IMMOBILIER OCCITANIE

Dataviz Reporting

GROUPE 3:

BILLARD Melissa
BRACCIALE-COMBAS Lola
CARRILHO LAMEIRA Rita



Sommaire

Introduction	3
Présentation des membres du groupe	3
Présentation de l'entreprise : Occitanie Immobilier	3
Contexte et objectifs du projet	3
Échelle géographique du projet	4
Compétences évaluées et attentes	5
Méthodologie	5
Source de données	5
Outils et technologies	6
Étapes de la méthodologie	7
Configuration Jasper	8
Analyse des données	10
Profilage des données collectées	10
Analyse statistique préliminaire	11
Identification des tendances et indicateurs pertinents	12
Résultats	14
Analyse des rendements locatifs en Occitanie	14
Présentation du tableau des départements avec le rendement locatif	14
Graphiques illustrant les rendements locatifs par zone	15
Corrélation entre rendement locatif, pauvreté et chômage	17
Discussion	17
Interprétation des résultats	17
Identification des zones à fort rendement locatif	17
Implications pour les investisseurs immobiliers	18
Limitations de l'étude et perspectives d'amélioration	18
Conclusion	18
Synthèse des principaux résultats	18
Recommandations pour les investisseurs	19
Réflexions sur les compétences acquises et les outils utilisés	19
Annexes	20
Tables BDD	20
Dataset Jasper	20
Modélisation rapport Jasper	21
Lien GIT	21

Introduction

Le projet de datavisualisation et de reporting que nous avons entrepris s'inscrit dans le cadre de la certification I1 EISI pour l'année **2023-2024**, au sein du module [BDOE832] Datavisualisation. Il a pour objectif d'aider un investisseur immobilier à identifier les zones à fort rendement locatif dans la *région Occitanie*, afin d'optimiser ses investissements.

Présentation des membres du groupe

Notre équipe est composé de trois personnes :

- BILLARD Mélissa
- BRACCIALE-COMBAS Lola
- CARRILHO LAMEIRA Rita

Chacune d'entre nous a apporté ses compétences spécifiques et son expertise pour mener à bien ce projet complexe et enrichissant.



Présentation de l'entreprise : Occitanie Immobilier



Pour les besoins de ce projet, nous avons créé une entreprise fictive nommée **Occitanie Immobilier**. Cette entreprise se spécialise dans <u>l'investissement immobilier</u> et vise à fournir des analyses détaillées et des conseils précieux aux investisseurs intéressés par le marché locatif de la région Occitanie.

Contexte et objectifs du projet

Le marché immobilier est un secteur en perpétuelle évolution, où les opportunités d'investissement peuvent *varier* considérablement d'une région à l'autre. Dans ce contexte, notre projet vise à fournir une analyse précise et exhaustive des rendements locatifs dans le région Occitanie, en <u>utilisant des données ouvertes</u> et des <u>outils de datavisualisation avancés</u>. Notre objectif principal est de mettre en lumière les **zones géographiques** offrant les **meilleurs rendements locatifs**, tout en tenant compte de divers facteurs socio-économiques comme la *démographie* et le *taux de chômage*.

Ce projet s'articule autour de plusieurs axes majeurs :

- La collecte et l'analyse de données provenant de différentes sources fiables
- La création de visualisations claires et informatives pour illustrer les rendements locatifs
- La mise en place d'une présentation professionnelle et esthétique, intégrant des éléments de storytelling pour une meilleure compréhension des données par les utilisateurs finaux.

Échelle géographique du projet

Notre étude se concentre sur la **région Occitanie**, une vaste région du sud de la France connue pour sa <u>diversité géographique et culturelle</u>. L'Occitanie se compose de **treize départements**, chacun présentant des caractéristiques uniques en termes de marché locatif. Pour cette analyse, nous avons choisi de nous pencher sur certains départements et villes clés qui présentent un intérêt particulier pour les investisseurs immobiliers.

La région Occitanie est riche en opportunités d'investissement immobilier, grâce à son <u>dynamisme économique</u> et à son <u>attrait touristique</u>. En analysant les données au *niveau des départements et des villes*, nous visons à fournir une vue détaillée et nuancée des rendements locatifs, permettant ainsi aux investisseurs de prendre des décisions éclairées.



