

# UQAC

Université du Québec  
à Chicoutimi

---

## EDA ZOMATO

---

Rapport Final

19 DÉCEMBRE 2025

8PRO408 – OUTILS DE PROGRAMMATION POUR LA SCIENCE DES DONNÉES

Remis à : Habba's Ngodjou Doukaga

## Table des matières

<b><i>Introduction</i></b> .....	<b>2</b>
<b><i>Méthodologie et Traitement des Données</i></b> .....	<b>2</b>
<b><i>Profils de Quartiers et Répartition Géographique</i></b> .....	<b>2</b>
<b><i>Analyse de la Popularité et du Coût</i></b> .....	<b>3</b>
<b><i>Limites et Recommandations</i></b> .....	<b>3</b>
<b><i>Conclusion</i></b> .....	<b>3</b>

## Introduction

Ce rapport présente les résultats d'une analyse exploratoire (EDA) effectuée sur le jeu de données Zomato, comprenant plus de 51 000 établissements à Bengaluru. L'objectif est d'identifier les tendances de consommation, les structures de prix et les zones géographiques clés afin de comprendre la dynamique de la "Silicon Valley de l'Inde".

## Méthodologie et Traitement des Données

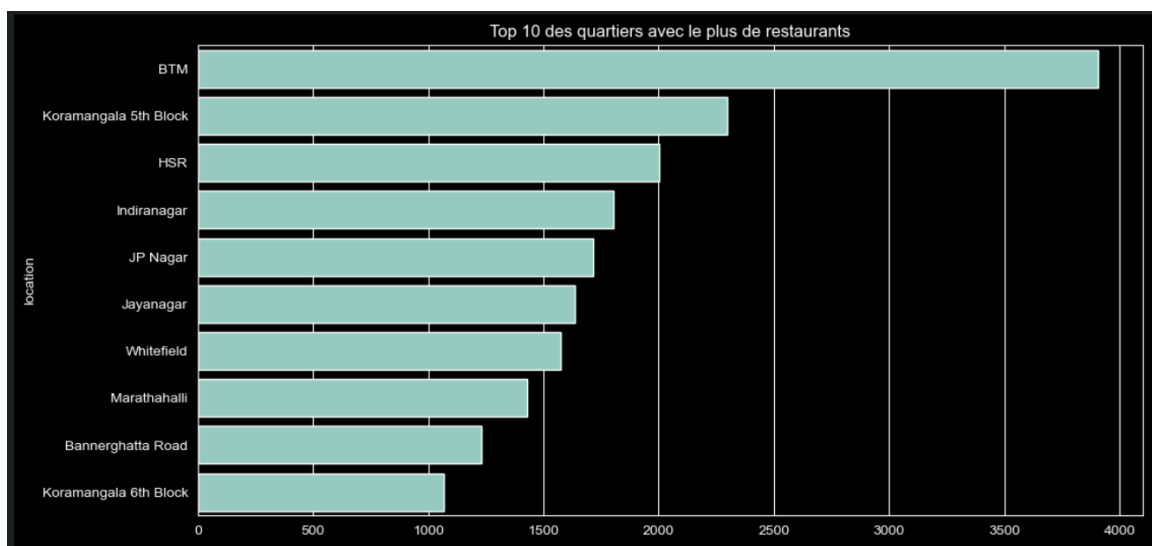
Le jeu de données a nécessité un nettoyage rigoureux pour permettre une analyse statistique fiable :

- **Normalisation des notes** : Extraction de la valeur numérique de la colonne `rate` (ex: "4.1/5" transformé en 4.1).
- **Conversion monétaire** : Nettoyage des caractères spéciaux et conversion du coût moyen (`approx_cost`) en format numérique.
- **Gestion des valeurs manquantes** : Imputation ou suppression des lignes incomplètes pour les variables critiques (localisation et prix).

