

Projet fil rouge: Airbnb paris

DataFrame: Filtrage et Transformation des Données

1. Charger depuis la base de données PostgreSQL les tables "listings" et "reviews" en tant que DataFrame
2. Afficher le schéma pour les deux DataFrames
3. Filtrer les listings avec une disponibilité supérieure à 30 jours dans les 365.
4. Calculer la moyenne des prix des listings pour chaque type de chambre
5. Extraire le mois à partir de la colonne "Date" dans les reviews.
6. Ajouter une nouvelle colonne "Year" qui extrait l'année à partir de la colonne "Date_review" dans le DataFrame "reviews".
7. Filtrer les listings avec disponibilité supérieure à 30 jours et prix inférieur à 100 euros.
8. Créer une nouvelle colonne "price_per_room" (prix par chambre) dans les listings.
9. Remplacer les valeurs manquantes dans "reviews_per_month" par 0 dans les listings.
10. Remplacer les valeurs nulles dans la colonne "review_scores_rating" par la moyenne calculée.

Analyse des Données et Statistiques:

11. Calculer la moyenne de la colonne "review_scores_rating".
12. Calculer la moyenne des prix pour chaque type de logement.
13. Trouver le nombre total de types de chambre dans chaque zone voisine.
14. Calculer la corrélation entre le nombre de nuits disponibles et le prix des listings.
15. Trouver les top 5 quartiers avec le plus grand nombre de listings.

PostgreSQL: Lire et Écrire:

16. Compter le nombre total d'annonces dans chaque ville.
17. Calculer la disponibilité et les prix moyens par quartier.
18. Obtenir le nombre total d'avis dans la table des avis.
19. Trouver la date de l'avis le plus récent dans la table des avis.
20. Identifier les cinq utilisateurs ayant publié le plus grand nombre d'avis.
21. Compter le nombre d'hôtes uniques dans la table des listings.

22. Ajouter une colonne pour indiquer si le listing est premium et ajouter les résultats dans une nouvelle table dans la base de données.

23. Calculer le nombre moyen de nuits disponibles pour chaque hôte et ajouter les résultats dans une nouvelle table dans la base de données.

1. Bar Plot (Graphique en barres)

Titre : Prix moyen par type de chambre

Objectif : Créer un graphique en barres montrant le prix moyen des locations pour chaque type de chambre.

- Groupez les données du DataFrame listings par type de chambre (room_type).
- Calculez le prix moyen (avg) pour chaque type de chambre.
- Visualisez les résultats sous forme de graphique en barres, où :
 - L'axe X représente les types de chambres (par exemple, "Private Room", "Entire Home/Apartment").
 - L'axe Y représente le prix moyen en euros.

2. Line Plot (Graphique en ligne)

Titre : Nombre de commentaires par mois

Objectif : Créer un graphique en ligne illustrant la variation du nombre total de commentaires au cours de l'année.

- Extrayez le mois à partir de la date de révision (date_review) dans le DataFrame reviews.
- Groupez les données par mois et comptez le nombre total de commentaires pour chaque mois.
- Visualisez les résultats sous forme de graphique en ligne, où :
 - L'axe X représente les mois de l'année (de janvier à décembre).
 - L'axe Y représente le nombre total de commentaires.