

Rapport final du Mini projet Zomato

Nom : Mouctar DAFFE

CODE PERMANENT : DAFM01119900

Analyse exploratoire des restaurants à Bengaluru (Zomato)

Ce rapport présente une analyse exploratoire du dataset Zomato Bangalore, qui regroupe plus de 51 000 restaurants situés dans différents quartiers de la ville de Bengaluru. L'objectif du projet est de comprendre la structure du paysage de la restauration, d'identifier les tendances principales et d'analyser les facteurs associés à la popularité des restaurants.

1. Description du dataset

Le dataset contient 51 717 restaurants et 17 variables. Les principales informations disponibles concernent le nom du restaurant, sa localisation, le type de restaurant, les cuisines proposées, le coût moyen pour deux personnes, la note attribuée par les clients et le nombre de votes.

Certaines variables, comme les avis détaillés ou les plats appréciés, présentent un nombre important de valeurs manquantes, ce qui est cohérent avec la nature du dataset.

2. Nettoyage et préparation des données

Un nettoyage minimal a été effectué conformément aux consignes du projet. Les notes des restaurants ont été converties en valeurs numériques, les coûts ont été uniformisés et les valeurs non exploitables ont été remplacées par des valeurs manquantes. Aucun doublon n'a été détecté dans le dataset.

3. Analyses descriptives

L'analyse descriptive montre que certains types de restaurants, notamment les restaurants dine-out et les cafés, sont particulièrement nombreux. Certaines cuisines dominent le paysage culinaire, reflétant la diversité alimentaire de la ville.

4. Analyse géographique et popularité

L'analyse géographique met en évidence une forte concentration de restaurants dans certains quartiers centraux. La popularité des restaurants, mesurée par les notes et le nombre de votes, varie selon le type de restaurant et le quartier.

Les résultats montrent que des prix plus élevés ne garantissent pas nécessairement de meilleures notations, ce qui suggère que d'autres facteurs, comme la qualité du service ou la diversité de l'offre, jouent un rôle important.

5. Dépôt Github et l'app Streamlit

Le jeu de données Zomato utilisé dans ce projet est volumineux (environ 574 MB), ce qui dépasse la limite de taille autorisée par GitHub pour le versionnement des fichiers. Pour cette raison, le dataset complet n'a pas été inclus dans le dépôt GitHub.

L'analyse exploratoire complète a été réalisée localement sur l'ensemble du dataset à l'aide de Python et Pandas. Afin de permettre le déploiement et l'exécution de l'application Streamlit, un échantillon représentatif du jeu de données a été créé et utilisé pour l'application interactive. J'ai opté pour cette option afin de concilier contraintes techniques et reproductibilité des analyses.

6. Conclusion

Cette analyse exploratoire a permis d'identifier les principales tendances de la restauration à Bengaluru. Elle met en évidence des différences notables entre les quartiers, les types de restaurants et les cuisines proposées.

Les limites principales du dataset concernent la présence de valeurs manquantes et l'absence d'informations détaillées sur la fréquentation réelle des restaurants. Des analyses futures pourraient inclure l'étude des avis clients ou des comparaisons temporelles.