

La revitalisation des petites villes en France



Quels facteurs socio-économiques, environnementaux et culturels influencent la dynamique démographique des petites villes en France ?

SOMMAIRE

Introduction

Données analysées et technologies

Définitions

Présentation approfondie du programme PVD

1. Le programme PVD
2. Corrélation entre PVD et revenus
3. Etude de cas

Analyse de l'Impact des facteurs socio-économiques, environnementaux et culturels sur la dynamique démographique des PVD.

Conclusion

Annexe

Dictionnaire de données

Introduction

Les petites villes jouent un rôle essentiel dans la structuration du territoire français, particulièrement dans les zones rurales. Elles assurent une centralité indispensable pour les habitants en leur offrant un accès aux services de santé, à l'éducation, aux équipements sportifs et culturels, et aux commerces. Cependant, ces petites villes font face à des défis multiples, notamment la désertification, le vieillissement de la population, et le déclin économique. Depuis 2020, l'État a lancé le programme « Petites Villes de Demain » (PWD) pour soutenir ces territoires et revitaliser leur cadre de vie.

Malgré les efforts déployés, les petites villes françaises demeurent hétérogènes : certaines connaissent un essor démographique, tandis que d'autres subissent un déclin marqué. Cette situation soulève la question des facteurs qui influencent cette dynamique démographique. **Quels sont les éléments socio-économiques, environnementaux et culturels qui favorisent ou freinent l'attractivité de ces territoires ?** Comprendre ces déterminants est crucial pour orienter les politiques publiques et optimiser les actions du programme PVD.

Ce projet vise à analyser et à représenter les spécificités des petites villes en France à travers une approche multi-thématische. Les objectifs principaux sont :

- Identifier les facteurs qui influencent la dynamique démographique des petites villes.
- Comparer les spécificités de certaines communes, notamment celles intégrées au programme PVD, avec d'autres ne bénéficiant pas de ce dispositif.
- Proposer des outils interactifs et des visualisations pour faciliter la compréhension des enjeux territoriaux.

Données analysées et technologies

Nous avons principalement utilisé les jeux de données disponibles sur la plateforme du défi **Revitalisation des petites villes**. Ces données couvrent des thématiques comme la population, les revenus, les équipements et services, les commerces, ainsi que les infrastructures énergétiques. En complément, le site **Observatoire des Territoires** nous a permis d'accéder à des données épurées et prêtes à l'emploi, facilitant l'analyse grâce à une interface intuitive.

Pour le traitement et l'analyse des données, nous avons utilisé **Python** comme outil principal, avec les bibliothèques Pandas pour la manipulation des données, Plotly Express pour les visualisations interactives, et **Streamlit** pour la création de notre application. Celle-ci permet de naviguer à travers les analyses et inclut des cartes interactives développées à l'aide de Streamlit Folium.

La collaboration au sein de l'équipe s'est organisée via **GitHub**, où nous avons centralisé le code et les fichiers, et **Notion**, où nous avons organisé tout notre travail et répertorié notre avancée. Pour la rédaction du rapport et la préparation de la soutenance, nous avons opté pour Canva.

Définitions

Programme Petites Villes de Demain (PVD)

Mis en place par l'Agence Nationale de la Cohésion des Territoires (ANCT) et la Banque des Territoires, ce programme accompagne 1 645 petites villes en France. L'objectif est de renforcer leur dynamisme économique, social et environnemental en leur fournissant des ressources pour revitaliser leurs centres et améliorer le cadre de vie des habitants.

Dynamique démographique

La "dynamique démographique" désigne l'évolution de la population d'une commune ou d'un territoire, en prenant en compte des facteurs tels que les naissances, les décès, les migrations, et les variations dans la structure par âge.

Facteurs socio-économiques

Les "facteurs socio-économiques" incluent des variables telles que les niveaux de revenus, le taux d'emploi, les inégalités sociales, et la répartition de la richesse qui influencent le développement et l'attractivité d'une commune.

Facteurs environnementaux

Les "facteurs environnementaux" regroupent des éléments liés à la qualité de vie dans une commune, tels que l'accessibilité, les infrastructures de transport, les espaces verts, et la gestion des ressources naturelles.

Facteurs culturels

Les "facteurs culturels" concernent les infrastructures et les activités liées à la culture, comme les équipements culturels (musées, cinémas, bibliothèques) ou les événements locaux qui participent à l'attractivité d'un territoire.

Hétérogénéité des petites villes

L'"hétérogénéité des petites villes" fait référence aux disparités entre ces communes, qu'il s'agisse de leur situation géographique, de leur dynamique démographique, de leur développement économique ou de leur accès aux services publics.

Attractivité territoriale

L'"attractivité territoriale" se définit comme la capacité d'un territoire à attirer et à retenir des populations, des activités économiques, et des investissements grâce à des infrastructures, des services publics de qualité, et un cadre de vie favorable.

Politiques publiques

Les "politiques publiques" désignent les actions et initiatives mises en œuvre par les institutions gouvernementales et les collectivités locales pour répondre aux besoins des citoyens et favoriser le développement des territoires.

Revitalisation

La revitalisation fait référence aux actions mises en œuvre pour redynamiser une ville ou un territoire, en particulier dans les domaines économique, social et culturel. Elle inclut le développement des commerces, des infrastructures, des services publics, et l'amélioration de l'environnement urbain.

Équipements et services

Ces termes regroupent l'ensemble des infrastructures et ressources disponibles pour répondre aux besoins des habitants d'une ville : écoles, collèges, commerces, médecins généralistes, pharmacies, infrastructures sportives, équipements culturels, etc. La présence de ces éléments est un indicateur clé de la vitalité d'une commune.

Présentation approfondie du programme PVD

1. Le programme PVD

La carte des communes bénéficiaires du programme **PVD** (Petites Villes de Demain) révèle une répartition géographique **étendue**, soulignant une présence dans l'ensemble du territoire français. Au total, **1 627 communes** participent actuellement à ce programme, ce qui représente un pourcentage significatif du total de **34 965 communes** que compte la France. Ce chiffre met en évidence l'ampleur des besoins en développement et en revitalisation des petites communes. Cependant, il convient de noter que ce programme fonctionne sur la base du **volontariat**. Par conséquent, certaines communes qui sont **éligibles** n'y participent pas, bien que leur situation pourrait bénéficier de ce soutien.

