CS381 Data Analytic Quiz 2

Instruction: For multiple choice questions, clearly circle one of the choice; for all other questions, write your answer right below the questions. All questions carry the same weights.

Name:

Question 1: Which of the following statements is/are true about Type-1 and Type-2 errors?

```
1. Type-1 error occurs when we rejects a null hypothesis when it is actually true
2. Type-2 error occurs when the prediction is positive while the actual case is
   negative
3. Type-2 error means the prediction is wrong from the actual case
```

    A. Only 1
    B. Only 2
    C. 1 and 2
    D. 1 and 3
    E. 2 and 3

Answer: 1 is saying Type 1 error is FP which is true, 2 is saying Type-2 error is FP (therefore it is wrong), 3 is just wrong so, the answer is A

Question 2: Fill in the blank in the following statements with one of the words among: "higher", "lower", "bigger", "smaller" or "similar"

    A. The household income distribution of a gated community should have a **smaller**_ standard deviation than a random sample of the whole population

    B. The Graduation Rate distribution among different colleges should have a **higher**_ kurtosis than the average height distrbution among different colleges.

explanation: The variation of the average height for different colleges are limited by physical laws. Chances to have extreme values are lower. There are no physical laws to dictate the variation of graduation rate of a college. It can varies more dramatically among different type of college and can change over time. So the chances to have extreme values (outliers) for graduation rate are higher.

Question 3: Given the following data sets

X	Y
2	5
3	6

X	Y
8	10
-1	0
1	3

    A. Calculate the standard deviation of X and Y. Show your work.

    B. Calculate the covariance between X and Y. Show your work.

Answer: 1. 3.01, 3.31 2. 9.72