Introduction à l'analyse spatiale dans le cadre du géomarketing

Serge Lhomme

Maître de conférences en géographie
http://sergelhomme.fr/
serge.lhomme@u-pec.fr

Introduction

Fondement épistémologique de l'analyse spatiale

La géographie s'appuie historiquement, en particulier en France, sur une démarche d'étude des phénomènes dans leur singularité. Cette démarche est qualifiée d'idiographique.

Ainsi, la géographie s'est développée en s'appuyant notamment sur de longues monographies concernant des territoires délimités par des caractéristiques physiques, culturelles, historiques ou sociales.

Tout le savoir géographique fut alors mis au service de cette unicité, en évacuant probablement ce qui est peut-être le plus fécond : ce qui est commun entre les territoires.

La démarche idiographique tend ainsi à s'opposer à une géographie « générale » qui étudierait les processus et les phénomènes pris par thème, classant les phénomènes et cherchant les règles générales qui les régissent.

Introduction

Fondement épistémologique de l'analyse spatiale

L'analyse spatiale est une démarche qui, à l'instar d'une géographie « générale », recherche les similarités entre des phénomènes spatiaux et tente d'établir des lois (des règles). Cette démarche est qualifiée de nomothétique.

En l'occurrence, comme son nom l'indique, l'analyse spatiale recherche ou applique des lois spatiales.

Ces lois qui reposent généralement sur des hypothèses simplificatrices peuvent dans certains cas être testées afin de déterminer leur validité. L'analyse spatiale est donc généralement présentée comme une démarche « hypothético-déductive ».

Certains modèles d'analyse spatiale sont très utilisés dans le domaine du géomarketing, notamment en matière de zone de chalandise.

Introduction

Fondement épistémologique de l'analyse spatiale

Définition

L'analyse spatiale met en évidence des structures et des formes d'organisation spatiale récurrentes.

Elle analyse les processus qui sont à l'origine de ces structures, à travers des concepts comme ceux de distance, d'interaction, de portée spatiale, de polarisation, de centralité, de stratégie, de territorialité...

Des lois de la spatialité relient ces formes et ces processus.

L'analyse spatiale est souvent associée à une géographie quantitative, car certains modèles (certaines lois, règles) issus de l'analyse spatiale sont quantitatifs.

Introduction

Un modèle d'interaction fondamental : le modèle gravitaire

Le modèle gravitaire est destiné à formaliser, étudier, reproduire et prévoir les interactions (les flux) entre des lieux.

En effet, la répartition des interactions dans un ensemble de lieux dépend de leur configuration, notamment des forces d'attraction et des difficultés de communication entre chaque lieu.

Le modèle gravitaire a été formulé par analogie avec la loi de la gravitation universelle de Newton : deux corps s'attirent en raison directe de leur masse et en raison inverse de la distance qui les sépare.

Le modèle gravitaire (Physique)

$$F_{ij} = \frac{k \times P_i \times P_j}{{d_{ij}}^2} = k \times P_i \times P_j \times {d_{ij}}^{-2}$$

Introduction

Un modèle d'interaction fondamental : le modèle gravitaire

Le modèle gravitaire (Géographie)

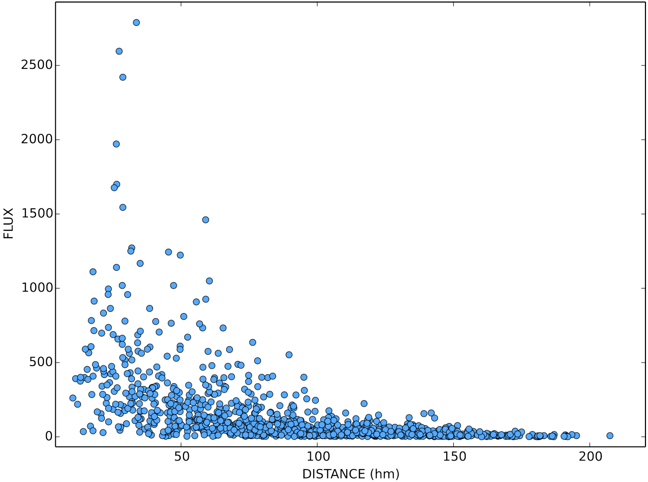
$$F_{ij} = \frac{k \times P_i \times P_j}{{d_{ij}}^n} = k \times P_i \times P_j \times {d_{ij}}^{-n}$$

Dans un espace de circulation relativement homogène, les échanges entre deux lieux (régions, villes...) seront d'autant plus importants que le poids des lieux est grand et d'autant plus faibles qu'ils seront éloignés.

Le modèle gravitaire résume bien l'essentiel des mouvements qui se produisent dans un milieu où la mobilité et l'accessibilité sont relativement homogènes. Il prédit par exemple assez bien l'ampleur des flux de déplacements domicile-travail dans un bassin d'emploi urbain, à partir de la répartition des zones de résidence et des zones d'emploi.

Introduction

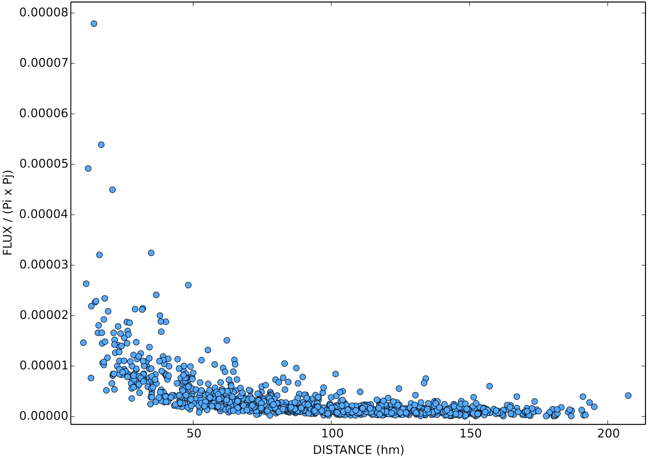
Un modèle d'interaction fondamental : le modèle gravitaire



Les flux domicile-travail entre les communes du Val-de-Marne en fonction de la distance.

Introduction

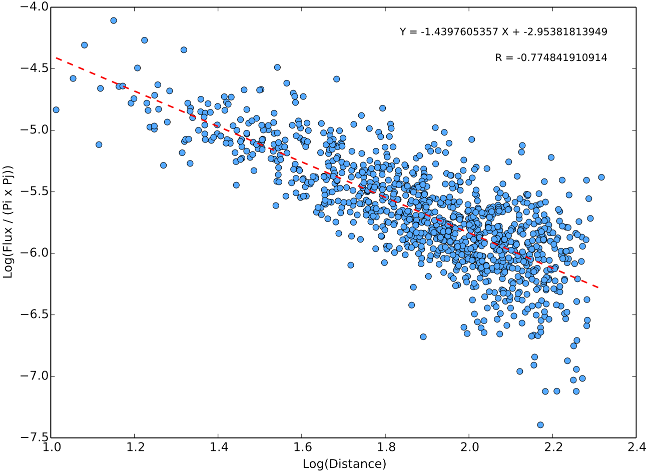
Un modèle d'interaction fondamental : le modèle gravitaire



Les flux domicile-travail rapportés par les forces d'attraction entre les communes du Val-de-Marne en fonction de la distance.

