**L’épidémie de choléra à Londres septembre 1854 : l’apport de l’analyse spatiale**

**Sommaire**

[Objectif 1](#_Toc211441981)

[Préparation de l’espace de travail 1](#_Toc211441982)

[1. Localiser et décrire pour poser une hypothèse 1](#_Toc211441983)

[2. Cartographier des morts de choléra en septembre 1854 : réponse à la question « Où et comment ? » 3](#_Toc211441987)

[3. Repérer de la structure spatiale des morts du choléra : réponse à la question « Comment et pourquoi ? » 5](#_Toc211441988)

[4. Expliquer et fournir une aide à la décision 9](#_Toc211441993)

[Annexe 1 : Créer un jeu de données vecteur de géométrie ponctuel à partir de coordonnées XY 10](#_Toc211441994)

[Annexe 2 : Cartographier une variable quantitative de stock 11](#_Toc211441995)

[Annexe 3 : Les polygones de Thiessen 12](#_Toc211441996)

# Objectif

Mobiliser des outils d’analyse spatiale pour comprendre la diffusion d’une épidémie. Il s’agit de suivre la même démarche que le médecin John Snow a entreprise il y a 160 ans pour comprendre les causes de la diffusion du choléra à Londres, mais avec l'aide de moyens et méthodes modernes. On considère souvent l’analyse de Dr Snow comme la première véritable analyse spatiale.

*Inspiré de l’analyse de Dr Snow https://www.gislounge.com/john-snows-cholera-map-gis-data/*

*A voir également : http://www.theguardian.com/news/datablog/2013/mar/15/john-snow-cholera-map*

# Préparation de l’espace de travail

1. Pensez à nommer fichiers, dossiers, champs, attributs avec des noms explicites, courts et surtout sans espaces et caractères particuliers (ex. accents, %, \*). D’une manière générale, le SEUL caractère spécial utilisable est l’underscore (« \_ »). Exemple :

- Nom d’un dossier correct : Data\_Cholera

- Nom d’un dossier à éviter : Données Choléra

1. Dans QGIS ouvrez un projet vierge que vous nommez DrSnow\_votreNom. N’ouvrez pas encore de fichiers. Modifiez le système de coordonnées du projet en *British\_National\_Grid* (EPSG 27700) via le menu Projet > Propriétés du projet. Les données que nous allons utiliser étant localisés à Londres, ce système de coordonnées est adapté.
2. Rappel : pour faciliter la navigation dans l’arborescence de vos fichiers, QGIS propose de mettre certains dossiers en favoris pour un accès rapide aux fichiers. Créez une marque-pages pour votre dossier (dans la fenêtre Explorateur : clic droit sur le dossier).
3. Télécharger les données à utiliser à partir du moodle du cours. Le dossier Data\_Origine contient :

* Deaths.csv - la localisation des morts de choléra pendant l’épidémie de septembre 1854
* Pumps.csv - la localisation des pompes à eau dans le quartier de Soho en 1854
* Snow\_1854.tif – raster géoréférencé à partir de la carte de Soho fait par John Snow en 1854
* ReseauRuesSoho.shp – réseau de rue digitalisé à partir de la carte de Soho fait par John Snow

1. Dans le dossier DrSnow\_cholera\_Sentinelles créez un Geopackage que vous nommez DrSnow. Vous enregistrerez l’ensemble des fichiers produites pendant la séance dans le DrSnow.gpkg.

# Localiser et décrire pour poser une hypothèse

*Au milieu du 19ème siècle, il y a eu une grande épidémie de choléra en Europe. Le choléra est une maladie infectieuse, due à une bactérie, dont on n'a pas su pendant longtemps comment elle se transmettait. Lors d'une épidémie de choléra dans le centre de Londres, le médecin John Snow a cherché la source de cette bactérie (l’épidémie est partie du quartier de Soho où habitait Snow). Pour pouvoir faire des hypothèses sur la source de l’épidémie Dr Snow a dressé une carte représentant la répartition des morts dans le quartier de Soho qui était le plus atteint par l’épidémie.*

Vous allez commencer par cartographier des données spatiales décrivant le nombre de mort à chaque immeuble du quartier de Soho.

## Comprendre les données spatiales décrivant les morts de choléra

Ouvrez le fichier *Deaths.csv,* puis observez les données*.* Ce fichier contient un tableau décrivant des adresses dans Soho où des morts de choléra ont été observés en septembre 1854 par Dr Snow.

