

# STATISTIQUES SPATIALES ET SIG

Projet Master 2 - 2021/2022



Anaël YAHY & Antoine Le Bleis



---

<b>1- PROCESSUS SPATIAL</b>	<b>3</b>
<b>2- MISE EN ÉVIDENCE D'UNE STRUCTURE SPATIALE</b>	<b>4</b>
<b>3- INTERPOLATIONS DÉTERMINISTES</b>	<b>5</b>
<b>A- INTERPOLATION IDW</b>	<b>6</b>
<b>B- INTERPOLATION TIN</b>	<b>7</b>
<b>C- INTERPOLATION DES PLUS PROCHES VOISINS</b>	<b>8</b>
<b>D- RETOUR SOUS R</b>	<b>9</b>
<b>4- INTERPOLATION PAR MÉTHODE PROBABILISTE OU KRIGEAGE</b>	<b>10</b>
<b>5- CONCLUSION</b>	<b>11</b>
<b>6- ANNEXE</b>	<b>12</b>



*Photographie de la ville de Montpellier*

---

---

# 1- PROCESSUS SPATIAL

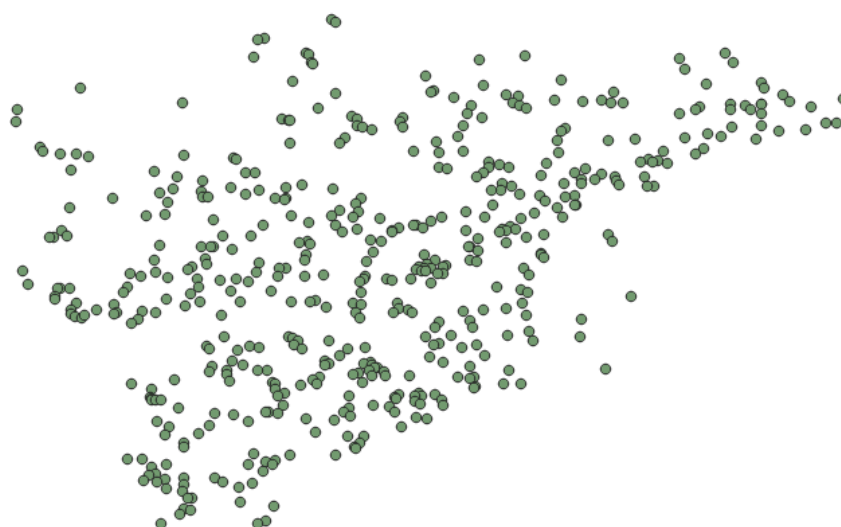
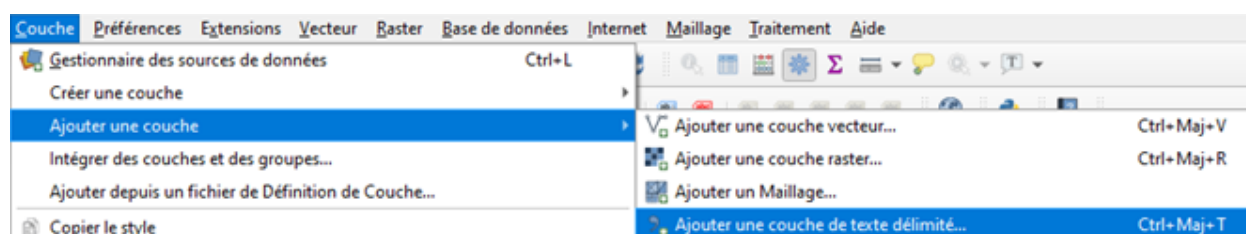
Notre groupe a été choisi pour traiter les données du fichier MontpellierE.csv.

Ce fichier est composé de plus de 2 000 points, tous situés dans la ville de Montpellier.

Nous avons pour chacun de ses points :

- Leurs coordonnées en WGS84.
- Le niveau sonore LDEN (LDEN est un indicateur de niveau sonore qui signifie Level Day-Evening-Night).
- Des données sur les zones où se trouvent les points.

