

Exercice 3 : La composition socio-démographique des villes selon leur taille

CORRIGÉ

La population de France métropolitaine a été distribuée selon la taille de la ville de résidence (tableau 1). Neuf groupes de villes ont été définis. Pour chacun de ces groupes de villes, on a calculé pour l'année 2006 les proportions moyennes :

- de personnes âgées de 50 ans ou plus ;
- d'actifs âgés de 25-54 ans appartenant à la catégorie des cadres et professions intermédiaires.

a) Représentez graphiquement ces données afin de mettre en évidence le lien éventuel entre la composition socio-démographique (part des 50 ans et plus, proportion de cadres et de professions intermédiaires) et la taille des villes. Commentez brièvement ces données.

b) Y-a-t-il un lien statistique entre la proportion de personnes âgées de 50 ans ou plus et celle de cadres et de professions intermédiaires ? Préciser l'intensité de ce lien statistique en calculant le coefficient de détermination et le coefficient de corrélation linéaire.

Tableau 1 : Proportions moyennes d'habitants âgés de 50 ans ou plus et de cadres et de professions intermédiaires selon la taille de la ville de résidence en 2006

Taille des villes	Cadres et professions intermédiaires	50 ans et +
Commune rurale	32%	37%
moins de 5 000 hab.	34%	38%
de 5 000 à 9 999 hab.	35%	38%
de 10 000 à 19 999 hab.	35%	38%
de 20 000 à 49 999 hab.	36%	37%
de 50 000 à 99 999 hab.	37%	36%
de 100 000 à 199 999 hab.	40%	35%
de 200 000 à 1 999 999 hab.	45%	33%
Unité urbaine de Paris	51%	29%

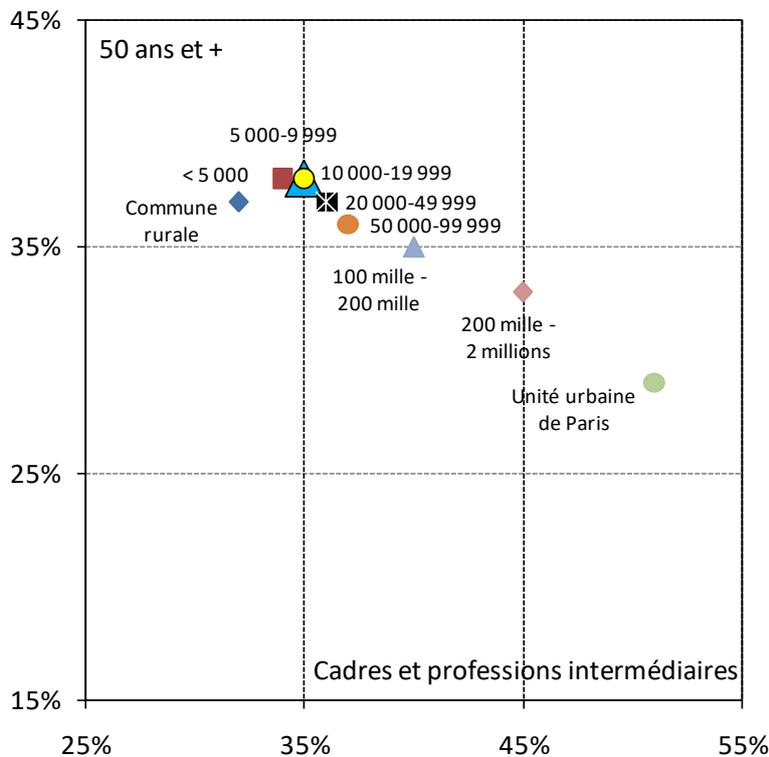
Lecture : la population résidant dans des villes de 100 000 à 199 999 habitants compte 40 % de cadres et de professions intermédiaires (parmi les actifs âgés de 25-54 ans) et 35 % de personnes âgées de 50 ans ou plus.

Définition : la ville est entendue ici au sens de l'agglomération ou unité urbaine (commune centre + communes assurant une continuité du bâti autour de la commune centre).

Source : calculs JFL d'après INSEE, RR 2006

Corrigé

a) Représentation graphique



La composition socio-démographique de la population des différents types de villes est ici définie par deux critères : la proportion de personnes âgées et celle des cadres (supérieurs et moyens) parmi la population active. Ces deux critères doivent figurer sur le graphique (cf. supra).

De cette manière, chaque point correspond à la population d'un type de ville. Il apparaît nettement que selon leur taille, les villes accueillent des populations dont la composition diffère sensiblement. Il est même possible de mettre en évidence un lien entre taille des villes et la composition socio-démographique.

