# Tutorial 6

1. Consider dataset `midterm_marks` (folder `Data` on Luminus) which contains midterm scores of students taking a statistical module, called this variable `mark`.

   (a) Derive mean and standard deviation of variable `mark` from the given sample.

   (b) Manually perform a test to test if the mean of midterm scores is 20 or less than 20. Report the test statistic and p-value of the test.

   (c) Manually form a 95% confidence interval for the mean of midterm scores.

   (d) Use a built-in function to perform the test and to derive the CI mentioned above.

   (e) Checking the normality assumption made for the test above. Do you think the result of the test performed above is reliable if the distribution of `mark` is not approximately normal? Explain.

   (f) Repeat all questions above in Python and SAS.

2. Dataset `glaucoma_dep.csv` consists of measurements of corneal thickness of 8 subjects affected with glaucoma in one eye. The difference of the thickness between the affected and non-affected eyes are obtained. Answer the question below using R, Python and SAS.

   (a) Perform a test on the difference to decide if glaucoma decreases the thickness of the corneal at 0.05 significance level.

   (b) Suppose that we ignore the fact that the data are dependent. What would the conclusion be, at the same significance level?

3. During election campaigns, there have been television debates between two candidates. A researcher was interested in determining whether or not a particular debate between two candidates was effective in changing viewer's preferences for the candidates. The researcher randomly selected a sample of voters before the debate and asked them to indicate their preferences for the two candidates. The same voters were asked for their preferences for the two candidates after the television debate. The results of the survey are give as follows.

| | | Preference after TV debate | |
|---|---|---|---|
| Voter's Gender | Preference before debate | Candidate A | Candidate B |
| Male | Candidate A | 67 | 28 |
| | Candidate B | 46 | 54 |
| Female | Candidate A | 58 | 42 |
| | Candidate B | 37 | 61 |

   (a) For each gender group, use SAS to test if the debate between two candidates was effective in changing viewers' preferences for the candidates.

   (b) Repeat question above using R and Python.