* Les colonnes sont composées d’un identifiant (Id) et trois variables décrivant chaque lieu :
  + le nombre de morts (NbDeaths)
  + les coordonnées X (longitudes) et Y (latitudes) qui permettent de localiser ces lieux dans Londres. Pour information ces coordonnées sont définies dans un système de projection qui s’appelle *British\_National\_Grid (EPSG 7405)*, qui est un système de coordonnées adapté pour afficher des données cartographiques en Grande Bretagne*.* Rappel :D’autres systèmes de coordonnées sont utilisés dans d’autres territoires (en France on utilise entre autres *Lambert-93*).

Vous ne disposez donc pas d’un jeu de données vecteur (par exemple, un fichier de forme, un *shapefile*) des entités de points représentant les adresses de morts de choléra.

Or, en disposant des coordonnées X et Y des adresses des morts de choléra il est possible d’afficher ces lieux sur la carte, puis d’exporter les données sous forme d’un jeu de données vecteur. C’est ce que vous allez faire en 1.2.

## Créer un jeu de données vecteur de géométrie ponctuel à partir de coordonnées XY

1. A partir du fichier *Deaths.csv* créez un jeu de données vecteur que vous enregistrez dans le geopackage DrSnow. Il s’agit pour commencer de visualiser les coordonnées sur une carte sous forme de points (on parle de projeter les coordonnées), puis les exporter dans le geopackage. Si nécessaire, reportez-vous à l’annexe 1 pour suivre les étapes.
2. Vérifiez le résultat en superposant la couche *Deaths* au scan géoréférencé de la carte conçue par John Snow : *snow\_1854.tif*. Sur cette carte, chaque trait correspond à un mort du choléra à cette adresse. Sauf erreur de votre part dans les étapes de création du jeu de données des adresses des morts de choléra les deux cartes se superposent.

## Cartographier des morts de choléra en septembre 1854 : réponse à la question « Où ? »

1. Quelle variable visuelle utiliserez-vous pour cartographier la répartition du nombre de morts à chaque adresse ? Pourquoi ?

1. Cartographiez la répartition du nombre de morts à chaque adresse. Si nécessaire reportez-vous à la démarche décrite en Annexe 2.

Pour la mise en forme :

* Veillez à ce que les cercles plus petits soient systématiquement placés au-dessus des plus grands.
* Pour améliorer le rendu graphique de vos cartes par cercles proportionnels pensez à détourer les cercles en blanc plutôt qu’en noir.
* Pour voir apparaître la légende, cliquez sur Avancé (Toujours dans la symbologie de la couche) > « Légende pour la Taille définie par des données », puis cliquez sur Légende repliée.

1. Combien de personnes sont mortes victimes de l'épidémie de choléra ? Quel outil utilisez-vous pour répondre à cette question ? Dans le tableau résultats, à quoi correspond *Compte* ?

**Pensez à enregistrer votre projet !**

# Cartographier des morts de choléra en septembre 1854 : réponse à la question « Où et comment ? »

Outre la cartographie thématique, il existe différentes méthodes qui permettent de visualiser les zones de concentration d’un phénomène pour lequel on dispose de données vectorielles (points, polygones ou lignes).

On peut notamment représenter la répartition des objets (ici les cas de choléra) sous forme d’une *surface continue* de données lissées. L’intérêt majeur de ce type de traitement spatial des données est de faciliter la lecture d’une distribution spatiale en mettant mieux en évidence les zones à forte concentration.

**Encadré 1 : Méthode de lissage**

Dans QGIS il existe différents outils ou extensions qui permettent d’effectuer un lissage.

L’outil *Carte de chaleur (Estimation par noyau)* (Kernel Density en anglais) permet d’effectuer ce lissage. Vous le trouverez dans la boîte à outils interpolation[[1]](#footnote-1).

Le lissage renforce la visibilité des zones de concentration et élimine des épiphénomènes. Il permet donc de souligner la structure spatiale du phénomène étudié.

Avec l’outil *Carte de chaleur* dans QGIS les valeurs affichées sur la carte résultat (en mode raster) sont calculées :

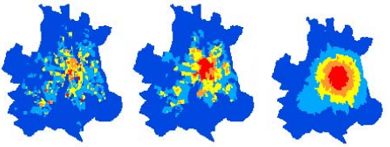
- soit en fonction du nombre de points (par exemple, le nombre de fermes touchées par la grippe aviaire) : si une ferme est entourée de nombreuses fermes touchées par la grippe aviaire, la cellule correspondant à cette ferme dans le raster résultat verra sa valeur renforcée. A l'inverse, une ferme touchée par la grippe aviaire isolée dans l’espace géographique (non entourée d’autres fermes touchées par la grippe aviaire) n’apparaîtra pas comme une ferme touchée par la grippe aviaire (la cellule du raster aura une valeur nulle ou très faible).