Les raisons de cette **non-participation** sont multiples. Certaines communes peuvent manquer de **ressources locales**, humaines ou financières, pour s'engager pleinement dans le programme. Par ailleurs, la **complexité administrative** ou le manque d'une vision à long terme peuvent freiner leur implication. Enfin, même des communes intéressées peuvent ne pas être retenues si leur projet n'est pas jugé **suffisamment structurant**.

Premièrement, on constate que les villes bénéficiaires ne se concentrent pas uniquement dans la fameuse "**diagonale du vide**" (zone s'étendant des Ardennes aux Pyrénées, souvent caractérisée par une faible densité de population). En effet, elles se retrouvent aussi bien dans les **zones littorales** qu'à proximité des **frontières**, notamment dans l'est, près de pays économiquement dynamiques comme la **Suisse**, le **Luxembourg** et **l'Allemagne**. Ces zones transfrontalières peuvent bénéficier d'une attractivité économique accrue grâce aux flux de **travailleurs transfrontaliers**.

Un autre point important est la présence de nombreuses communes proches **des grandes métropoles**. Ces villes, bien qu'influencées par la proximité d'agglomérations majeures comme **Paris**, **Lyon** ou **Marseille**, montrent un besoin de soutien pour répondre aux défis spécifiques d'une attractivité locale qui peut être éclipsée par celle des métropoles voisines.

Une **densité plus élevée** de communes bénéficiaires est visible dans le **nord-ouest de la France**, notamment en **Bretagne**, **Normandie**, et **Pays de la Loire**. Ces régions, historiquement rurales et marquées par une tradition agricole, bénéficient aujourd'hui d'un regain d'attention pour leur développement économique, notamment grâce au **tourisme**, à **l'agriculture durable** et aux **initiatives culturelles**.



Figure 1: Commune bénéficiaires du programme PVD

2. Corrélation entre PVD et revenus

Dispersion des revenus médians (PWD vs Non-PVD)

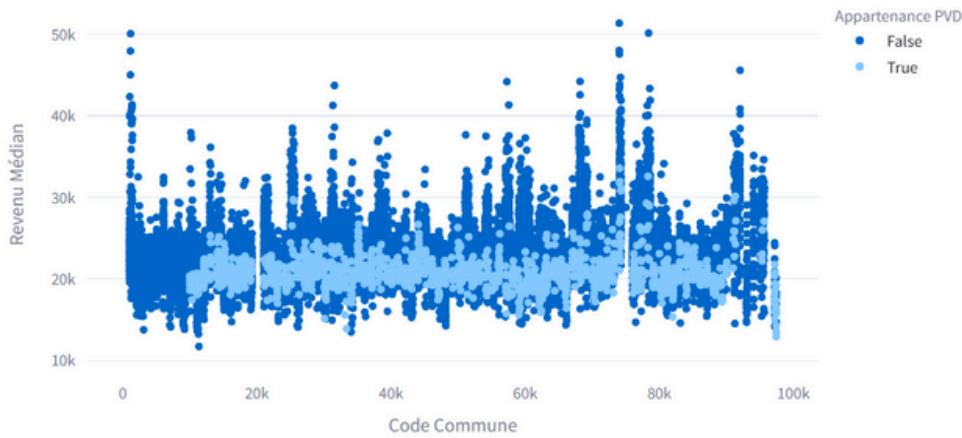


Figure 2 : Revenus médians de toutes les communes de France, PVD ou non

Ce diagramme montre la dispersion des **revenus médians** des communes françaises, en distinguant les communes participant au programme Petites Villes de Demain (PWD) de celles qui n'y participent pas. Les points bleu clair représentent les communes PVD, et les points bleu foncé, les non-PVD. Les revenus médians se concentrent majoritairement **entre 20 000 et 30 000 euros** pour les deux groupes, bien que les communes PVD soient davantage présentes dans les tranches de revenus modestes, reflétant l'objectif du programme de soutenir les communes en difficulté.

Une variabilité importante est observable sur l'ensemble des revenus, de moins de 10 000 euros à plus de 50 000 euros. Les communes non-PVD montrent une **dispersion plus large**, incluant des communes à hauts revenus, ce qui indique qu'elles ne partagent pas les mêmes priorités ou **besoins** que celles ciblées par le programme. Par ailleurs, certaines communes non-PVD économiquement fragiles pourraient être **éligibles** mais n'ont pas intégré le programme pour des raisons administratives, stratégiques ou politiques.

Dispersion des revenus médians (PWD vs Non-PVD)

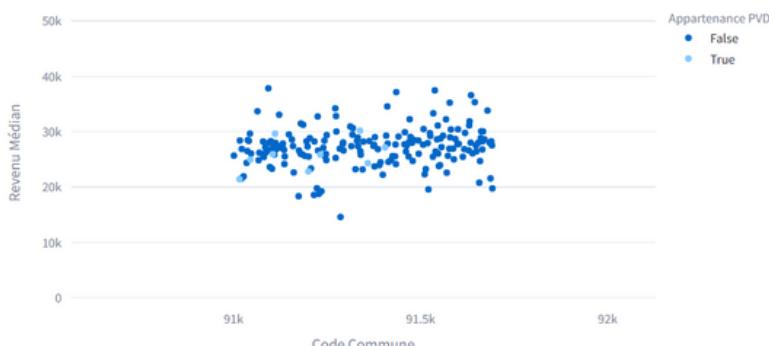


Figure 3 : Revenus médians des communes du département de l'Essonne (91)

Dispersion des revenus médians (PVD vs Non-PVD)