Introduction

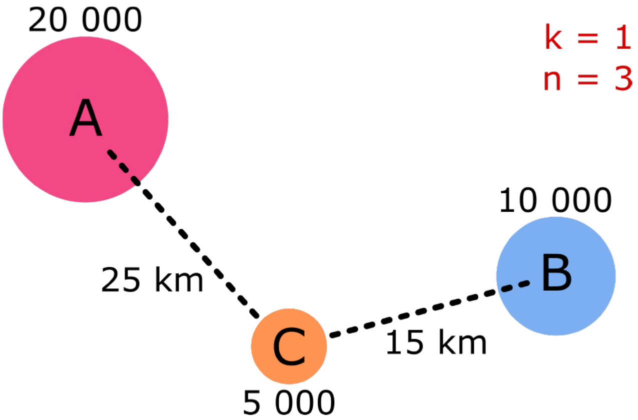
Un modèle d'interaction fondamental : le modèle gravitaire



Les flux domicile-travail rapportés par les forces d'attraction en fonction de la distance dans un diagramme bi-logarithmique.

Introduction

Un modèle d'interaction fondamental : le modèle gravitaire



Introduction

La loi de Reilly : le rapport des attractions commerciales

En 1931, Reilly formula une loi concernant l'organisation des zones d'attraction commerciale qui reprend la formulation du modèle gravitationnel de Newton. Cette attraction se traduit concrètement par des volumes d'achats.

«*Deux centres attirent les achats des populations situées entre elles en proportion directe du nombre total d'habitants des villes considérées et en proportion inverse du carré de la distance qu'il faut parcourir pour s'y rendre.*»

Le modèle gravitaire (Attraction commerciale)

$$V_{ij} = \frac{k \times P_i \times P_j}{{d_{ij}}^2} = k \times P_i \times P_j \times {d_{ij}}^{-2}$$

Introduction

La loi de Reilly : le rapport des attractions commerciales

Selon Reilly, cette loi est fondée sur des observations empiriques qui quantifient les attractions commerciales entre différentes villes.

Dans ce cadre, si l'on prend un agent économique situé entre deux centres A et B, alors les attractions de A et de B sur l'agent économique sont les suivantes :

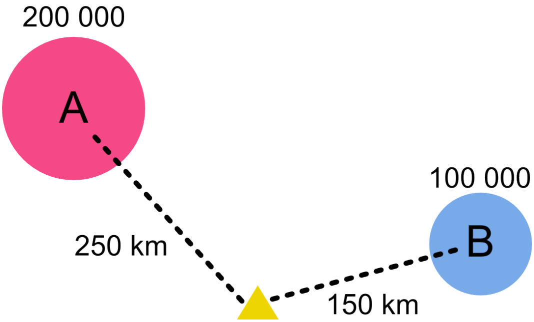
$$V_A = \frac{P_A}{{d_A}^2} \times cste \; \& \; V_B = \frac{P_B}{{d_B}^2} \times cste$$

Loi de Reilly (Rapport des attractions)

$$\frac{V_A}{V_B} = \frac{P_A}{P_B} \times (\frac{d_B}{d_A})^2$$

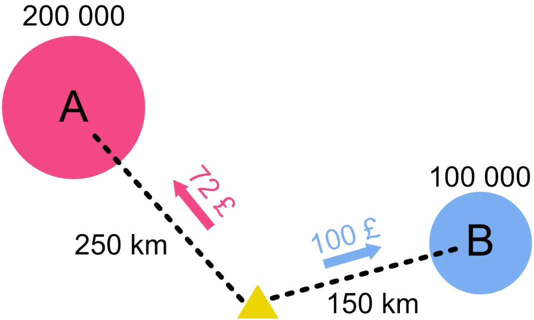
Introduction

La loi de Reilly : le rapport des attractions commerciales



Introduction

La loi de Reilly : le rapport des attractions commerciales



Loi de Reilly (Conséquence)

$$V_A = \frac{V_{TOT} \times LR}{1 + LR} \text{ et } V_B = \frac{V_{TOT}}{1 + LR}$$

Introduction

La loi de Converse : déterminer les points de partage

Par la suite, Converse (1951) établit un modèle permettant de délimiter les frontières des aires d'influence entre deux pôles commerciaux. Pour cela, il s'appuya sur le modèle gravitaire et les travaux de Reilly.

Plus précisément, ce modèle permet de déterminer un point d'équilibre (breaking point) entre les zones de desserte de deux pôles commerciaux.

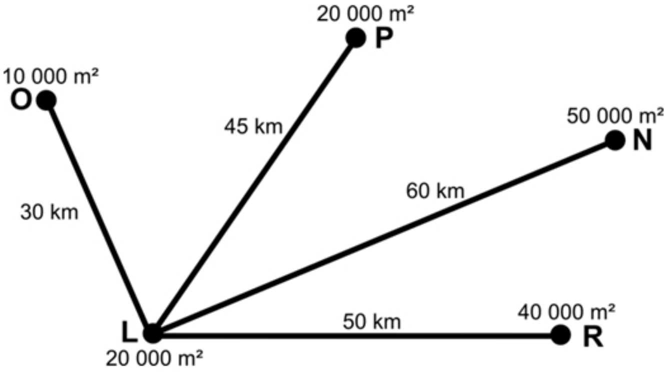
Ce point de partage (ou d'équilibre) définit la limite entre les aires d'influence de deux pôles de taille Pa et Pb séparées par une distance Dab, car au niveau de cette frontière Va et Vb sont équivalents, on a donc Va/Vb $= 1$.

Loi de Converse (Point de partage)

$$D_{xb} = \frac{D_{ab}}{1 + \sqrt{(P_a/P_b)}}$$

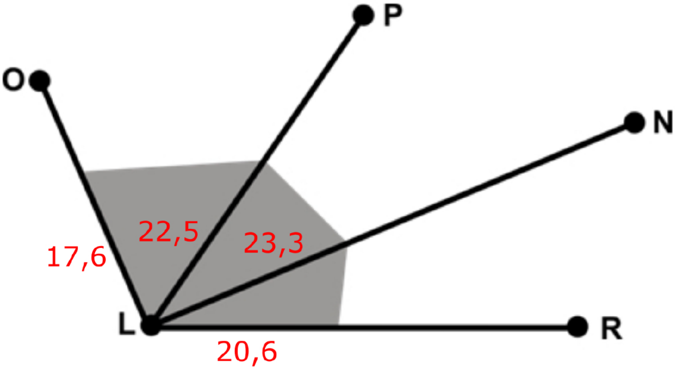
Introduction

La loi de Converse : déterminer les points de partage



Introduction

La loi de Converse : déterminer les points de partage



Introduction

Le modèle de Huff : une approche probabiliste

Certains reprochent à la loi de Reilly (Converse) son caractère déterministe et préfèrent le modèle de Huff qui propose une formulation probabiliste de la loi de Reilly.