- soit en fonction d’un champ de la table attributaire (par exemple, le nombre de volailles touchées par la grippe aviaire dans une ferme). Dans ce cas, si une ferme avec un grand nombre de volailles touchées par la grippe aviaire est entourée de fermes également très fortement atteintes par le la grippe aviaire (nombre de volailles touchées est élevé), la cellule correspondant à cette ferme dans le raster, ainsi que les cellules voisines, auront une valeur renforcée.

Il est nécessaire de définir un rayon de lissage qui délimite le périmètre à l’intérieur duquel les entités sont utilisées pour calculer la valeur affectée à chaque cellule. Le choix du rayon influence le niveau de généralisation du résultat, plus le rayon de recherche est important, plus le lissage est important (exemple cartes ci-dessous) :

**Lissage de la population à Toulouse en 1999 avec 3 rayons de lissage différents**

50 m 150 m 1000 m



Source : Certu, 2005

1. Comme vous l’avez compris en lisant le texte ci-dessus, il est possible d’effectuer un lissage directement à partir de la localisation des points ou à partir d’un champ numérique décrivant les points. Est-ce que vous pouvez faire le lissage à partir d’un des champs qui décrit les entités du jeu de données *Deaths* ? Si oui, lequel et pourquoi ce choix peut être intéressant ?

1. Réalisez la carte lissée de la répartition des morts dans le quartier de Soho en paramétrant comme suit :

* Rayon : 50 m
* Weight from field : Deaths (pour effectuer le lissage sur le nombre de morts)
* Choisissez une taille des pixels de 1 m x 1 m (pour vérifier quelle est l’unité de mesure de la distance du projet en cours, regardez les propriétés du projet dans menu projet)
* Laissez les autres paramètres par défaut
* Pensez à enregistrer le résultat
  1. Quel est le mode de représentation des données produites ?
  2. Quelle est la résolution des données produites ?
  3. A quoi correspondent les valeurs des cellules du raster à votre avis ?

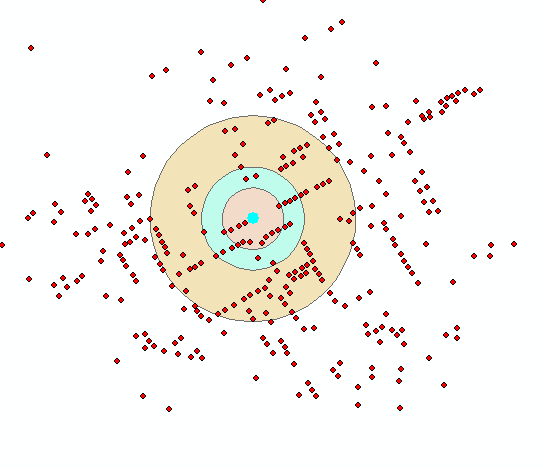
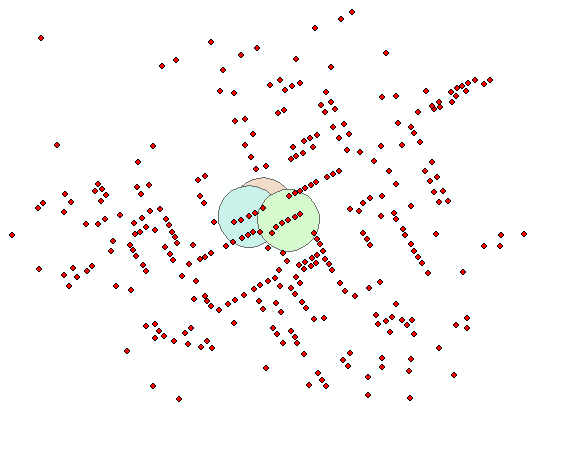
* 1. Quelle distance est utilisée pour caractériser ce rayon de recherche et à quelle famille de distance appartient-elle ?

