



Impact des Éoliennes sur les Prix des Maisons

Imene Zebiri

Université Paris Nanterre

Date : 27 janvier 2025



Sommaire

1	Introduction	4
2	Présentation des Données	5
2.1	Données des Eoliennes	5
2.1.1	Données des Turbines Éoliennes	5
2.1.2	Données des Parcs Éoliens	5
2.2	Données des Maisons	6
2.3	Prétraitement des Données	7
2.3.1	Jointure des Données Éoliennes	7
2.3.2	Filtrage des Données Éoliennes	7
2.3.3	Filtrage des Données des Maisons	8
2.4	Conclusion	8
3	Difficultés rencontrées et traitement des données	8
3.1	Problèmes liés à RStudio	8
3.2	Gestion des Coordonnées GPS des Éoliennes	8
3.3	Traitement des données <code>maisons_{eo}.csv</code>	10
3.4	Solutions envisagées pour optimiser le temps d'exécution	11
3.5	Conclusion	11
4	PARTIE I : SIG	12
4.1	Difficultés détaillées	12
5	PARTIE II : ECONOMETRIE	14
5.1	Statistiques descriptives des éoliennes	14
5.1.1	Puissance des Éoliennes	15
5.1.2	Hauteur des Éoliennes	16
5.1.3	Durée de Construction	17
5.1.4	Nombre d'Éoliennes par Parc	18
5.1.5	Éoliennes Construites par Année	19
5.1.6	Conclusion	19
5.2	Statistiques descriptives des biens immobiliers	19
5.2.1	Mutations par Année	19
5.2.2	Fréquence des Mutations par Mois	20
5.2.3	Valeur Foncière	20
5.2.4	Analyse des villes avec le plus de ventes immobilières	21
5.2.5	Surface Réelle Bâtie	21
5.2.6	Distances aux Infrastructures	22
5.2.7	Conclusion	23
5.3	Regression linéaire et Statistiques des biens immobiliers	23
6	Résultats du modèle de régression	23
7	Analyse des résidus	23

8 Graphiques de régression	24
8.0.1 Conclusion	24
8.1 Régressions linéaires et Statistiques sur les deux jeux de données	24
8.1.1 Statistiques Descriptives et Matrice de Corrélation	24
8.1.2 Statistiques Descriptives	27
8.1.3 Matrice de Corrélation	28
8.1.4 Interprétation des Résultats	28
8.1.5 Régression de la Surface et du Nombre d'Éoliennes	28
8.1.6 Régression de la Valeur Foncière Moyenne et de la Hauteur des Éoliennes	29
8.1.7 Régression de la Valeur Foncière Moyenne et du Nombre d'Éoliennes	29
8.1.8 Régression de la Valeur Foncière Moyenne et de la Puissance Moyenne des Éoliennes	30
8.2 Conclusion des Régressions	31
8.3 Synthèse Globale de la Partie II	31
9 Conclusion Globale du Projet	32
10 Réponse à la Problématique	32
11 Remerciements	33
12 Méthode de Travail	33

1 Introduction

The development of éoliennes in France is inscrit in an ambitious energy transition strategy, aimed at reducing the dépendance to fossil energy and increasing the part of renewable energy. Depuis les années 2000, la construction de parcs éoliens a considérablement augmenté, marquant l'engagement de la France envers les objectifs climatiques, notamment ceux fixés par l'Accord de Paris. Cependant, cette expansion soulève des interrogations quant à son impact sur les dynamiques locales, en particulier dans les zones rurales où se concentrent la majorité des installations.

Une question centrale émerge : **Les éoliennes ont-elles un impact sur le prix de l'immobilier en zone rurale ?** Cette interrogation reflète des enjeux à la fois environnementaux, sociaux et économiques. D'un côté, la proximité d'éoliennes peut engendrer des nuisances visuelles et sonores, ainsi que des inquiétudes sur leur impact environnemental, ce qui pourrait entraîner une dépréciation des biens immobiliers. De l'autre, elles symbolisent l'engagement environnemental et peuvent renforcer l'attractivité de certains territoires grâce à la création d'infrastructures locales et à la valorisation des énergies renouvelables.

Les zones rurales constituent un terrain d'étude particulièrement pertinent. Caractérisées par des marchés immobiliers souvent moins dynamiques, elles permettent d'analyser avec précision l'influence potentielle des éoliennes sur la valeur des propriétés. Cette problématique, qui se situe à l'intersection des enjeux de transition énergétique et d'aménagement territorial, revêt une importance croissante pour les décideurs locaux et les habitants concernés.

Pour répondre à cette question, ce projet s'appuie sur une démarche méthodologique en plusieurs étapes. Nous travaillerons tout d'abord sur les bases de données géographiques (SIG) afin de produire des cartes illustrant la répartition des éoliennes et des transactions immobilières dans les zones rurales étudiées. Cette phase permettra de visualiser les interactions spatiales entre les deux phénomènes. Ensuite, des statistiques descriptives seront réalisées sur ces données pour examiner les tendances générales, telles que la répartition géographique des prix, la densité des éoliennes ou les caractéristiques des maisons à proximité.

Enfin, une approche économétrique sera utilisée pour approfondir l'analyse. À travers des modèles statistiques, nous chercherons à déterminer si la proximité des éoliennes a un effet significatif sur les prix immobiliers, en tenant compte de variables explicatives telles que la surface, le nombre de pièces, la localisation géographique ou encore l'année de transaction.

L'objectif de ce projet est donc de déterminer si la présence d'éoliennes en zone rurale entraîne des variations significatives des prix immobiliers. Les résultats obtenus permettront de mieux comprendre les interactions entre la transition énergétique et les marchés immobiliers, tout en fournissant des éléments concrets pour éclairer les politiques publiques et les décisions locales. En définitive, ce projet cherche à répondre à une problématique clé : l'installation d'éoliennes constitue-t-elle une opportunité ou un frein pour les marchés immobiliers en zone rurale ?

2 Présentation des Données

Dans ce projet, nous avons utilisé plusieurs ensembles de données pour analyser la production d'énergie éolienne en France. Les données proviennent du site : https://files.georisques.fr/bdeole/production/csv/eolien_national.zip. Nous avons principalement les fichiers `aerogenerateur_national_20241221.csv`, `parc_national_20241221.csv` et `maisons_geo.csv`.

2.1 Données des Eoliennes

2.1.1 Données des Turbines Éoliennes

Le fichier `aerogenerateur_national_20241221.csv` contient les colonnes suivantes :

- **id_aerogenerateur** : Identifiant unique de l'aérogénérateur.
- **nom_eolienne** : Nom de l'éolienne.
- **id_parc** : Identifiant du parc éolien auquel l'éolienne appartient.
- **code_insee** : Code INSEE de la commune.
- **nom_commune** : Nom de la commune.
- **code_dept** : Code du département.
- **code_reg** : Code de la région.
- **puissance** : Puissance de l'éolienne en kW.
- **hauteur_totale** : Hauteur totale de l'éolienne.
- **hauteur_mat_nacelle** : Hauteur du mat de la nacelle.
- **diametre_rotor** : Diamètre du rotor.
- **cote_ngf** : Cote NGF (niveau général de référence).
- **periode_allumage** : Période d'allumage de l'éolienne.
- **type_feu** : Type de feu de l'éolienne.
- **constructeur** : Constructeur de l'éolienne.
- **reference_modele** : Référence du modèle de l'éolienne.
- **x_aerogenerateur** : Coordonnée X de l'éolienne.
- **y_aerogenerateur** : Coordonnée Y de l'éolienne.
- **epsg** : Code EPSG pour le système de projection.
- **libelle** : Libellé descriptif de l'éolienne.
- **date_maj** : Date de mise à jour des données.

Choix de cette base de données : Ces colonnes ont été choisies car elles fournissent des informations essentielles sur chaque aérogénérateur, permettant d'analyser leur répartition géographique, leur puissance et d'autres caractéristiques techniques.

2.1.2 Données des Parcs Éoliens

Le fichier `parc_national_20241221.csv` contient les colonnes suivantes :

- **id_parc** : Identifiant unique du parc éolien.
- **nom_usuel** : Nom usuel du parc.
- **puissance_parc** : Puissance totale du parc en kW.

- **code_insee** : Code INSEE de la commune.
- **nom_commune** : Nom de la commune.
- **code_dept** : Code du département.
- **code_reg** : Code de la région.
- **date_depot_demande_autorisation** : Date de dépôt de la demande d'autorisation.
- **date_delivrance_avis_autorite_environnementale** : Date de délivrance de l'avis.
- **bfdate_delivrance_autorisation** : Date de délivrance de l'autorisation.
- **date_debut_construction** : Date de début de construction du parc.
- **date_mise_en_service** : Date de mise en service du parc.
- **date_debut_demantlement** : Date de début de démantèlement du parc.
- **etat_parc** : État actuel du parc.
- **statut_admin** : Statut administratif du parc.
- **statut_parc** : Statut opérationnel du parc.
- **date_arrete_prefectoral_complementaire_renouvellement** : Date de l'arrêté préfectoral.
- **renouvellement** : Indicateur de renouvellement.
- **nom_exploitant** : Nom de l'exploitant du parc.
- **siret** : Numéro SIRET de l'exploitant.
- **date_maj** : Date de mise à jour des données.

Choix de cette base de données : Ces colonnes sont cruciales pour comprendre la structure et l'historique des parcs éoliens, ainsi que leur impact environnemental et administratif.

2.2 Données des Maisons

Le fichier `maisons_geo.csv` contient les colonnes suivantes :

- **jour_mutation** : Jour de la mutation.
- **mois_mutation** : Mois de la mutation.
- **annee_mutation** : Année de la mutation.
- **valeur_fonciere** : Valeur foncière de la propriété.
- **code_postal** : Code postal de la propriété.
- **nom_commune** : Nom de la commune.
- **code_departement** : Code du département.
- **id_parcelle** : Identifiant de la parcelle.
- **surface_carrez_total** : Surface totale en loi Carrez.
- **surface_reelle_bati_total** : Surface réelle bâtie totale.
- **surface_terrain_total** : Surface totale du terrain.
- **nombre_lots** : Nombre de lots dans la propriété.
- **nombre_maisons** : Nombre de maisons sur la parcelle.
- **surface_carrez_maisons** : Surface Carrez des maisons.
- **surface_reelle_bati_maisons** : Surface réelle bâtie des maisons.
- **surface_terrain_maisons** : Surface du terrain des maisons.

