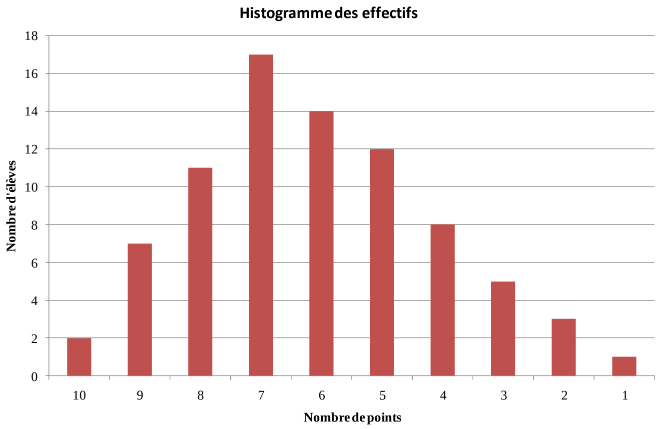
Les statistiques

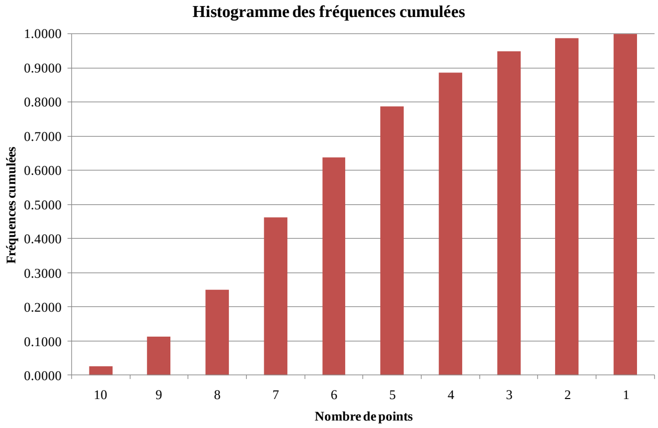
Exercice 1

1. Tableau statistique

Résultats obtenus dans une épreuve commune de mathématique						
Points	Effectifs	Fréquence	Fréquence cumulée	Fréquence cumulée inv	Fréquence pondérée	Fréquence pond. carrée
x_i	n_i	f_i	F_i	F'_i	f_ix_i	$f_ix_i^2$
10	2	0.0250	0.0250	1.0000	0.2500	2.5000
9	7	0.0875	0.1125	0.9750	0.7875	7.0875
8	11	0.1375	0.2500	0.8875	1.1000	8.8000
7	17	0.2125	0.4625	0.7500	1.4875	10.4125
6	14	0.1750	0.6375	0.5375	1.0500	6.3000
5	12	0.1500	0.7875	0.3625	0.7500	3.7500
4	8	0.1000	0.8875	0.2125	0.4000	1.6000
3	5	0.0625	0.9500	0.1125	0.1875	0.5625
2	3	0.0375	0.9875	0.0500	0.0750	0.1500
1	1	0.0125	1.0000	0.0125	0.0125	0.0125
	80	1			6.100	41.175

2. Histogrammes





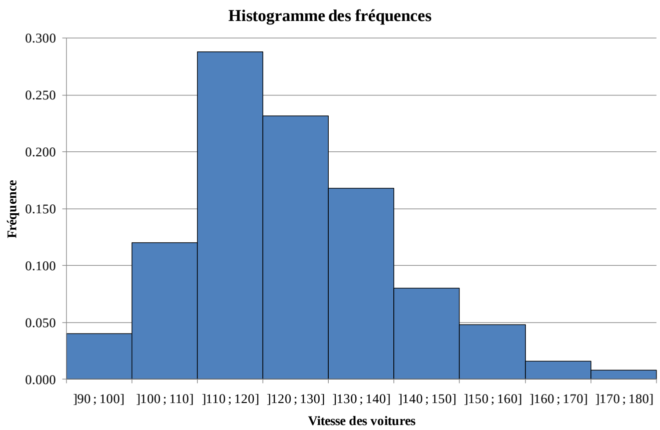
3. Mode : 7, médiane : 6
4. Moyenne : 6.1, écart-type : 1.991
5. 63.75 % ont obtenus 6 points, donc la moyenne de 4

