



Nathalie Villa-Vialaneix

Livret d'exercices de
Statistique Descriptive I (M1102A)

Année scolaire 2013/2014



Université de Perpignan Via Domitia, IUT
STatistique et Informatique Décisionnelle (STID)

Table des matières

1	Généralités sur la statistique descriptive	5
1.1	Population, variable	5
1.2	Population, variable, effectifs, classes	5
1.3	Population, variable	7
2	Représentations graphiques	11
2.1	Représentation graphique d'un tableau de fréquences	11
2.2	Diagramme cumulatif	11
2.3	Diagramme chronologique	12
2.4	Histogramme, polygone cumulatif	12
2.5	Histogramme, polygone des effectifs, polygone cumulatif	12
3	Caractéristiques numériques	13
3.1	Comparaison de populations	13
3.2	Moyenne, variance, quantiles	13
3.3	Moyenne, variance, quantiles	14
3.4	Problème	14

1 Généralités sur la statistique descriptive

1.1 Population, variable

Les États les plus étendus

Pays	Superficie (en km ²)
URSS	22 400 000
Canada	9 975 000
Chine	9 600 000
États-Unis	9 364 000
Brésil	8 512 000
Australie	7 700 000
Inde	3 268 000
Argentine	2 780 000
Soudan	2 506 000
Algérie	2 380 000

Les États les moins étendus

Pays	Superficie (en km ²)
Vatican	0,4
Tuvalu	2
Nauru	21
Saint-Marin	24
Liechtenstein	61
Monaco	160
Andorre	261
Saint Christopher and Nevis	300
Kiribati	311
Seychelles	316

Source : « Théma Encyclopédie, Le Monde d'Aujourd'hui (1991), Larousse, France »

Dans les deux tableaux ci-dessus, quelle est (quelles sont) la (les) population(s) étudiée(s), la (les) variable(s) étudiée(s) ? Donner la taille de la (des) population(s) ainsi que le type de la (des) variable(s).

1.2 Population, variable, effectifs, classes

1. Dans le tableau ci-dessous ¹, quelle est la population étudiée, les variables étudiées ? Donner la taille de la population ainsi que le type des variables.
2. Pour les deux variables quantitatives étudiées, construire les tableaux d'effectifs et d'effectifs cumulés à partir d'un regroupement en 5 classes de même amplitude. Ce regroupement en classes est-il pertinent ? (Justifier) Si non, en proposer un meilleur.