Le modèle de Huff propose une généralisation de la loi de Reilly et un changement de perspective en prenant comme point de départ de la formulation les clients et non les centres :

- Il est alors possible d'étudier simplement la concurrence entre plusieurs centres et ce sans recours à des traitements géométriques (la formulation sous la forme d'un duopole de Reilly peut en effet mener à des apories).
- Il est aussi possible de prendre en considération plus finement les effets de la distance sur certains types de biens (n n'est plus obligatoirement égal à deux).

Introduction

Le modèle de Huff : une approche probabiliste

Chaque centre d'achats j représente pour le consommateur i une opportunité de destination que l'on peut évaluer par la formule suivante : $O_{ij} = P_j / {D_{ij}}^n$

Le potentiel d'opportunités pour un consommateur situé en i est égal à la somme de toutes les opportunités de destination : $\bar{O}_i = \sum O_{ij}$

La probabilité de choisir une destination est égale à l'opportunité de cette destination divisée par la somme totale des opportunités de destination.

Modèle de Huff

$$PR_{ij} = \frac{O_{ij}}{\sum_j O_{ij}} = \frac{P_j / {D_{ij}}^n}{\sum_j P_j / {D_{ij}}^n}$$

Introduction

Le modèle de Huff : une approche probabiliste

A noter que le poids d'un centre d'achats est souvent, en première approximation, évalué en se fondant sur la superficie commerciale du centre d'achats.

Temps de déplacement			Hypermarchés	Surfaces des rayons		
				Alimentation $\lambda = 1$	Vestimentaire $\lambda = 2$	Mobilier $\lambda = 3$
TA	TB	TC				
15'	30'	10'	H1 2 500 m²	1 500 m²	500 m²	500 m²
10'	10'	15'	H2 2 000 m²	1 000 m²	400 m²	600 m²
20'	10'	18'	H3 2 300 m²	1 300 m²	700 m²	300 m²

Introduction

Le modèle de Huff : une approche probabiliste

Temps de déplacement			Hypermarchés	Surfaces des rayons		
				Alimentation $\lambda = 1$	Vestimentaire $\lambda = 2$	Mobilier $\lambda = 3$
TA	TB	TC				
15'	30'	10'	H1 2 500 m²	1 500 m²	500 m²	500 m²
10'	10'	15'	H2 2 000 m²	1 000 m²	400 m²	600 m²
20'	10'	18'	H3 2 300 m²	1 300 m²	700 m²	300 m²

$$P_{H1A} = \frac{1500/15}{1500/15 + 1000/10 + 1300/20} = 0,377$$

Introduction

Le modèle de Huff : une approche probabiliste

	A	B	C	A (1500)	B (2000)	C (1600)	Total
H1	0.377	0.178	0.520	565	356	832	1753
H2	0.377	0.357	0.229	565	714	366	1645
H3	0.245	0.464	0.250	367	928	400	1695

Les résultats pour l'alimentation en tenant compte du nombre de clients dans chaque ville.

Introduction

Les modèles d'optimisation de l'espace

Derrière la très riche variété des paysages concrets que nous observons à la surface du globe, il existe des organisations structurantes résultant du désir des habitants ou de leurs dirigeants d'aménager le sol pour leur plus grande commodité.

Ce processus d'adaptation à un milieu donné peut s'apparenter à un principe d'optimisation. En effet, les sociétés cherchent très souvent à maximiser une fonction d'utilité qui varie en fonction des lieux et des périodes.

Cette fonction d'utilité peut être purement économique (maximisation économique de l'utilisation du sol) ou plus complexe (maximisation du bien-être individuel ou social). Ainsi, il n'est pas si surprenant qu'un des premiers modèles d'analyse spatiale, le modèle de « Von Thünen », repose sur un principe de maximisation : celui de la maximisation des rentes foncières.

Introduction

Les modèles d'optimisation de l'espace

Von Thünen (1826) se donne comme objectif d'expliquer la localisation des activités agricoles. Selon lui, les activités agricoles ne sont pas disposées de manière aléatoire dans l'espace, mais répondent à une distribution spécifique.

Dans sa théorie, Von Thünen a repris de l'idée développée par Adam Smith, selon laquelle les producteurs cherchent avant toute chose à maximiser le profit de leur terre.

Von Thünen, lui-même propriétaire terrien, savait qu'un tel profit repose sur l'utilisation des surfaces agricoles et sur les coûts de transport des produits agricoles.

$$R_i(d) = p_i - c_i - T_i \times d = b_i - T_i \times d$$

Introduction

Les modèles d'optimisation de l'espace

	Cultures maraîchères	Maïs	Elevage
Prix de vente (Pi)	3000	1000	1500
Coûts de production (Ci)	1800	660	850
Coûts de transport au km (Ti)	25	1.5	4

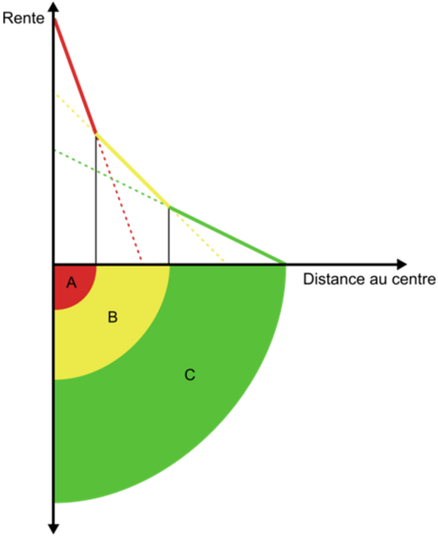
Introduction

Les modèles d'optimisation de l'espace

	Cultures maraîchères	Maïs	Elevage
Bénéfice brut (ordonnée à l'origine)	3000 – 1800 = 1200	1000 – 660 = 340	1500 – 850 = 650
Distance maximale de rentabilité	1200 / 25 = 48	340 / 1.5 = 227	650 / 4 = 162.5
Pente de la rente	- 25	- 1.5	- 4

Introduction

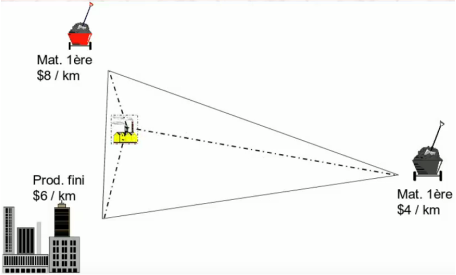
Les modèles d'optimisation de l'espace



Introduction

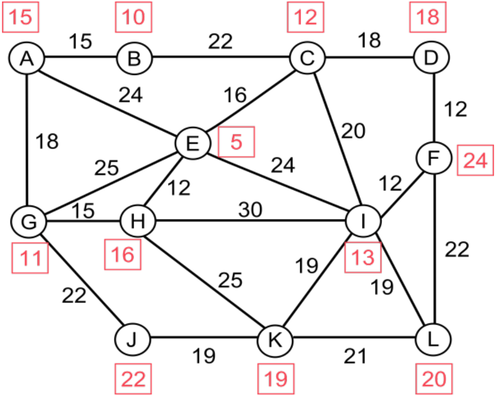
Les modèles d'optimisation de l'espace

Tandis que le modèle de Von Thunen détermine la meilleure production pour une localisation donnée, Werber se questionne sur la meilleure localisation pour une industrie donnée (un produit donné). Il fonde sa réflexion sur les coûts de transport.