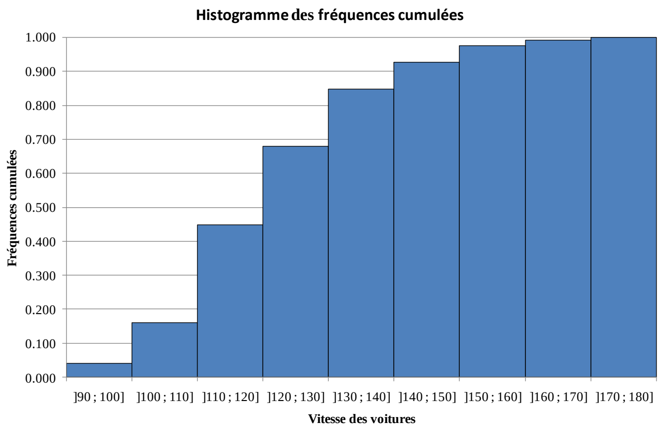
EXERCICE 2

1. Tableau statistique

Vitesse mesurée (en km/h) de 125 véhicules sur l'autoroute							
Classes (en km/h)	Centre de classe	Effectifs	Fréquence	Fréquence cumulée	Fréquence cumulée inv	Fréquence pondérée	Fréquence pond. carrée
]90 ; 100]	95	5	0.040	0.040	1.000	3.800	361.000
]100 ; 110]	105	15	0.120	0.160	0.960	12.600	1323.000
]110 ; 120]	115	36	0.288	0.448	0.840	33.120	3808.800
]120 ; 130]	125	29	0.232	0.680	0.552	29.000	3625.000
]130 ; 140]	135	21	0.168	0.848	0.320	22.680	3061.800
]140 ; 150]	145	10	0.080	0.928	0.152	11.600	1682.000
]150 ; 160]	155	6	0.048	0.976	0.072	7.440	1153.200
]160 ; 170]	165	2	0.016	0.992	0.024	2.640	435.600
]170 ; 180]	175	1	0.008	1.000	0.008	1.400	245.000
		125	1.0			124.28	15695.40

2. Histogrammes





3. Classe modale :]110 ; 120], classe médiane :]120 ; 130]
4. Vitesse moyenne : 124.28 km/h, écart-type :15.808
5. Vitesse autorisée en Suisse : 120km/h. 55.2 % des voitures dépassent cette vitesse

EXERCICE 3

Lundi après-midi							
Classe	Centre de	Effectif	Fréquence	Fréquence cumulée	Fréquence cum. Inverse	Fréquence pondérée	Fréquence pond. carrée
]0 ; 20]	10	7	0.070	0.070	1.000	0.700	7.000
]20 ; 40]	30	16	0.160	0.230	0.930	4.800	144.000
]40 ; 60]	50	32	0.320	0.550	0.770	16.000	800.000
]60 ; 80]	70	24	0.240	0.790	0.450	16.800	1176.000
]80 ; 100]	90	10	0.100	0.890	0.210	9.000	810.000
]100 ; 120]	110	7	0.070	0.960	0.110	7.700	847.000
]120 ; 140]	130	3	0.030	0.990	0.040	3.900	507.000
]140 ; 160]	150	1	0.010	1.000	0.010	1.500	225.000
Totaux		100	1			60.400	4516.000

Samedi après-midi							
Classe	Centre de	Effectif	Fréquence	Fréquence cumulée	Fréquence cum. Inverse	Fréquence pondérée	Fréquence pond. carrée
]0 ; 20]	10	5	0.020	0.020	1.000	0.200	2.000
]20 ; 40]	30	12	0.048	0.068	0.980	1.440	43.200
]40 ; 60]	50	23	0.092	0.160	0.932	4.600	230.000
]60 ; 80]	70	48	0.192	0.352	0.840	13.440	940.800
]80 ; 100]	90	61	0.244	0.596	0.648	21.960	1976.400
]100 ; 120]	110	34	0.136	0.732	0.404	14.960	1645.600
]120 ; 140]	130	29	0.116	0.848	0.268	15.080	1960.400
]140 ; 160]	150	14	0.056	0.904	0.152	8.400	1260.000
]160 ; 180]	170	12	0.048	0.952	0.096	8.160	1387.200
]180 ; 200]	190	6	0.024	0.976	0.048	4.560	866.400
]200 ; 220]	210	4	0.016	0.992	0.024	3.360	705.600
]220 ; 240]	230	2	0.008	1.000	0.008	1.840	423.200
Totaux		250	1			98.000	11440.800