Nous les avons donc ajoutés sur notre projet QGIS, afin de pouvoir les visualiser :



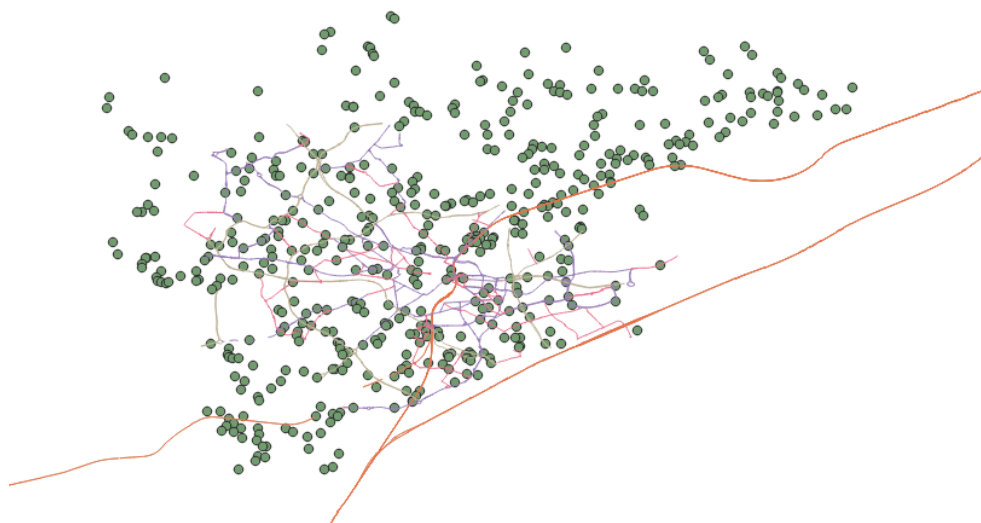
Nous avons donc voulu savoir où ces points se situaient dans la ville de Montpellier.

Nous avons donc installé l'extension QuickOSM et nous avons, par la suite, fait apparaître les routes principales/secondes/tertiaires, ainsi que le chemin de fer de la ville de Montpellier.

---

---

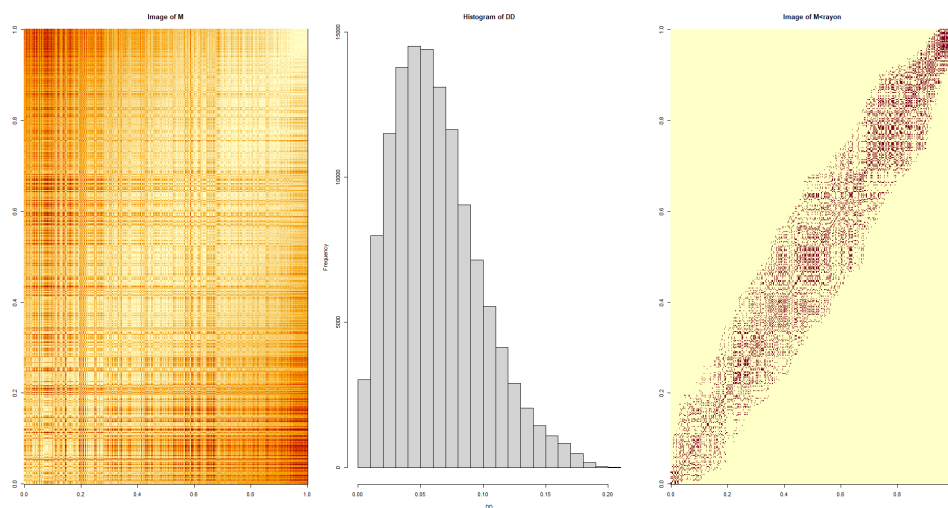
Nous obtenons une carte comme celle-ci :



Nos points ne sont donc pas regroupés dans une partie spécifique de la ville. Nous pouvons maintenant passer à l'analyse.

## 2- MISE EN ÉVIDENCE D'UNE STRUCTURE SPATIALE

Pour commencer, nous avons calculé plusieurs matrices de voisinage.





---

Nous avons ensuite, grâce à celles-ci, calculé le coefficient de Geary et de Moran.

Nous avons enfin effectué un test de validité, sous l'hypothèse  $H_0$  : indépendance spatiale.

```
Moran I test under randomisation

data: df$LDEN
weights: Mlistw

Moran I statistic standard deviate = -9.948, p-value = 1
alternative hypothesis: greater
sample estimates:
Moran I statistic      Expectation      Variance
-1.448811e-02      -2.004008e-03      1.574877e-06

> |
```

### 3- INTERPOLATIONS DÉTERMINISTES

Les distributions spatiales et spatio-temporelles des phénomènes physiques et socio-économiques peuvent être approximées par des fonctions dépendant de l'emplacement dans un espace multidimensionnel. Des exemples typiques sont les élévations, les phénomènes climatiques, les propriétés des sols, densités de population etc.