Introduction

Les modèles d'optimisation de l'espace : le problème p-median



Où dois-je implanter mon magasin pour minimiser la distance parcourue par l'ensemble de mes futurs clients potentiels ? Qu'en est-il si je peux implanter 2 magasins ? 3 ? 4 ?

L’analyse de semis de points

Présentation

Cette partie traite de la répartition d’ensemble de lieux qui correspondent aux différentes localisations d’un phénomène.

Ces lieux peuvent être des habitations, des commerces, des personnes, des clients...

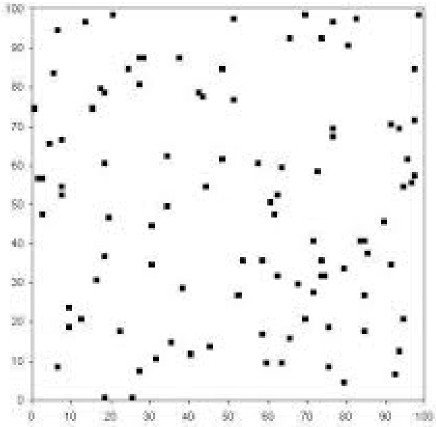
Ces lieux peuvent être traités comme des points à un certain degré de généralisation. On parlera donc de semis de points.

Pour comparer des semis de points ou pour mettre en exergue certaines de leurs spécificités, on va être amené à étudier leur forme.

Dans ces analyses, l’espace est souvent considéré comme homogène.

L'analyse de semis de points

Présentation



L'analyse de semis de points

Identifier le centre d'un semis de points : la position moyenne

Le point moyen :

$$\bar{X} = \frac{1}{N} \times \sum_{i=1}^{N} X_i \text{ et } \bar{Y} = \frac{1}{N} \times \sum_{i=1}^{N} Y_i$$

Le point moyen pondéré :

$$\bar{X_p} = \frac{\sum_{i=1}^{N}(P_i \times X_i)}{\sum_{i=1}^{N} P_i} \text{ et } \bar{Y_p} = \frac{\sum_{i=1}^{N}(P_i \times Y_i)}{\sum_{i=1}^{N} P_i}$$

Le point médian :

Le point médian M est le point le plus accessible, c'est-à-dire celui qui minimise la somme des distances à l'ensemble de tous les points.

L'analyse de semis de points

Identifier le centre d'un semis de points : la position moyenne

D'après le tableau ci-dessous, quel est le centre de ce semis de points ?

i	Xi	Yi	Pi
1	10	40	500
2	60	10	200
3	70	50	100
4	80	30	100
5	90	40	100

L'analyse de semis de points

Identifier le centre d'un semis de points : la position moyenne

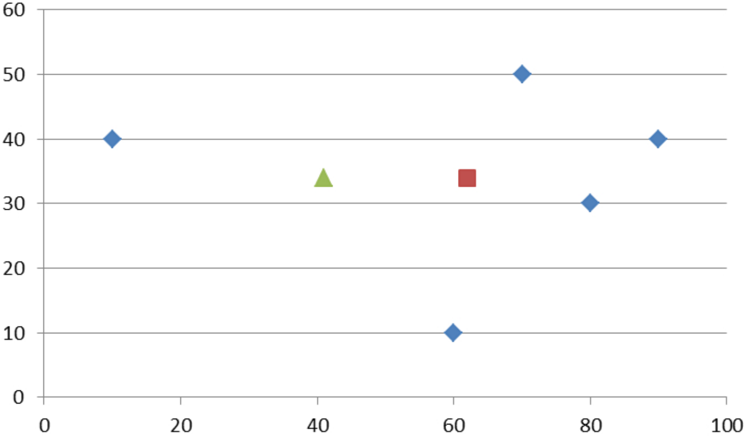
i	Xi	Yi	Pi
1	10	40	500
2	60	10	200
3	70	50	100
4	80	30	100
5	90	40	100

Le point moyen a donc pour coordonnées : $\bar{X} = 62$ et $\bar{Y} = 34$

Le point moyen pondéré a donc pour coordonnées : $\bar{X_p} = 41$ et $\bar{Y_p} = 34$

L'analyse de semis de points

Identifier le centre d'un semis de points : la position moyenne



L'analyse de semis de points

Mesurer la dispersion d'un semis de points

Ayant déterminé le point moyen, on peut chercher à mesurer la dispersion des lieux autour de ce point central. On parle de distance-type :

$$\sigma_D = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(X_i - \bar{X})^2 + (Y_i - \bar{Y})^2} = \sqrt{\sigma_X^2 + \sigma_Y^2}$$

Dans l'exemple précédent, la distance type est de 31.

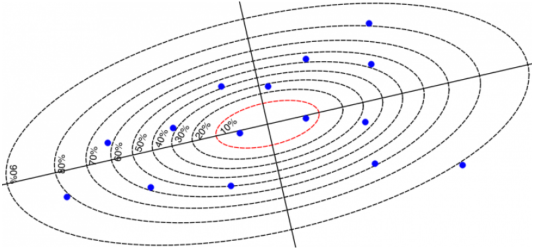
On pourra alors tracer un cercle du rayon de la distance-type centré sur le point moyen et ainsi obtenir un cercle de dispersion. On retrouvera à l'intérieur de ce cercle une proportion de points pouvant être considérés comme plutôt centraux.

L'analyse de semis de points

Ellipse de dispersion, centrographie

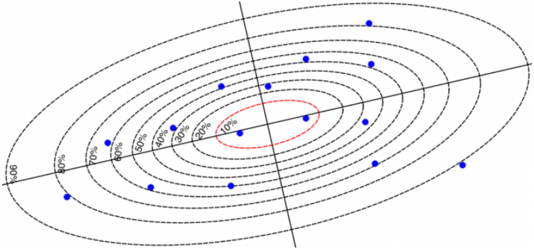
On parle de centrographie, d'analyse centrographique et dans les faits on préférera calculer des ellipses de dispersion plutôt que des cercles.

L'ellipse de dispersion est un outil statistique utilisé en analyse spatiale pour représenter la dispersion d'un ensemble de points autour d'un centre. Elle permet de visualiser à la fois la direction et l'étendue de cette dispersion.