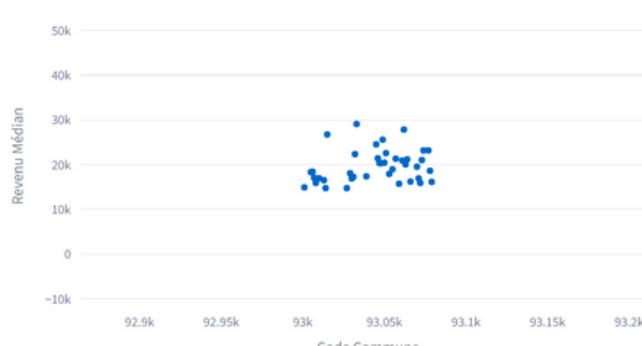


Figure 4 : Revenus médians des communes du département de Seine-Saint-Denis (93)

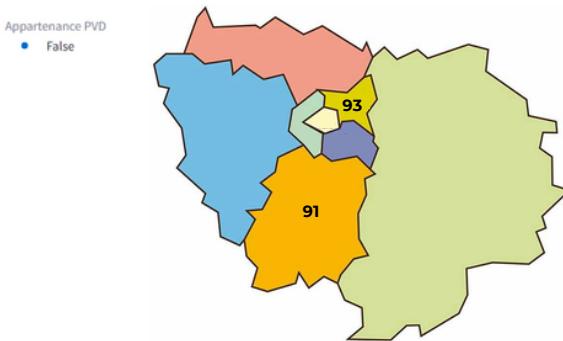


Figure 5 : Carte des départements d'Île de France

Dans le département de **l'Essonne** (91), représenté par la Figure 3, les revenus médians des communes se situent globalement autour de **30 000 euros**, avec une faible dispersion en dessous de 20 000 euros. Cela montre une relative homogénéité dans la distribution des revenus médians, avec peu de communes présentant des écarts significatifs vers des valeurs plus basses. Par ailleurs, l'absence de points très dispersés suggère une **stabilité économique** relative entre les communes de ce département.

En revanche, dans la **Seine-Saint-Denis** (93), représentée par la Figure 4, les revenus médians des communes sont **nettement plus bas**, avec une majorité des valeurs situées entre 10 000 et 20 000 euros. La dispersion est plus prononcée vers les valeurs inférieures, reflétant une situation économique plus précaire. Ce département est connu pour sa forte proportion de communes classées comme **défavorisées** sur le plan socio-économique, ce qui se traduit clairement dans le graphique.

Une différence marquante est également **l'absence** d'appartenance au programme **PVD** (indiquée par les points clairs) dans les deux cas, mais cette absence est particulièrement notable en Seine-Saint-Denis, où les communes pourraient théoriquement bénéficier d'un tel programme en raison des revenus médians faibles.

Par ailleurs, parmi les 20 communes ayant les **taux de pauvreté** les plus élevés en France, on trouve Roubaix (46 %), Saint-Benoît (46 %), Saint-André (46 %), Grigny (44 %) et Saint-Louis (43 %). Étonnamment, aucune de ces communes ne participe au programme Petites Villes de Demain (PVD).

Plusieurs hypothèses expliquent cette absence : le programme repose sur le **volontariat**, et ces communes peuvent manquer de ressources pour candidater ou avoir d'autres priorités. De plus, PVD cible les petites villes structurantes, tandis que ces communes, souvent intégrées à de **grandes agglomérations**, bénéficient déjà **d'autres dispositifs** (zones prioritaires, contrats de ville). Enfin, certaines autorités locales pourraient privilégier des solutions plus adaptées à leurs besoins spécifiques.

Concernant les données, les taux de pauvreté issus des open data ou de l'observatoire des territoires étaient indisponibles ou inexploitables, mais le site inegalites.fr a fourni un classement fiable pour ces analyses.

3. Etude de cas

Étude de cas : Saint-Clément-de-Rivière et La Grand-Combe

Localisation et situation des départements

Saint-Clément-de-Rivière, avec une population de 5 140 habitants en 2022, est située dans le département de l'Hérault (34), dans la région Occitanie. Ce département se distingue par sa croissance économique soutenue grâce au dynamisme de la métropole de Montpellier. L'Hérault bénéficie d'un fort attrait touristique, d'un climat méditerranéen, et d'une économie diversifiée.

La Grand-Combe, quant à elle, comptait 4 891 habitants en 2021 et est située dans le département du Gard (30), également en région Occitanie. Le Gard repose sur une économie plus traditionnelle, centrée sur l'agriculture et l'artisanat, bien qu'il bénéficie également du tourisme grâce à des sites emblématiques comme le Pont du Gard. Cependant, certaines zones, notamment La Grand-Combe, subissent encore les conséquences d'un passé industriel marqué. Ces disparités font de La Grand-Combe une commune prioritaire pour des initiatives telles que le programme PVD.

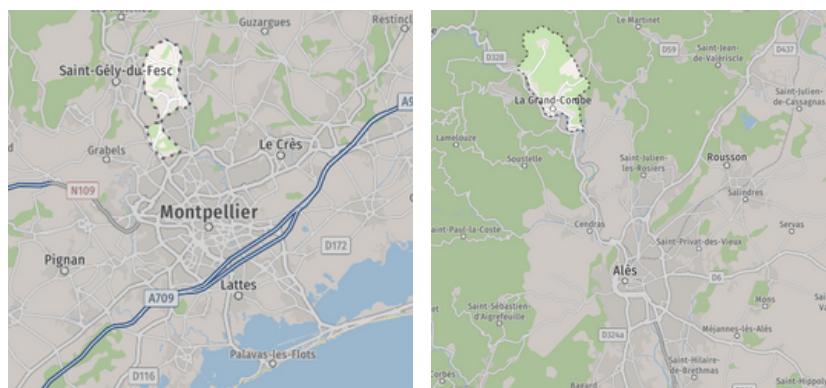


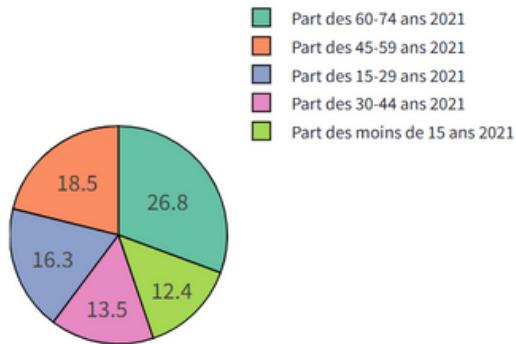
Figure 6 : Localisation des communes étudiées

Analyse démographique

Les diagrammes des tranches d'âge montrent une population relativement âgée à Saint-Clément-de-Rivière, avec une proportion significative de personnes âgées de 60 à 74 ans (26,8 %). En revanche, La Grand-Combe a une population légèrement plus jeune, avec une répartition plus équilibrée entre les différentes tranches d'âge, bien que les 60-74 ans y soient également majoritaires (21,1 %).