	Lundi	Samedi
Mode	]40 ; 60]	]80 ; 100]
Médiane	]40 ; 60]	]80 ; 100]
Moyenne	60.40	98.00
Écart-type	29.459	42.858
% $\leq$ à Fr. 60.—	55 %	16 %
% $>$ à Fr. 100.—	11 %	40.4 %

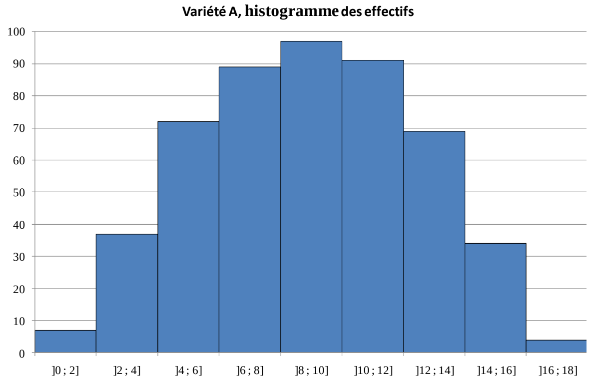
Lundi après-midi : Il y a moins de clients, ils font des achats moins onéreux. En général les clients effectuent les courses pour le lendemain.

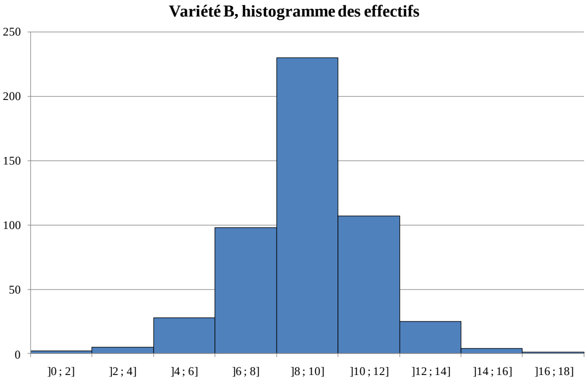
Samedi après-midi : Il y a beaucoup plus de clients qui font des achats plus onéreux. Ils les font les courses pour la semaine et achètent des marchandises plus importantes que le lundi (mobilier, jardinage, vêtements, …)

EXERCICE 4

Variété A							
Taille des grains de riz	Centre de classe	Effectifs	Fréquence	Fréquence cumulée	Fréquence cumulée inv	Fréquence pondérée	Fréquence pond. carrée
]0 ; 2]	1	7	0.014	0.014	1.000	0.014	0.014
]2 ; 4]	3	37	0.074	0.088	0.986	0.222	0.666
]4 ; 6]	5	72	0.144	0.232	0.912	0.720	3.600
]6 ; 8]	7	89	0.178	0.410	0.768	1.246	8.722
]8 ; 10]	9	97	0.194	0.604	0.590	1.746	15.714
]10 ; 12]	11	91	0.182	0.786	0.396	2.002	22.022
]12 ; 14]	13	69	0.138	0.924	0.214	1.794	23.322
]14 ; 16]	15	34	0.068	0.992	0.076	1.020	15.300
]16 ; 18]	17	4	0.008	1.000	0.008	0.136	2.312
		500	1.000			8.900	91.672