L'analyse de semis de points

Ellipse de dispersion, centrographie



La règle des 3 sigmas de proportion des points dans les ellipses de dispersion dépend des méthodes de calculs retenues : 39,35%, 84,47% et 98,89% ou 68,27%, 95,45% et 99.73%.

Si l'ellipse calculée est proche d'un cercle (excentricité $= 1$), alors la répartition des points semble caractéristique d'un espace isotrope.

L'analyse de semis de points

Mesurer la concentration d'un semis de points

Une distribution est aléatoire si :

1. Tous les emplacements de l'espace ont la même probabilité d'accueillir un point.
2. La position d'un nouveau point est indépendante de la position des points précédents.

Une distribution aura tendance à être concentrée si :

1. Certains emplacements de l'espace ont plus de chances d'accueillir un point.
2. La localisation d'un premier point favorise l'apparition d'autres points à proximité.

L'analyse de semis de points

Mesurer la concentration d'un semis de points

Une distribution aura tendance à être régulière si :

1. Tous les emplacements de l'espace ont la même probabilité d'accueillir un point
2. La localisation d'un premier point défavorise l'apparition d'autres points à proximité.

L'analyse de semis de points

Mesurer la concentration d'un semis de points

La méthode des quadrats permet de mesurer des concentrations (des densités) dans un semis de points :

1. Soit un semis de N points distribués sur un espace E.
2. On recouvre l'espace E d'un ensemble de K mailles d'une forme régulière (carré, rectangle, cercle).
3. Le nombre moyen de points théorique par maille est égale à $D=N/K$.
4. On associe à chaque maille i le nombre de points qu'elle contient, puis on calcule la variance du nombre de points par maille V(D) et on en déduit un indice de concentration(Ic). $IC=V(D)/D$.

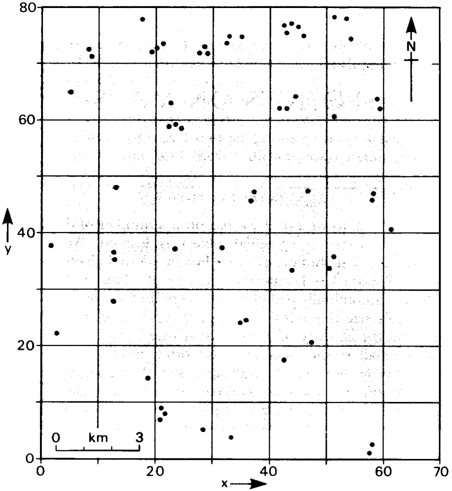
Si $IC=1$, la distribution est aléatoire.

$IC>1$, la distribution est plutôt concentrée.

$IC<1$, la distribution est plutôt régulière.

L'analyse de semis de points

Mesurer la concentration d'un semis de points



L'analyse de semis de points

Mesurer la concentration d'un semis de points

Nombre de points n	Nombre de quadrats K	Nombre de points n.K
0	25	0
1	15	15
2	8	16
3	5	15
4	1	4
5	2	10
Total	56	60

L'analyse de semis de points

Mesurer la concentration d'un semis de points

Nombre de points **n**	Nombre de quadrats **K**	Nombre de points **n.K**	Ecarts à la moyenne **(n-D)**	**K(n - D)²**
0	25	0	-1.071	28.676
1	15	15	-0.071	0.076
2	8	16	0.929	6.904
3	5	15	1.929	18.605
4	1	4	2.929	8.579
5	2	10	3.929	30.874
Total	56	60		93.714

Densité moyenne **D** = nb. de points / nb. de quadrats = 60/56 **= 1.071**
Variance **V(D)** = 93.714 / 55 = **1.704**
Indice de concentration IC = V(D)/D= **1.590**

L'analyse de semis de points

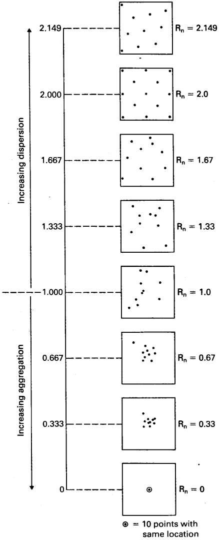
Mesurer la forme d'un semis de points

La méthode du plus proche voisin permet aussi d'étudier la dispersion, donc la forme d'un semis de points.

1. Soit un semis de N points distribués sur un espace de surface S. On note D la densité moyenne de points par unité de surface (D=N/S).
2. On calcule pour chaque point i la distance Dmin(i) qui le sépare de son voisin le plus proche.
3. On calcule ensuite la moyenne des distances observées au plus proche voisin D0.
4. On détermine la distance théorique moyenne au plus proche voisin DT dans le cas d'une distribution aléatoire (DT=0.5/$\sqrt{D}$).
5. On calcule l'indice de dispersion qui est le rapport : R=D0/DT.

L'analyse de semis de points

Mesurer la forme d'un semis de points



L'analyse de semis de points

Mesurer la forme d'un semis de points

i	Xi	Yi
1	1,5	7
2	1	7
3	1,5	6,8
4	0,5	5,8
5	2,2	7,5
6	0,3	7
7	0,6	4,8
8	1,8	4,1
9	2,1	5,2
10	4,3	5,8
11	1,6	7,2
12	3,1	6,4
13	0,7	2,9
14	0,1	2,6
15	1,5	4,4
16	3,1	5,3
17	5,2	6,2
18	5,1	7,9
19	1,7	1
20	2,4	1,8
21	4,2	5
22	7	6,1
23	6,8	3,8
24	7,2	0,3

L'analyse de semis de points

Mesurer la forme d'un semis de points

i	dmin
1	0,2
2	0,5
3	0,2
4	1,0
5	0,7
6	0,7
7	1,0
8	0,4
9	1,0
10	0,8
11	0,2
12	1,1
13	0,7
14	0,7
15	0,4
16	1,0
17	1,0
18	1,7
19	1,1
20	1,1
21	0,8
22	1,8
23	2,3
24	3,5

L'analyse de semis de points

Mesurer la forme d'un semis de points

$\mathrm{D0} = 0.99$

Comme la surface est égale 64 et l'effectif est égal à 24, on obtient une densité de 0,375 et par conséquent $\mathrm{DT} = 0.816$

$\mathrm{R} = \mathrm{D0}/\mathrm{DT} = 1{,}22$

L'analyse de semis de points

Exercice

L'objectif de cet exercice est d'analyser l'évolution des emplacements des magasins IKEA en France. En effet, depuis l'ouverture du premier magasin dans les années 1980 en région parisienne, de nombreux autres magasins ont ouvert leur porte dans toute la France. Quelle est la stratégie d'IKEA ? Elargir au maximum ses implantations dans toute la France ? L'entreprise IKEA est-elle plutôt rentrée dans une phase de concentration ?

1. Etudiez l'évolution du point moyen des emplacements des magasins IKEA.
2. Etudiez l'évolution de la distance-type des emplacements des magasins IKEA.
3. Etudiez la dispersion des emplacements des magasins IKEA.