- **nombre_dependances** : Nombre de dépendances.
- **surface_carrez_dependances** : Surface Carrez des dépendances.
- **surface_reelle_bati_dependances** : Surface réelle bâtie des dépendances.
- **surface_terrain_dependances** : Surface du terrain des dépendances.
- **nombre_pieces_principales** : Nombre de pièces principales.
- **nature_culture** : Nature de la culture.
- **nature_culture_speciale** : Nature de la culture spéciale.
- **id_mutation** : Identifiant de la mutation.
- **geometry** : Géométrie de la parcelle.
- **distance.gare** : Distance à la gare.
- **distance.mairie** : Distance à la mairie.
- **distance.ecoles** : Distance aux écoles.
- **distance.magasins** : Distance aux magasins.
- **distance.primaire** : Distance à l'école primaire.
- **distance.college** : Distance au collège.
- **distance.littoral** : Distance au littoral.

Choix de cette base de données : Ces colonnes ont été sélectionnées pour leur pertinence dans l'analyse des propriétés foncières, en particulier dans les zones rurales, et pour comprendre l'impact des éoliennes sur le marché immobilier.

2.3 Prétraitement des Données

Pour préparer les données pour l'analyse, plusieurs étapes de prétraitement ont été effectuées :

2.3.1 Jointure des Données Éoliennes

Les données des aérogénérateurs et des parcs ont été reliées à l'aide d'une jointure gauche, en supprimant certaines colonnes inutiles pour simplifier l'analyse :

```
eolienne <- left_join(aerogenerateur_data, parc_data, by = "id_parc") %>%
  select(-siret, -nom_exploitant, -date_depot_demande_autorisation,
         -date_delivrance_avis_autorite_environnementale, -date_delivrance_autorite,
         -date_arrete_prefectoral_complementaire_renouvellement, -date_debut_demande)
```

2.3.2 Filtrage des Données Éoliennes

Les aérogénérateurs situés dans les DOM-TOM ont été exclus, et les coordonnées ont été ajustées pour garantir leur précision :

```
aerogenerateur_data <- aerogenerateur_data %>%
  filter(!code_dept %in% c("971", "974")) %>%
  mutate(
    x_aerogenerateur = ifelse(x_aerogenerateur > 1.1e7, x_aerogenerateur / 10, x_aerogenerateur),
    y_aerogenerateur = ifelse(y_aerogenerateur > 1.1e7, y_aerogenerateur / 10, y_aerogenerateur))
  select(id_aerogenerateur, id_parc, code_dept, code_insee, puissance, x_aerogenerateur, y_aerogenerateur)
```

2.3.3 Filtrage des Données des Maisons

Les zones non rurales ont été exclues de l'ensemble de données des maisons pour se concentrer sur les propriétés pertinentes :

```
maison_data <- maison_data %>%
  filter(code_departement %in% c('75', '92', '93', '94', '91',
  '78', '69', '13', '31', '33', '44',
  '59', '62', '77', '68', '67', '83'))
```

2.4 Conclusion

Ce projet a permis de rassembler et d'analyser des données sur les éoliennes et les propriétés foncières en France. Les étapes de prétraitement et de sélection des colonnes ont été essentielles pour garantir la qualité et la pertinence des données utilisées dans l'analyse.

3 Difficultés rencontrées et traitement des données

La gestion des coordonnées a été particulièrement difficile, ce qui a considérablement entravé le bon déroulement de mon projet. De plus, j'ai rencontré d'autres problèmes, tels que la lourdeur des données et des difficultés liées aux bibliothèques, qui ont également impacté l'avancement de mon analyse.

3.1 Problèmes liés à RStudio

Une des premières difficultés rencontrées concernait le fonctionnement de RStudio mais aussi avec QGIS. Lors de l'ajout de nouvelles librairies, le logiciel insistait souvent pour effectuer des mises à jour, ce qui ralentissait considérablement les sessions de travail. À cela s'est ajouté un bug persistant, nécessitant la désinstallation et la réinstallation de l'application, ce qui a conduit à une perte de temps significative (environ une journée de travail).

3.2 Gestion des Coordonnées GPS des Éoliennes

La gestion des coordonnées GPS des éoliennes a présenté plusieurs défis. Tout d'abord, il était nécessaire de diviser les valeurs des coordonnées x par 10 ou 100 selon les cas, afin de les adapter au système de référence approprié. Les coordonnées devaient également être transformées en Lambert93, un système de projection couramment utilisé en France.

Pour ce faire, j'ai consulté une carte disponible sur le site geofree.fr, qui indiquait que les valeurs de x devaient se situer entre 100000 et 1300000, tandis que les valeurs de y devaient être comprises entre 5900000 et 7300000.

Cependant, une erreur est survenue, indiquant qu'il n'existe pas de colonnes `x_aerogenerateur` et `y_aerogenerateur` dans `aerogenerateur_sf`. Cela a nécessité une extraction des coordonnées à partir de l'objet `geometry` :

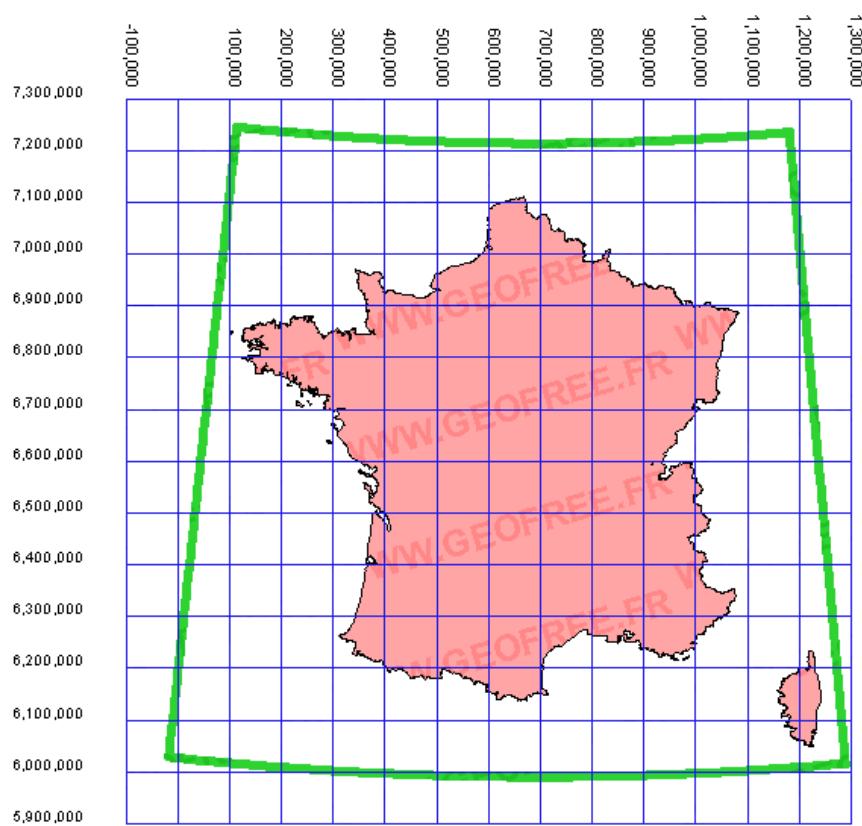


FIGURE 1 – Carte Lambert 93 indiquant les plages de coordonnées acceptables.

```

aerogenerateur_sf <- aerogenerateur_sf %>%
  mutate(
    coords = st_coordinates(geometry), # Extraire les coordonnées
    x_aerogenerateur = coords[, 1], # Coordonnée x
    y_aerogenerateur = coords[, 2] # Coordonnée y
  )

```

Il est important de noter que, bien que les coordonnées de latitude et de longitude semblent correctes, elles ne représentent pas fidèlement la position des éoliennes en France. En effet, si j'utilise uniquement les coordonnées de la France, il n'y aura aucun point représentant les éoliennes sur la carte. L'objectif était d'étendre la carte de la France pour inclure ces données, mais les coordonnées initiales n'étaient pas correctes.

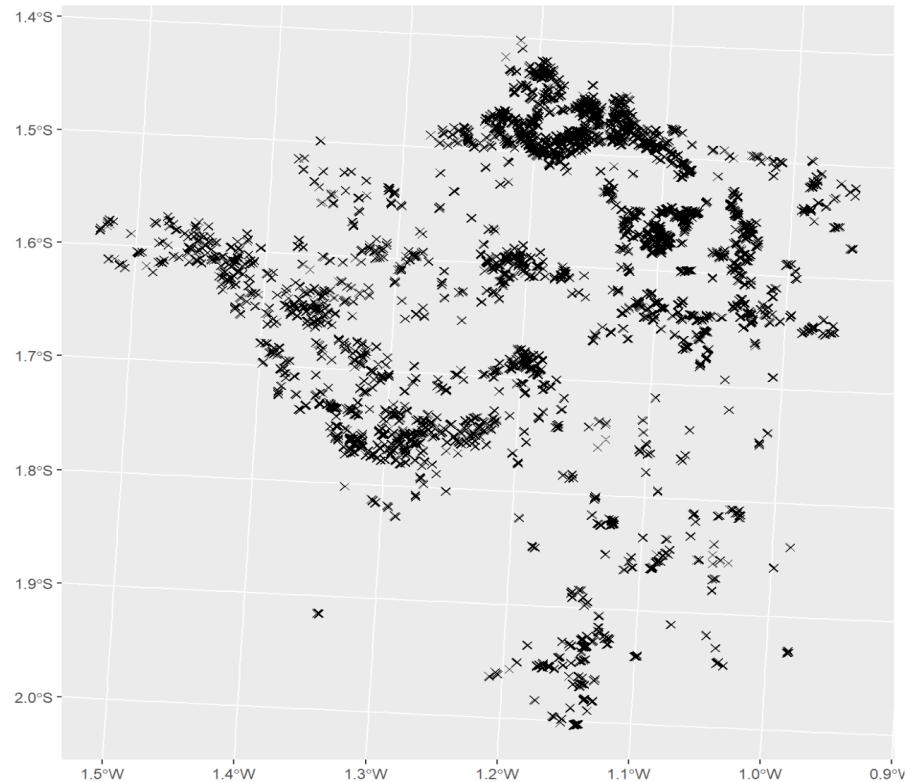


FIGURE 2 – Visualisation des éoliennes sur la carte.