Variété B							
Taille des grains de riz	Centre de classe	Effectifs	Fréquence	Fréquence cumulée	Fréquence cumulée inv	Fréquence pondérée	Fréquence pond. carrée
]0 ; 2]	1	2	0.004	0.004	1.000	0.004	0.004
]2 ; 4]	3	5	0.010	0.014	0.996	0.030	0.090
]4 ; 6]	5	28	0.056	0.070	0.986	0.280	1.400
]6 ; 8]	7	98	0.196	0.266	0.930	1.372	9.604
]8 ; 10]	9	230	0.460	0.726	0.734	4.140	37.260
]10 ; 12]	11	107	0.214	0.940	0.274	2.354	25.894
]12 ; 14]	13	25	0.050	0.990	0.060	0.650	8.450
]14 ; 16]	15	4	0.008	0.998	0.010	0.120	1.800
]16 ; 18]	17	1	0.002	1.000	0.002	0.034	0.578
		500	1.000			8.984	85.080





	Variété A	Variété B
Classe modale	]8 ; 10]	]8 ; 10]
% de la classe modale	19.4 %	46 %
$\% \leq 8$ mm	41 %	26.6 %
$\% > 10$ mm	39.6 %	27.4 %
Classe médiane	]8 ; 10]	]8 ; 10]
Moyenne	8.9	8.984
Écart-type	3.53	2.09

Pourcentage des grains dont la taille est de $6 < \text{taille} \leq 14$ pour la variété A : 69.2 %
Pourcentage des grains dont la taille est de $6 < \text{taille} \leq 14$ pour la variété B : 92 %
La variété B se prête le mieux à commercialisation.

EXERCICE 5

La première distribution statistique montre une courbe de Gauss classique. Le mode, la médiane et la moyenne sont proches des valeurs centrales de la distribution (ici 15). L'écart-type est faible, cela signifie que les valeurs extrêmes influencent peu la moyenne et que la majorité des valeurs sont proches de la moyenne. Le caractère de la population n'est pas influencé par des critères externes à celui-ci. On peut donner comme exemple les notes à un examen de mathématique, la taille des garçons de 10 ans dans une école, etc.

La deuxième distribution statistique montre une courbe à deux bosses type chameau. La médiane et la moyenne sont voisines de 15, mais le mode est fortement décentré, il peut même y avoir deux modes très différents, ils mettent en évidence les deux bosses de la courbe. L'écart-type est très important, il signifie que les représentants des valeurs externes sont très nombreux. Ce type de courbe montre qu'un facteur externe influence fortement le caractère étudié. On rencontre cette distribution statistique si on étudie par exemple le poids d'une population de lions sans distinguer les mâles des femelles. La première bosse rassemble en majorité les femelles et la deuxième les mâles qui sont plus lourds que les femelles. On obtiendra le même type de courbe en étudiant la longueur des cheveux des élèves d'une école sans différencier les filles des garçons.

La troisième distribution statistique est caractérisée par une constance des mesures du caractère étudié et ceci pour l'ensemble de la population. La médiane et la moyenne restent centrées, mais on constate la présence de plusieurs modes, répartis sur toute la population. L'écart-type est important, ce qui montre que les valeurs ne sont pas centrées autour de la moyenne. On obtient ce genre de distribution statistique lorsque l'on étudie un caractère qui varie très peu dans une population. Par exemple, la longueur des cheveux des filles selon leur âge ou la hauteur moyenne des maisons dans différentes villes de Suisse,

EXERCICE 6

i	x_i	y_i	x_i^2	y_i^2	x_iy_i
1	1	12	1	144	12
2	2	10	4	100	20
3	3	13	9	169	39
4	4	12	16	144	48
5	5	11	25	121	55
6	6	14	36	196	84
7	7	20	49	400	140
8	8	13	64	169	104
9	9	8	81	64	72
10	10	15	100	225	150
11	11	16	121	256	176
12	12	15	144	225	180
13	13	17	169	289	221
14	14	25	196	625	350
15	15	15	225	225	225
16	16	18	256	324	288
17	17	20	289	400	340
18	18	16	324	256	288
19	19	22	361	484	418
20	20	20	400	400	400
21	21	31	441	961	651
22	22	22	484	484	484
23	23	24	529	576	552
24	24	15	576	225	360
25	25	23	625	529	575
26	26	24	676	576	624
27	27	21	729	441	567
28	28	33	784	1089	924
29	29	19	841	361	551
30	30	24	900	576	720
Somme	**465**	**548**	**9455**	**11034**	**9618**
Moyenne	**15.500**	**18.267**	**315.167**	**367.800**	**320.600**