Si la plupart de ces phénomènes sont caractérisés par des données ponctuelles mesurées ou numérisées, ils sont souvent irrégulièrement répartis dans l'espace. De plus, ils sont mesurés à l'aide de diverses méthodes, et cela conduit à des jeux de données hétérogènes, avec des représentations numériques différentes qui doivent être combinées afin de créer un modèle spatial unique.

De nombreuses méthodes d'interpolation et d'approximation ont été développées pour prédire les valeurs des phénomènes spatiaux dans des emplacements non échantillonnés. Dans les applications SIG, ces méthodes ont été conçues pour prendre en charge les transformations entre différentes représentations discrètes et continues de champs spatiaux et spatio-temporels, généralement pour transformer des points irréguliers ou des données de ligne à une représentation raster, ou pour rééchantillonner entre différentes résolutions raster.

Ici nous allons effectuer 3 types d'interpolations différents. Commençons par l'interpolation IDW.

---

## A- INTERPOLATION IDW

IDW veut dire Inverse Distance Weighted. On peut le traduire par l'inverse de la distance pondérée. Dans cette méthode, les points d'échantillonnage sont pondérés pendant l'interpolation, de sorte que l'influence d'un point par rapport à un autre diminue avec la distance par rapport au point inconnu qu'on a généré.

La pondération est attribuée aux points d'échantillonnage grâce à l'utilisation d'un coefficient de pondération qui contrôle la manière dont l'influence de la pondération diminuera à mesure que la distance par rapport au nouveau point augmente.

Plus le coefficient de pondération est élevé, moins les points auront d'effet s'ils sont éloignés du point généré pendant le processus d'interpolation. Lorsque le coefficient augmente, la valeur du point généré se rapproche de la valeur du point d'observation le plus proche.

### Ses avantages :

Ce type d'interpolation peut augmenter ou diminuer le nombre de points d'échantillonnage, afin d'influencer les valeurs cherchées.

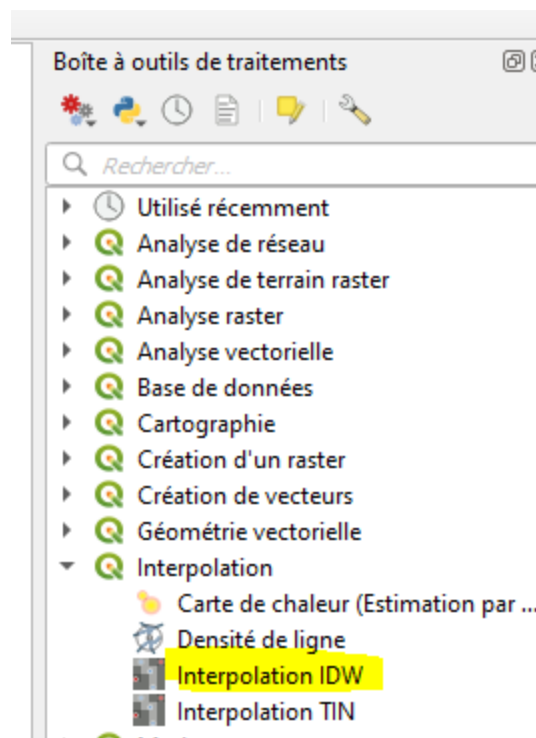
Les points espacés uniformément denses sont bien interpolés.

Ce type d'interpolation peut estimer les changements extrêmes du terrain.