Ces différences peuvent être liées à l'attractivité résidentielle de Saint-Clément-de-Rivière pour les retraités grâce à sa proximité avec Montpellier et son cadre de vie calme. À La Grand-Combe, une population active plus jeune persiste, probablement en raison d'un coût de la vie plus abordable.

Saint-Clément-de-Rivière : Tranches d'âge



La Grand-Combe : Tranches d'âge

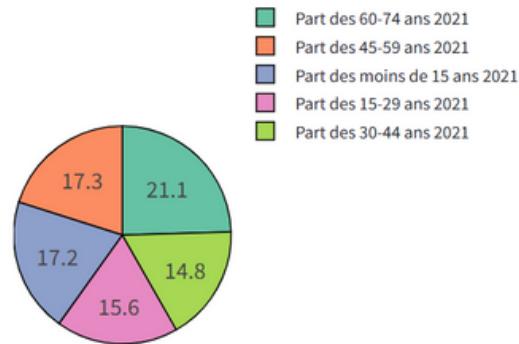


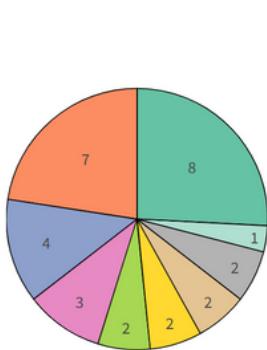
Figure 7 : Diagrammes de répartition des tranches d'âges des communes étudiées

Analyse des services

Saint-Clément-de-Rivièr e offre une couverture impressionnante en termes de services de santé, avec 8 dentistes et 7 médecins généralistes, un chiffre bien supérieur aux 2 dentistes et 3 médecins de La Grand-Combe. Cette disparité est un indicateur clair de la différence d'attractivité économique et de moyens entre ces deux communes.

Concernant les commerces et infrastructures, Saint-Clément dispose de 4 boulangeries, 2 grandes surfaces, et 2 supérettes, tandis que La Grand-Combe compte 3 boulangeries, 1 grande surface, et 1 supérette. Les équipements éducatifs sont similaires, avec 2 collèges et 1 lycée dans chaque commune. Cependant, La Grand-Combe se distingue par un plus grand nombre d'écoles primaires (5 contre 3), reflétant peut-être une population plus jeune.

Saint-Clément-de-Rivièr e : Services



La Grand-Combe : Services

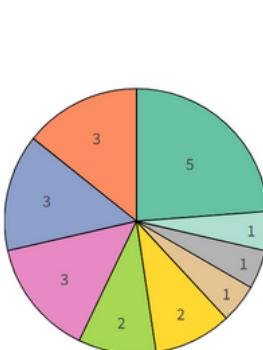


Figure 8 : Diagrammes de répartition des services possédés par les communes étudiées

Analyse économique

La médiane du revenu disponible par unité de consommation en 2020 montre un écart important : 34 310 € à Saint-Clément-de-Rivièr e contre 15 180 € à La Grand-Combe. Ce contraste illustre les différences marquées dans le niveau de vie des habitants. Saint-Clément, plus proche de Montpellier, attire des ménages aisés et bénéficie d'une économie locale prospère. À l'inverse, La Grand-Combe continue de ressentir les conséquences du déclin industriel.

Comparaison des communes pour Médiane du revenu disponible par UC 2020

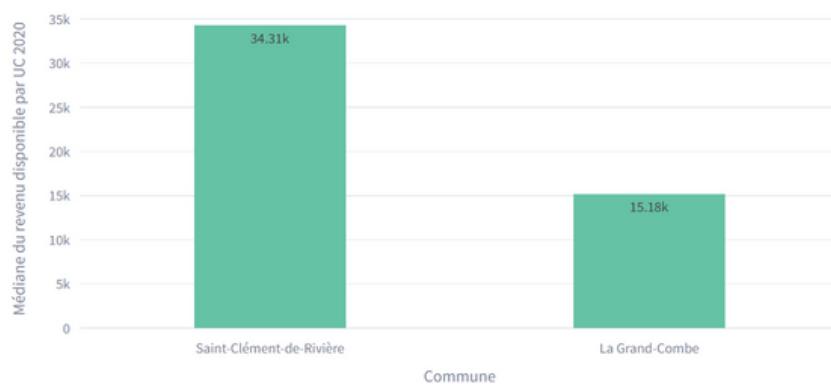


Figure 9 : Diagramme des revenus médians des communes étudiées

Analyse des équipements sportifs

Saint-Clément-de-Rivièr e dispose de 19 équipements sportifs, contre 16 à La Grand-Combe. Cependant, en rapportant ces chiffres à leur population, les taux d'équipements pour 1 000 habitants sont proches : 3,8 pour Saint-Clément contre 3,2 pour La Grand-Combe. Cela montre que, malgré ses contraintes économiques, La Grand-Combe s'efforce de maintenir un certain niveau d'infrastructures sportives.