$a = 0.500,\ b = 10.715,\ \ r = 0.741$

$y = 0.5x + 10.715$

On peut considérer que le nombre de repas est en augmentation constante, le coefficient de corrélation, qui est assez élevé le confirme. On remarque des journées dans le mois où le nombre de repas est plus important, le 7^e, 14^e, 21^e, et 28^e jour du mois. Ces dates correspondent toutes à un même jour de la semaine, peut-être un dimanche ?. Cela pourrait expliquer la fréquentation plus forte du restaurant à ces dates.

EXERCICE 7

i	x_i	y_i	x_i^2	y_i^2	$x_i y_i$
1	1	146	1	21316	146
2	2	206	4	42436	412
3	3	146	9	21316	438
4	4	133	16	17689	532
5	5	156	25	24336	780
6	6	202	36	40804	1212
7	7	124	49	15376	868
Somme	**28**	**1113**	**140**	**183273**	**4388**
Moyenne	**4.000**	**159.000**	**20.000**	**26181.857**	**626.857**

$a = -2.286,\ \ b = 168.143,\ \ r = -0.152$

$y = -2.286x + 168.143$

Le coefficient de corrélation est proche de zéro, il indique que les variables sont très peu dépendantes l'une de l'autre. L'exploitant a donc eu une mauvaise intuition !

EXERCICE 8

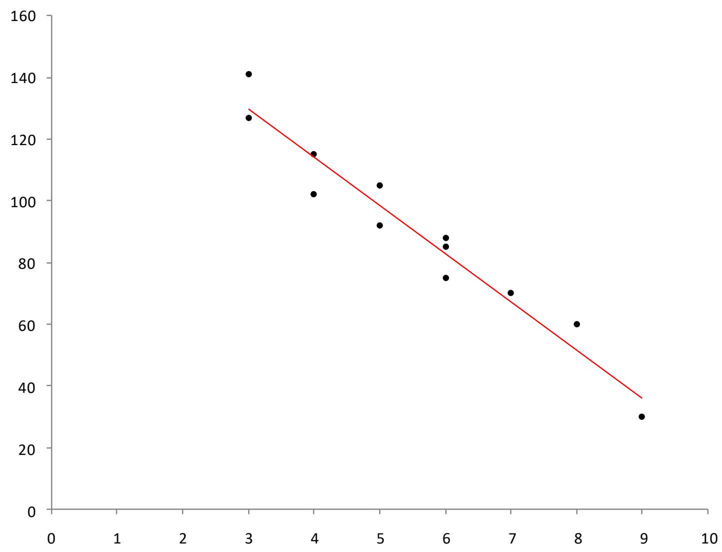
1. Calcul des coefficients a, b et r

i	x_i	y_i	x_i^2	y_i^2	$x_i y_i$
1	6	75	36	5625	450
2	5	92	25	8464	460
3	4	102	16	10404	408
4	6	85	36	7225	510
5	8	60	64	3600	480
6	4	115	16	13225	460
7	7	70	49	4900	490
8	5	105	25	11025	525
9	6	88	36	7744	528
10	9	30	81	900	270
11	3	141	9	19881	423
12	3	127	9	16129	381
Somme	**66**	**1090**	**402**	**109122**	**5385**
Moyenne	**5.500**	**90.833**	**33.500**	**9093.500**	**448.750**

$a = -15641,\ \ b = 176.859,\ \ r = 0.971$

$y = -15.641x + 176.859$

2. Graphique



3. Selon ce modèle de régression, quel serait la quantité d'objets vendus pour un prix de vente de Fr.10.— ?

   En remplaçant x par 10 dans l'équation, on trouve : $y = -15.641 \cdot 10 + 176.859 \cong 20.45$

4. Selon ce modèle de régression, à quel prix correspondrait une vente de 100 objets ?

   Il faut résoudre l'équation : $100 = -15.641x + 176.859 \Rightarrow x = 4.91$ francs