En conclusion, bien que la carte Lambert93 représente bien la France, les coordonnées GPS des éoliennes nécessitaient des ajustements significatifs pour être correctement intégrées dans le système de référence approprié.

3.3 Traitement des données `maisonsgeo.csv`

Le jeu de données `maisonsgeo.csv`, contenant des informations sur les ventes immobilières, a été

Colonnes et séparateurs : Le fichier Excel était mal structuré, avec des colonnes mal délimitées et utilisant des virgules au lieu de points-virgules comme séparateurs. Il a fallu corriger manuellement cette structure en définissant clairement les noms des colonnes et en préparant un code R capable de séparer correctement les données.

Variables utilisées : Parmi les nombreuses variables disponibles (surface, distances aux littoraux, distances aux gares, jour et année de mutation, valeur foncière, code postal, nom des communes, etc.), seules celles directement pertinentes pour l'analyse ont été conservées, telles que la surface habitable, le nombre de pièces, et les coordonnées géographiques.

Problèmes liés à la géométrie : La colonne contenant des données géométriques était particulièrement problématique, car elle nécessitait une reformulation pour être exploitée correctement dans les outils d'analyse géographique. Elles possédaient énormément de caractères spéciaux, lorsque nous essayons de les séparer la latitude est bonne mais la longitude aura comme valeur NA.

3.4 Solutions envisagées pour optimiser le temps d'exécution

Afin de limiter ces temps d'exécution, plusieurs ajustements ont été mis en œuvre :

- **Réduction du volume de données** : Des colonnes jugées inutiles pour l'analyse ont été supprimées. Par exemple, dans MaisonGeo, les colonnes relatives aux distances littorales, gares, écoles, et commerces ont été exclues pour alléger le traitement.
- **Filtrage géographique préalable** : Les données ont été restreintes à la France métropolitaine, excluant les Dom-Tom et se concentrant uniquement sur les départements métropolitains d'intérêt. Cela a permis de réduire le nombre de lignes dans les fichiers.
- **Optimisation des scripts R** : Les scripts ont été ajustés pour minimiser les opérations redondantes. Les corrections des coordonnées Lambert 93, par exemple, ont été regroupées en une seule étape au lieu de multiples itérations.
- **Exécution par lots** : Les tâches les plus longues ont été découpées en sous-tâches exécutées séparément, permettant un suivi plus précis et évitant les blocages prolongés en cas d'erreur.

Malgré ces efforts, le temps d'exécution reste un facteur limitant pour des analyses à grande échelle comme celle-ci. Ce constat met en évidence l'importance de disposer d'outils performants et de bases de données bien structurées dès le départ pour assurer une progression fluide du projet.

3.5 Conclusion

En conclusion, l'absence de coordonnées précises pour les éoliennes a considérablement entravé la réalisation d'une carte fiable et l'exécution d'analyses de ré-

gression valables. Malgré mes efforts pour ajuster les données et les intégrer dans le système de référence approprié, les coordonnées initiales se sont révélées inexactes, ce qui a conduit à des visualisations erronées et à des résultats d'analyse peu concluants. En utilisant les bases de données imposées pour ce projet, il est devenu évident que la qualité des données géographiques est cruciale pour obtenir des résultats significatifs. Sans des coordonnées correctes, il n'a pas été possible de tirer des conclusions pertinentes sur l'impact des éoliennes sur les prix de l'immobilier ou d'autres variables d'intérêt. Ce projet souligne l'importance de la précision des données géographiques dans toute analyse spatiale.

4 PARTIE I : SIG

Dans le cadre de ce projet, nous avons utilisé des outils de Système d'Information Géographique (SIG) pour visualiser l'impact des éoliennes sur les prix de l'immobilier. La cartographie choroplèthe, une technique couramment utilisée dans ce domaine, permet de représenter des données géographiques à l'aide de couleurs ou de dégradés. Chaque zone géographique, telle qu'une commune ou un département, se voit attribuer une couleur en fonction de la valeur d'une variable, ce qui facilite la visualisation des différences spatiales.

Les cartes choroplèthes réalisées dans ce projet illustrent la répartition des prix de l'immobilier et l'implantation des parcs éoliens, permettant ainsi de mettre en évidence d'éventuelles corrélations géographiques entre la présence d'éoliennes et l'évolution des prix dans différentes zones. Cette approche cartographique offre une vue d'ensemble claire des dynamiques économiques et géographiques du département étudié.

Cependant, comme expliqué dans la section *Difficultés rencontrées*, la réalisation de ces cartes n'a pas été possible. Les raisons techniques et méthodologiques seront détaillées dans cette section.

4.1 Difficultés detaillées

Tout d'abord, j'ai constaté que la manipulation des coordonnées fournies dans mes bases de données n'étaient pas correctes. Cela a rendu impossible tout lien entre les éoliennes, les départements ou les maisons, car les données géographiques ne correspondaient pas. Ce problème était d'autant plus frustrant. Lorsque j'ai tenté de diviser ou de corriger la colonne `geometry` en R, cela n'a donné aucun résultat exploitable. J'ai donc cherché des solutions en regardant des vidéos sur YouTube et en consultant des sites spécialisés, mais malgré mes efforts et de nombreux tests, je n'ai pas réussi à résoudre ce problème.

Par ailleurs, j'ai essayé d'utiliser les TP fournis par le professeur en atelier économétrique. Cependant, les données disponibles dans ces TP étaient soit limitées à la région Occitanie, soit non directement liées à mon projet. Cette restriction ne permettait pas d'analyser correctement mes bases de données. J'ai également tenté de travailler avec l'application QGIS pour produire une carte choroplèthe. J'ai essayé d'utiliser des outils comme la graduation, mais j'ai rencontré encore plus de difficultés qu'en R, rendant le résultat final inutilisable.

En R, j'ai malgré tout pu générer une carte, bien qu'elle ne soit absolument pas correcte. Que ce soit pour les maisons ou pour les éoliennes, les données représentées étaient incohérentes. J'ai tout fait pour éviter de m'appuyer exclusivement sur des outils comme ChatGPT et souhaitais réaliser ce travail par moi-même, en m'appuyant sur les TP de classe. Cependant, je suis restée bloquée malgré le temps important consacré à ces tentatives.

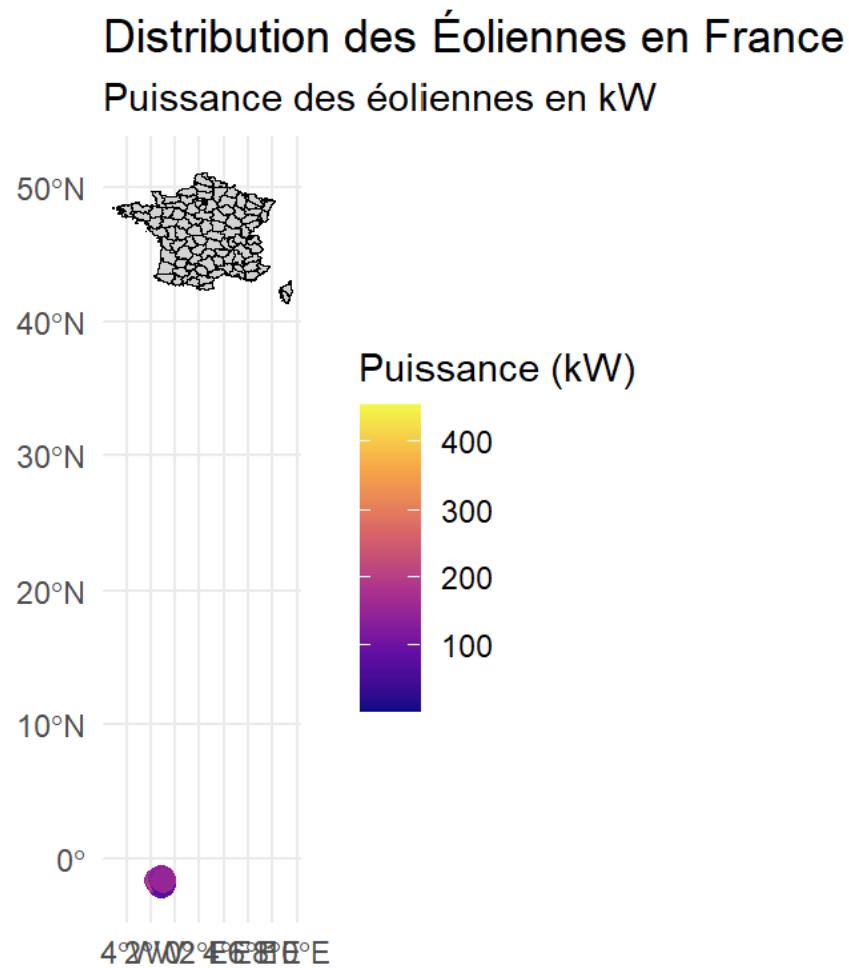


FIGURE 3 – Distribution de la puissance des éoliennes.

Dans la carte suivante, nous pouvons remarquer que les éoliennes ont toutes une puissance de moins de 20kW.