**Explication cartographique des rayons de différentes distances :**

Rayon de **100 m**

**98** points voisins sont utilisés dans le calcul du nombre de morts lissé

Rayons de 30 m pour trois points. En réalité le calcul du nombre de morts lissé (donc la moyenne) est effectué pour chaque cellule dans le raster en sortie, à partir de ces voisins (déterminé en fonction du rayon de recherche choisi (cf figure à gauche)

Rayon de **50 m**

**25** points voisins sont utilisés dans le calcul du nombre de morts lissé

Rayon de **30 m**

**11** points voisins sont utilisés dans le calcul du nombre de morts lissé

Point de référence dans l’exemple

1. Améliorez le rendu de la carte lissée :

* dans *Type de rendu* choisissezune *Pseudocouleurs à bande unique*, puis une palette de couleur dégradé
* laissez *Type d’interpolation* linéaire
* sélectionnez *Classer* pour faire apparaître les valeurs des cellules.
* si besoin inversez la palette de couleur
* ne mettez aucune couleur dans les premières classes (valeurs très faibles) –> opacité nulle
* modifiez la transparence (onglet transparence) du lissage pour rendre visible la carte de Dr Snow

1. Refaites l’analyse en faisant varier le rayon de recherche à 70 m, puis de 100 m. Qu’observez-vous sur les cartes ?

1. Proposez une rédaction de légende qui aide le lecteur à facilement comprendre la carte et modifiez la légende (remplacer les chiffres). Pour faire cette légende il est nécessaire de choisir Type de rendu : pseudo-couleur à bande unique. Cette option permet de choisir un nombre de de classes et une discrétisation. Vous pouvez l’utiliser simplement pour faire la légende en choisissant un faible nombre de classes (par ex 5), puis en modifiant les étiquettes de la légende. C’est un peu du bricolage et pour un rendu final il faudra ajouter cette légende à la mise en page, soit via un logiciel de dessin (comme Inkscape ou Illustrator), soit simplement pour un « découpage manuelle et collage » dans la mise en page

1. Commentez la répartition spatiale des victimes de l’épidémie dans le quartier de Soho.

.

1. Au milieu du 19e siècle la théorie communément acceptée était la transmission du choléra par l’air, pourtant d’après l’observation de la répartition des morts John Snow a déduit que l'air n’était pas en cause. Arrivez-vous à la même conclusion ? Argumentez.

Aujourd’hui on sait que le vecteur de transmission du choléra est l’eau (et le contact physique entre personnes), mais c’est la carte de la répartition des morts de choléra qui a donnée l’idée à Snow de vérifier l’hypothèse de l’eau comme vecteur de transmission.

*Dans cette première étape vous avez déterminé la localisation des objets étudiés (les morts) et vous vous êtes posé la question de la répartition géographique du phénomène. En effet, la première question que Dr Snow s’est posée était « où ? ». Il a répondu par l’étude de la carte des victimes de choléra. Puis il s’est posé une deuxième question, « où et comment ? ». J. Snow s’est ainsi intéressé à la localisation relative des victimes (les unes par rapport aux autres) et a cherché à déterminer comment les morts se sont répartis à l’intérieur du quartier (concentration).*

**Pensez à enregistrer votre projet**

# Repérer de la structure spatiale des morts du choléra : réponse à la question « Comment et pourquoi ? »

*Avec l’hypothèse de l’eau comme vecteur de transmission et dans l’objectif de comprendre la cause de l’épidémie, ainsi que les causes de sa diffusion par la suite, Dr Snow a localisé très précisément sur la carte précédente les* ***pompes à eau*** *dans le quartier de Soho.*

D’après l’observation de la carte, Dr Snow s’est demandé s’il n’y avait pas une concentration plus importante de cas de choléra autour de certaines pompes à eau. Il n’avait pas à l’époque les moyens pour vérifier de façon scientifique cette deuxième hypothèse. Aujourd’hui vous avez ces moyens et pour analyser la structure spatiale des victimes autour des différentes pompes à eau, vous allez utiliser différentes méthodes d’analyse spatiale pour vérifier si certaines pompes à eau pourraient être les sources de contamination.

## Identifier les pompes à eau les plus proches de chaque adresse de mort avec une distance mathématique

1. Créez un jeu de données vecteur (dans le geopackage) des pompes à eau à partir du fichier Pumps.csv et rajoutez-le au projet en cours.
2. Vous allez commencer par utiliser la méthode des polygones de Thiessen (= diagramme ou cellules de Voronoï). Quel est le principe de cette méthode ? Que définit-elle ? Sur quelle distance se base-t-elle ? (pour répondre à ces questions consultez l’annexe 3 de ce document et d’autres sources si nécessaire)

1. Quelle est l’hypothèse sous-jacente à cette méthode en ce qui concerne l’utilisation des pompes à eau par les habitants. Quelle est la principale limite de cette hypothèse ?