L'autocorrélation spatiale

Présentation

Compte tenu de la répartition inégalitaire de certaines variables, des objets géographiques se ressemblent plus que d'autres. Une question d'analyse spatiale que l'on est alors en droit de se poser est la suivante :

Est-ce que les objets géographiques qui sont proches se ressemblent plus que les objets géographiques qui sont éloignés ? C'est la question de l'autocorrélation spatiale.

Mesurer l'autocorrélation spatiale d'un phénomène (d'une distribution) revient à déterminer s'il semble exister une organisation spatiale sous-jacente à ce phénomène (à cette distribution) et donc qu'il (qu'elle) ne se repartit pas de façon aléatoire au sein du territoire étudié.

Par exemple : Les personnes riches se regroupent-elles ? Les communes très peuplées côtoient-elles des communes très peu peuplées ?

L'autocorrélation spatiale

Présentation

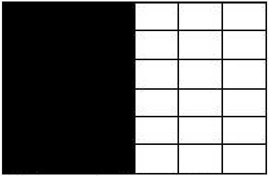
L'autocorrélation spatiale est positive si les lieux proches ont tendance à se ressembler davantage que les lieux éloignés.

Elle est négative si les lieux proches ont tendance à être plus différents que les lieux éloignés.

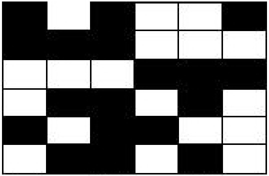
L'autocorrélation est nulle quand aucune relation n'existe entre la proximité des lieux et leur degré de ressemblance.

L'autocorrélation spatiale

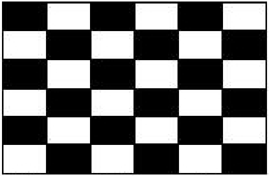
Présentation



Autocorrélation spatiale positive



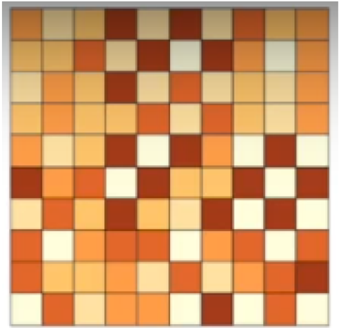
Autocorrélation spatiale nulle

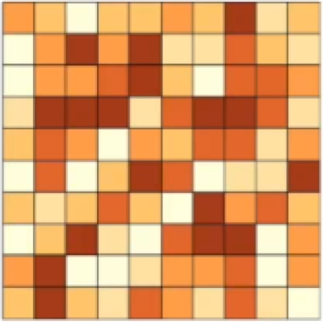


Autocorrélation spatiale négative

L'autocorrélation spatiale

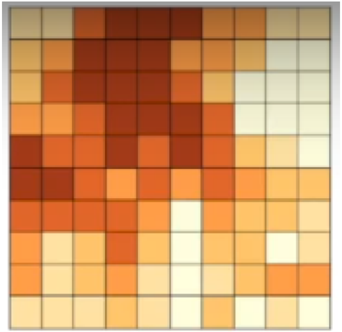
Présentation : aléatoire ou pas

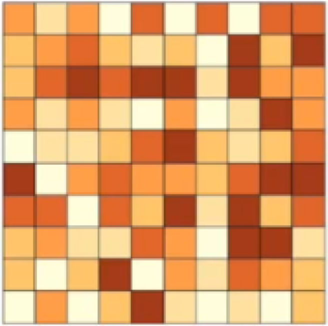




L'autocorrélation spatiale

Présentation : aléatoire ou pas





L'autocorrélation spatiale

Présentation statistique

Les coefficients d'autocorrélation spatiale sont alors construits statistiquement de telle manière qu'il soit possible de répondre à la question suivante :

La variation d'un caractère entre unités contigües (proches) est-elle plus ou moins grande que la variation de ce même caractère pour l'ensemble du territoire ou plus précisément entre unités non-contigües (éloignées) ?

Il convient dès lors de définir ce qui est proche, de définir ce qui est voisin. Le plus simple est de le déterminer de manière binaire en s'appuyant par exemple sur la notion de contiguïté.

Il existe plusieurs indicateurs pour mesurer l'autocorrélation spatiale. Les deux principaux, c'est-à-dire les plus couramment utilisés, sont les indices de Moran et de Geary.

L'autocorrélation spatiale

Les indices de Moran et Geary

$$G = \frac{N-1}{2L} \times \frac{\sum_{i,j} l_{ij} \times (X_i - X_j)^2}{\sum_i (X_i - \bar{X})^2}$$

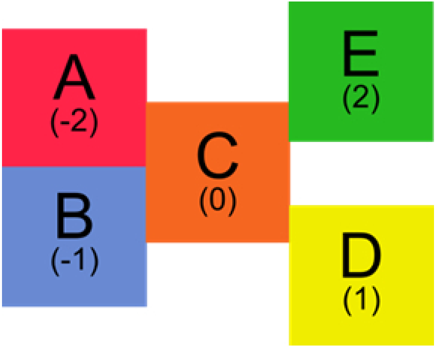
Les valeurs de l'indice de Geary s'étendent de 0 à 2. La valeur 1 signifie qu'aucune autocorrélation spatiale n'est présente dans les mesures effectuées. Une valeur plus petite que 1 correspond à une autocorrélation spatiale positive.

$$M = \frac{N}{L} \times \frac{\sum_{i,j} l_{ij} \times (X_i - \bar{X}) \times (X_j - \bar{X})}{\sum_i (X_i - \bar{X})^2}$$

Les valeurs de l'indice de Moran s'étendent de -1 (corrélation négative) à $+1$ (corrélation positive). Une valeur nulle correspond à un modèle spatial parfaitement aléatoire.

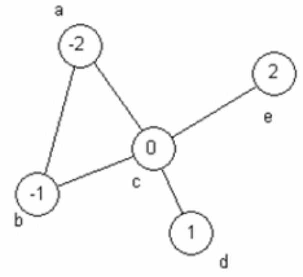
L'autocorrélation spatiale

Exemple



L'autocorrélation spatiale

Exemple



L'autocorrélation spatiale

Exemple

$$\bar{X} = \frac{(-2-1+0+2+1)}{5} = 0$$

$$\sum_i (X_i - \bar{X})^2 = (-2)^2 + (-1)^2 + (0)^2 + (1)^2 + (2)^2 = 10$$

$$N = 5 \text{ et } L = 10$$

L'autocorrélation spatiale

Exemple

$A \rightarrow B$	$(-2-(-1))^2 = 1$	$A \rightarrow C$	$(-2-0)^2 = 4$	$B \rightarrow A$	$(-1-(-2))^2 = 1$
$B \rightarrow C$	$(-1-0)^2 = 1$	$C \rightarrow A$	$(0-(-2))^2 = 4$	$C \rightarrow B$	$(0-(-1))^2 = 1$
$C \rightarrow D$	$(0-1)^2 = 1$	$C \rightarrow E$	$(0-2)^2 = 4$	$D \rightarrow C$	$(1-0)^2 = 1$
$E \rightarrow$	$(2-0)^2 = 4$	**Total**		*22*	

$$G = \frac{(5-1)\times 22}{2\times 10\times 10} = 0,44$$

L'autocorrélation spatiale

Exercice

En région parisienne, les prix de l'immobilier sont en partie conditionnés par l'accessibilité à Paris. Ainsi, la proximité (la contiguïté) vis-à-vis de Paris fait grimper les prix de l'immobilier.