1. Source : « Théma Encyclopédie, Le Monde d'Aujourd'hui (1991), Larousse, France ».

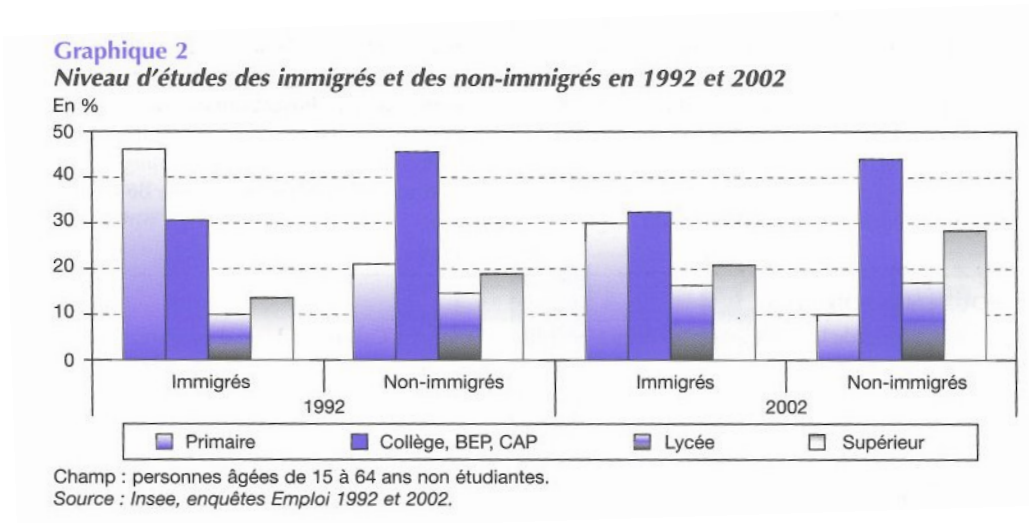
États	Superficie (en km ²)	Population	Capitale
Alabama (1)	133 915	3 894 000	Montgomery
Alaska (2)	1 530 000	407 000	Juneau
Arizona (3)	295 000	2 718 000	Phoenix
Arkansas (4)	138 000	2 286 000	Little Rock
Californie (5)	411 000	23 668 000	Sacramento
Caroline du Nord (6)	136 413	5 882 000	Raleigh
Caroline du Sud (7)	80 582	3 122 000	Columbia
Colorado (8)	270 000	3 139 000	Denver
Connecticut (9)	13 000	3 138 000	Hartford
Dakota du Nord (10)	183 119	653 000	Bismarck
Dakota du Sud (11)	199 730	691 000	Pierre
Delaware (12)	5 295	595 000	Dover
Floride (13)	151 670	9 740 000	Tallahassee
Géorgie (14)	152 488	5 463 000	Atlanta
Hawaï (15)	16 600	965 000	Honolulu
Idaho (16)	216 412	944 000	Boise
Illinois (17)	146 075	11 419 000	Springfield
Indiana (18)	94 000	5 499 000	Indianapolis
Iowa (19)	146 000	2 913 000	Des Moines
Kansas (20)	213 063	2 364 000	Topeka
Kentucky (21)	104 623	3 661 000	Frankfort
Louisiane (22)	125 674	4 204 000	Baton Rouge
Maine (23)	86 156	1 125 000	Augusta
Maryland (24)	27 092	4 216 000	Annapolis
Massachusetts (25)	21 500	5 737 000	Boston
Michigan (26)	150 780	9 258 000	Lansing
Minnesota (27)	217 735	4 077 000	Saint Paul
Mississippi (28)	123 500	2 521 000	Jackson
Missouri (29)	180 500	4 917 000	Jefferson
Montana (30)	381 000	787 000	Helena
Nebraska (31)	200 000	1 597 000	Lincoln
Nevada (32)	295 000	891 000	Carson
New Hampshire (33)	24 000	959 000	Concord
New Jersey (34)	20 169	7 468 000	Trenton
New York (35)	128 400	17 667 000	Albany
Nouveau-Mexique (36)	315 000	1 399 000	Santa Fe
Ohio (37)	107 000	10 797 000	Columbus
Oklahoma (38)	181 000	3 298 000	Oklahoma City
Oregon (39)	251 000	2 662 000	Salem
Pennsylvanie (40)	117 400	11 895 000	Harrisburg
Rhode Island (41)	3 144	955 000	Providence
Tennessee (42)	109 152	4 685 000	Nashville-Davidson
Texas (43)	690 000	15 724 000	Austin
Utah (44)	220 000	1 619 000	Salt Lake City
Vermont (45)	24 887	525 000	Montpelier
Virginie (46)	105 586	5 550 000	Richmond
Virginie Occidentale (47)	62 759	1 965 000	Charleston
Washington (48)	176 500	4 132 000	Olympia
Wisconsin (49)	145 348	4 751 000	Madison
Wyoming (50)	253 500	514 000	Cheyenne

1.3 Population, variable

Pour les graphiques ci-dessous, répondre aux questions suivantes :

- Quelle est (quelles sont) la (les) population(s) étudiée(s) ? Si possible, donner leur(s) taille(s).
- Quelle est (quelles sont) la (les) variable(s) étudiée(s) ? Leur(s) type(s) ?

