A) A parking lot has an average of 2,362 free spaces per hour. If we assume that the number of free spaces during this interval follows a Poisson distribution, what is the probability that a driver will find a parking space in an interval of one hour?

Let X be the number of free spaces per hour in the parking lot. We try to find the probability that there is at least one free place, that is to say $P(X \geq 1)$.

$$P(X \geq 1) = 1 - P(X < 1) = 1 - P(X = 0) = 1 - \frac{e^{-2.362} 2.362^0}{0!} = 0.906$$

B) the speed of 457 vehicles on one lane is measured over a period of 5 hours and the following results are obtained :

Speed interval (km / h)	%
0-20	0
20-40	6
40-60	10
60-80	33
80-100	39
100-120	12

a) Calculate the average speed of vehicles

The first step is to calculate the number of staff for each speed category, that is to say the number of vehicles associated with the given percentages. Then take an average value for each interval (10, 30, 50, 70, and 110) :

Speed interval	middle	%	Numbre
0-20	10	0	0
20-40	30	6	27
40-60	50	10	46
60-80	70	33	151
80-100	90	39	178
100-120	110	12	55

The average is calculated with :

$$\overline{x} = \frac{10 * 0 + 30 * 27 + 50 * 46 + 70 * 151 + 90 * 178 + 110 * 55}{0 + 27 + 46 + 151 + 178 + 55} = ¿$$

78.23 km/h

b) If the distribution followed a Normal law, what should be the expected number of vehicles associated with each of the classes if the standard deviation is 20 km / h?

Let X be the speed of a vehicle. We want to calculate $P(0 \le X \le 20)$, $P(20 \le X \le 40)$, *etc. if X follows the normal law (centered and reduced) :*

$$P(0 \le X \le 20) = P\left(\frac{0 - 78.23}{20} \le Z \le \frac{20 - 78.23}{20}\right) = \Phi(-2.91) - \Phi(-3.91) = 0.00176$$

Calculation of the numbers if the law is normal

Speed interval	Probability	Numbe
0-20(-∞-20)	0.0018	0.82
20-40	0.0262	11.96
40-60	0.1530	69.94
60-80	0.3542	161.89
80-100	0.3265	149.23
100-120 (100-∞)	0.1381	63.15

c) Check if the distribution follows a normal law (null hypothesis, the speed distribution follows the normal law)

Perform a chi-square goodness-of-fit test, after combining the first two categories.

$$X^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} = 31.34$$

Compare with $\chi^2_{5-2-1}(0.05) = 5.99$

We reject the hypothesis that the distribution follows a Normal law because

$$X^2 > \chi^2_2(0.05)$$

d) Calculate a confidence interval on the mean with a threshold of 5%

At a 95% confidence level, the average speed at this point is in the range :
$$\left[78.23 - \frac{1.96 * 20}{\sqrt{457}}, 78.23 + \frac{1.96 * 20}{\sqrt{457}}\right]$$

f) Is the sample taken large enough to obtain the speed at ± 2 km / h with a risk α = 5% and a standard deviation of 20 km / h?

$n = \frac{k^2 \sigma^2}{e^2} = \frac{1.96^2 \, 20^2}{2^2} = 384.16$ □ 385 □ *The sample of 457 samples is quite large.*

C) We decide to install stop signs in a residential area. The speeds of vehicles passing on a street, bounded by two of these signs, are recorded.

Speed Interval	Without	With
0-10	0	1
10-20	0	1
20-30	0	5
30-40	2	20
40-50	3	13
50-60	3	24
60-70	39	27
70-80	41	8
80-90	12	1

Check if the average speed has been significantly decreased at a 95 % confidence level.

The hypotheses of the test to be carried out are :

H_0: the means of the average speeds before (sample 1) and after (sample 2) are equal ($\mu_1 = \mu_2$)

H_1: the average is lower than the average $\mu_2 < \mu_1$

The mean and standard deviation of the samples must be calculated for the two distributions ::

$$\overline{X}_1 = \frac{0*5+0*15+0*25+2*35+3*45+3*55+39*65+41*75+12*85}{100} = 70.0$$

$$\overline{X}_2 = \frac{1*5+1*15+5*25+20*35+13*45+24*55+27*65+8*75+1*85}{100} = 51.9$$

$$\sigma_1 = \sqrt{\frac{0*5^2+0*15^2+0*25^2+2*35^2+3*45^2+3*55^2+39*65^2+41*75^2+12*85^2}{100} - \overline{X}_1^2} = 9.85$$

$$\sigma_2 = \sqrt{\frac{1*5^2+1*15^2+5*25^2+20*35^2+13*45^2+24*55^2+27*65^2+8*75^2+1*85^2}{100} - \overline{X}_1^2} = 15.41$$

For the test to be performed, the decision variable is calculated with

$$Z=\frac{\overline{X}_1-\overline{X}_2}{\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{n}}}=\frac{70.0-51.9}{\sqrt{\frac{9.85^2}{100}+\frac{15.41^2}{100}}}=9.90$$

The statistic Z *should be compared to* $z_\alpha=z_{0.05}=1.645$. *We will reject* H_0 *if*

$Z>z_\alpha=1.645$

Since $9.90>1.645$, *the average speed of sample 2 is lower than that of sample 1.*

D) A cordon survey carried out among 700 vehicles revealed that the average number of occupants per vehicle passing through a certain study point is 1.488. Determine the 95 % confidence interval on the mean assuming the standard deviation is 0.500 passengers.

The tolerance of the confidence interval is calculated with

$$1.488\pm1.96\cdot\frac{0.500}{\sqrt{700}}$$

The mean is therefore in the interval $[1.451, 1.525]$ *with a confidence level of 95%.*

E) The number of bikes available at a BIXI station in the city of Montreal is counted every hour of the day.