Compétences évaluées et attentes

Ce projet vise à évaluer <u>plusieurs compétences</u> essentielles en **datavisualisation** et en **reporting** :

- La capacité à créer un projet de datavisualisation en utilisant un outil de reporting
- La compétence d'acquisition et de traitement de données à partir de sources variées
- La création de représentations graphiques et tabulaires précises et esthétiques
- La mise en page et le storytelling, pour une communication claire et engageante des résultats obtenus

En répondant à ces attentes, notre projet permettra de démontrer notre <u>maîtrise des outils et des techniques de data visualisation</u>, ainsi que notre capacité à appliquer ces compétences à des problématiques réelles et concrètes dans le domaine de l'investissement immobilier.

Méthodologie

Pour mener à bien ce projet de datavisualisation et de reporting, nous avons adopté une méthodologie structurée et rigoureuse, reposant sur l'utilisation de sources de données fiables et d'outils technologiques adaptés.

Source de données

La première étape de notre méthodologie a consisté à *identifier et collecter des données* pertinentes provenant de <u>différentes sources ouvertes</u>. Nous avons sélectionnée les bases de données suivantes :

- 1. Statistiques DVF (Demandes de Valeurs Foncières)
 - Source : data.gouv.fr
 - Description: Cette base de données est produite par le Ministère de l'Économie, des Finances et de la Souveraineté industrielle et numérique. Elle contient des informations détaillées sur les transactions immobilières en France, permettant ainsi d'analyser les prix d'acquisition et les loyers mensuels pour calculer les rendements locatifs.
- 2. Données sur le chômage en Occitanie
 - Source : DREETS Occitanie

- Description: Ces données sont produites par l'Acoss et la DARES et fournissent des informations détaillées sur le taux de chômage dans la région Occitanie. Elles sont essentielles pour comprendre <u>l'impact de</u> <u>l'économie locale</u> sur les rendements locatifs.
- 3. Données démographiques en Occitanie
 - Source : <u>DREETS Occitanie</u>
 - Description : Ces données, également produites par la DREETS, fournissent des informations sur la démographie de la région Occitanie, incluant la population, la structure par âge, et d'autres indicateurs démographiques importants.

Outils et technologies

Pour traiter, analyser et visualiser les données collectées, nous avons utilisé plusieurs outils et technologies.

 Jasper: Cet outil de reporting a été utilisé pour créer le rapport final, incluant des tableaux et des graphiques interactifs. Jasper permet une mise en page professionnelle et une intégration facile des données provenant de différentes sources.



 Python: Nous avons utilié Python pour automatiser certaines tâches de traitement des données. Les scripts Python nous ont permis de charger facilement les données dans notre base de données.



 WAMP et MySQL: Pour stocker et gérer les données en local, nous avons utilisé le serveur WAMP et la base de données MySQL. Cela nous a permis de centraliser les données et de les rendre facilement accessibles pour les analyses et les visualisations.



 PowerBI: Cet outil a été utilisé pour le profilage initial des données, permettant de visualiser rapidement les distributions et d'identifier les valeurs aberrantes ou les tendances préliminaires.



Étapes de la méthodologie

1. Collecte des données

Nous avons commencé par *rassembler les données* à partir des sources mentionnées ci-dessus. Chaque jeu de données a été téléchargé et vérifié pour s'assurer de sa complétude et de sa qualité.

2. Prétraitement des données

Nous avons *nettoyé les données* pour éliminer les **doublons**, corriger les **erreurs** et **harmoniser les formats**. Cette étape inclut également la transformation des données en formats compatibles avec <u>Jasper</u> et <u>MySQL</u>.

3. Intégration des données

Les données nettoyées ont été *importées* dans une base de données <u>MySQL</u>, où elles ont été structurées de manière à **faciliter les requêtes** et les analyses ultérieures.

4. Analyse et profilage

Nous avons utilisé <u>PowerBI</u> pour effectuer un *profilage initial des données*, identifiant les **tendances principales** et les **valeurs aberrantes**. Cette étape a permis de guider les analyses statistiques plus détaillées.

5. Création des visualisations

En utilisant <u>Jasper</u>, nous avons créé divers *tableaux et graphiques* pour visualiser les **rendements locatifs**, les **taux de chômage** et les **indicateurs démographiques**. Les visualisations ont été conçues pour être claires, informatives et esthétiques.