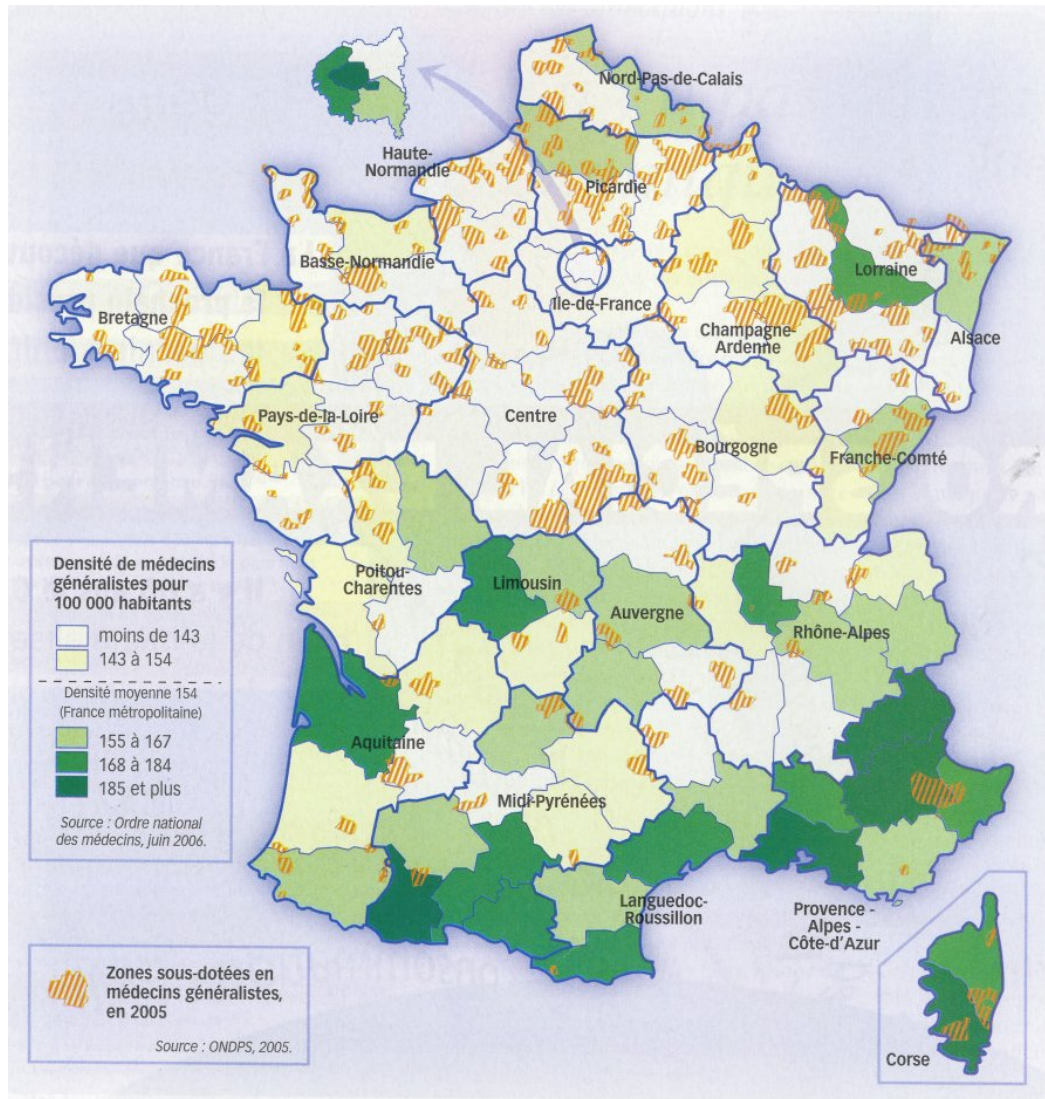
1. a) À partir de ce graphique², construire des tableaux de fréquences.
- b) Le tableau de fréquences cumulées a-t-il un sens ? Si oui, le déterminer.



2. a) À partir du graphique ci-dessous³, construire un tableau d'effectifs, d'effectifs cumulés, de fréquences cumulées.
- b) Quel est le pourcentage de départements français dont la densité de médecins généralistes est inférieure à la moyenne nationale ?

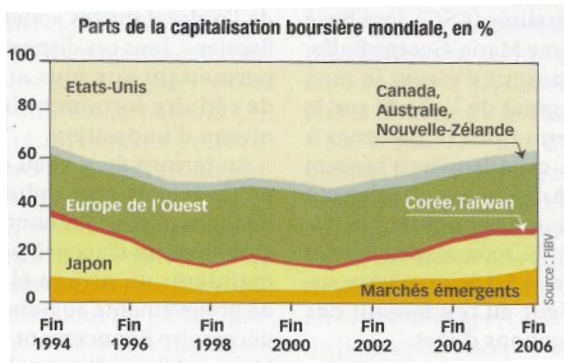
2. Source : « France, Portrait Social », Collection INSEE Références.