Pour conclure, je vais tout de même inclure la carte que j'ai réalisée, même si elle

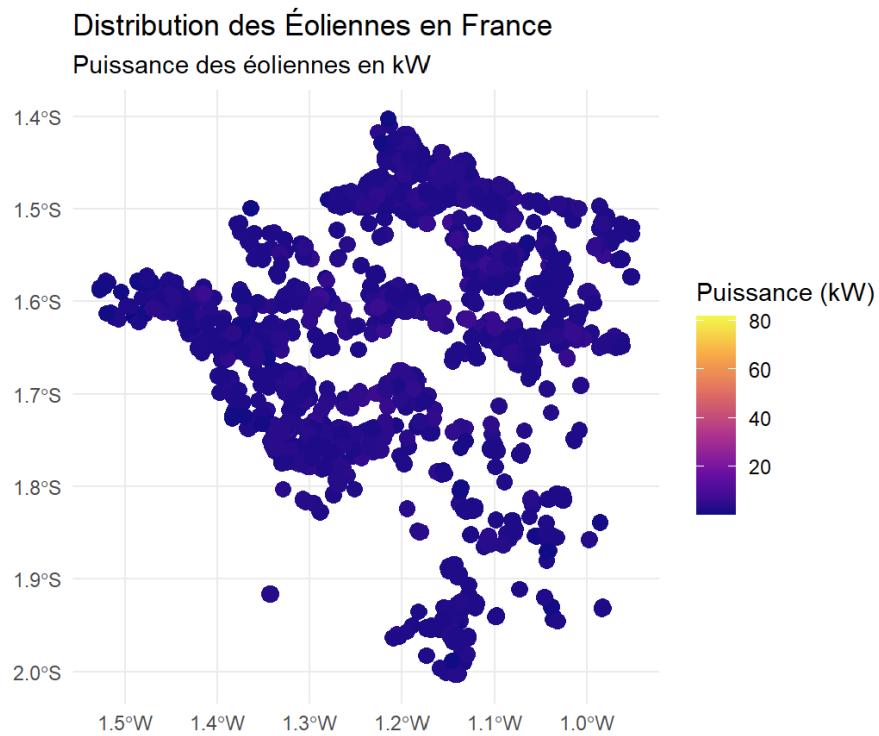


FIGURE 4 – Distribution de la puissance des éoliennes.

est incorrecte, afin de montrer que je maîtrise la méthode pour produire des cartes choroplèthes. Les problèmes proviennent uniquement des coordonnées incorrectes dans mes données et non de ma capacité à utiliser les outils SIG.

5 PARTIE II : ECONOMETRIE

5.1 Statistiques descriptives des éoliennes

Ces statistiques descriptives présentent une analyse des éoliennes en France à travers leurs caractéristiques physiques et de construction. Les données incluent des statistiques sur la puissance, la hauteur, et la durée de construction, ainsi que des graphiques pour visualiser ces éléments. L'objectif est d'explorer l'impact des éoliennes sur les prix de l'immobilier.

5.1.1 Puissance des Éoliennes

Statistique Valeur (kW)

Moyenne	2.74
Médiane	2.3
Minimum	0.03
Maximum	84
Écart-type	1.55

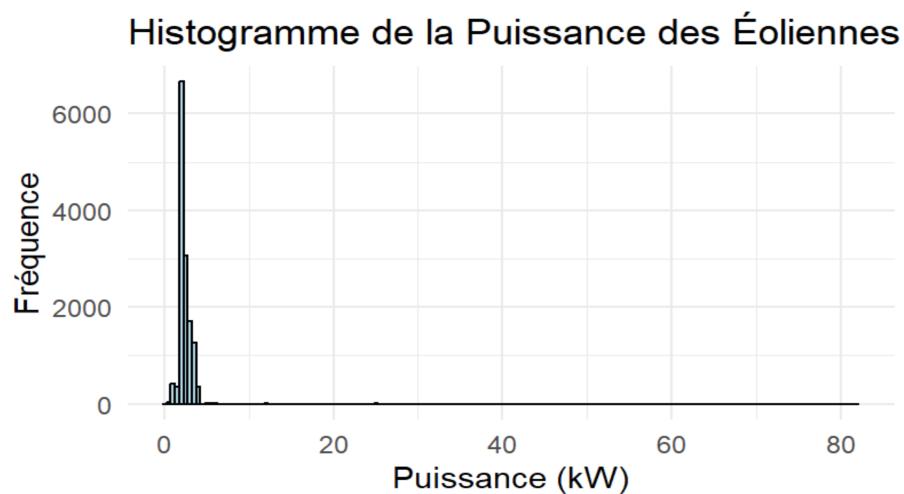


FIGURE 5 – Distribution de la puissance des éoliennes.

La puissance moyenne de 2,74 kW par éolienne illustre une capacité énergétique significative, adaptée aux besoins des installations modernes. La puissance maximale de 84 kW démontre le potentiel des technologies actuelles pour atteindre des niveaux élevés de production. La faible puissance minimale de 0,03 kW pourrait refléter la présence d'anciennes installations ou de modèles spécifiques adaptés à des usages réduits.

5.1.2 Hauteur des Éoliennes

Statistique Valeur (m)

Moyenne	145.98
Médiane	149
Minimum	0.8
Maximum	454
Écart-type	31.26

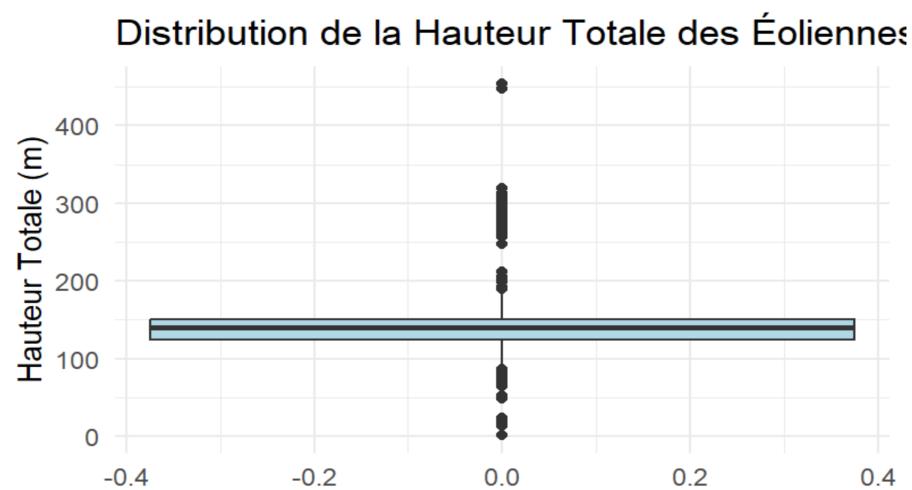


FIGURE 6 – Boîte à moustache représentant la répartition des hauteurs des éoliennes.

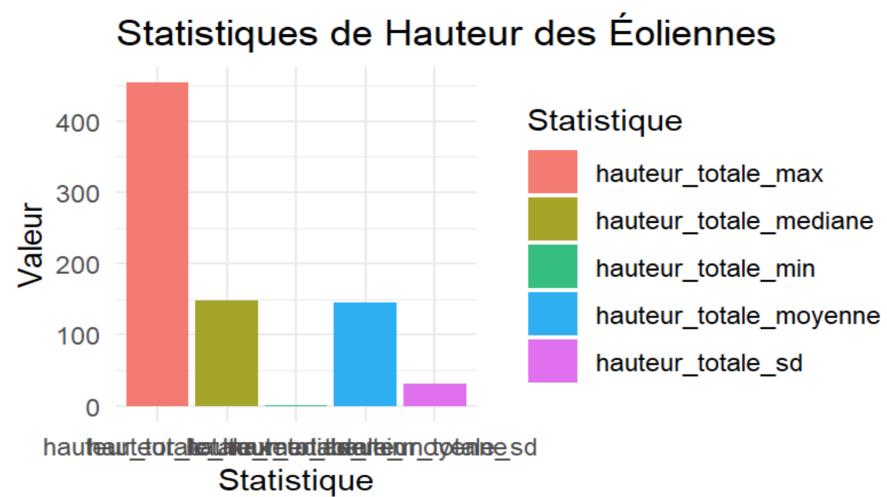


FIGURE 7 – Histogramme de la répartition des hauteurs des éoliennes.

Avec une hauteur moyenne de 146 mètres, les éoliennes étudiées montrent une tendance à augmenter en taille, ce qui correspond à la recherche d'une meilleure capture des vents en altitude. La hauteur maximale atteignant 454 mètres reflète des innovations significatives dans le secteur, visant à améliorer la production énergétique. La variation notable des hauteurs, avec un écart-type de 31,26 mètres, met en évidence la diversité des éoliennes utilisées selon les spécificités des sites.

5.1.3 Durée de Construction

Statistique Durée (jours)

Moyenne	365
Médiane	350
Minimum	0
Maximum	5938
Écart-type	257.38

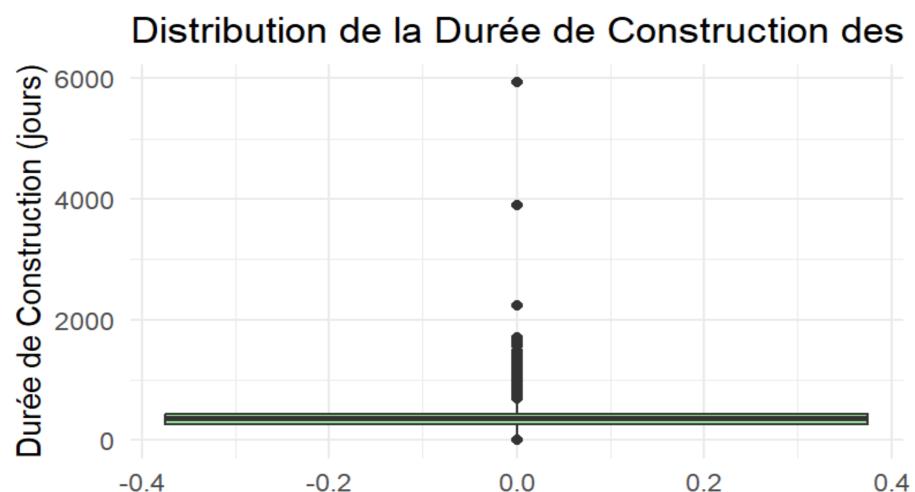


FIGURE 8 – Boîte à moustache représentant la durée moyenne de construction des éoliennes.

Avec une durée moyenne de 365 jours, la construction des éoliennes s'inscrit dans des projets de long terme nécessitant une planification rigoureuse. La durée maximale de 5938 jours illustre les défis logistiques et techniques rencontrés sur certains sites complexes. La large dispersion des durées de construction, allant de 0 à 5938 jours, souligne les différences selon les tailles, les technologies et les conditions locales.