C'est pourquoi, on peut s'attendre à ce que les prix de l'immobilier en région parisienne soient autocorrélés spatialement. Cette assertion est-elle vraie ?

Exercice : Calculez l'indice de Geary pour déterminer s'il y a autocorrélation spatiale.

Le quotient de localisation

Présentation

Le quotient de localisation est un indicateur de « concentration », de spécialisation.

Il donne une mesure de l'importance relative d'une modalité (d'une valeur) dans une unité spatiale comparée à son poids dans les autres unités spatiales.

Le quotient de localisation est un outil d'analyse spatiale, car il permet de caractériser le degré de concentration d'une sous-population dans une unité spatiale en le comparant à toutes les autres unités spatiales d'un même ensemble territorial.

Il permet de mener cette comparaison en faisant abstraction des inégalités de poids entre les unités spatiales et les différentes catégories.

Le quotient de localisation

Présentation : tableau de contingence

Candidats	Agriculteurs	Artisans, Commerçants et chefs d'entreprise	Professions libérales, Cadres Supérieurs	Professions intermédiaires	Employés	Ouvriers	Étudiants	Chômeurs	Total
Schivardi	0	0	0	0	12	10	0	0	21*
Laguiller	0	0	0	9	23	29	0	4	65
Besancenot	0	0	4	53	58	78	24	32	249
Buffet	5	0	4	9	23	20	3	24	87
Bové	5	0	7	18	12	10	8	4	64
Royal	14	40	110	276	289	205	85	128	1148
Voynet	4	0	7	36	12	10	5	4	77
Nihous	0	0	0	9	12	20	0	4	44
Bayrou	32	64	103	178	185	157	59	88	865
Sarkozy	64	117	103	231	335	205	56	76	1189
Villiers	20	5	7	18	35	10	5	0	100
Le Pen	34	40	11	53	162	225	21	36	582
Total	178	267	356	889	1156	978	267	400	4492

* Avertissement : Le tableau donne les effectifs de vote aux dix millièmes (4492 au lieu de 44.920.000 individus). Les effectifs sont exprimés sans aucune décimale ce qui conduit à des approximations quant aux calculs des effectifs marginaux. Par exemple, le nombre de votes pour le candidat Schivardi a été estimé à 21 (soit 210.000) électeurs et non à 22 (soit 22.000) électeurs (12+10). Ce constat est généralisable à l'ensemble des tableaux de résultats. Cette approximation n'interfère en aucun cas sur le résultat de l'AFC.

Le quotient de localisation

Présentation

	C1	C2	C3	Total
B1	x11	x12	x13	x1●
B2	x21	x22	x23	x2●
B3	x31	x32	x33	x3●
Total	x●1	x●2	x●3	x●●

$$Q(x[ij]) = (x[ij]/x[.j])/(x[i.]/x[..]) = (x[ij] \times x[..])/(x[i.] \times x[.j])$$

Le quotient de localisation

Exemple

Branche / Zone	B1	B2	B3	Total
Z1	48	325	287	660
Z2	27	185	148	360
Z3	45	90	45	180
Total	120	600	480	1200

Le quotient de localisation

Exemple

Branche / Zone	B1	B2	B3	Total
Z1	48	325	287	660
Z2	27	185	148	360
Z3	45	90	45	180
Total	120	600	480	1200

$$Q(X[z2b1]) = \frac{(27/120)}{(360/1200)} = \frac{0{,}225}{0{,}300} = 0,75$$

Le quotient de localisation

Exemple

Branche / Zone	B1	B2	B3
Z1	0,727	0,985	1,087
Z2	0,750	1,028	1,028
Z3	2,500	1,000	0,625

Le quotient de localisation

Compréhension

Effectifs observés (Nij)									
1963	*ALIM*	*TEXT*	*BOIS*	*EDIT*	*CHIM*	*CONS*	*META*	*EQUIP*	**Total**
BULGARIE	130	128	39	14	29	47	21	151	**559**
HONGRIE	144	241	53	28	77	61	91	423	**1118**
POLOGNE	380	612	164	84	222	199	147	881	**2689**
R.D.A.	206	451	119	118	308	142	109	1056	**2509**
ROUMANIE	136	305	244	41	76	114	106	366	**1388**
TCHECO.	185	412	130	63	139	151	177	883	**2140**
YOUGOSL.	126	223	132	58	76	78	69	307	**1069**
Total	**1307**	**2372**	**881**	**406**	**927**	**792**	**720**	**4067**	**11472**

Le quotient de localisation

Compréhension

Effectifs observés (Nij)									
1963	*ALIM*	*TEXT*	*BOIS*	*EDIT*	*CHIM*	*CONS*	*META*	*EQUIP*	**Total**
BULGARIE	130	128	39	14	29	47	21	151	**559**
HONGRIE	144	241	53	28	77	61	91	423	**1118**
POLOGNE	380	612	164	84	222	199	147	881	**2689**
R.D.A.	206	451	119	118	308	142	109	1056	**2509**
ROUMANIE	136	305	244	41	76	114	106	366	**1388**
TCHECO.	185	412	130	63	139	151	177	883	**2140**
YOUGOSL.	126	223	132	58	76	78	69	307	**1069**
Total	**1307**	**2372**	**881**	**406**	**927**	**792**	**720**	**4067**	**11472**

Avec un tableau de contingence, on peut obtenir la valeur totale des effectifs concernés (par exemple un nombre de personnes). Par exemple : E = 11472.

On peut aussi obtenir la taille d'une catégorie vis-à-vis des autres. Par exemple : ALIM = 1307 / 11472 = 0.11

On peut aussi obtenir la taille de chaque entité géographique vis-à-vis des autres : BULGARIE = 559 / 11472 = 0.05

Le quotient de localisation

Compréhension

Si l'on multiplie l'ensemble de ces valeurs, on obtient une modalité théorique, qui correspond à ce que l'on pourrait attendre en cas d'une répartition égalitaire des modalités en fonction de la taille de l'entité géographique et de la variable.