6. Mise en page et storytelling

La mise en page finale du rapport a été réalisée dans <u>Jasper</u>, en intégrant des éléments de storytelling pour faciliter la compréhension des résultats par les utilisateurs finaux. Le rapport a été conçu pour être **interactif** et **facile à naviguer**.

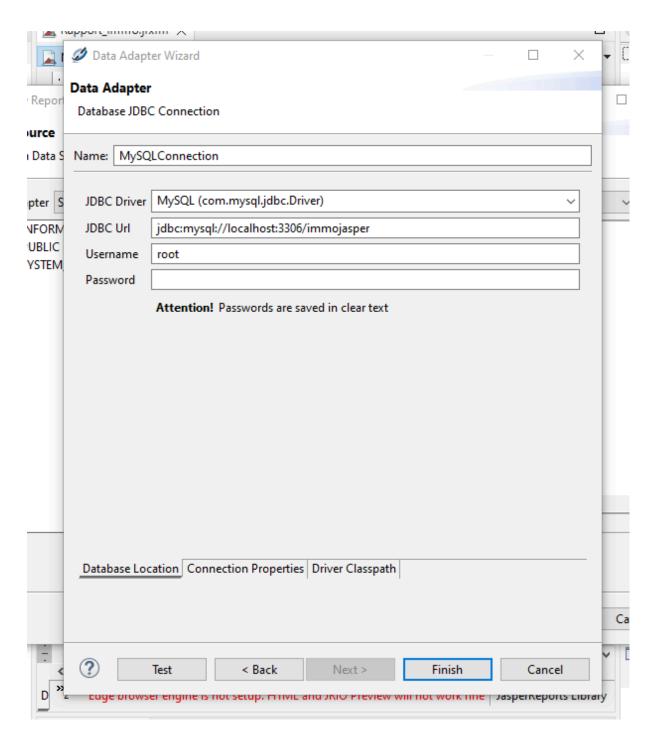
Cette méthodologie nous a permis de traiter et d'analyser de manière efficace les données collectées, tout en garantissant la **qualité** et la **fiabilité** des résultats obtenus.

Configuration Jasper

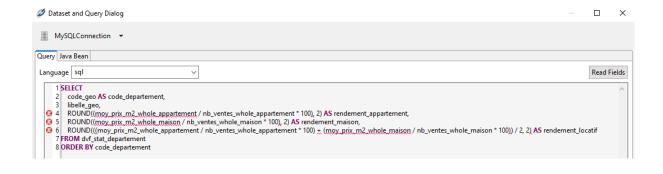
Avant de modéliser notre rapport, nous avons commencé par **créer une base de données**. Dans un premier temps, nous avons réalisé un script Python qui nous a permis de créer la base de données et les tables en local.

script python (voir GIT)

Une fois cela fait, nous avons ajouté sur Jasper un data adapter pour pouvoir nous connecter à nos données stockées en local sur MySQL (Wamp).



Après avoir testé et confirmé la connexion, nous avons créé nos datasets. Nous avons créé en tout trois datasets : un pour les données de la table commune, un autre pour les données de la table département qui comprend les indicateurs chômage, emploi et démographiques, et un dataset dans lequel, en SQL, nous avons directement calculé les rendements locatifs par maison, appartement et département. Une fois cela fait, nous avons pu commencer à assembler notre rapport.



Analyse des données

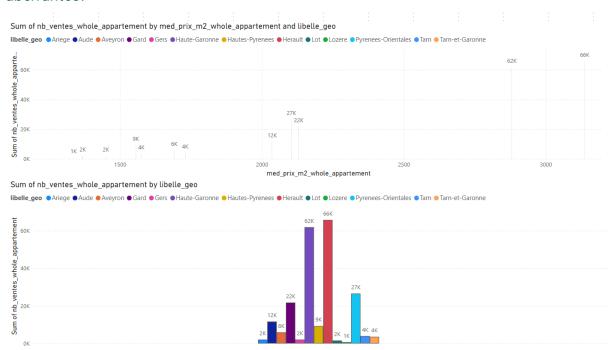
Nous avons procédé de manière méthodique afin de *profiler*, analyser et identifier les tendances pertinentes dans les données que nous avons collectées.