3. Source : « Alternatives économiques », Numéro 257, Avril 2007.

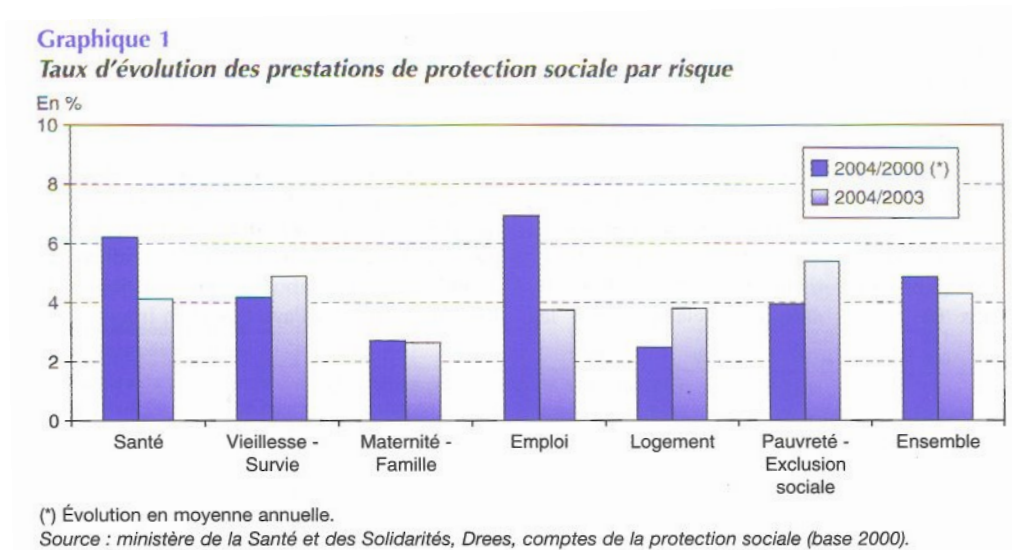


3. a) À partir du graphique ci-dessous⁴, quelle est la part de la capitalisation boursière de chacun des grands ensembles présentés fin 2004 ?
- b) Que peut-on dire de l'évolution de la part de la capitalisation boursière en Europe de l'Ouest entre 1994 et 2006 ? Même question pour les États-Unis.

4. Source : « Alternatives économiques », Numéro 257, Avril 2007.



4. a) À partir du graphique ci-dessous⁵, quel est le taux annuel moyen d'évolution des prestations santé entre 2000 et 2004? Quel est le taux global d'évolution entre ces deux dates?
- b) Peut-on dire que les prestations santé ont diminué entre 2003 et 2004? Si oui, de quelle quantité?

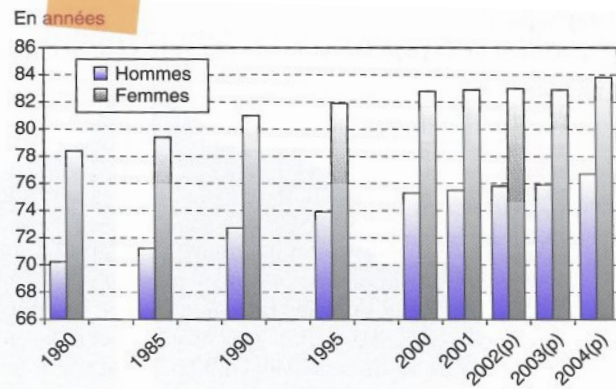


5. a) Le graphique ci-dessous⁶ peut induire une erreur d'interprétation : pourquoi?
- b) Proposez une meilleure représentation graphique de ces données.

5. Source : « France, Portrait Social », Collection INSEE Références.

6. Source : « France, Portrait Social », Collection INSEE Références.

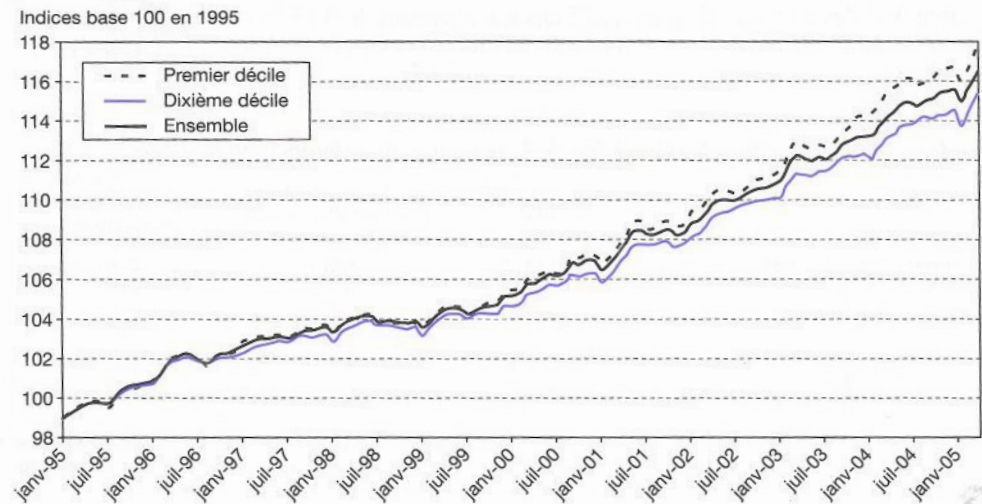
Graphique 1
Évolution de l'espérance de vie à la naissance