Comparaison des communes pour Nombre d'équipements sportifs 2023

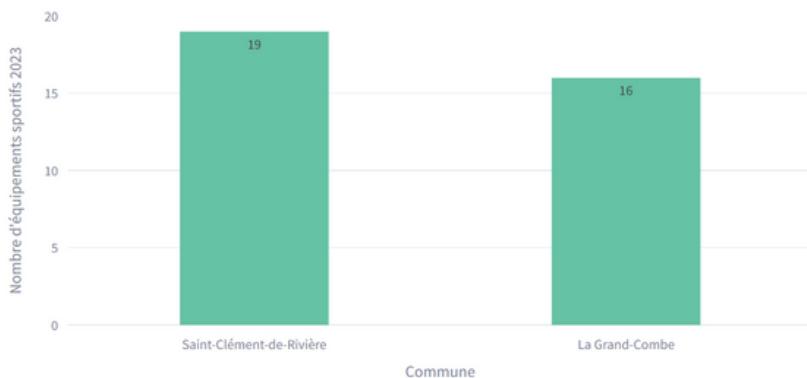


Figure 10 : Diagramme du nombre d'équipements sportifs des communes étudiées

Comparaison des communes pour Taux d'équipements sportifs pour 1 000 habitants 2023



Figure 11 : Diagramme des taux d'équipements sportifs des communes étudiées

Conclusion

Cette étude de cas illustre de manière frappante les disparités entre deux communes de taille similaire, Saint-Clément-de-Rivière, non incluse dans le programme Petites Villes de Demain, et La Grand-Combe, qui en bénéficie. Saint-Clément-de-Rivière, avec son cadre résidentiel privilégié, son **revenu médian élevé**, et ses services variés, n'entre pas dans les critères d'éligibilité du PVD. Cette commune, grâce à sa proximité avec Montpellier et à sa population aisée, ne fait pas face aux mêmes défis structurels que d'autres petites villes.

À l'inverse, La Grand-Combe, marquée par un revenu médian bien inférieur et une **économie locale plus fragile**, est un exemple typique des communes que le programme PVD vise à soutenir. Les infrastructures et services y sont **modestes**, malgré un effort visible pour maintenir un niveau acceptable d'équipements sportifs et éducatifs. L'inclusion de La Grand-Combe dans le programme PVD répond à des besoins réels : **revitalisation économique**, amélioration des conditions de vie, et soutien aux initiatives locales pour relancer son attractivité.

Cette comparaison met en lumière l'objectif du PVD : concentrer les ressources et l'accompagnement sur les communes en difficulté, souvent confrontées à un déclin économique ou démographique, afin de **réduire les inégalités territoriales**. Cependant, elle révèle également que certaines communes prospères comme Saint-Clément-de-Rivière, bien qu'exclues du programme, n'en ont pas besoin grâce à leur dynamisme et à leur attractivité économique naturelle. En ce sens, le PVD agit comme un levier indispensable pour rééquilibrer les disparités entre territoires, en focalisant son soutien là où les besoins sont les plus pressants.

Méthodologie de Collecte et Traitement et des Données

1. Présentation des Sources de Données

- **Projet Défi-Data-Gouv :** Les données issues du Projet Défi-Data-Gouv ont constitué une base précieuse pour ce projet, bien que des limitations aient été rencontrées. Les jeux de données étaient souvent **lourds** et difficiles à télécharger complètement, ce qui a nécessité une sélection rigoureuse des informations à extraire. De nombreuses valeurs étaient **manquantes** ou **incohérentes**, compliquant leur exploitation. Ainsi, une part significative du travail a été consacrée à **l'épuration des données** pour en garantir la fiabilité et la pertinence.
- **Observatoire des Territoires :** En revanche, les données fournies par l'Observatoire des Territoires se sont révélées particulièrement utiles. La possibilité de **pré-sélectionner** les catégories et les colonnes souhaitées avant le téléchargement a permis de constituer des jeux de données plus **ciblés** et exploitables. Ces fichiers comportaient très peu de valeurs manquantes ou incohérentes, réduisant considérablement les efforts de nettoyage préalable. Cette qualité des données a été essentielle pour alimenter les analyses complexes liées à la redynamisation des petites villes.

2. Analyse Préliminaire des Jeux de Données

- **Répartition des Tâches au Sein de l'Équipe :** Afin d'assurer une organisation optimale, un espace collaboratif sur Notion a été mis en place pour centraliser les informations, les liens vers les jeux de données, et les tâches. Chaque membre a ensuite reçu des responsabilités précises. Voici une synthèse de la répartition :
 - **Épuration et Analyse initiale des Données :** Chaque membre a été chargé d'explorer un jeu de données spécifique pour en extraire les informations pertinentes.

VAR_ID	VAR_LIB	VAR_LIB_LONG
IRIS	IBS	Code du département suivi du numéro de commune ou du numéro d'arrondissement municipal suivi du numéro d'IRIS
UBRIS	Ubella de l'IRIS	Libellé de l'IRIS suivi de la commune ou de l'arrondissement municipal pour Paris, Lyon et Marseille
COM	Commune ou ARM	Code du département suivi du numéro de commune ou du numéro d'arrondissement municipal pour Paris, Lyon et Marseille
UBCOM	Ubella commune ou ARM	Libellé de la commune ou de l'arrondissement municipal pour Paris, Lyon et Marseille
DEC_PMPIB	Part des ménages fiscaux imposés (%)	Part des ménages fiscaux imposés (%)
DEC_TP601B	Taux de bas revenus déclarés au	Taux de bas revenus déclarés au

3. Approche Méthodologique pour le Traitement des Données

- **Méthodes de Traitement des Données**

- Nous avons utilisé pour le traitement des données, des scripts Python utilisant la bibliothèque pandas qui permet la manipulation et l'analyse de données structurées en DataFrames. Ainsi nous avons pu facilement:

```
# Chargement des données à partir d'un fichier CSV
import pandas as pd
df = pd.read_csv('chemin/vers/fichier.csv', sep=';', encoding='utf-8')

# Filtrer les données
df_filtre = df[df['colonne'] > valeur]

# Renommer les colonnes
df.rename(columns={'ancien_nom': 'nouveau_nom'}, inplace=True)

# Supprimer les colonnes non pertinentes
df = df.drop(columns=['colonne_a_supprimer'])

# Sélectionner des colonnes spécifiques
df = df[['colonne1', 'colonne2', 'colonne3']]

# Fusionner deux DataFrames
# basés sur une clé commune, telle que code_insee dans notre cas
df_fusionne = pd.merge(df1, df2, on='code_insee', how='inner')

# Enregistrer les résultats dans un fichier CSV
df.to_csv('chemin/vers/fichier_modifie.csv', sep=';', index=False, encoding='utf-8')
```

