Quantitative Tools - Level II **October 18, 2017**
Time Limit: 60 Minutes

First midterm

Instructions

- The use of a mobile phone, or any other means of communication, is forbidden.
- Give all the results with **a two decimals** precision.
- Write answers on a **separate** sheet of paper. **Underline** your results.

Section are independent and can be treated separately.

1 Exercice 1 (8 points)

Seven employees of Leafy Lakes, an amusement park in British Columbia, Canada, are comparing their wages and years of study at the coffee break. The following table reproduce their wages and years of study.

	Wages	Years of study
Alice	42000	8
Bob	17000	4
Craig	10000	1
Dave	22000	5
Eve	20000	5
Frank	36000	7
Wendy	70000	11

1. Find the median, the arithmetic mean, and standard deviation of this set of wages.
2. We are interested in the link between years of study and wages. Looking at the data and coefficients of correlation, what seems to be the appropriate model (linear, semi-log, log-log) ?
3. Estimate your best model and report its coefficients.
4. What is the impact of wages of one more year of study ?
5. If you were to improve the model, which variable(s) will you add ?
6. Oscar, 9 years of study, and Wallet, 3 years of study, are joining the conversation at the coffee break. What is the forecast of their wages using the model ?

In order to prepare the annual meeting, an union want to assess what is the level of inequality among Leafy Lakes employees. The following table presents the wages distribution

Wages (in k$)	Frequencies
10 - 40	30
40 - 70	54
70 - 100	16

7. Draw the Lorenz Curve.
8. What is the Gini coefficient ?

2 Exercice 2 (6 points)

Sales per quarter of a clothing company in 2014, 2015, 2016 are reproduced below.

	2014	2015	2016
Q1	105	106	107
Q2	101	102	102
Q3	105	106	107
Q4	102	102	103

1. Plot the timeseries (on Excel). Is the timeseries stationnary ? Is it seasonal ?
2. Estimate a trend.
3. Find the seasonal coefficients.
4. Forecast the sales in 2017.

3 Exercice 3 (6 points)

1. How many ways are there to split a dozen people into 2 teams, where each team has 6 people?
2. Spam[1] filter is designed by looking at commonly occurring phrases in spam. Suppose that 15% of email is not spam. In 10% of the spam emails, the phrase "fast cash" is used, whereas this phrase is only used in 2% of non-spam emails. A new email has just arrived, which does mention "fast cash". What is the probability that it is spam?
3. Two baskets contains balls of different colours. The first basket B_1 contains ten red balls and five white balls. B_2 contains four red balls and six white balls. We pick one ball in B_1 and put in B_2. We pick a ball in B_2
   - What is the probability that the ball picked in B_2 in white ?

[1]Hint: For questions 2 & 3, use well-defined events and recall that for any A,B : $P(A) = P(A|B)*P(B)+P(A|\neg B)*P(\neg B)$

- If the ball in B_2 is white, what is the probability that the one previously picked in B_1 was white ?

***