Par exemple : $11472 \times 0.11 \times 0.05 = 63.096$

La comparaison entre la valeur réelle (130) et la valeur théorique (63), nous permet de dire s'il y a sur-représention ou sous-représentation d'une variable au sein d'une entité. On retrouve le coefficient de localisation ou coefficient de spécialisation : $130 \ / \ 63 = 2.063$

$$\hat{X}_{ij} = X_{..} \times \frac{X_{i.}}{X_{..}} \times \frac{X_{.j}}{X_{..}} = \frac{X_{i.} \times X_{.j}}{X_{..}} \Rightarrow \frac{X_{ij}}{\hat{X}_{ij}} = \frac{X_{ij} \times X_{..}}{X_{i.} \times X_{.j}} = Q(X_{ij})$$

Le quotient de localisation

Deux profils possibles : le profil en ligne

Profils en ligne (Nij/Ni.)									
1963	*ALIM*	*TEXT*	*BOIS*	*EDIT*	*CHIM*	*CONS*	*META*	*EQUIP*	**Total**
BULGARIE	23%	23%	7%	3%	5%	8%	4%	27%	**100%**
HONGRIE	13%	22%	5%	3%	7%	5%	8%	38%	**100%**
POLOGNE	14%	23%	6%	3%	8%	7%	5%	33%	**100%**
R.D.A.	8%	18%	5%	5%	12%	6%	4%	42%	**100%**
ROUMANIE	10%	22%	18%	3%	5%	8%	8%	26%	**100%**
TCHECO.	9%	19%	6%	3%	6%	7%	8%	41%	**100%**
YOUGOSL.	12%	21%	12%	5%	7%	7%	6%	29%	**100%**
Total	**11%**	**21%**	**8%**	**4%**	**8%**	**7%**	**6%**	**35%**	**100%**

On appelle INDICE DE SPECIALISATION (Si) l'écart entre le profil d'une unité spatiale et le profil général de l'ensemble de référence.

$$S_i = \sum_{j=1}^{n} \left| \frac{N_{ij}}{N_{i.}} - \frac{N_{.j}}{N_{..}} \right|$$

Le quotient de localisation

Deux profils possibles : le profil en ligne

Profils en ligne (Nij/Ni.)									
1963	*ALIM*	*TEXT*	*BOIS*	*EDIT*	*CHIM*	*CONS*	*META*	*EQUIP*	**Total**
BULGARIE	23%	23%	7%	3%	5%	8%	4%	27%	**100%**
HONGRIE	13%	22%	5%	3%	7%	5%	8%	38%	**100%**
POLOGNE	14%	23%	6%	3%	8%	7%	5%	33%	**100%**
R.D.A.	8%	18%	5%	5%	12%	6%	4%	42%	**100%**
ROUMANIE	10%	22%	18%	3%	5%	8%	8%	26%	**100%**
TCHECO.	9%	19%	6%	3%	6%	7%	8%	41%	**100%**
YOUGOSL.	12%	21%	12%	5%	7%	7%	6%	29%	**100%**
Total	**11%**	**21%**	**8%**	**4%**	**8%**	**7%**	**6%**	**35%**	**100%**

$S_{(BULGARIE)} = \mid 0.23 - 0.11 \mid + \mid 0.23 - 0.21 \mid + \mid 0.07 - 0.08 \mid + \mid 0.03 - 0.04 \mid + \mid 0.05 - 0.08 \mid + \mid 0.08 - 0.07 \mid + \mid 0.04 - 0.06 \mid + \mid 0.03 - 0.04 \mid + \mid 0.27 - 0.35 \mid = 0.30$

Le quotient de localisation

Deux profils possibles : le profil en ligne

BULGARIE	30%
HONGRIE	15%
POLOGNE	11%
R.D.A.	24%
ROUMANIE	28%
TCHECO.	17%
YOUGOSL.	13%

L'indice de spécialisation est pertinent d'un point de vue cartographique.

Le quotient de localisation

Deux profils possibles : le profil en colonne

Profils en colonne (Nij/N.j)									
1963	ALIM	TEXT	BOIS	EDIT	CHIM	CONS	META	EQUIP	Total
BULGARIE	10%	5%	4%	3%	3%	6%	3%	4%	5%
HONGRIE	11%	10%	6%	7%	8%	8%	13%	10%	10%
POLOGNE	29%	26%	19%	21%	24%	25%	20%	22%	23%
R.D.A.	16%	19%	14%	29%	33%	18%	15%	26%	22%
ROUMANIE	10%	13%	28%	10%	8%	14%	15%	9%	12%
TCHECO.	14%	17%	15%	16%	15%	19%	25%	22%	19%
YOUGOSL.	10%	9%	15%	14%	8%	10%	10%	8%	9%
Total	100%	100%	100%	100%	100%	100%	100%	100%	100%

On appelle INDICE DE LOCALISATION (Lj) l'écart entre le profil d'une catégorie et le profil général de l'ensemble de référence.

$$L_j = \sum_{j=1}^{n} \mid \frac{N_{ij}}{N_{.j}} - \frac{N_{i.}}{N_{..}} \mid$$

Le quotient de localisation

Deux profils possibles : le profil en colonne

Profils en colonne (Nij/N.j)									
1963	ALIM	TEXT	BOIS	EDIT	CHIM	CONS	META	EQUIP	Total
BULGARIE	10%	5%	4%	3%	3%	6%	3%	4%	5%
HONGRIE	11%	10%	6%	7%	8%	8%	13%	10%	10%
POLOGNE	29%	26%	19%	21%	24%	25%	20%	22%	23%
R.D.A.	16%	19%	14%	29%	33%	18%	15%	26%	22%
ROUMANIE	10%	13%	28%	10%	8%	14%	15%	9%	12%
TCHECO.	14%	17%	15%	16%	15%	19%	25%	22%	19%
YOUGOSL.	10%	9%	15%	14%	8%	10%	10%	8%	9%
Total	100%	100%	100%	100%	100%	100%	100%	100%	100%

$$L_{(ALIM)} = \mid 0.10 - 0.05 \mid + \mid 0.11 - 0.10 \mid + \mid 0.29 - 0.23 \mid + \mid 0.16 - 0.22 \mid + \mid 0.10 - 0.12 \mid + \mid 0.14 - 0.19 \mid + \mid 0.10 - 0.09 \mid = 0.26$$

Le quotient de localisation

Deux profils possibles : le profil en colonne

ALIM	TEXT	BOIS	EDIT	CHIM	CONS	META	EQUIP
26%	9%	43%	24%	25%	12%	25%	13%

L'indice de localisation n'est pas pertinent d'un point de vue cartographique, mais a du sens d'un point de vue géographique.

Le quotient de localisation

Exercice

Vous disposez d'un fichier Excel qui contient le nombre d'habitants appartenant à une catégorie socioprofessionnelle donnée pour chaque département français. On est alors en droit de se poser des questions du type : Quelles sont les « spécialités »de chacun des départements ? Le département du Nord est-il un département « ouvrier » ? Paris est-elle une ville de cadres ?

1. Pour chaque département, calculez les quotients de localisation pour les ouvriers et pour les cadres.
2. A l'aide d'un profil en ligne, déterminez le département français le plus "spécialisé".
3. A l'aide d'un profil en colonne, déterminez la CSP la plus "localisée".