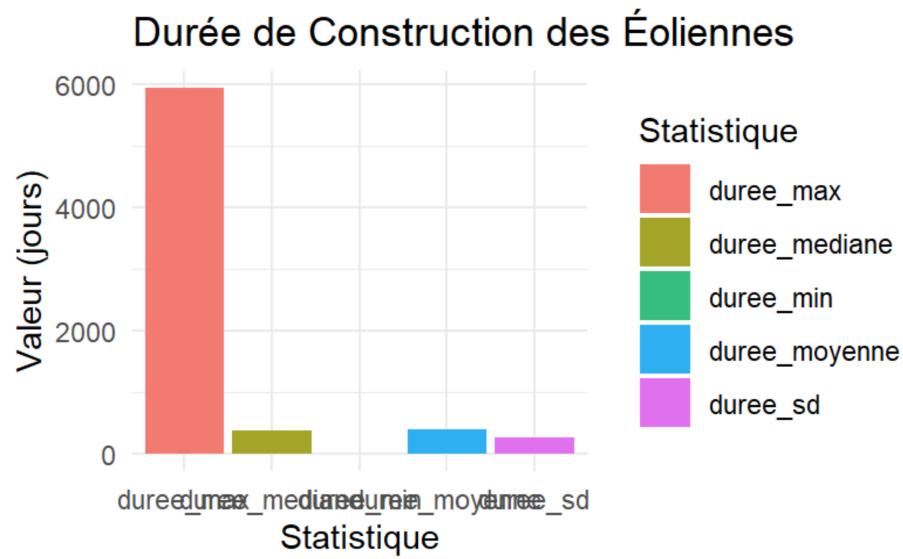


FIGURE 9 – Histogramme de la durée moyenne de construction des éoliennes.

5.1.4 Nombre d'Éoliennes par Parc

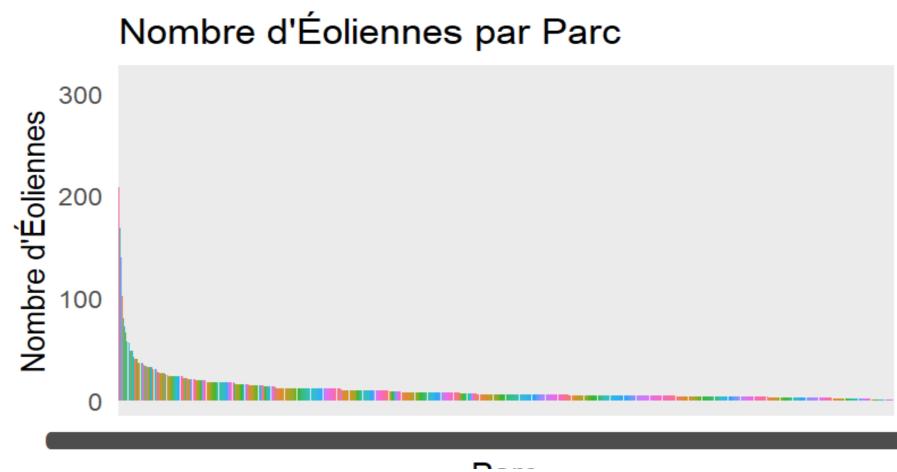


FIGURE 10 – Nombre moyen d'éoliennes par parc.

En moyenne, les parcs éoliens possèdent approximativement 5 à 6 éoliennes, seuls quelques parcs possèdent plus de 100 éoliennes en un parc.

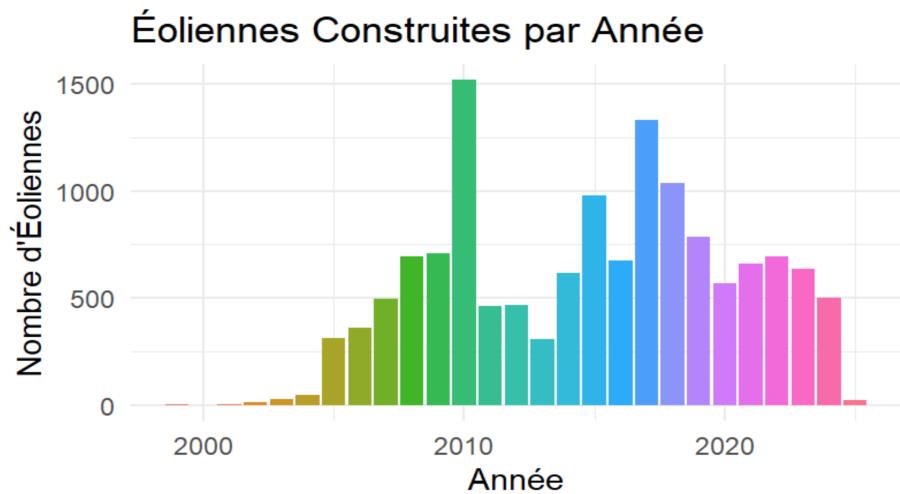


FIGURE 11 – Nombre d'éoliennes construites par année.

5.1.5 Éoliennes Construites par Année

En 2010, on observe une augmentation du nombre de construction des éoliennes avec plus de 1500 éoliennes alors que dans les années 2000, il n'y a eu très peu d'éoliennes construites. En 2024, nous avons construit, en France, environ 500 éoliennes.

5.1.6 Conclusion

Les graphiques et les statistiques présentés dans ce document permettent d'explorer les caractéristiques principales des éoliennes. Ces informations serviront à évaluer leur impact potentiel sur les prix de l'immobilier, dans le cadre de ce projet en cours.

5.2 Statistiques descriptives des biens immobiliers

Ces statistiques descriptives présentent une analyse des biens immobiliers, en mettant en évidence leurs caractéristiques principales : valeur foncière, surface, nombre de pièces principales, et distances aux infrastructures (gare, mairie, école). Les graphiques illustrent les tendances des données.

5.2.1 Mutations par Année

Année	Nombre de mutations
2014	192,207
2015	220,080
2016	234,888
2017	92,551

Le graphique montre une augmentation des mutations (ventes, donations...) entre 2014 et 2016, suivie d'une forte baisse en 2017. Cela peut indiquer des facteurs économiques ou des changements réglementaires.

5.2.2 Fréquence des Mutations par Mois

Mois	Jan	Fév	Mar	Avr	Mai	Juin	JUIL
Nombre	48,012	57,384	51,546	54,258	53,336	68,113	82,675

Le nombre totale de mutations est de 739726.

5.2.3 Valeur Foncière

Statistique	Valeur (€)
Moyenne	169,257.60
Médiane	142,000
Minimum	0
Maximum	3,300,000,000
Écart-type	3,862,403

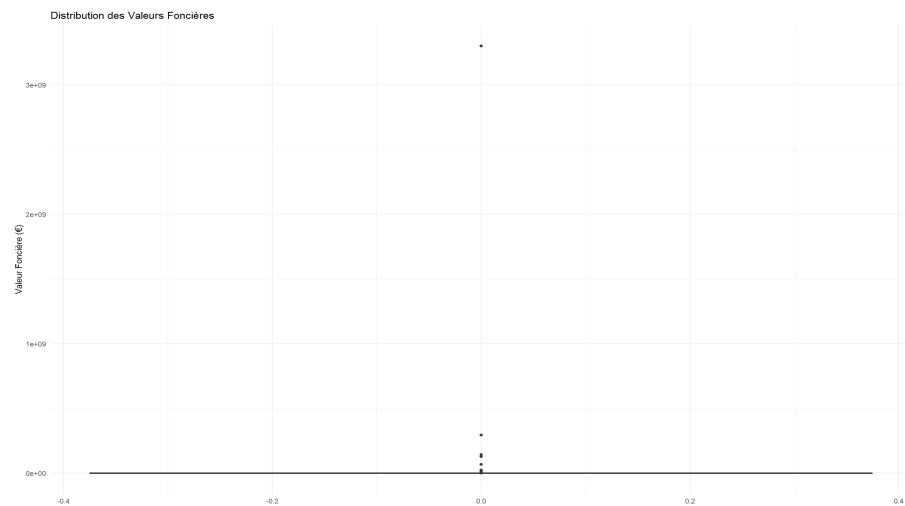


FIGURE 12 – Boîte à moustache des valeurs foncières.

La médiane des valeurs foncières est de 142,000 €, avec une distribution très

dispersée, comme l'illustre la boîte à moustache.

5.2.4 Analyse des villes avec le plus de ventes immobilières

Les données montrent les villes avec le plus grand nombre de transactions immobilières enregistrées :

- Le Mans : 2 947 ventes
- Brest : 2 295 ventes
- Amiens : 2 232 ventes
- La Rochelle : 2 003 ventes
- Quimper : 1 924 ventes
- Les Sables-d'Olonne : 1 897 ventes
- Nîmes : 1 797 ventes
- Angers : 1 755 ventes
- Saint-Quentin : 1 655 ventes
- Niort : 1 650 ventes

Ces villes, bien réparties sur le territoire français, illustrent une forte dynamique des marchés immobiliers locaux, en particulier dans des zones où la demande semble active. Les facteurs tels que l'attractivité économique, la localisation géographique et la présence d'infrastructures peuvent contribuer à expliquer ces volumes de transactions.

5.2.5 Surface Réelle Bâtie

Statistique	Valeur (m ²)
Moyenne	96.83
Médiane	90
Minimum	0
Maximum	7,626

La surface réelle bâtie des biens immobiliers varie considérablement, avec une moyenne de 96,83 m² et une médiane légèrement inférieure de 90 m², ce qui indique que la majorité des biens ont des surfaces proches de cette valeur. Cependant, l'écart entre le minimum de 0 m² (probablement des anomalies ou des données manquantes) et le maximum de 7 626 m² révèle la présence de biens exceptionnellement grands, susceptibles de tirer la moyenne vers le haut.