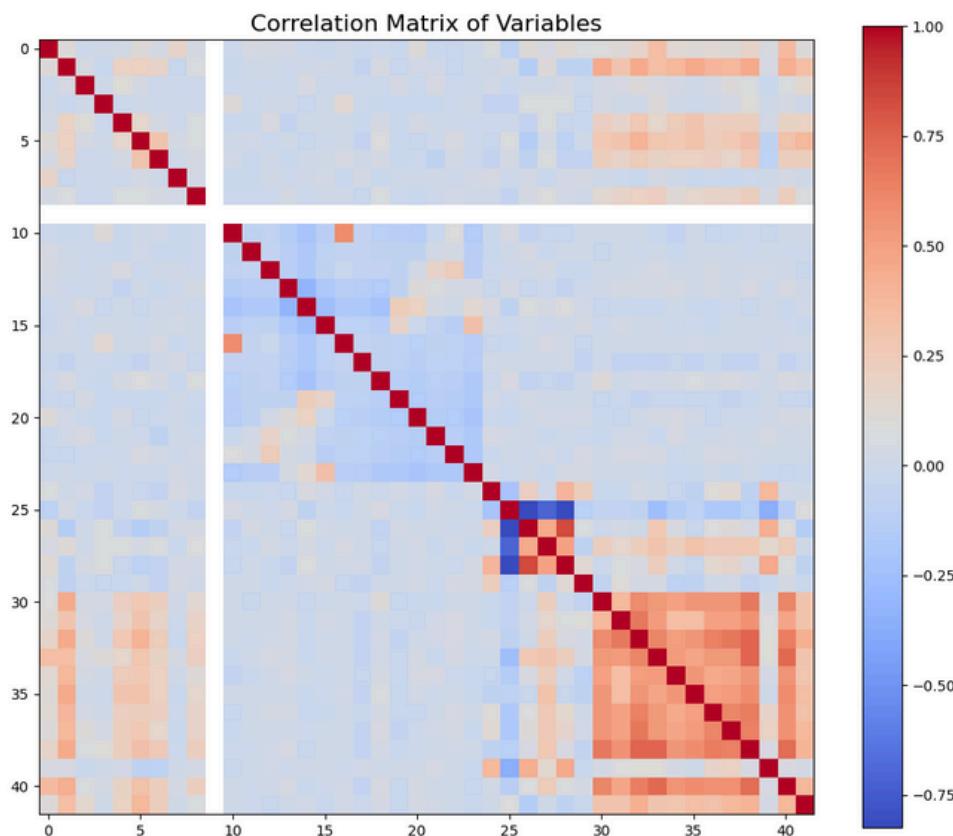
- Ces scripts nous ont permis d'élaborer un ensemble de données pertinent pour analyser les facteurs influençant la dynamique démographique des petites villes en France.
- En nous concentrant sur les communes bénéficiaires du programme "Petites Villes de Demain" (PWD), nous avons intégré diverses variables reflétant les taux d'évolution démographique (net, solde migratoire, solde naturel), les caractéristiques environnementales (telles que la taille de la commune et son appartenance à une aire d'influence urbaine), les équipements culturels et sportifs (par exemple, le nombre de cinémas par commune) et les indicateurs socio-économiques (comme la médiane du revenu disponible par unité de consommation). Cette compilation de données nous a permis de mener des analyses approfondies pour tenter de répondre au mieux à notre problématique.

4. Analyse de l'Impact de de facteurs socio-économiques, environnementaux et culturels sur la dynamique démographique des PVD.

Étapes préliminaires

Après notre épuration des données nous nous retrouvons avec un échantillon de 1397 communes pour lesquelles nous avons un jeu de données complet, au lieu des 1627 communes que compte le PVD au total.

Nous avons choisi comme variable de réponse le taux d'évolution annuel de la population. Avant de construire le modèle, nous avons créé une matrice de corrélation comportant toutes nos variables explicatives. Toute variable qui aurait une corrélation avec une autre supérieure à 0,85 serait retiré des observations afin d'éviter au maximum le risque de multicolinéarité. Lors de cette étape, aucune de nos variables explicatives n'a été retirée.



Construction du modèle

Au vu du grand nombre de prédicteurs, nous doutions que certains d'entre eux ne seraient pas significatifs, nous avons donc procédé à une sélection backward.

Cela signifie que nous avons commencé par générer un modèle comportant toutes les variables explicatives pour ensuite retirer une à une celles ayant la plus grande p-value supérieure au seuil de 0,05. Cela veut dire que nous retirons les variables pour lesquelles nous ne pouvons pas prouver qu'il existe d'effet significatif sur la variable à expliquer. Cette sélection a ainsi permis de retirer 28 des 41 variables que nous avions choisies. Faisant ainsi passer l'Akaike Information Criterion (AIC) du modèle de 3067 à 3035.