1. Créez des polygones de Thiessen (aussi appelés polygones de Voronoï) en fonction des pompes. Dans QGIS l’outil s’intitule Polygones de Voronoï et permet d’élargir la surface couverte des polygones en ajoutant une zone tampon à la limite des zones couvertes des points en entrés (ici le pompes). Testez une fois sans zone tampon, puis une deuxième fois avec une zone tampon de 10 % en plus de l’emprise initiale. Comparez les deux résultats. Selon vous, quel est le résultat le plus pertinent ? Pourquoi ?

1. Affichez sur chaque polygone la localisation des morts et les noms des pompes à eau. Que constatez-vous grâce aux polygones de Thiessen ?

1. Quel pourcentage des décès de choléra se sont produits dans les zones d’influence des pompes telles qu’elles ont été définies par les polygones de Thiessen ? Différentes méthodes permettent de répondre à cette question. Vous pouvez soit faire une jointure spatiale avec un résumé des attributs (Joindre les attributs par localisation (résumé)), soit simplement sélectionner la zone concernée, calculer les statistiques automatiquement, puis manuellement le pourcentage en question.
2. Quel pourcentage des décès de choléra se sont produits dans la zone d’influence de la pompe à Broad street telle qu’elle a été définie par les polygones de Thiessen ?

**Pensez à enregistrer votre projet**

## Identifier les pompes à eau les plus proches de chaque adresse de mort avec la distance par le réseau viaire

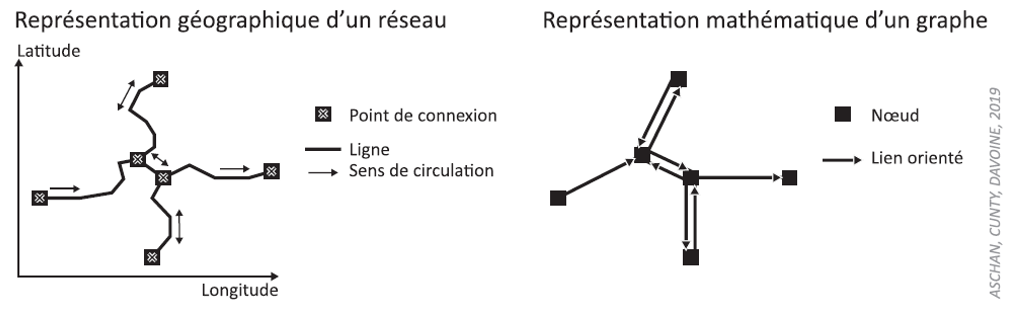
L’objectif est ici de déterminer quel est la pompe la plus proche de chaque lieu dans le quartier de Soho, en tenant compte d’un déplacement par les rues (= distance empirique) et non par un déplacement à vol d’oiseau (= distance euclidienne) comme dans l’étape précédente. Il s’agit donc à nouveau d‘identifier les zones de proximité (zones potentielles d’influence) de chacune des pompes, mais cette fois-ci en se fondant sur une distance empirique qui est la distance kilométrique sur le réseau de rues.

### Préparer les jeux de données

1. Vérifiez que le jeu de données d’entités linéaires est structuré sous forme d’un réseau

Pour effectuer des analyses de réseau sur un jeu de données, il est important que les entités linéaires qui représentent les tronçons (liens) du réseau soient bien « coupés » à chaque point de connexion (nœuds) du réseau (figure 1). Vérifiez que c’est bien le cas pour ReseauRuesSoho.shp en sélectionnant quelques tronçons, un à un.

**Figure 1** : Représentation géographique et mathématique



***Commentaire figure 1 :*** *La structuration des données sous la forme d’un réseau implique de définir des relations topologiques qui lient les entités points (nœuds) et les entités lignes (liens) entre elles. On simplifie alors le réseau sous la forme d’un objet mathématique appelé graphe ou réseau logique. Les liens peuvent être orientés (figure 1), ce qui signifie qu’ils ont des nœuds de départ et d’arrivée selon un sens de circulation. Cette information est indispensable pour pouvoir ensuite calculer la distance sur le réseau.*

1. Pour pouvoir faire une analyse de l’accessibilité sur un réseau il est indispensable de connecter au réseau les lieux pour lesquels les distances seront calculées. Ces lieux sont appelés des sommets.
2. A quoi correspondent les sommets ici ?
3. Est-ce que les pompes sont connectés au réseau ? Expliquez pourquoi.