## Profils de Quartiers et Répartition Géographique

L'analyse par quartier révèle une segmentation très claire du marché :

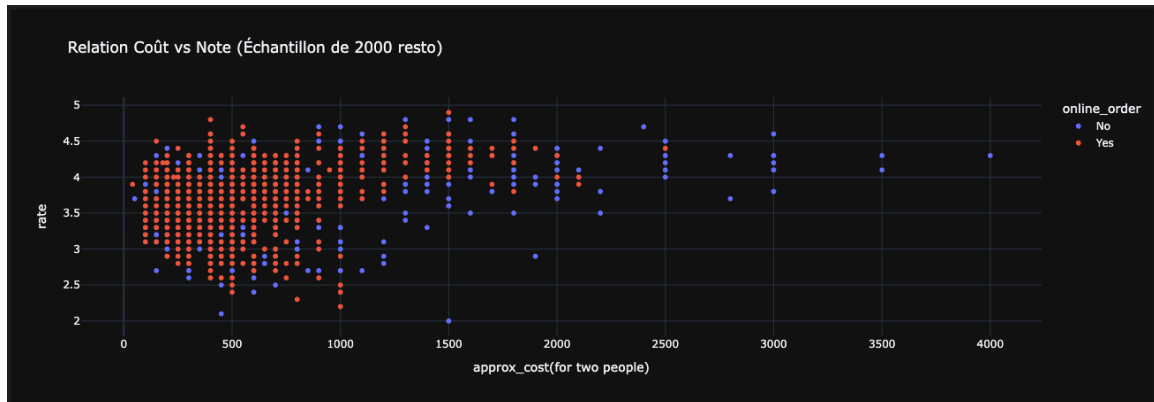
- **Les Hubs de Livraison (BTM, HSR, Koramangala)** : Ces quartiers affichent la plus haute densité de restaurants. Ils sont caractérisés par une dominance du service "Delivery" et des prix abordables.
- **Les Centres Premium (Lavelle Road, Indiranagar)** : Ces zones regroupent les établissements les plus coûteux. On y observe une forte proportion de réservations de table ("Book Table") et une concentration de "Fine Dining".
- **Zones de Vie Nocturne (Brigade Road, MG Road)** : Spécialisées dans les catégories "Pubs and Bars", ces zones reçoivent un volume de votes très élevé, indiquant une forte popularité auprès d'un public jeune.



## Analyse de la Popularité et du Coût

L'étude croisée entre les tarifs et les notations permet de tirer plusieurs conclusions :

1. **Corrélation Prix/Note** : Il existe une corrélation positive modérée. Les restaurants avec un coût plus élevé ont tendance à avoir des notes plus constantes (souvent au-dessus de 4.0), ce qui s'explique par un meilleur service et un cadre plus soigné.
2. **Accessibilité de la Qualité** : Malgré la tendance précédente, le quartier de **Banashankari** présente de nombreux restaurants à bas prix avec d'excellentes notes, ce qui en fait une zone d'opportunité pour les consommateurs à petit budget.



## Limites et Recommandations

- **Limites** : Le dataset contient beaucoup de valeurs manquantes pour la colonne `dish_liked`, ce qui empêche une analyse précise des plats préférés par quartier.
- **Recommandations** : Pour un investisseur, les quartiers comme **Bellandur** présentent un fort potentiel car la demande (votes) est élevée mais l'offre de restaurants "Dine-out" est moins saturée que dans le centre-ville.

## Conclusion

Bengaluru offre un paysage hôtelier diversifié où la livraison à domicile domine le volume, mais où l'expérience en salle (Dine-out) reste le moteur des meilleures notations. Cette analyse a permis de segmenter la ville en pôles stratégiques, facilitant ainsi la compréhension des besoins locaux.

### Liens :

Dépôt Github :

<https://github.com/pllarouch2/ZomatoProjectFinal>

Application Streamlit :

<https://zomatoprojec-vawzk9ear2fzfm5bzsqsyx.streamlit.app>