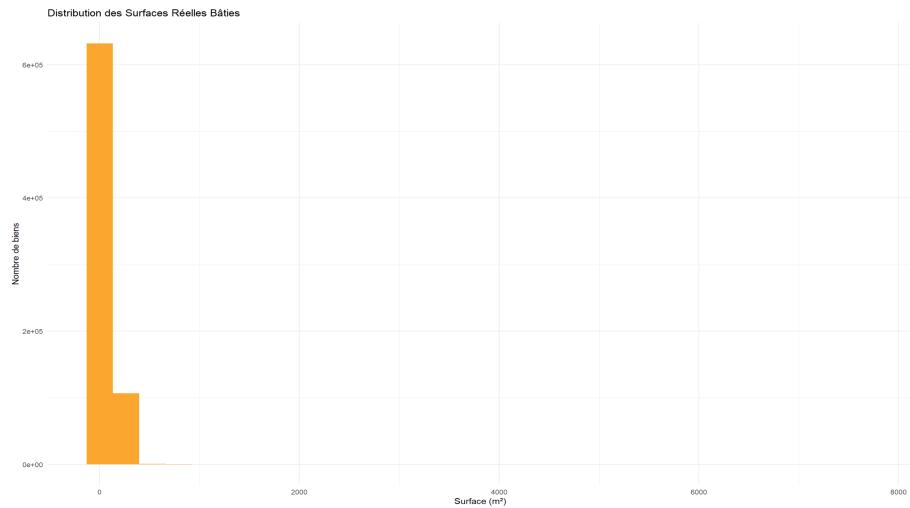


FIGURE 13 – Histogramme de la surface réelle bâtie.

5.2.6 Distances aux Infrastructures

Statistique	Distance Gare (km)	Distance Mairie (km)	Distance École (km)
Moyenne	NaN	7.32	1.09
Médiane	NA	4.61	0.80

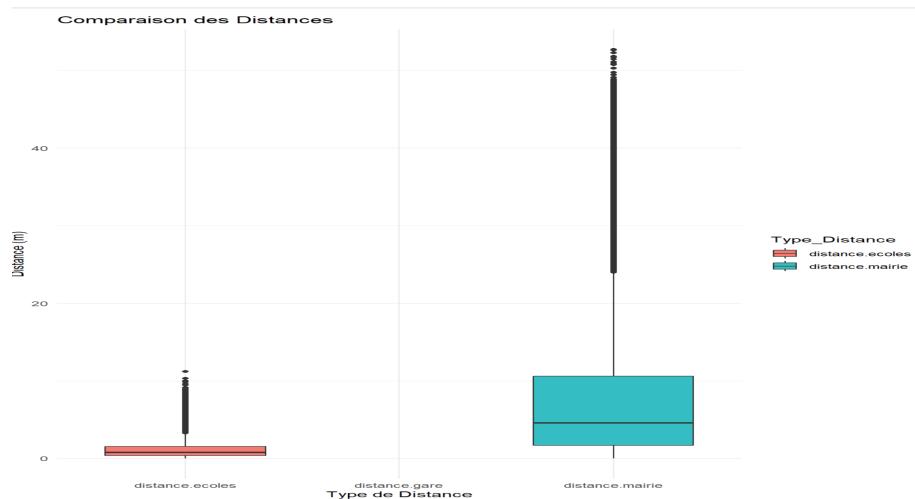


FIGURE 14 – Boîte à moustache comparant les distances des maisons aux écoles et aux gares.

Les distances moyennes des biens immobiliers aux principales infrastructures révèlent des écarts intéressants :

- Distance moyenne à la gare : 7,32 km (médiane : 4,61 km)
- Distance moyenne à la mairie : 1,09 km (médiane : 0,80 km)

- Distance moyenne aux écoles : NaN (médiane : NA)

Ces données mettent en évidence une relative proximité des biens immobiliers aux mairies et aux écoles, ce qui reflète une organisation urbaine où ces infrastructures sont bien réparties et accessibles. En revanche, les distances moyennes aux gares, plus importantes, peuvent s'expliquer par le caractère plus centralisé ou éloigné des réseaux ferroviaires pour certains biens situés dans des zones rurales ou périurbaines. Les valeurs manquantes (*NaN*) pour certaines distances aux écoles indiquent que des données doivent être vérifiées ou complétées pour une analyse plus précise.

5.2.7 Conclusion

Les données montrent des disparités importantes entre les biens immobiliers en termes de valeur, de surface et de distances aux infrastructures. Ces informations fournissent une base solide pour examiner l'impact des caractéristiques immobilières sur d'autres variables, telles que les prix ou la satisfaction des habitants.

5.3 Regression linéaire et Statistiques des biens immobiliers

Dans cette analyse, nous avons réalisé une régression linéaire sur le jeu de données `maison_geo.csv` pour expliquer la variable `log(valeur_fonciere)` en fonction de plusieurs variables explicatives, telles que le département, l'année de mutation, la surface du terrain et le nombre de pièces principales.

Le modèle de régression utilisé est le suivant :

$$\log(valeur_fonciere) = \beta_0 + \sum_{i=1}^n \beta_i \cdot X_i + \epsilon \quad (1)$$

où X_i sont les variables explicatives et ϵ l'erreur aléatoire.

6 Résultats du modèle de régression

Nous avons ajusté un modèle de régression linéaire avec les variables suivantes :

- `factor(code_departement)`
- `annee_mutation`
- `surface_terrain_maisons`
- `nombre_pieces_principales`

Les résultats du modèle sont résumés dans le tableau suivant :

Le modèle montre que toutes les variables sont significativement liées à la `log(valeur_fonciere)` avec des p-valeurs très faibles, ce qui indique une forte significativité statistique.

Le R^2 multiple est de 0.3048, ce qui signifie que 30.48% de la variabilité de la `log(valeur_fonciere)` est expliquée par les variables du modèle. L'erreur standard des résidus est de 0.6931.

7 Analyse des résidus

Les résidus du modèle sont présentés ci-dessous. Les statistiques des résidus montrent que la médiane est proche de zéro et que les résidus sont plutôt symétriques.

triques, avec une légère concentration autour de zéro.

- Minimum : -20.1363
- Premier quartile : -0.2585
- Médiane : 0.0885
- Troisième quartile : 0.3792
- Maximum : 9.7737

8 Graphiques de régression

Voici les graphiques des nuages de points pour certaines des variables du modèle, en relation avec la `log(valeur_fonciere)` :

8.0.1 Conclusion

L'analyse de régression montre que plusieurs facteurs, notamment le département, l'année de mutation, la surface du terrain et le nombre de pièces principales, influencent de manière significative la valeur foncière des biens immobiliers. Les résultats indiquent une forte relation entre la superficie et la valeur foncière, ainsi qu'une variabilité importante selon les départements.

8.1 Régressions linéaires et Statistiques sur les deux jeux de données

8.1.1 Statistiques Descriptives et Matrice de Corrélation

Nous présentons ici les statistiques descriptives des variables utilisées dans l'analyse, ainsi que la matrice de corrélation pour évaluer les relations entre ces variables.

Variable	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-4.577	1.709	-2.678	0.007416**
factor(code_departement)10	-0.463	0.0101	-45.754	< 2e-16***
factor(code_departement)11	-0.408	0.0081	-50.527	< 2e-16***
:	:	:	:	:
annee_mutation	0.007819	0.000848	9.220	< 2e-16***
surface_terrain_maisons	0.00005371	0.000000777	69.104	< 2e-16***
nombre_pieces_principales	0.2063	0.000567	363.754	< 2e-16***

TABLE 1 – Résumé des résultats du modèle de régression

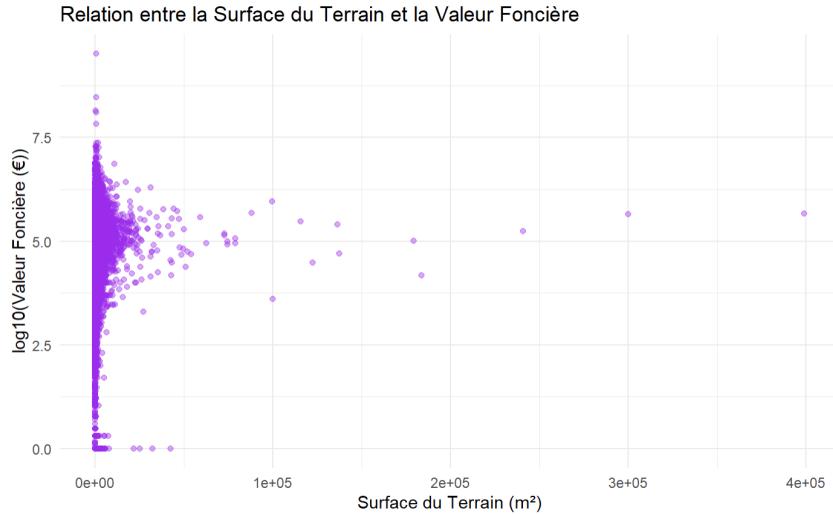


FIGURE 15 – Relation entre $\log(\text{valeur_fonciere})$ et la surface du terrain. La relation est légèrement positive.

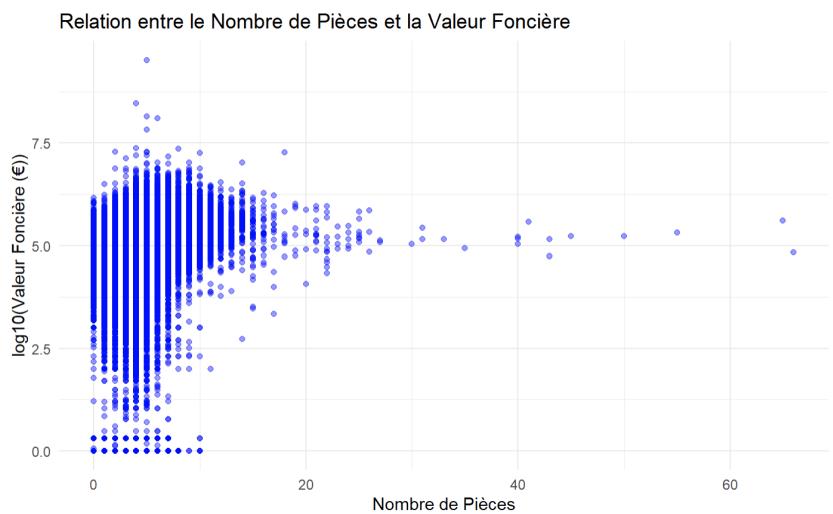


FIGURE 16 – Relation entre le nombre de pièces et $\log(\text{valeur_fonciere})$. La relation est légèrement positive.