### Ses inconvénients :

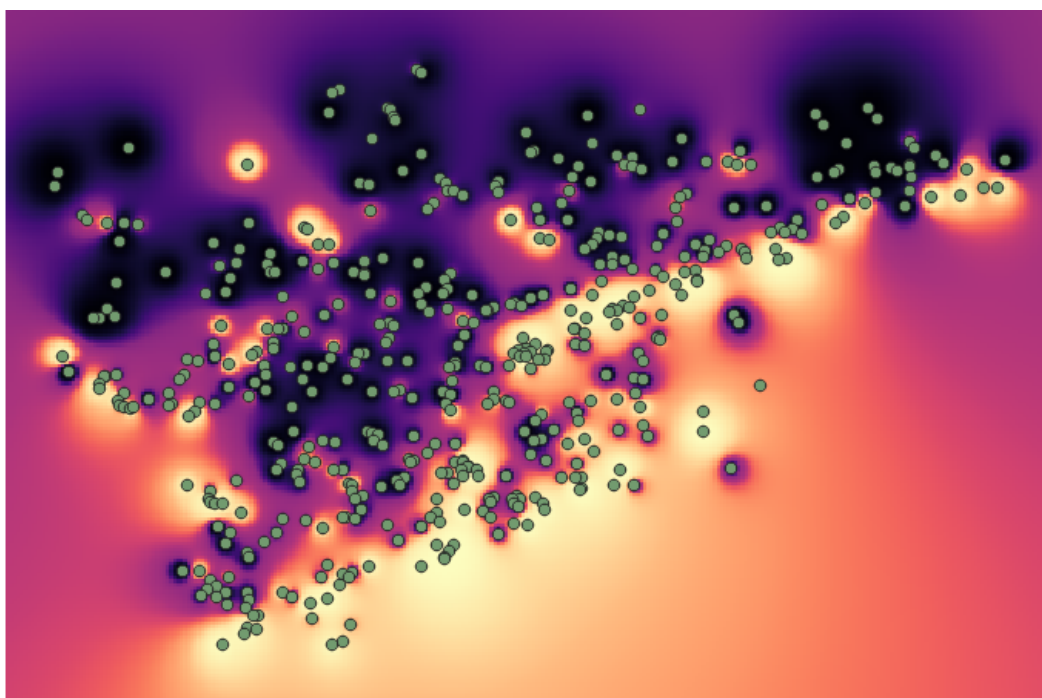
La qualité du résultat de l'interpolation peut diminuer si la distribution des points de données de l'échantillon est inégale.

Il est impossible d'estimer au-dessus des valeurs maximales, ou en dessous des valeurs minimales.



Sous QGIS, cela donne :