Interpretation des résultats

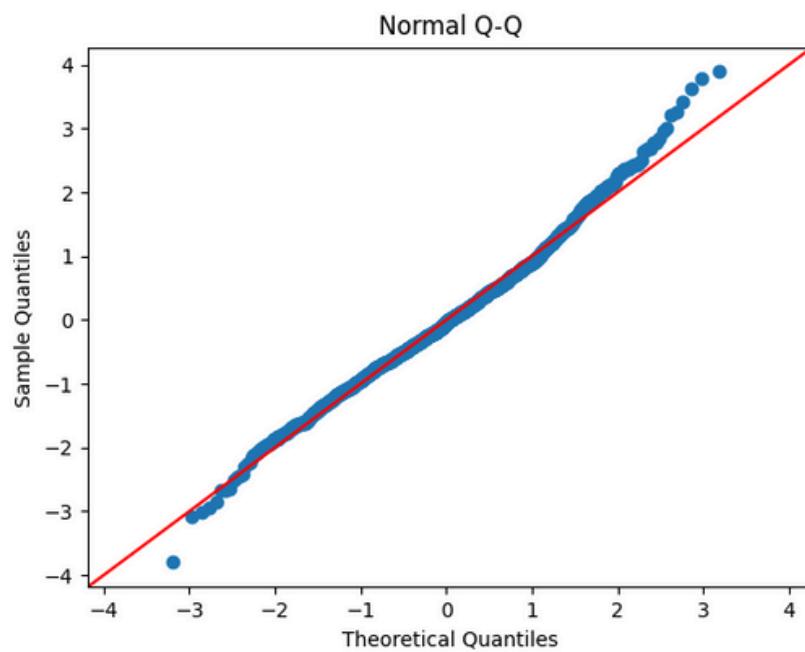
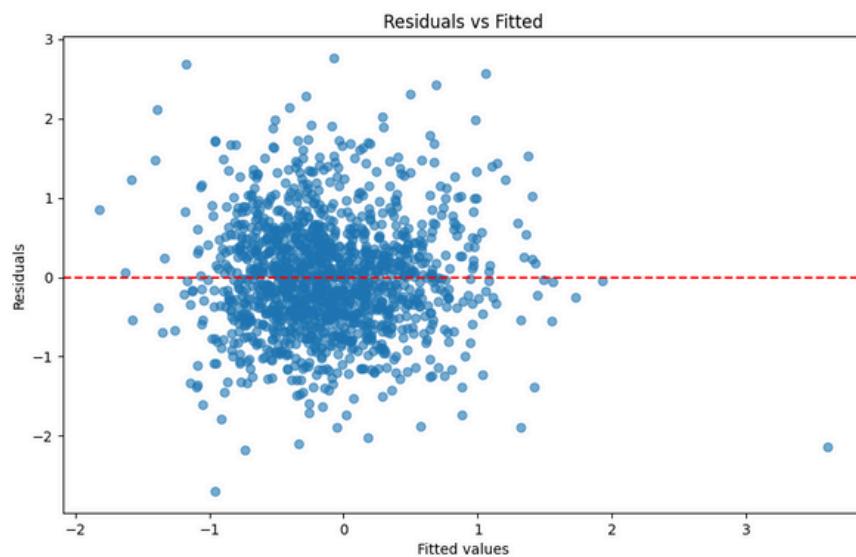
- B0 : La constante du modèle est de -2,6937, indiquant une valeur de base lorsque toutes les autres variables sont nulles.
- Variables significatives (Bj):
 - Nombre de cinémas : Positivement associé (+0,1088). Plus il y a de cinémas, plus le taux d'évolution est élevé.
 - Nombre de conservatoires : Négativement associé (-1,5047). Les conservatoires semblent réduire le taux d'évolution.
 - Nombre de musées : Impact négatif (-0,2256). Plus il y a de musées, plus le taux d'évolution diminue.
 - Médiane du revenu : Positivement associé (+0,0001). Une hausse du revenu médian est liée à une augmentation du taux d'évolution.
 - Part des moins de 15 ans : Effet positif (+0,0914). Une population jeune stimule l'évolution.
 - Part des 15-19 ans : Effet négatif (-0,0382). Une proportion plus élevée de cette tranche d'âge est associée à un taux d'évolution plus faible.
 - Part des 30-44 ans : Impact positif (+0,0509).
 - Part des 45-59 ans : Effet négatif (-0,0678).
 - Nombre de supérettes et épiceries : Positivement associé (+0,0242). La présence de supérettes favorise l'évolution.
 - Nombre d'écoles primaires, maternelles et élémentaires : Effet négatif (-0,0603). Plus d'écoles primaires semblent réduire le taux d'évolution.
 - Nombre de médecins généralistes : Positivement associé (+0,0185). Plus de médecins généralistes sont associés à une augmentation du taux d'évolution.
 - nombre_de_dentistes : Effet positif (+0,0253).
 - nombre_de_festivals : Impact positif (+0,0380).
- R² : le R² du modèle est de 0,35 ce qui indique que notre modèle n'explique qu'un peu plus d'un tiers de la variance du taux d'évolution de la démographie des communes du PVD. Cela n'est pas des plus étonnant, compte tenu de la complexité de cette variable, qui est souvent analysée et dont les causes sont souvent remises en question.
- F-test : la p-value pour le Model Significance Test est extrêmement faible et bien en dessous de 0,05 ce qui confirme que le modèle est significatif.

OLS Regression Results						
	coef	std err	t	P> t	[0.025	0.975]
Dep. Variable:	taux_evolution	R-squared:	0.347			
Model:	OLS	Adj. R-squared:	0.340			
Method:	Least Squares	F-statistic:	56.42			
Date:	mer., 15 janv. 2025	Prob (F-statistic):	7.30e-118			
Time:	01:33:16	Log-Likelihood:	-1503.7			
No. Observations:	1397	AIC:	3035.			
Df Residuals:	1383	BIC:	3109.			
Df Model:	13					
Covariance Type:	nonrobust					
const	-2.6937	0.295	-9.128	0.000	-3.273	-2.115
Cinema	0.1088	0.044	2.456	0.014	0.022	0.196
Conservatoire	-1.5047	0.513	-2.934	0.003	-2.511	-0.499
Musee	-0.2256	0.060	-3.734	0.000	-0.344	-0.107
mediane_du_revenu_disponible_par_uc_2020	0.0001	1.06e-05	9.675	0.000	8.15e-05	0.000
part_des_moins_de_15_ans_2021	0.0914	0.012	7.921	0.000	0.069	0.114
part_des_15-29_ans_2021	-0.0382	0.009	-4.436	0.000	-0.055	-0.021
part_des_30-44_ans_2021	0.0509	0.014	3.636	0.000	0.023	0.078
part_des_45-59_ans_2021	-0.0678	0.011	-5.951	0.000	-0.090	-0.045
nombre_de_superettes_et_epieries	0.0242	0.009	2.675	0.008	0.006	0.042
nombre_ecoles_primaire_maternelles_elementaires	-0.0603	0.012	-5.079	0.000	-0.084	-0.037
nombre_de_medicins_generalistes	0.0185	0.007	2.843	0.005	0.006	0.031
nombre_de_dentistes	0.0253	0.008	3.352	0.001	0.010	0.040
nombre_de_festivals	0.0380	0.017	2.197	0.028	0.004	0.072

4. Analyse de l'Impact de de facteurs socio-économiques, environnementaux et culturels sur la dynamique démographique des PVD.