Profilage des données collectées

Le profilage des données est une étape cruciale qui consiste à **examiner les données** pour en comprendre *la structure, les distributions, et identifier d'éventuelles anomalies*. Voici les principales étapes de profilage que nous avons suivies :

1. Examen des distributions

Nous avons analysé **la distribution des variables** principales telles que <u>les prix</u> <u>d'acquisition, les loyers mensuels, et les taux de chômage</u>. Cela nous a permis de comprendre la répartition des données et d'identifier *les valeurs extrêmes ou aberrantes*.



2. Vérification de la complétude

Nous avons vérifié que toutes les variables essentielles étaient présentes dans les jeux de données et que *les valeurs manquantes* étaient traitées de manière appropriée. Les valeurs manquantes ont été imputées ou éliminées selon le contexte.

3. Qualité des données

Nous avons effectué des vérifications de cohérence pour nous assurer que les données étaient *logiques et conformes* aux attentes. Par exemple, les prix d'acquisition ont été comparés aux loyers pour s'assurer qu'ils étaient réalistes.

4. Transformations nécessaires

Les données ont été *normalisées et transformées* en formats compatibles pour les analyses et les visualisations ultérieures.

Analyse statistique préliminaire

L'analyse statistique préliminaire permet d'<u>explorer les relations entre différentes variables</u> et de commencer à *identifier des modèles* ou des *tendances intéressantes*. Voici quelques-unes des analyses que nous avons effectuées :

1. Statistiques descriptives

Nous avons calculé des *statistiques descriptives* telles que les moyennes, médianes, écarts-types, etc., pour chaque variable principale. Cela nous a aidés à avoir une vue d'ensemble des données.

2. Analyse de corrélation

Nous avons examiné les *corrélations entre différentes variables* pour identifier des relations potentielles. Par exemple, la <u>corrélation entre les rendements locatifs et les taux de chômage</u>, ou entre les rendements locatifs et la démographie.

3. Visualisations initiales

Nous avons créé des *visualisations préliminaires*, comme des histogrammes, des boxplots, et des scatter plots, pour visualiser les distributions et les relations entre les variables.

Identification des tendances et indicateurs pertinents

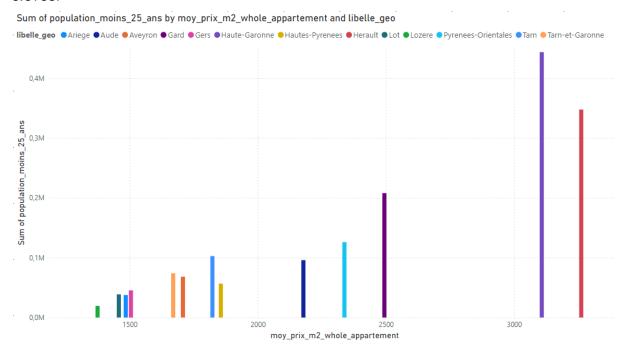
Sur la base des analyses statistiques préliminaires, nous avons *identifié les tendances* et les **indicateurs** les plus pertinents pour le rendement locatif. Voici quelques-uns des points clés :

1. Rendements locatifs par département

Nous avons observé que **certains départements** de la région Occitanie présentent des rendements locatifs significativement plus élevés que d'autres. Ces départements seront analysés plus en détail pour comprendre les facteurs sous-jacents.

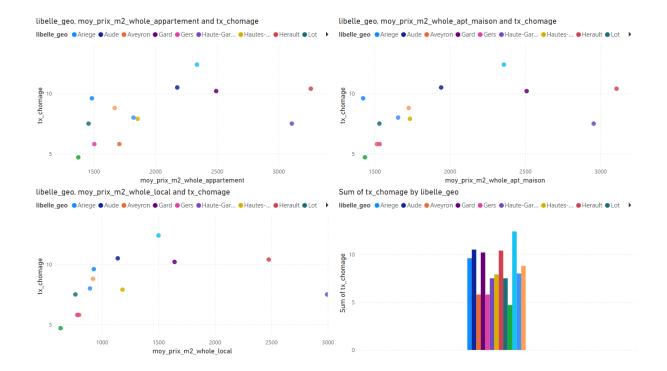
2. Impact de la démographie

Nous avons identifié que la **structure démographique des villes** (par exemple, la proportion de jeunes adultes) peut avoir un impact sur les rendements locatifs. Les villes avec une population plus jeune tendent à avoir des rendements locatifs plus élevés.