1. Mesurez les distances (outil « mesurer une longueur » dans la barre d’outils) entre quelques sommets et le réseau. Précisez l’ordre de grandeur des distances observées (avec l’unité de mesure).
2. Dans QGIS, le module V.net (vecteur network de Grass) permet de relier les sommets au point le plus proche sur ce réseau. Ouvrez l’outil et sélectionnez les paramètres suivants :
   1. Opération à effectuer : *Connect*
   2. « Threashold for connecting centers to the network » : il s’agit ici de spécifier le seuil de distance maximal à utiliser pour connecter les pompes au tronçon du réseau de rue la plus proche. Quelle distance proposez-vous ici (en fonction de ce que vous avez observé en question 3 ci-dessus)?
   3. Les autres paramètres n’ont pas besoin d’être modifies.
   4. Spécifier le nom du fichier résultat et le dossier d’enregistrement
3. Exécutez l’outil V.net
4. Quel est le résultat graphique de cette opération ? Expliquez et insérez une copie d’écran du résultat ici

### Déterminer la pompe la plus proche de chaque tronçon de route et donc l’« aire d’attraction » de chaque pompe en fonction d’une distance empirique

1. Ici il s’agit d’identifier pour chaque tronçon du réseau de rues quelle est la pompe la plus proche selon la distance calculée à partir du réseau. Pour effectuer cette opération on utilise l’outil v.net.alloc qui va affecter chaque tronçon du réseau à la pompe la plus proche parmi l’ensemble des dix pompes.
   1. Effectuez ce traitement en veillant à bien choisir les données en entrée.
   2. Observez le résultat (carte et table attributaire) et expliquez-le.
   3. A quoi correspond le champ *cat* crée ?

* 1. Affichez les aires d’attraction des pompes sur la carte. Quelle variable visuelle choisissez-vous et pourquoi ?

* 1. Modifiez l’épaisseur des traits (pour l’ensemble des catégories en même temps via le symbole) pour améliorer la lisibilité de la carte et changez les couleurs pour chaque aire d’influence si nécessaire.

1. Quelle est l’hypothèse sous-jacente sur les pratiques de déplacement des habitants de Soho pour s’alimenter en eau ?

1. Comment pourrait-on affiner la modélisation des tronçons des rues pour ensuite pouvoir calculer une distance coût au lieu d’une simple distance kilométrique (ce n’est pas possible ici mais c’est une réflexion dans l’absolu) ?

1. Quel pourcentage des décès de choléra se sont produits dans la zone d’influence de la pompe de Broad Street (quand l’aire d’influence est identifiée par le chemin le plus court à chaque pompe par les rues) ? Plusieurs étapes sont nécessaires ici :
2. La première étape consiste à joindre à la couche Deaths les champs de la table attributaire de la couche des tronçons de rue associés à la pompe la plus proche que vous venez de créer.
3. Quel type de jointure réalisez-vous ici : spatiale ou attributaire ?
4. Aucun des opérateurs spatiaux (appelés prédicats géométriques dans QGIS) ne permettent de faire une jointure basée sur la distance. Cherchez dans la boîte à outils l’outil qui permet de faire cette jointure
5. Réalisez la jointure, sachant qu’il s’agit de joindre le champ *cat* à la table de la couche Deaths (qui sera donc la couche source). NB. Le nombre maximum de voisins les plus proche que vous devez choisir est 1.
6. Rappelez à quoi correspond le champ *cat.*

1. Donnez des précisions sur le résultat de la jointure spatiale :

* Quelles sont les entités de la couche ?
* Combien d’entités contient la couche ?
* Est-ce que la géométrie des entités a été modifiée ? Expliquez la raison.

* La jointure spatiale par le plus proche ajoute cinq nouveaux champs. A quoi correspondent le champ distance ? .

1. Ensuite, vous devez utiliser un outil permettant de calculer le nombre de morts qui sont localisées le plus près de chaque pompe. Pour cela vous pouvez par exemple utiliser l’outil Statistiques par catégories.
2. Calculez la part des morts (en %) dont l’adresse est située dans la zone d’influence de chacun des pompes.
3. Quelle est la part dans l’aire d’attraction de la pompe de Broad Street ?
4. Est-ce que les résultats obtenus par les polygones de Thiessen et par la méthode d’allocation des coûts prenant en compte la distance réelle par les rues entre le domicile et chaque pompe convergent ? Pour répondre à la question, comparez les deux cartes et les parts de morts (%) dans les aires d’attraction identifiées avec les deux méthodes.

1. Ces résultats appuient-ils la supposition que le point de départ de l'épidémie de choléra est de l'eau contaminée ? Expliquez.

1. Que pouvez-vous en conclure ?

Dans cette étape vous avez comparé et analysé conjointement les répartitions de deux phénomènes dans l’objectif d’expliquer une localisation par d’autres localisations. Il s’agissait donc de déceler en quoi la localisation des pompes à eau a pu avoir une influence sur la diffusion de l’épidémie de choléra.

