

MATHS. Biochemistry degree

SHEET 1 : DESCRIPTIVE STATISTICS

1. We measure the pH in 18 samples of a certain aqueous solution, obtaining the data

7'4 6'7 7'0 7'2 6'7 6'1 7'5 6'3 7'0 6'5 7'8 6'0 7'5 6'9 7'3 8'1 7'2 5'1

- (a) Represent the data with a stem-plot
- (b) Represent the data with a box-plot
- (c) Compute the median and quartiles, and determine possible outliers.

2. The scores obtained by two student groups, A and B, in a Math exam have been:

Group A: 9'1 6'8 10 7'5 9 10 8'4 6 9'2 9'4 9'4 9'6 8'6 7
 Group B: 9'1 9'2 8'6 7'7 7'3 6 8'6 7'3 7'1 7 6 7'4 8'5 7'5

- (a) Represent the data in a *joint* stemplot (draw the stem in the middle, and the leaves on the left or right hand sides, depending in the group).
- (b) Represent jointly the bloxplot of each group.
- (c) From the previous representations, explain which group did better in the exam.

3. We investigate the effect of a certain enzyme *Z* in the speed of a biochemical reaction. After performing 20 experiments we obtain the following data (in mMol/min):

Speed without enzyme Z

380 321 366 356 283 349 402 462 356 410329 399 350 384 316 272 345 455 360 431

Speed with enzyme Z

361 447 401 375 434 403 393 426 406 318 467 407 427 420 477 392 430 339 410 326

- (a) Draw a joint box-plot of these data, and explain what you see.
- (b) Compute the median and quartiles for each group, and explain the differences you observe.

4. According to UE regulations, car manufacturers must provide the fuel consumption for the vehicles they intend to sell. The following data give the fuel consumptions (in liters/100 Km) in a highway test of 30 automobile models:

8,3 5,1 6,1 6,4 5,7 7,3 7,3 6,8 5,7 6,1 6,8 6,5 7,3 5,7 6,8
 6,5 8,8 8,3 6,8 7,3 5,1 7,8 7,8 10,1 6,8 14,8 6,8 6,4 6,1 6,8

- (a) Use the statical graphics you have learned to represent these data. Explain the main properties you observe (shape, outliers, etc).
- (b) What is the median consumption? What would be the largest admissible consumption to be in the 25% lower group. Find the proportion of data above 7'5 l/100 km?
- (c) Find the mean, and explain how it compares to the median.

5. The following data represent the cranial perimeters (in cm) of 35 newborns

33 | 04
 33 | 67789
 34 | 01222233
 34 | 556667789
 35 | 11223
 35 | 68
 36 | 01
 36 | 6

- (a) Draw a histogram with 6 intervals for these data
- (b) Compute the mean and standard deviation of these data.

6. The glucose concentration in blood is measured in 70 diabetic patients. The grouped data are given in the following table

concentr (mg/dl)	47'5–52'5	52'5–57'5	57'5–62'5	62'5–67'5	67'5–72'5	72'5–77'5
abs freq	1	8	24	21	15	1

- (a) Draw a histogram for these data
 (b) Find \bar{x} and σ
 (c) Find the proportion of data in the intervals $\bar{x} \pm \sigma$ and $\bar{x} \pm 2\sigma$.
 (d) Find the median and quartiles, and draw a boxplot.
7. The following table gives the number of newborns in Spain during 2011, grouped according to the age of their mothers.

mother age	15–19	20–24	25–29	30–34	35–39	40–44	45–49
number of newborns	10.589	38.615	92.158	178.517	123.909	24.526	1.580

- (a) Draw a histogram for these data.
 (b) Find the mean \bar{x} and the standard deviation SD.
 (c) Find the proportion of individuals in the intervals $\bar{x} \pm DT$ and $\bar{x} \pm 2DT$.
 (d) In what percentile would be a 27 year old mother? How old are mothers in the percentile 90?
8. In a new vaccine experiment, a lab studies the time for the first local reactions. After testing in a large sample, the obtained data are presented in the following form :

Percentile	10	30	50	70	100
Time (in hours)	12	18	22	24	30

- (a) Draw histogram of densities for these data (you should first make a correct frequency table for these data).
 (b) Find the mean time and standard deviation for a local reaction to the vaccine.
 (c) Find the *median* time for local reaction. Find also a time interval for the middle 50% of individuals.