(p) Résultats provisoires.
Champ : France métropolitaine.
Sources : Insee, statistiques de l'état civil et enquête Villes.

6. a) À partir du graphique ci-dessous⁷, que peut-on dire du niveau de vie des ménages en janvier 2005 par rapport au niveau de vie moyen des ménages en 1995 ?

Graphique 1
Indices catégoriels des premiers et dixième déciles de niveau de vie et indice d'ensemble des ménages



7. Source : « France, Portrait Social », Collection INSEE Références.

2 Représentations graphiques

2.1 Représentation graphique d'un tableau de fréquences

Le tableau ci-dessous¹ donne, pour les couples ayant divorcé moins de 5 ans après leurs mariages, le nombre d'années (entières) entre leurs mariages et leurs divorces.

Nb d'années	1 an	2 ans	3 ans	4 ans	5 ans
Fréquence	3,96 %	18,02 %	26,92 %	25,38 %	25,72 %

1. Déterminer la population, la variable étudiée et son type.
2. Quelles sont toutes les représentations graphiques vues en cours de cette variable ? (Les faire)

2.2 Diagramme cumulatif

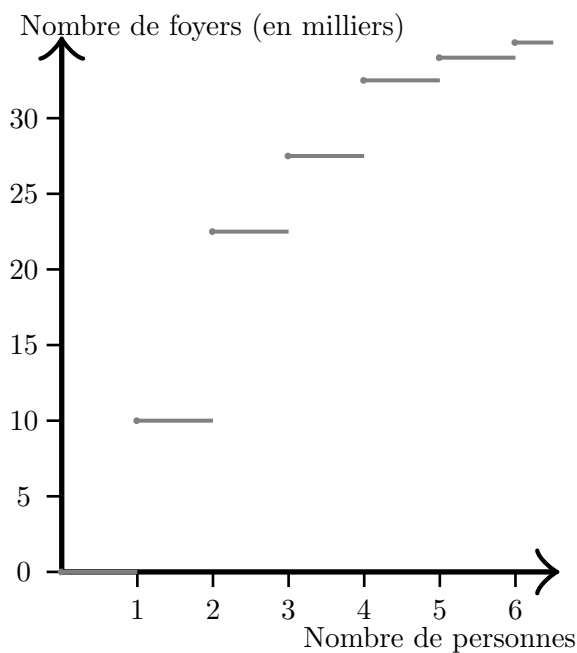


Diagramme cumulatif du nombre de personnes par foyer dans le département de la Corrèze en 1999

1. Source : Données INSEE.

Le graphique ci-dessus est le diagramme cumulatif du nombre de personnes par foyer (effectifs en nombre de foyers) dans le département de la Corrèze. *Source : Données INSEE, recensement 1999*

1. Identifier la population, la variable étudiée et son type.
2. À partir du graphique précédent, reconstruire le tableau des effectifs cumulés puis le tableau des effectifs. Quelle est la taille de la population ?
3. Construire le diagramme en bâtons des effectifs de cette variable.

2.3 Diagramme chronologique

Reprendre l'exercice 1.3 page 7, question 5. Effectuer le diagramme chronologique en lignes des variables « Espérance de vie des hommes » et « Espérance de vie des femmes ».

2.4 Histogramme, polygone cumulatif

Reprendre l'exercice 1.2 page 5. À partir du regroupement en classes de la variable population ci-dessous :

Pop. ($\times 10^3$)	[400; 1 000[[1 000; 3 000[[3 000; 5 000[[5 000; 10 000[[10 000; 23 700]
n_i	12	11	13	8	6

construire :

1. l'histogramme de la variable « Population » ;
2. le polygone cumulatif de cette variable.