Hour	Nb of available bikes
0:00	33
1:00	33
2:00	33
3:00	33
4:00	32
5:00	32
6:00	32
7:00	26
8:00	25
9:00	25
10:00	26
11:00	24
12:00	23
13:00	23
14:00	25
15:00	28

16:00	31
17:00	33
18:00	33
19:00	32
20:00	30
21:00	31
22:00	33
23:00	33

a) Calculate the average number of bikes available per hour in a 24 hour period.

$\bar{x} = ¿ 29.54$ *bikes*

b) What is the median of this distribution?

Since there are 24 data, the median is between 12 e and 13 e values, once they are sorted in ascending order.

The 12 th value is 31 and the 13 th , 32. The median is thus 31.5 bicycles.

c) Calculate the standard deviation of the distribution.

The standard deviation is calculated with the formula

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n} - \left(\frac{\sum_{i=1}^{n} x_i}{n}\right)^2}$$

The standard deviation is 3.71 bikes.

d) If the terminal contains 33 spaces and the number of free spaces follows a Poisson distribution, what is the probability that a member will be able to return his bike to this terminal during the day?

Let X be the number of bikes available at the terminal. Since the terminal has 33 spaces, the average number of vacant spaces Y (Y = 33-X) during the day is 3.46.

The probability that there is at least one free space is calculated as follows :

$$P(Y \geq 1) = \sum_{i=1}^{i=33} P(Y=i) = 1 - P(Y=0) = 1 - \frac{e^{-3.46} 3.46^0}{0!} = 0.97$$

F) Residents of a residential area file a complaint against the speeding tickets made by vehicles circulating there. They ask to install donkey rides or stop signs to force vehicles to slow down in their neighborhood. A pneumatic tube is installed a priori in order to collect the speeds on a street before any intervention. Temporary retarders are then installed and the speed of the vehicles is measured again. Consider a confidence threshold at α = 5 % for the statistical tests.

	Speed before	Speed after
MEAN	47.259	36.097
STD DEVIATION	9.311	7.093
NUMBER of OBSERVATIONS	52	51

a) For which sample size would the error made on the AVANT mean be less than 2 km / h?

$$n = \frac{1.96^2 9.311^2}{2^2} = 83.26$$, *or n = 84 observations*

b) For which sample size would the error made on the average AFTER be less than 2 km / h?

We get n=49 observations

c) Are both samples large enough?

The BEFORE sample is not large enough
The AFTER sample is large enough

d) Would it be effective to install permanent speed bumps there?

This is to check whether the speed bumps have had an impact on the traffic at that location. It is therefore necessary to carry out a test of hypotheses on the means.

The hypotheses of the test to be carried out are :
H_0 : mean before $\overline{X}_1$ and after $\overline{X}_2$ are equal
H_1 : mean«after» $\overline{X}_2$ is smaller than average before $\overline{X}_1$

Calculation of the decision variable:

$$Z = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}}} = \frac{47.259 - 36.097}{\sqrt{\frac{9.311^2}{52} + \frac{7.093^2}{51}}} = 6.625$$

We reject H_0 if $Z > z_\alpha$

In this case, $z_\alpha = 1.645$ is less than 6.625. We reject the hypothesis H_0; the "after" average is significantly lower than the "before" average. The probability that a variable according to the normal distribution is less than 6,625 exceeds 95%. According to this experiment, installing permanent donkeys at this location is an effective intervention to reduce vehicle speed.

G) Electromagnetic loops installed in the roadway measure the speed of vehicles at a point. The following table shows the distribution of the speeds of vehicles that have passed through it in a week.

Speed intervals (km/h)	**Number of vehicles**
0-15	13
15-30	97
30-45	428
45-60	441
60-75	343
75-90	127
90-105	62
105+	1

a) Calculate the mean, median and standard deviation of the speed at this point.

AVERAGE : Average speed is calculated using an average value for each interval (7.5, 22.5, 37.5, 52.5, 67.5, 82.5, 97.5, 112.5).

$$\overline{X}=\frac{13*7.5+97*22.5+428*37.5+\ldots}{1512}=53.75\,km/h$$

MEDIAN : Since there are 1512 rates collected in total, the median of the distribution is between the 756 th and 757 th data.

Since these two elements are in the 45-60 km / h range, the median is between 45 and 60 km / h.

STANDARD DEVIATION : The standard deviation is calculated with the formula :

$$\sigma=\sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}-\left(\frac{\sum_{i=1}^{n} x_i}{n}\right)^2}$$ *= 18.69 km/h.*

b) Are there reasons to consider that this distribution obeys a normal law?

It is a question of finding the theoretical frequencies for each of the intervals (center and reduce to use the tables of the normal distribution).

Intervals	*Number of vehicles*	*Theoretical Frequency*	*Expected (Normal Law)*
0-15	*13*	*0.017*	*25.79*
15-30	*97*	*0.083*	*125.25*
30-45	*428*	*0.218*	*329.50*
45-60	*441*	*0.311*	*470.43*
60-75	*343*	*0.241*	*364.79*
75-90	*127*	*0.102*	*153.55*
90-105	*62*	*0.023*	*35.03*
105+	*1*	*0.003*	*4.62*

We then perform a chi-square goodness-of-fit test :

$$\text{The decision variable is}$$
$$X^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} = 73.50$$

For the law to be accepted, it is necessary that $X^2 < \chi^2_{8-p-1}$ *with p = 2 for the normal law*

We find $\chi^2_5(0.05) = 11.07$

We reject the hypothesis that the velocity distribution follows a Normal law at a confidence level of at least 95 %.