Hypothèses du modèle

- Homoscédasticité : les résidus ont bien une moyenne nulle ainsi qu'une variance constante et indépendante du taux d'évolution de la population, nous pouvons donc bien utiliser les tests de nullité sur les coefficients et le F-test sur le modèle.
- Normalité des résidus : le QQ plot montre que les résidus ne suivent pas une distribution normale.



Conclusion

Ce projet nous a offert une grande liberté pour explorer et approfondir les dynamiques démographiques des petites villes en France. Il nous a permis d'utiliser des outils modernes tels que Python et ses bibliothèques (Pandas, Plotly Express, Streamlit) pour traiter et visualiser les données, ainsi que Notion pour organiser et suivre nos progrès de manière collaborative. En outre, nous avons pu appliquer concrètement les notions vues en travaux pratiques de statistique, notamment en adaptant et en traduisant les analyses réalisées en R vers Python. Ce cadre de travail a enrichi notre compréhension des problématiques territoriales tout en consolidant nos compétences en analyse de données et en gestion de projet.

Il serait très intéressant de pouvoir explorer à nouveau ce sujet plus en profondeur afin de pouvoir expliquer davantage la dynamique démographique des communes appartenant au PVD.

Annexe

URL de l'application déployée : projet-data-science.streamlit.app

Lien vers la page du Défi DataGouv associé à ce projet:
<https://defis.data.gouv.fr/defis/revitalisation-des-petites-villes>

Dictionnaire de données

code_insee	Code Insee de la commune d'implantation de l'équipement, sans distinction des arrondissements pour Paris, Lyon et Marseille (respectivement 75056, 69123 et 13055)	Défi-Data-Gouv
Bibliotheque	Nombre de bibliothèques dans la commune.	Défi-Data-Gouv
Cinema	Nombre de cinémas dans la commune.	Défi-Data-Gouv
Conservatoire	Nombre de conservatoires dans la commune.	Défi-Data-Gouv
Espace_protege	Nombre de d'espaces protégés dans la commune.	Défi-Data-Gouv
Librairie	Nombre de librairies dans la commune.	Défi-Data-Gouv
Monument	Nombre de monuments dans la commune.	Défi-Data-Gouv
Musee	Nombre de musées dans la commune.	Défi-Data-Gouv
Parc_et_jardin	Nombre de parcs et jardins dans la commune.	Défi-Data-Gouv
Theatre	Nombre de théâtres dans la commune.	Défi-Data-Gouv
Etablissement_d'enseignement_superieur	Nombre d'établissements d'enseignement supérieur dans la commune.	Défi-Data-Gouv
GCD	Valeurs possibles : - [6 - Rural à habitat dispersé', '5 - Bourgs ruraux', '4 - Ceintures urbaines', '2 - Centres urbains intermédiaires', '1 - Grands centres urbains', '7 - Rural à habitat très dispersé', '3 - Petites villes']	Défi-Data-Gouv
AAV	Valeurs possibles: -[30 - Hors attraction des villes', '22 - Couronnes de 50 000 à moins de 200 000 hab.', '11 - Pôles de moins de 50 000 hab.', '23 - Couronnes de 200 000 à moins de 700 000 hab.', '21 - Couronnes de moins de 50 000 hab.', '12 - Pôles de 50 000 à moins de 200 000 hab.', '24 - Couronnes de 700 000 hab. ou plus', '13 - Pôles de 200 000 à moins de 700 000 hab.', '14 - Pôles de 700 000 hab. ou plus', nan]	Défi-Data-Gouv
mediane_du_revenu_disponible_par_uc_2020	La médiane du revenu disponible correspond au niveau au-dessous duquel se situent 50 % de ces revenus. C'est de manière équivalente le niveau au-dessus duquel se situent 50 % des revenus. Le revenu disponible par unité de consommation (UC), également appelé "niveau de vie", est le revenu disponible par "équivalent adulte". Il est calculé en rapportant le revenu disponible du ménage au nombre d'unités de consommation qui le composent.	Observatoire des territoires
part_des_60-74_ans_2021	Population de cette tranche d'âge rapportée à la population de la commune	Observatoire des territoires
part_des_moins_de_15_ans_2021	Population de cette tranche d'âge rapportée à la population de la commune	Observatoire des territoires
part_des_15-29_ans_2021	Population de cette tranche d'âge rapportée à la population de la commune	Observatoire des territoires
part_des_30-44_ans_2021	Population de cette tranche d'âge rapportée à la population de la commune	Observatoire des territoires
part_des_45-59_ans_2021	Population de cette tranche d'âge rapportée à la population de la commune	Observatoire des territoires
nombre_grandes_surfaces_2023	Nombre de grandes surfaces dans la commune.	Observatoire des territoires
nombre_de_superettes_et_epieries	Nombre de supérettes et épiceries dans la commune.	Observatoire des territoires
nombre_boulangeries_et_patisseries	Nombre de boulangeries et de pâtisseries dans la commune.	Observatoire des territoires
nombre_ecoles_primaire_maternelle_elementaire	Nombre d'écoles (primaires, maternelles, élémentaires) dans la commune.	Observatoire des territoires
nombre_de_colleges	Nombre de collèges dans la commune.	Observatoire des territoires
nombre_de_lycees	Nombre de lycées dans la commune.	Observatoire des territoires
nombre_de_medecins_generalistes	Nombre de médecins généralistes dans la commune.	Observatoire des territoires
nombre_de_dentistes	Nombre de dentistes dans la commune.	Observatoire des territoires
nombre_de_pharmacies	Nombre de pharmacies dans la commune.	Observatoire des territoires
taux_evolution	Taux d'évolution annuel moyen de la population communale sur une période donnée.	Observatoire des territoires
taux_evolution_due_solde_migratoire	Part du taux d'évolution annuel de la population attribuable au solde migratoire apparent.	Observatoire des territoires
taux_evolution_due_solde_naturel	Part du taux d'évolution annuel de la population attribuable au solde naturel.	Observatoire des territoires
nombre_d_equipements_sportifs	Nombre total d'infrastructures sportives accessible au public dans la commune. (installations intérieures et extérieures)	Observatoire des territoires
nombre_de_festivals	Nombre de festivals ayant eu lieu dans la commune en 2019	Défi-Data-Gouv