Relation entre le Département et la Valeur Foncière

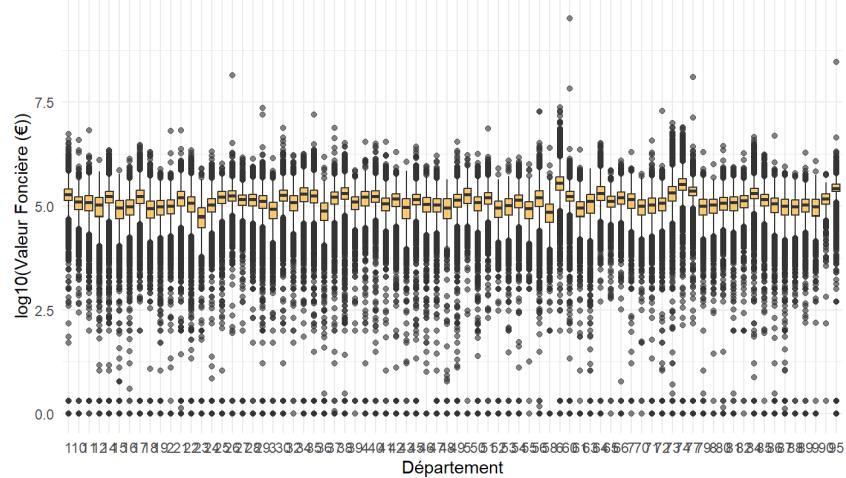


FIGURE 17 – Relation entre $\log(\text{valeur_fonciere})$ et les départements. Les graphiques sous forme de boîtes à moustaches montrent que la valeur foncière est concentrée autour de 5.0 log euros.

Relation entre l'Année de Mutation et la Valeur Foncière

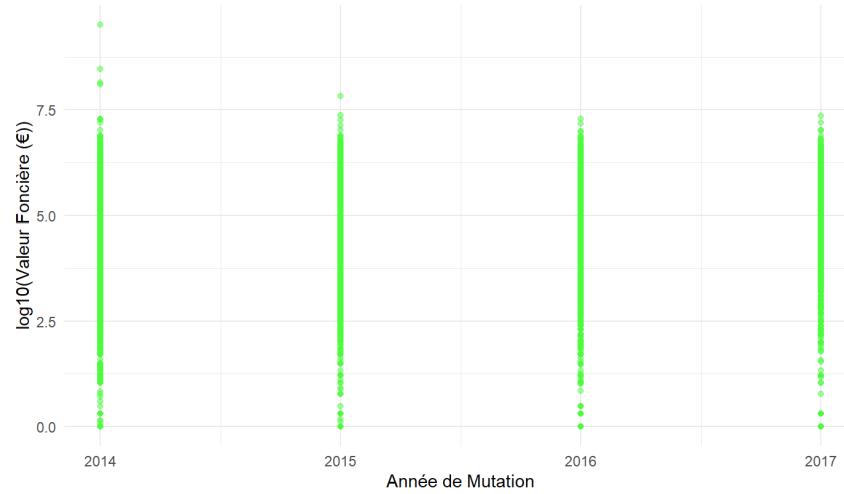


FIGURE 18 – Relation entre $\log(\text{valeur_fonciere})$ et l'année de mutation (2014-2017). Les points sont serrés et montrent une faible dispersion.

8.1.2 Statistiques Descriptives

Les statistiques descriptives des principales variables utilisées dans notre analyse sont les suivantes :

Variable	Statistique
valeur_fonciere_moyenne	Min. : 67271 1er Quart. : 108005 Médiane : 134562 Moyenne : 155805 3e Quart. : 176028 Max. : 463651
surface_carrez_moyenne	Min. : 0.1926 1er Quart. : 0.7037 Médiane : 1.2412 Moyenne : 2.2998 3e Quart. : 3.1155 Max. : 19.9813
nombre_mutations	Min. : 1302 1er Quart. : 5491 Médiane : 8794 Moyenne : 9733 3e Quart. : 12120 Max. : 27802
puissance_moyenne	Min. : 1.000 1er Quart. : 2.002 Médiane : 2.241 Moyenne : 2.245 3e Quart. : 2.486 Max. : 3.125
hauteur_totale_moyenne	Min. : 95.38 1er Quart. : 123.89 Médiane : 136.52 Moyenne : 135.36 3e Quart. : 147.11 Max. : 193.00
nombre_eoliennes	Min. : 2.0 1er Quart. : 41.0 Médiane : 101.0 Moyenne : 197.5 3e Quart. : 233.0 Max. : 1597.0

TABLE 2 – Statistiques descriptives des variables

8.1.3 Matrice de Corrélation

La matrice de corrélation ci-dessous montre les relations entre les différentes variables étudiées. Les coefficients de corrélation sont compris entre -1 et 1, où des valeurs proches de 1 ou -1 indiquent une relation forte, et des valeurs proches de 0 indiquent une absence de relation.

	valeur_fonciere_moyenne	puissance_moyenne	surface_carrez_moyenne	nombre_eoliennes
valeur_fonciere_moyenne	1.0000	-0.1446		
puissance_moyenne		1.0000		
surface_carrez_moyenne			0.5852	-0.2093
nombre_eoliennes				0.2363

TABLE 3 – Matrice de corrélation des variables

8.1.4 Interprétation des Résultats

Les résultats obtenus dans la matrice de corrélation indiquent les relations suivantes :

- Il existe une corrélation modérée (0.5852) entre la *valeur foncière moyenne* et la *surface Carrez moyenne*, suggérant qu'en moyenne, les propriétés plus grandes tendent à avoir des valeurs foncières plus élevées.
- La *puissance moyenne des éoliennes* présente une corrélation négative avec la *valeur foncière moyenne* (-0.1446), ce qui pourrait suggérer qu'une puissance plus élevée des éoliennes est associée à une diminution légère de la valeur foncière.
- Le *nombre d'éoliennes* est faiblement corrélé à la *valeur foncière moyenne* (0.0623) et à la *surface Carrez moyenne* (-0.0772), indiquant une faible relation entre ces variables.
- La *puissance moyenne des éoliennes* présente une faible corrélation positive avec le *nombre d'éoliennes* (0.2363), ce qui indique que dans les zones avec plus d'éoliennes, leur puissance moyenne tend à être légèrement plus élevée.

8.1.5 Régression de la Surface et du Nombre d'Éoliennes

Dans cette régression, nous avons exploré la relation entre la *surface* des biens immobiliers et le *nombre d'éoliennes*. Le coefficient de régression pour cette combinaison est légèrement négatif. Cela suggère une légère relation inverse entre la surface des biens et le nombre d'éoliennes dans les zones concernées. Cependant, ce résultat ne permet pas de conclure à une relation forte, et il est important de noter que d'autres variables non prises en compte dans ce modèle pourraient influencer cette relation.

L'axe des ordonnées représente la valeur foncière moyenne et l'axe des abscisses représente la surface des biens immobiliers et le nombre d'éoliennes. La légère pente négative indique qu'avec l'augmentation du nombre d'éoliennes, la surface moyenne des propriétés tend à diminuer légèrement, bien que la corrélation reste faible.

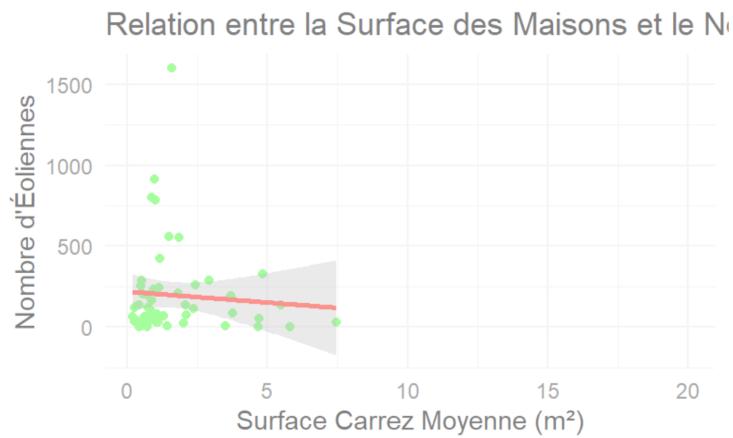


FIGURE 19 – Régression de la Surface et du Nombre d’Éoliennes

8.1.6 Régression de la Valeur Foncière Moyenne et de la Hauteur des Éoliennes

La régression entre la *valeur foncière moyenne* et la *hauteur des éoliennes* a montré un coefficient de régression négatif substantiel. Cela indique que, dans les données analysées, une augmentation de la hauteur des éoliennes semble être associée à une diminution de la valeur foncière des biens. Ce résultat est surprenant, car intuitivement, une plus grande hauteur pourrait être perçue comme un indicateur de plus grande performance ou efficacité des éoliennes, ce qui pourrait, dans certains cas, entraîner une augmentation de la valeur foncière. Cependant, dans notre cas, le modèle montre l'inverse, ce qui pourrait être dû à des facteurs locaux, comme la présence d'éoliennes massives ou visibles de manière perturbante, ce qui pourrait affecter négativement la perception des acheteurs potentiels.

La pente négative marquée suggère une relation inverse entre la hauteur des éoliennes et la valeur foncière.

8.1.7 Régression de la Valeur Foncière Moyenne et du Nombre d’Éoliennes

La régression entre la *valeur foncière moyenne* et le *nombre d'éoliennes* n'a montré aucune relation significative. Le coefficient de régression est proche de zéro, ce qui suggère qu'il n'existe pas de lien direct entre le nombre d'éoliennes et la valeur foncière des biens immobiliers dans les zones étudiées. Ce résultat pourrait indiquer que les éoliennes, bien qu'elles soient présentes en grand nombre dans certaines zones, n'ont pas d'impact mesurable sur les prix des propriétés. Il est possible que d'autres

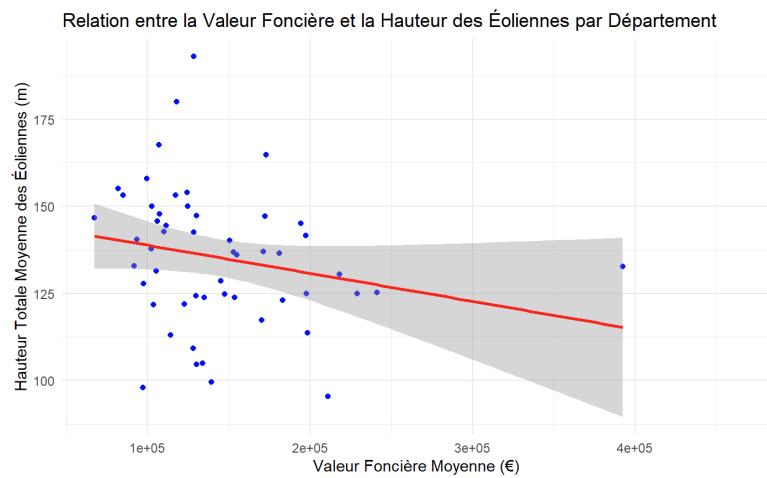


FIGURE 20 – Régression de la Valeur Foncière Moyenne et de la Hauteur des Éoliennes

facteurs, comme la proximité de services ou d'infrastructures, jouent un rôle plus important dans la détermination de la valeur foncière.