**Pensez à enregistrer votre projet**

# Expliquer et fournir une aide à la décision

Affichez sur la carte le nom des compagnies de distribution des eaux pour chaque pompe ? Que remarquez-vous ?

***Explication****: A l’époque, la distribution de l’eau à Londres était faite par plusieurs compagnies. Il y avait à la fois des puits publics peu profonds où les habitants venaient chercher leur eau et des conduites qui puisaient l’eau de la Tamise. Le puits de Broad Street, exploité par « Southwark and Vauxhall Water Company » était un puits public où l’eau provenait de la Tamise à l’aval de Londres. En outre le puits de Broad Street était creusé à un mètre d’une fosse d’aisance (pas d’égout à l’époque). La fosse d’aisance était vraisemblablement infectée par le Vibrio cholerae.*

***Aide à la décision****: C'est le 7 septembre que, suite aux observations de Snow, le puit fut mis hors service (on a enlevé la poignée de la pompe !). Le 11 du même mois, l'épidémie disparaissait pratiquement. Malgré l'évidence, les autorités médicales ne voulurent pas croire que les latrines pouvaient être à l'origine de la contamination. Même si un fervent partisan du docteur Snow, le pasteur Whitehead, découvrit que certaines de celle-ci fuyaient réellement dans le puit qui alimentait le puit, ce n'est qu'en 1866, alors que le choléra frappa à nouveau Londres, que la théorie de Snow fut acceptée, quelques années après sa mort.*

# Mettez en page une carte au choix que vous avez créé pendant ce TD et insérez la carte ici.

**Déposez sur moodle ce document de TD complété par les réponses aux questions et des copies d’écran de l’ensemble des traitements et résultats avant le 15/11/2025.**

# Annexe 1 : Créer un jeu de données vecteur de géométrie ponctuel à partir de coordonnées XY

Le fichier *Deaths* est enregistré en format .csv (avec séparateur : point-virgule) qui est un format texte délimité. Ajoutez ce fichier dans QGIS en passant par le *Menu Couche > Ajouter une couche > Ajouter une couche de texte délimité…*

1. Donnez le chemin d’accès au fichier csv
2. Bien garder le codage UTF-8
3. Indiquez comme format de fichier : nous avons un fichier enregistré en csv mais avec un séparateur de point-virgule. Il faut donc sélectionner *Délimiteur personnalisé* et cocher l’option *point-virgule*.
4. Cochez les options des champs et des enregistrements :

- *en-têtes en première ligne* et *Détecter les types de champs.*

*- précisez le séparateur décimal des valeurs de champs* (vérifiez dans le tableau)

1. Définition de la géométrie

* Indiquer qu’il s’agit de créer des points
* Précisez les champs contenant les coordonnées X et Y
* Précisez le système de coordonnées de référence (SCR) de la géométrie (des points). Pour pouvoir afficher ces données actuellement sous format texte, il faut préciser à quel système de coordonnées (SC) correspondent les coordonnées des deux champs X et Y. Ouvrez la fenêtre sélectionneur de SCR avec puis utilisez le filtre pour pouvoir trouver le SCR correspondant aux coordonnées X et Y (voir plus haut).
* Il est donc indispensable de connaître puis, d’indiquer le « système des coordonnées en entrée ». **TRES IMPORTANT**: Ceci une étape cruciale pour pouvoir afficher les données « au bon endroit » et éviter qu’elles soient localisées, par exemple, au milieu de l’Atlantique. Pour cela il est impératif de choisir le système de coordonnées dans lequel sont définies les coordonnées X et Y dans le fichier que l’on veut transformer en jeu de données vecteur.
* Les coordonnées X et Y des lieux où il y eu des morts de choléra en septembre 1854, sont définis dans le système appelé *British\_National\_Grid.* Pour « aller chercher » ce système des coordonnées, faites « *Modifier »*, puis, dans la fenêtre *« Propriétés de la référence spatiale »,* effectuez une recherche sur le terme « British » pour retrouver ce système de coordonnées. Pensez à mettre ce SC dans vos favoris (vous en aurez besoin plus tard). Validez vos choix.

1. Enfin, vérifiez que l’échantillon de données présente correctement les données dans les champs.
2. Les données XY ajoutées ne sont pas automatiquement transformées en un jeu de données vecteur, elles sont simplement affichées dans le bloc de données. Pour pouvoir les utiliser, il est donc nécessaire de les exporter en un format en mode vecteur (par exemple, shapefile ou Geopackage).