Commenter ces graphiques.

2.5 Histogramme, polygone des effectifs, polygone cumulatif

Reprendre l'exercice 1.3 page 7, question 2. À partir du regroupement en classes de la variable « Densité de médecins généralistes » que l'on rappelle ci-dessous :

Densité	[100; 143[[143; 155[[155; 168[[168; 185[[185; 250]
n_i	44	17	17	12	6

construire :

1. l'histogramme et le polygone des effectifs de la variable « Densité de médecins généralistes » ;
2. le polygone cumulatif de cette variable.

3 Caractéristiques numériques

3.1 Comparaison de populations

On reprend les données de l'exercice 1.1 page 5 :

Les États les plus étendus		Les États les moins étendus	
Pays	Superficie (en km ²)	Pays	Superficie (en km ²)
URSS	22 400 000	Vatican	0,4
Canada	9 975 000	Tuvalu	2
Chine	9 600 000	Nauru	21
États-Unis	9 364 000	Saint-Marin	24
Brésil	8 512 000	Liechtenstein	61
Australie	7 700 000	Monaco	160
Inde	3 268 000	Andorre	261
Argentine	2 780 000	Saint Christopher and Nevis	300
Soudan	2 506 000	Kiribati	311
Algérie	2 380 000	Seychelles	316

Comparativement, respectivement, aux 10 États les plus étendus et aux 10 États les moins étendus, lequel de ces deux pays, Brésil ou Monaco, a la plus grande superficie relative ?

3.2 Moyenne, variance, quantiles

On reprend les données de l'exercice 1.2 page 5 :

1. Dans le regroupement en classes ci-dessous, l'Alaska (1 530 000 km²) a été exclue.

Superficie (en milliers de km ²)	[0; 100[[100; 150[[150; 200[[200; 300[[300; 700]
n_i	13	14	8	10	4

Calculer la moyenne puis l'écart type de la variable « Superficie » pour ces 49 États à partir du regroupement en classes proposé. Que deviennent la moyenne et l'écart type si on ajoute l'Alaska ? Commentez la différence.

2. Déterminer, graphiquement puis numériquement, la médiane et les quartiles de la variable « Superficie » à partir du regroupement en classes proposé à la première question. Construire la boîte à moustaches.
Pourquoi ces valeurs évoluent peu si on ajoute l'Alaska ?

- À partir des caractéristiques numériques déterminées dans les questions 1 et 2, faire des commentaires sur la distribution de la variable « Superficie » sur l'ensemble des États américains.

3.3 Moyenne, variance, quantiles

On reprend les données de l'exercice 2.2 page 11 :

Nb d'années avant le divorce	1 an	2 ans	3 ans	4 ans	5 ans
Fréquence	3,96 %	18,02 %	26,92 %	25,38 %	25,72 %

- Calculer la moyenne et l'écart type de la variable « Nombre d'années avant le divorce ».
- Déterminer la médiane de cette variable : que nous apprend la comparaison de la médiane et de la moyenne ?
- Déterminer les quartiles de cette variable.

3.4 Problème

On donne, ci-dessous, les résultats du recensement 2004/2005 pour des communes de la Corrèze (résultats partiels portant sur un cinquième des communes de moins de 10 000 habitants)¹.

1. Source : INSEE.

Nom de la commune	Recensement 2004/2005
Albignac	252
Allasac	3 561
Altillac	839
Argentat	3 111
Arnac-Pompadour	1 281
Astailac	223
Aubazines	798
Ayen	669
Bassignac-Le-Bas	95
Bassignac-Le-Haut	184
Beaumont	127
Beyssac	782
Beysсенac	360
Bonnefond	134
Brignac-La-Plaine	776
Camps-Saint-Marthurin-Léobazel	243
Chabrignac	467
Chamboulive	1 253
Champagnac-La-Prune	167
Chanac-Les-Mines	503
La Chapelle-Aux-Brocs	348
La Chapelle-Saint-Géraud	221
Chartrier-Ferrière	325
Chasteaux	510
Chauffour-Sur-Vell	354
Chaumeil	171
Chavagnac	59
Chenailler-Mascheix	164
Clergoux	374
Concèze	403
Confolent-Port-Dieu	30
Cublac	1 534
Darazac	159
Darnets	336
Donzenac	2 310
Égletons	4 424
Espagnac	341
Espartignac	370
Estivaux	363
Eyburie	491
Eyrein	510
Gimel-Les-Cascades	670
Gourdon-Murat	118
Hautefage	300
Le Jardin	73
Jugeals-Nazareth	776
Lafage-Sur-Sombre	123
Liginiac	611