L'absence de tendance claire est évidente dans le graphique suivant.

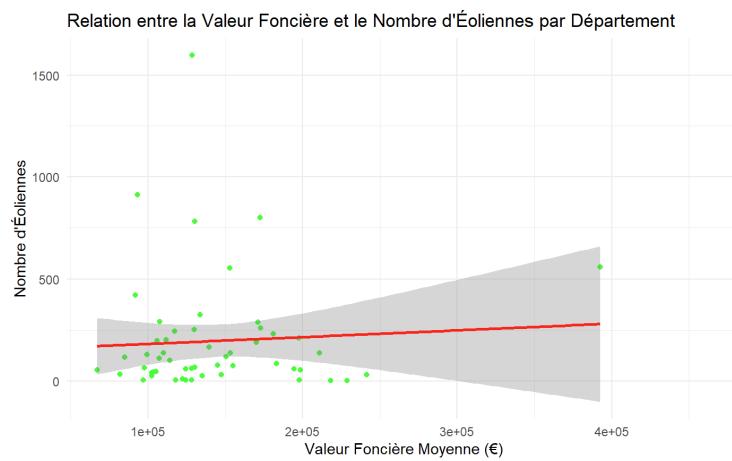


FIGURE 21 – Régression de la Valeur Foncière Moyenne et du Nombre d'Éoliennes

8.1.8 Régression de la Valeur Foncière Moyenne et de la Puissance Moyenne des Éoliennes

Enfin, la régression entre la *valeur foncière moyenne* et la *puissance moyenne* des éoliennes a également montré un coefficient négatif. Cela suggère que dans les zones où la puissance moyenne des éoliennes est plus élevée, la valeur foncière des biens tend à être plus faible. Ce résultat peut être lié à des perceptions négatives vis-à-vis

de l'impact visuel ou sonore des éoliennes de grande puissance. De plus, il pourrait y avoir une corrélation avec d'autres facteurs comme la densité de l'installation des éoliennes ou des considérations environnementales qui influencent la décision d'achat des propriétés dans ces zones.

La relation négative entre la puissance des éoliennes et la valeur foncière, indiquant une baisse de la valeur des propriétés à mesure que la puissance des éoliennes augmente.

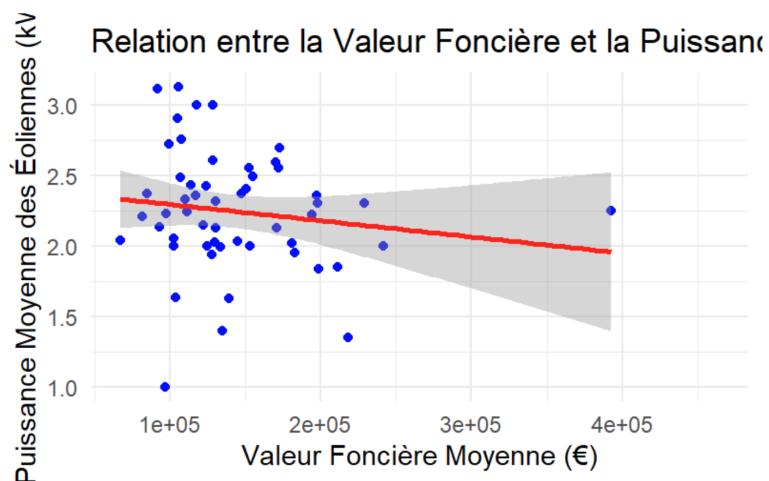


FIGURE 22 – Régression de la Valeur Foncière Moyenne et de la Puissance Moyenne des Éoliennes

8.2 Conclusion des Régressions

Les résultats des régressions indiquent des relations variables entre les caractéristiques des éoliennes et la valeur foncière des biens immobiliers. En résumé :

- La régression de la surface et du nombre d'éoliennes montre une légère relation négative, mais faible.
 - La régression de la valeur foncière moyenne et de la hauteur des éoliennes présente une relation inverse marquée, ce qui suggère que des éoliennes plus grandes pourraient avoir un effet négatif sur la perception des propriétés.
 - La régression de la valeur foncière et du nombre d'éoliennes n'indique pas de relation significative.
 - La régression entre la valeur foncière et la puissance des éoliennes montre également une relation négative.

8.3 Synthèse Globale de la Partie II

Dans l'ensemble, les analyses économétriques menées dans ce projet ont permis de dégager plusieurs tendances, mais ont également mis en lumière la complexité

des relations entre les variables. Bien que certaines corrélations et régressions suggèrent des liens entre les caractéristiques des éoliennes et les prix de l'immobilier, ces relations sont souvent faibles et doivent être interprétées avec prudence. De plus, il est probable que d'autres facteurs, non pris en compte dans ce modèle, puissent influencer les prix de l'immobilier de manière significative.

Ainsi, notre travail montre qu'il existe des liens entre l'implantation des éoliennes et les prix de l'immobilier, mais ces liens sont modérés et complexes. Il serait intéressant de poursuivre cette analyse en prenant en compte d'autres variables, telles que les politiques locales, les infrastructures disponibles ou la perception des habitants vis-à-vis des éoliennes.

En conclusion, bien que les résultats obtenus soient intéressants, des recherches supplémentaires et une modélisation plus approfondie sont nécessaires pour mieux comprendre l'impact réel des éoliennes sur le marché immobilier. Cela pourrait permettre de mieux informer les décisions des autorités locales et des promoteurs immobiliers dans les zones concernées.

9 Conclusion Globale du Projet

Ce projet a permis d'explorer l'impact des éoliennes sur les prix des maisons en zone rurale, en utilisant une approche méthodologique variée comprenant l'analyse des données et des tentatives de cartographie SIG. Nous avons cherché à établir des liens entre la présence d'éoliennes et les variations des prix immobiliers dans les zones étudiées. Les résultats montrent que, bien que des corrélations existent, elles sont souvent faibles, et les variables contextuelles jouent un rôle important. En effet, des facteurs tels que la perception des habitants vis-à-vis des éoliennes, la proximité des infrastructures, et les caractéristiques des propriétés influencent largement les tendances observées.

Il apparaît donc que l'impact des éoliennes sur le marché immobilier est complexe et mérite des recherches supplémentaires pour mieux comprendre ces dynamiques. L'intégration de nouvelles variables, comme les politiques locales, les tendances économiques régionales et les perceptions des résidents, permettrait de nuancer et d'approfondir les conclusions de ce projet.

10 Réponse à la Problématique

La question centrale de ce projet était : **Les éoliennes ont-elles un impact sur le prix de l'immobilier en zone rurale ?** Les quelques résultats que j'ai pu avoir, suggèrent que l'impact est modéré, mais qu'il varie selon le contexte local. Les éoliennes peuvent, dans certains cas, faire augmenter les prix des propriétés proches grâce à la perception d'un environnement plus durable et énergétiquement responsable. Cependant, elles peuvent aussi entraîner une baisse des prix dans d'autres cas, notamment en raison des nuisances visuelles et sonores perçues par certains habitants.

Les résultats de ce projet sont cohérents avec ceux de plusieurs études récentes. Par exemple, une étude de l'Ademe a trouvé que les maisons situées à moins de 5

km d'éoliennes subissent une baisse moyenne de 1,5

En somme, les résultats de ce projet sont en accord avec ces études, mais ils indiquent également que l'impact des éoliennes sur les prix immobiliers en zone rurale est modéré et doit être interprété en fonction des spécificités locales.

11 Remerciements

Je tiens à exprimer ma sincère gratitude envers mon professeur pour son soutien et ses conseils tout au long de ce projet. Ses cours ont été une ressource précieuse qui m'a permis de mieux comprendre les enjeux liés à l'énergie renouvelable, à l'immobilier et à l'analyse de données géographiques.

12 Méthode de Travail

Dans le cadre de ce projet, j'ai utilisé plusieurs outils et méthodes pour analyser les données et produire des résultats pertinents :

- **Analyse de données** : J'ai utilisé RStudio pour traiter et analyser les données, en effectuant des jointures et des filtrages pour garantir la qualité des informations. J'ai également exploré les relations entre les variables, comme la proximité des éoliennes et les prix de l'immobilier.
- **Système d'Information Géographique (SIG)** : J'ai tenté de visualiser les données géographiques à l'aide de QGIS, bien que des difficultés techniques aient entravé cette partie, notamment à cause de problèmes liés aux coordonnées géographiques.
- **Utilisation de ChatGPT** : J'ai eu recours à ChatGPT pour des clarifications de concepts et pour m'aider dans la rédaction de certaines parties du projet.

En conclusion, ce projet a été une expérience d'apprentissage enrichissante, malgré les défis techniques rencontrés. Les résultats obtenus peuvent contribuer à une meilleure compréhension des interactions complexes entre les éoliennes et le marché immobilier en zone rurale, tout en soulignant la nécessité de recherches plus approfondies.

Références

- [1] ADEME, *Éolien terrestre : quel réel impact sur les prix de l'immobilier*, 2022.
- [2] Connaissance des Énergies, *Éolien terrestre : quel réel impact sur les prix de l'immobilier*, 2022.
- [3] ADEME, *Éoliennes et immobilier : aucun impact significatif sur le prix des maisons*, 2022.