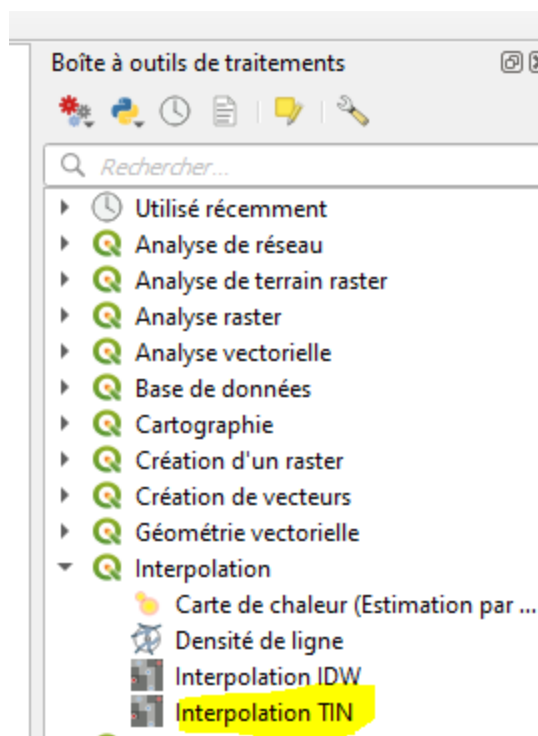
---



## B- INTERPOLATION TIN

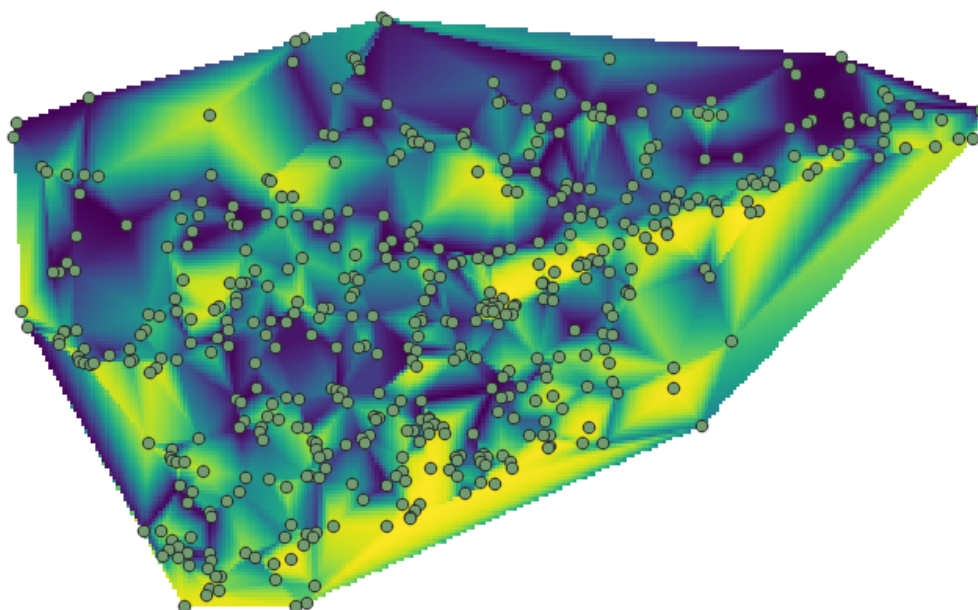
L'interpolation TIN est un autre outil populaire dans les SIG. Un algorithme TIN courant est appelé triangulation de Delaunay . Il essaie de créer une surface formée par des triangles de points voisins les plus proches. Pour ce faire, des cercles circonscrits autour de points d'échantillonnage sélectionnés sont créés et leurs intersections sont reliées à un réseau de triangles non superposés et aussi compacts que possible

Le principal inconvénient de l'interpolation TIN est que les surfaces ne sont pas lisses et peuvent donner un aspect irrégulier. Cela est dû à des pentes discontinues aux bords du triangle et aux points de données d'échantillon. De plus, la triangulation n'est généralement pas adaptée à l'extrapolation au-delà de la zone avec des points de données d'échantillon collectés.

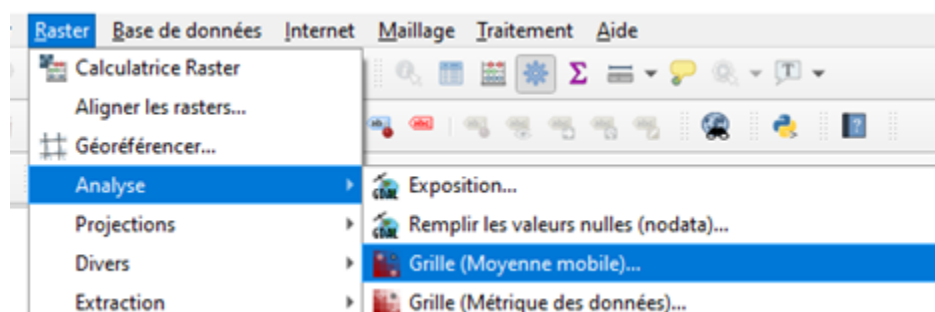


---

Sous QGIS, cela donne :



## C- INTERPOLATION DES PLUS PROCHES VOISINS



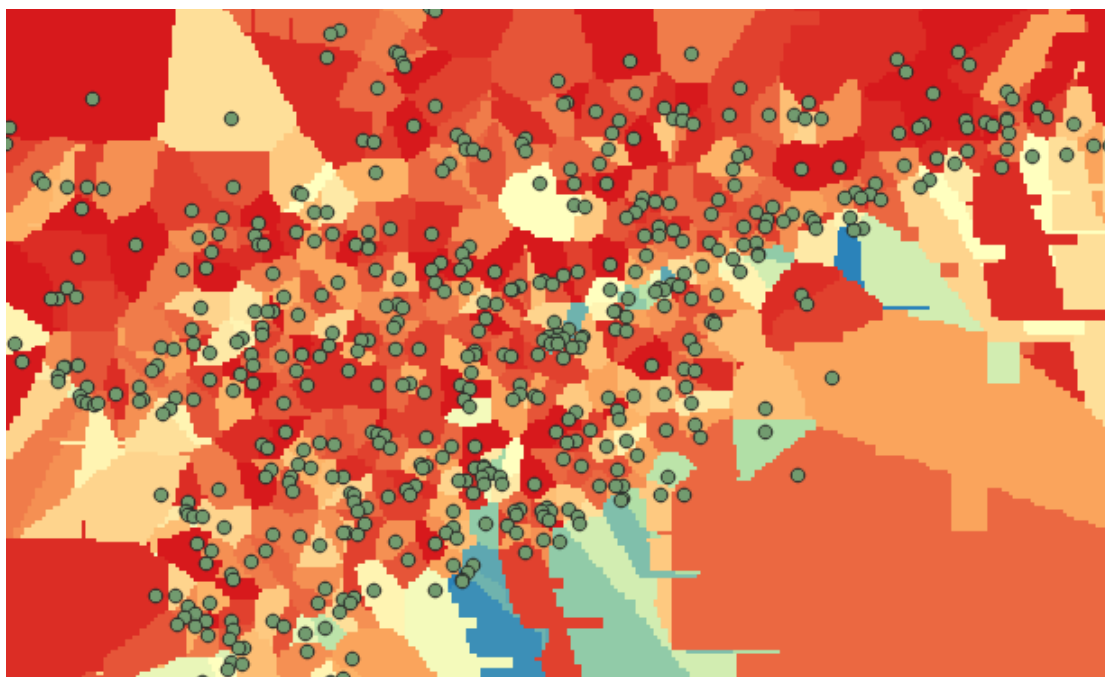
Également appelée interpolation arrondie, l'interpolation des plus proches voisins est la méthode d'interpolation la plus simple. Plutôt que de calculer une valeur moyenne, selon certains critères, ou de générer une valeur basée sur des règles complexes, cette méthode localise la valeur du point de données le plus proche, et lui attribue cette valeur.