Nom de la commune	Recensement 2004/2005
Lignareix	154
Le Lonzac	835
Malemort-Sur-Corrèze	7 019
Marc-La-Tour	151
Marcillac-La-Croze	210
Meilhards	530
Menoire	81
Mestes	297
Meyrignac-L'Église	54
Montgibaud	234
Moustier-Ventadour	434
Neuville	194
Noailhac	323
Noailles	744
Orgnac-Sur-Vézère	303
Orliac-De-Bar	242
Palisse	230
Péret-Bel-Air	96
Pérols-Sur-Vézère	183
Peyrelevade	806
Rilhac-Treignac	115
Rilhac-Xaintrie	317
La Roche-Canillac	173
Sadroc	737
Saint-Angel	670
Saint-Augustin	434
Saint-Bonnet-Avalouze	216
Saint-Bonnet-Elvert	185
Saint-Cyprien	254
Saint-Cyr-La-Roche	399
Saint-Étienne-Aux-Clos	229
Saint-Étienne-La-Geneste	71
Saint-Fréjoux	269
Saint-Geniez-ô-Merle	103
Saint-Germain-Les-Vergnes	912
Saint-Julien-Aux-Bois	490
Saint-Julien-Le-Pèlerin	145
Saint-Julien-Près-Bort	407
Saint-Martial-De-Gimel	498
Saint-Martin-La-Méanne	359
Saint-Mexant	1062
Saint-Pardoux-Corbier	346
Saint-Pardoux-Le-Vieux	273
Saint-Paul	233
Saint-Priest-De-Gimel	461
Saint-Privat	1 108
Saint-Sylvain	140
Saint-Victour	167
Sarran	292

Nom de la commune	Recensement 2004/2005
Ségur-Le-Château	233
Seilhac	1 724
Sexcles	229
Soudaine-Lavinadière	204
Thalamy	77
Treignac	1 389
Troche	504
Turenne	770
Vars-Sur-Roseix	283
Végennes	170
Veix	69
Vigeois	1 185
Vignols	578
Voutezac	1 174

Parallèlement, on donne la répartition suivante pour la même enquête concernant des communes des Yvelines :

Nombre d'habitants	[0; 400[[400; 600[[600; 800[[800; 1 200[[1 200; 2 500[[2 500; 10 000[
Nombre de communes	17	17	10	12	11	22

1. Quelle est (quelles sont) la (les) population(s) étudiée(s) ? Leur(s) taille(s) ? Quelle est (quelles sont) la (les) variable(s) étudiée(s) ? Leur(s) type(s) ?
2. Effectuer, pour les communes de la Corrèze, un regroupement en classes de la variable « Population ».
3. Construire, pour la Corrèze et pour les Yvelines, les histogrammes des séries statistiques regroupées en classes, à la même échelle (pour cela, on calculera les densités à partir des fréquences des deux séries statistiques). Commenter.
4. Calculer la moyenne et l'écart type de ces deux séries statistiques regroupées en classes. Commenter.
5. La commune de Essarts-Le-Roi, dans les Yvelines, a une population de 6 191 habitants. Comparativement au département dans lequel elle se situe, est-elle plus grande que la commune d'Égletons ?
6. Déterminer, pour ces deux séries statistiques, les quartiles (à partir des regroupements en classes). Quels commentaires vous amène à faire la comparaison moyenne / médiane pour ces deux séries ? Construire sur le même graphique les deux boîtes à moustaches correspondantes. Que pouvez-vous dire à la vue de ce graphique ? Cela rejoint-il les conclusions déjà tirées ?
7. Calculer les coefficients d'asymétrie et d'aplatissement de ces deux séries statistiques. Commenter. Pouvait-on prévoir ces résultats au vu de la question précédente ?
8. Construire, sur un même graphique, les courbes de Lorenz de ces deux séries statistiques. Calculer les indices de Gini correspondants puis commenter.