222	0.237	0.245	0.250	0.256
0.258	0.258	0.260	0.261	0.263	0.267	0.268	0.271	0.275
0.277	0.279	0.279	0.281	0.283	0.284	0.286	0.288	0.294
0.299	0.305	0.311	0.313	0.320	0.333	0.340		

- (a) Find the five-number summary for these data.
- (b) Construct a quantile boxplot. What appears to be the shape of the distribution of batting averages?
- (c) Construct an outlier boxplot. What is the actual shape of the distribution if we disregard outliers?

3. The distances travelled (in km) by a taxi in its first seven trips of the day are shown below:

9.7 3.1 10.5 5.6 14.9 6.7 8.3

- Calculate the mean and standard deviation of the distances of these seven trips.
 - If a variable Y can be written as a linear function of a variable X , i.e. if $y = a + bx$, then show that $\bar{y} = a + b\bar{x}$ and $s_y = |b|s_x$.
 - The taxi company charges a base fare of \$3.30 just to get into the taxi, plus \$1.25 for each kilometer travelled. Use the results from (b) to find the mean and standard deviation of the fares of the seven taxi trips.
 - After the eighth trip of the day, the mean distance travelled remains unchanged. What is the distance of the eighth trip?
 - When we add the eighth trip to the data set, what happens to the value of the standard deviation? (You don't need to do any calculations, just say whether the standard deviation increases, decreases or stays the same, and why).
4. The GPAs for a class of twelve females and fifteen males are recorded. The mean GPA for the females is 3.22 and the standard deviation is 0.62. The mean GPA for the males is also 3.22 and the standard deviation is 0.78. What is the standard deviation of GPAs for all students in the class?
5. A student took five university classes last semester. We have the following facts:
- The student received 87% in biology.
 - The student received 74% in history.
 - The student received the same score in geography and chemistry.
 - The student's lowest score was in calculus.
 - The student's average score was 78%.
 - The variance of the student's scores was $47.2(\%)^2$.

Find the student's score in each of his five classes last semester.

6. The number of points scored by a basketball team in their 24 games this season are shown below:

Game:	1	2	3	4	5	6	7	8	9	10	11	12
Points:	82	74	91	85	88	100	90	95	87	92	92	111

Game:	13	14	15	16	17	18	19	20	21	22	23	24
Points:	105	98	103	116	104	122	108	99	105	120	110	109

- Make a timeplot of the data.
- Construct a histogram for this dataset.
- Construct a stemplot for this dataset.
- Rank the three graphs you made in order of how much information you can get from them.