Globalement, plus les villes sont importantes sur le plan démographique, plus la part de cadres est élevée et plus celle de personnes âgées de 50 ans et plus est faible. Cela est lié au fait que les grandes métropoles, au premier rang desquelles figure Paris et son agglomération, concentrent la plupart des emplois de cadres (cf. le n° 1278 de *INSEE Première* consacrée à la Répartition géographique des emplois – en pièce jointe sur l'EPI). Elles accueillent en outre une population étudiante et active importante, ce qui réduit la part des retraités et donc des 50 ans et plus. Ces spécificités sociodémographiques des villes sont de plus en plus marquées (cf. en pièce jointe sur l'EPI un article sur l'évolution socio-démographique des villes moyennes. À propos de ce document : il se fonde sur des outils de statistique descriptive qui ont tous été présentés en cours cette année. Il s'agit donc d'un type de travail que vous êtes dorénavant en mesure de réaliser).

Enfin, on peut noter qu'en 2007, du point de vue des proportions de personnes âgées de 50 ans et plus et de cadres au sein de la population active, les villes de moins de 20 000 habitants forment un ensemble homogène. La population des villes comptant de 20 000 à 100 000 habitants se distingue encore un peu de cet ensemble. Mais ces caractéristiques socio-démographiques paraissent

toutefois plus proches de celles de la population des petites villes que de celles de la population résidant dans des villes comptant au moins 100 000 habitants.

b) Le graphique montre clairement qu'à partir d'un certain seuil, plus la proportion de cadres parmi la population active augmente, plus celle de personnes âgées de 50 ans et plus est faible. Le calcul des coefficients de détermination (R^2) et de corrélation linéaire (r) confirme sans surprise cette liaison statistique.

$$R^2 = \frac{\text{COV}(x, y)^2}{\sigma_x^2 \times \sigma_y^2}$$

$$r = \frac{\text{COV}(x, y)}{\sigma_x \times \sigma_y}$$

Avec :

$$\text{COV}(x, y) = \frac{1}{n} \times \sum_{i=1}^9 [(x_i - \bar{x}) \times (y_i - \bar{y})]$$

$$\sigma_x^2 = \frac{1}{n} \times \sum_{i=1}^9 (x_i - \bar{x})^2$$

$$\sigma_y^2 = \frac{1}{n} \times \sum_{i=1}^9 (y_i - \bar{y})^2$$

On considèrera que x correspond à la proportion de personnes âgées de 50 ans et plus et que y correspond donc à la proportion de cadres et professions intermédiaires au sein de la population active.

Application numérique

Taille des villes	i	% de 50 ans et + (x)	% de Cadres et prof. intermédiaires (y)	$(x_i - \bar{x}_{\text{moy}}) * (y_i - \bar{y}_{\text{moy}})$	$(x_i - \bar{x}_{\text{moy}})^2$	$(y_i - \bar{y}_{\text{moy}})^2$
Commune rurale	1	37	32	-8,4	1,8	40,1
< 5 000 hab.	2	38	34	-10,1	5,4	18,8
5 000 - 9 999 hab.	3	38	35	-7,8	5,4	11,1
10 000 - 19 999 hab.	4	38	35	-7,8	5,4	11,1
20 000 - 49 999 hab.	5	37	36	-3,1	1,8	5,4
50 000 - 99 999 hab.	6	36	37	-0,4	0,1	1,8
100 000 - 199 999 hab.	7	35	40	-1,1	0,4	2,8
200 000 - 1 999 999 hab.	8	33	45	-17,8	7,1	44,4
Unité urbaine de Paris	9	29	51	-84,4	44,4	160,4
Total				-141,0	72,0	296,0

$$\bar{x} = \frac{37 + 38 + \dots + 29}{9} = 36$$

$$\bar{y} = \frac{32 + 34 + \dots + 51}{9} = 38$$

$$\text{Cov}(x, y) = \frac{-141,0}{9} = -15,7$$

$$\text{Var}(x) = \frac{72,0}{9} = 8,0$$

$$\text{Var}(y) = \frac{296,0}{9} = 32,9$$

$$R^2 = \frac{(-15,7)^2}{8,0 \times 32,9} = 0,933 \cong 93 \%$$

$$r = \frac{-15,7}{\sqrt{8,0} \times \sqrt{32,9}} = -0,966 \cong -97 \%$$

La corrélation est bien très significative et négative.