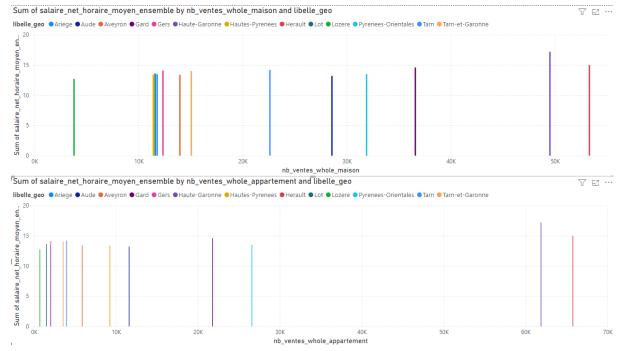
3. Taux de chômage et rendements locatifs

l existe une corrélation entre les **taux de chômage et les rendements locatifs**. Les zones avec des taux de chômage plus bas tendent à offrir de meilleurs rendements locatifs.



4. Facteurs socio-économiques

D'autres facteurs socio-économiques, comme le **niveau de revenu moyen** jouent également un rôle important dans les rendements locatifs.



Résultats

Analyse des rendements locatifs en Occitanie

L'analyse des rendements locatifs en Occitanie s'est concentrée sur les départements et les principales villes de la région. Les résultats montrent une variabilité significative des rendements locatifs selon les zones géographiques, influencée par divers facteurs économiques et démographiques.

Département	Nombre de ventes d'appartements	Prix moyen par mètre carré (appartement)	Nombre de ventes de maisons	Prix moyen par mètre carré (maison)
Ariege (9)	2034	1485.0€	11780	1413.0 €
Aude (11)	11608	2178.0€	28576	1846.0 €
Aveyron (12)	5884	1708.0 €	13974	1436.0 €
Gard (30)	21791	2494.0€	36586	2519.0 €
Haute-Garonne (31)	61889	3108.0 €	49514	2764.0 €
Gers (32)	2048	1505.0 €	12345	1539.0 €
Herault (34)	65774	3262.0 €	53304	2916.0 €
Lot (46)	1521	1458.0 €	11588	1541.0 €
Lozere (48)	708	1375.0€	3802	1447.0 €
Hautes-Pyrenees (65)	9260	1856.0 €	11418	1638.0 €
Pyrenees- Orientales (66)	26596	2338.0 €	31902	2376.0 €
Tarn (81)	3964	1823.0 €	22616	1627.0 €
Tarn-et-Garonne (82)	3551	1670.0 €	15067	1740.0 €

Présentation du tableau des départements avec le rendement locatif

Le tableau ci-dessous résument les <u>rendements locatifs moyens par département et par type de bien</u> (appartements et maisons) dans la région Occitanie. Ces rendements ont été calculés à partir des données sur *les prix d'acquisition et les loyers mensuels moyens*.

Département	Rendement appartement	Rendement maison	Rendement locatif
Aude (11)	18.76 %	6.46 %	12.61 %
Aveyron (12)	29.03 %	10.28 %	19.65 %
Gard (30)	11.45 %	6.89 %	9.17 %
Haute-Garonne (31)	5.02 %	5.58 %	5.3 %
Gers (32)	73.49 %	12.47 %	42.98 %
Herault (34)	4.96 %	5.47 %	5.21 %
Lot (46)	95.86 %	13.3 %	54.58 %
Lozere (48)	194.21 %	38.06 %	116.13 %
Hautes-Pyrenees (65)	20.04 %	14.35 %	17.19 %
Pyrenees-Orientales (66)	8.79 %	7.45 %	8.12 %
Tarn (81)	45.99 %	7.19 %	26.59 %
Tarn-et-Garonne (82)	47.03 %	11.55 %	29.29 %
Ariege (9)	73.01 %	11.99 %	42.5 %

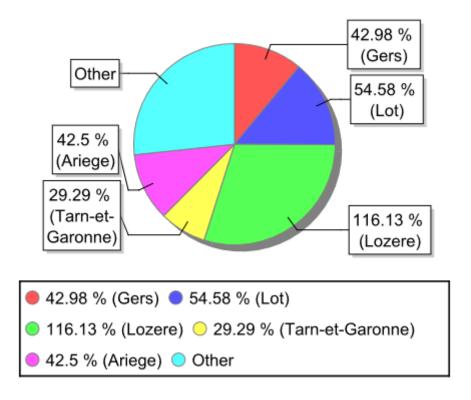
Ces données montrent des **disparités notables entre les départements**, avec certains offrant des rendements locatifs exceptionnellement élevés, comme **la Lozère et le Lot**, tandis que d'autres, comme *la Haute-Garonne et l'Hérault*, présentent des rendements plus modérés.