# Annexe 2 : Cartographier une variable quantitative de stock

1. Une image contenant texte, meubles, fichier

   Description générée automatiquementOuvrez les paramètres de symbologie dans les propriétés de la couche. Sélectionnez Symbole unique> Symbole simple et ouvrez « l’assistant taille » en cliquant sur le bouton à droite de la ligne « Taille ».

* Dans la fenêtre taille du symbole qui s’ouvre, renseignez dans la partie Saisie (cette partie concerne les valeurs de la variable utilisée.)
  + Dans Source : le nom du champ contenant la variable quantitative de stock à représenter
  + Valeurs depuis... à ... : cliquez sur le bouton pour récupérer automatiquement les modalités minimum et maximum de la variable.
* Dans la partie Sortie (cette partie concerne la manière dont les valeurs seront représentées)
  + Taille depuis... à ... : choisissez ici les surfaces correspondant aux valeurs minimale et maximale. Vous pouvez tester différentes valeurs, le résultat dépendra de l'échelle à laquelle la carte sera lue
  + Méthode de calcul : choisissez Surface car on cherche ici à ce que la surface des cercles soit proportionnelle à la population. NB : Les surfaces correspondant aux valeurs intermédiaires sont interpolées entre les tailles min et max indiquées… la proportionnalité n’est donc pas tout à fait parfaite !

**Attention à respecter le rapport de proportion entre le plus grand et le plus petit cercle**

* + - Calcul de la valeur du plus petit cercle en fonction de la taille du plus grand cercle
    - Taille du petit cercle = minimum de la série x taille du plus grand cercle / maximum de la série
    - Exemple de calcul pour une taille choisie pour le plus grand cercle = 30 et avec une valeur minimale de la série = 27 et une valeur maximale 41 248 : taille du petit cercle = 27 x 30 / 41 248
  + **Attention** à ne surtout pas faire varier la taille proportionnellement au rayon (les variations des surfaces seront alors exagérées et probablement mal interprétées).

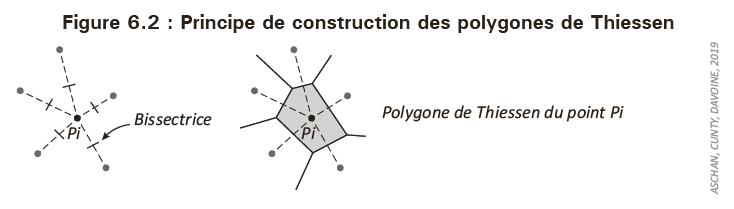
1. Avant d’appliquer la symbologie, il faut déterminer l’ordre de représentation des cercles pour que les cercles plus petits soient systématiquement placés au-dessus des plus grands. Pour cela en bas de la fenêtre Symbologie, ouvrez (clic sur la flèche) « Rendu de couche », cochez « Contrôle de l’ordre de rendu des entités » puis cliquez sur et dans la fenêtre « Définir l’ordre », indiquez le champ contenant la variable quantitative de stock à représenter dans « Expression » et « Descendant » dans « Croissant/Descendant ».
2. Pour améliorer le rendu graphique de vos cartes par cercles proportionnels pensez à détourer les cercles en blanc plutôt qu’en noir.
3. Enfin pour voir apparaître la légende, cliquez sur Avancé (Toujours dans la symbologie de la couche) > « Légende pour la Taille définie par des données », puis cliquez sur Légende repliée.

# Annexe 3 : Les polygones de Thiessen

(extrait de l’ouvrage Aschan C., Cunty C., Davoine PA., 2023, Les systèmes d’information géographique, Armand Colin, collection Cursus)

Une méthode simple pour définir les aires d’attraction respectives de chaque pompe, modélisées sous la forme d’un semis de points, est basée sur la construction de polygones de Thiessen qui utilisent la distance euclidienne.

Les polygones de Thiessen sont créés à partir d’un maillage de semis de points, en joignant les points les plus proches par des lignes pour effectuer une triangulation (appelée triangulation de Delaunay), puis en identifiant les bissectrices qui constituent les arcs des polygones (figure 6.2).



Les polygones de Thiessen divisent donc la surface couverte par le semis de points en polygones, chaque point se trouvant à l’intérieur d’un polygone et d’un seul, et chaque polygone contenant un point et un seul. Ainsi, tout emplacement dans un polygone de Thiessen est plus proche du point localisé dans ce polygone que de tout autre point du semis.

1. NB. Il ne s’agit pas d’une interpolation, mais d’un lissage des données (méthode différente de l’interpolation) [↑](#footnote-ref-1)