L'inconvénient de cette méthode est justement sa force : l'interpolation se fait de manière très directe. Cela peut porter à confusion et ne peut pas être utilisé pour des problèmes complexes.

Sous QGIS, cela donne :

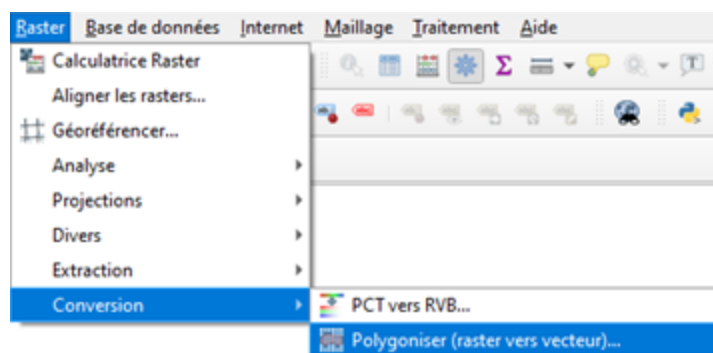
---





## D- RETOUR SOUS R

Ces méthodes génèrent des fichiers raster, soit des images. Pour les exporter sous R, il nous faut un fichier de nombres type CSV. Pour se faire, on commence par convertir les rasters en polygones.



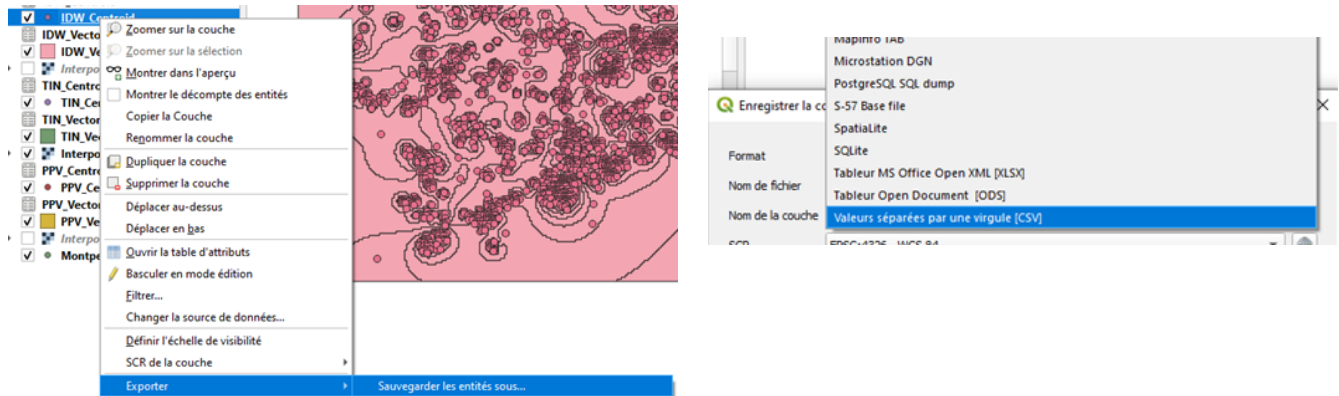
On peut ensuite les convertir en points, via centroïdes.