Graphiques illustrant les rendements locatifs par zone

Pour visualiser ces résultats, plusieurs types de graphiques ont été utilisés :

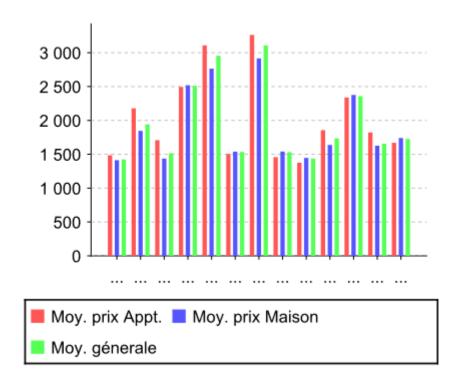
1. Camembert des rendements locatifs par département

Le graphique en camembert ci-dessous illustre <u>la répartition des rendements</u> <u>locatifs moyens par département</u>. Chaque segment représente *le rendement moyen d'un département*, permettant une visualisation claire des zones offrant les meilleurs rendements.



2. Graphiques à barres

Des graphiques à barres ont été utilisés pour <u>comparer les prix moyens de chaque</u> <u>département pour les appartements et les maisons</u>. Ces graphiques permettent de voir rapidement quelles zones géographiques offrent *les prix les plus élevés* pour chaque type de bien.



Corrélation entre rendement locatif, pauvreté et chômage

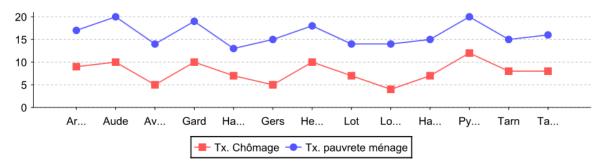
L'analyse des données a révélé <u>plusieurs corrélations intéressantes</u> entre les rendements locatifs et divers facteurs socio-économiques :

1. Corrélation entre rendement locatif et taux de chômage

Une **corrélation négative** a été observée entre le *taux de chômage* et le *rendement locatif*. Les départements avec un taux de chômage plus faible tendent à offrir des rendements locatifs plus élevés, suggérant que l'emploi stable et le pouvoir d'achat élevé favorisent la demande locative.

2. Corrélation entre rendement locatif et taux de pauvreté

Il existe une **corrélation inverse** entre le *taux de pauvreté* et les *rendements locatifs*. Les zones avec un taux de pauvreté plus faible tendent à offrir des rendements locatifs plus élevés, probablement en raison d'une meilleure qualité de vie et d'une demande locative plus forte.



Discussion

Interprétation des résultats

Lozère et le Lot, offrent des rendements locatifs <u>exceptionnellement élevés</u>. Ces départements, bien que *moins densément peuplés*, présentent des opportunités d'investissement intéressantes en raison de leurs *prix d'acquisition relativement bas* et de *loyers compétitifs*. À l'inverse, des départements plus urbains comme la Haute-Garonne et l'Hérault offrent des rendements <u>plus modérés</u>, bien que toujours attractifs pour des investissements *sûrs et stables*.

Identification des zones à fort rendement locatif

Les zones identifiées comme offrant les meilleurs rendements locatifs incluent :

- La Lozère (48): Avec un rendement locatif moyen de 116.13%, ce département se distingue par des <u>prix d'acquisition très bas</u> par rapport aux loyers perçus.
- Le Lot (46) : Offrant un rendement moyen de 54.58%, ce département présente également une excellente opportunité pour les investisseurs.
- L'Ariège (09) et le Gers (32) : Ces départements montrent des rendements moyens supérieurs à 40%, attirant les investisseurs cherchant des rendements élevés.

Implications pour les investisseurs immobiliers

Les investisseurs immobiliers peuvent tirer parti de ces analyses pour cibler leurs investissements dans les zones offrant les meilleurs rendements locatifs. Les départements ruraux et moins densément peuplés, malgré leurs <u>faibles taux de croissance économique</u>, présentent des rendements attractifs grâce à des *prix d'acquisition bas*. Les investisseurs doivent également considérer les tendances démographiques et les infrastructures locales pour maximiser leurs rendements.

Limitations de l'étude et perspectives d'amélioration

Bien que cette étude fournisse des informations précieuses, certaines limitations doivent être reconnues :

- **Données limitées :** Les données disponibles ne couvrent pas toutes les variables potentielles influençant les rendements locatifs.
- Variabilité temporelle : Les rendements locatifs peuvent fluctuer avec le temps en raison des conditions économiques et du marché immobilier.
- Facteurs non quantifiables : Certains aspects qualitatifs, comme la perception de la qualité de vie, ne sont pas inclus dans l'analyse.

Pour améliorer cette étude, il serait bénéfique d'intégrer davantage de *données qualitatives* et de réaliser **une analyse longitudinale** pour observer les tendances sur plusieurs années.

Conclusion

Synthèse des principaux résultats

L'analyse a révélé que certains départements de l'Occitanie, en particulier la Lozère et le Lot, offrent des rendements locatifs exceptionnellement élevés. Les départements urbains, bien que présentant des rendements plus modérés, restent attractifs pour des investissements stables. Les facteurs socio-économiques, tels

que <u>le taux de chômage, le taux de pauvreté et la démographie</u>, jouent un rôle crucial dans la détermination des rendements locatifs.

Recommandations pour les investisseurs

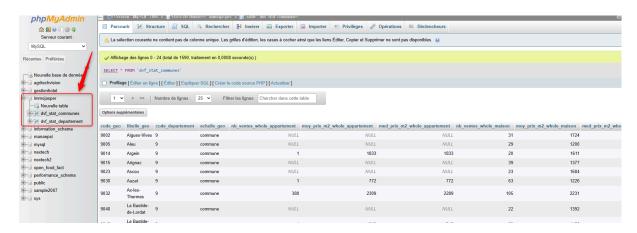
Les investisseurs devraient cibler les départements identifiés pour maximiser leurs rendements locatifs. Il est conseillé de <u>surveiller les tendances du marché immobilier</u> et de tenir compte des facteurs socio-économiques et des infrastructures locales lors de la prise de décision d'investissement.

Réflexions sur les compétences acquises et les outils utilisés

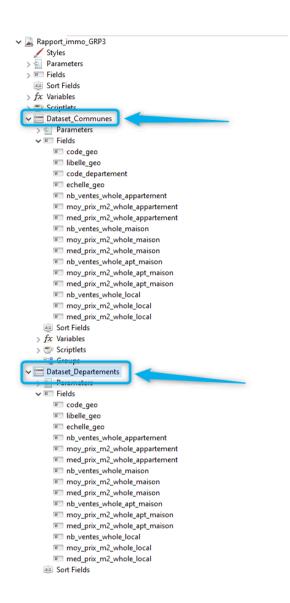
Ce projet a permis de développer des compétences essentielles en datavisualisation et en reporting. L'utilisation de <u>PowerBI</u> et de <u>Jasper</u> nous a aidés à **traiter**, **analyser** et visualiser les données de manière efficace, renforçant ainsi notre capacité à interpréter les données et à fournir des recommandations basées sur des analyses solides.

Annexes

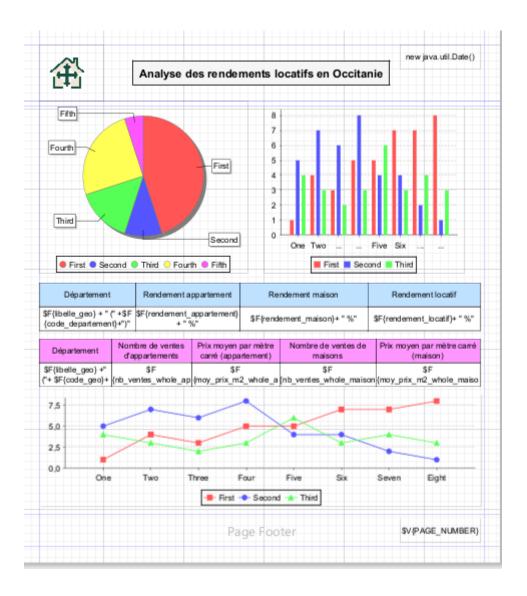
Tables BDD



Dataset Jasper



Modélisation rapport Jasper



Lien GIT

https://github.com/melissabillard/TP_jaspersoft