Ici, il manque les coordonnées X et Y de latitude et longitude dans les tables de données. On peut les ajouter manuellement via des entrées supplémentaires et les formules \$x

et \$y, ou ici via un script Python présent dans le projet qui ajoute immédiatement les colonnes aux trois couches de points ainsi créées.

Finalement, on peut enregistrer les couches générées de cette manière en CSV.



## 4- INTERPOLATION PAR MÉTHODE PROBABILISTE OU KRIGEAGE

Le krigeage est une méthode stochastique d'interpolation spatiale qui prévoit la valeur d'un phénomène naturel par une combinaison linéaire sans biais et à variance minimale des observations du phénomène en des sites voisins.

L'une des spécificités des méthodes de krigeage est qu'elles ne considèrent pas seulement la distance entre les observations, mais elles visent également à capturer la structure spatiale des données, en comparant deux à deux des observations séparées par des distances spatiales spécifiques.

On peut l'implémenter sur R à l'aide de la fonction `krige()` du package `gstat`. Pour l'appliquer, nous utilisons un modèle variogramme grâce au package `sp`, et une grille, en plus des données exportées depuis QGIS après des interpolations.

Nous obtenons donc 3 krigeages, adaptés à chaque interpolation déterministe. Par exemple, voici à quoi ressemble le krigeage adapté aux données TIN :

---

```
> summary(krigeTIN)
Object of class SpatialPixelsDataFrame
Coordinates:
      min      max
x 3.79166 3.99166
y 43.64328 43.68328
Is projected: NA
proj4string : [NA]
Number of points: 80
Grid attributes:
  cellcentre.offset cellsize cells.dim
x          3.79666      0.01         20
y          43.64828      0.01          4
Data attributes:
  var1.pred      var1.var
Min.   :49.35   Min.    : 1.781
1st Qu.:52.43   1st Qu.: 4.932
Median :53.69   Median : 7.486
Mean   :53.75   Mean    : 7.239
3rd Qu.:54.32   3rd Qu.:10.168
Max.   :63.17   Max.    :10.940
> |
```

En comparant notre krigeage aux interpolations, nous pouvons dire que le krigeage a l'air plus souple que les modèles déterministes.

## 5- CONCLUSION

Pour conclure, nous pouvons dire que ce projet nous a bien plu. En effet, il était assez intéressant de découvrir le logiciel QGIS durant tout ce semestre, et d'appliquer ce que nous avons appris en cours à des données réelles. Nous avons également appris plus en détails les méthodes d'interpolation vues en cours.

Le travail entre deux plateformes, R et QGIS, était aussi un challenge intéressant à relever, à la fois sur la gestion des formats et optimiser les fonctionnalités de chacun des logiciels.

---

---

## 6- ANNEXE

Le projet complet et les notes peuvent être trouvés sur Github :

[https://github.com/Aliquanto3/M2-DSMS\\_Systemes-Information-Geographiques](https://github.com/Aliquanto3/M2-DSMS_Systemes-Information-Geographiques)

Le cours de SIG de A. Antoni

<https://moodle.univ-ubs.fr/course/view.php?id=2006>

[https://docs.qgis.org/2.8/fr/docs/user\\_manual/processing\\_algs/saga/kriging.html](https://docs.qgis.org/2.8/fr/docs/user_manual/processing_algs/saga/kriging.html)

<http://web.pdx.edu/~jduh/courses/geog493f12/Week06.pdf>

[https://gro-1.itrcweb.org/simple-geospatial-methods/#inverse\\_distance](https://gro-1.itrcweb.org/simple-geospatial-methods/#inverse_distance)

<https://www.aspexit.com/spatial-data-interpolation-tin-idw-kriging-block-kriging-co-kriging-what-a-re-the-differences/#Kriging>

[https://gml.noaa.gov/webdata/aero/net/cpd3/doc/textconfig/type\\_model.html](https://gml.noaa.gov/webdata/aero/net/cpd3/doc/textconfig/type_model.html)

<https://gis.stackexchange.com/questions/359940/adding-x-y-coordinates-to-attribute-table-in-qgis>

[https://docs.qgis.org/2.14/en/docs/training\\_manual/complete\\_analysis/raster\\_to\\_vector.html](https://docs.qgis.org/2.14/en/docs/training_manual/complete_analysis/raster_to_vector.html)

[https://gisresources.com/types-interpolation-methods\\_3/](https://gisresources.com/types-interpolation-methods_3/)

“*Spatial interpolation*” de L. Mitas et H. Mitasova

---