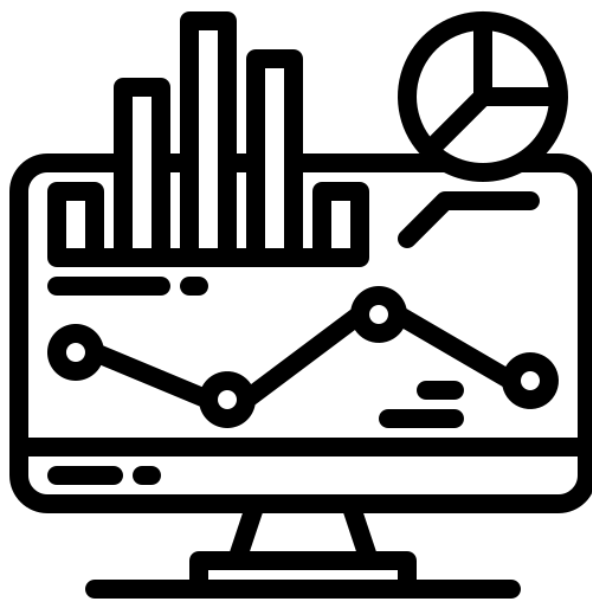


ANALYSE EXPLORATOIRE ET VISUALISATION DE DONNÉES

EYNAUD MATHIS & SAAD BEN SALAH MONGI

5 IR-SDBD

Janvier 2025



1 Introduction

Vous êtes-vous déjà demandé pourquoi certaines livraisons sont plus rapides que d'autres ? Quels sont les facteurs cachés qui influencent ces délais ? Dans un contexte où la rapidité et l'efficacité des services de livraison sont essentielles pour la satisfaction client, comprendre les éléments qui impactent ces délais devient une priorité. Ce rapport s'appuie sur un jeu de données conçu pour modéliser les temps de livraison de nourriture, en prenant en compte divers facteurs réels. L'objectif est d'explorer les relations entre des variables clés telles que la distance, le trafic, les conditions météorologiques, et bien d'autres, afin d'identifier leur influence sur les délais de livraison.

Le jeu de données utilisé, accessible sur Kaggle : <https://www.kaggle.com/denkuznetz/food-delivery-time-prediction>

Le code source est accessible sur le dépôt suivant : https://github.com/Mathis05000/data_visualisation

2 Présentation du Jeu de Données

Nous avons utilisé un jeu de données de Kaggle, contenant 1000 lignes, pour modéliser les temps de livraison de repas. Les variables incluent : **Order_ID** (identifiant de la commande), **Distance_km** (distance de livraison), **Weather** (conditions météorologiques), **Traffic_Level** (niveau de trafic), **Time_of_Day** (moment de la journée), **Vehicle_Type** (type de véhicule), **Preparation_Time_min** (temps de préparation), **Courier_Experience_yrs** (expérience du livreur) et **Delivery_Time_min** (temps total de livraison). L'objectif est d'analyser les relations entre ces variables pour comprendre les facteurs influençant les délais de livraison.

3 Conditions Imprévisibles : Météo et Trafic

Pluie, neige ou embouteillages ? Les conditions extérieures ont un impact direct sur la rapidité des livraisons. Grâce à nos données, nous pouvons faire une analyse approfondie qui nous permettra de quantifier cet impact.

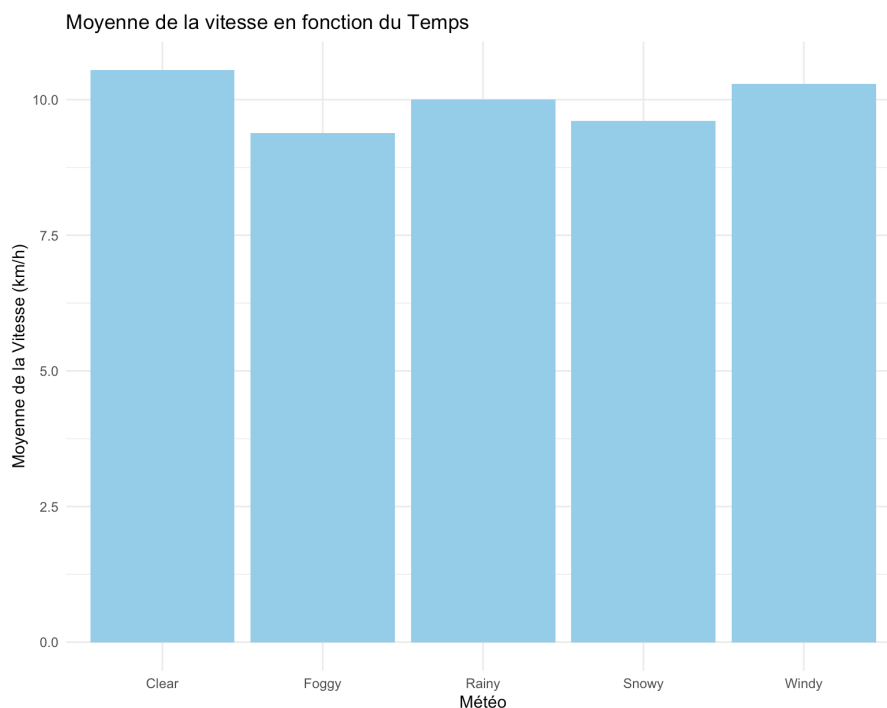


FIGURE 1 – Moyenne de la vitesse en fonction du Temps

Les conditions météorologiques exercent une influence sur la vitesse des livraisons. La pluie, la neige ou le brouillard réduisent la visibilité, alourdissent le trafic et dégradent l'état des routes, compliquant ainsi la tâche des livreurs. Ces facteurs extérieurs entraînent inévitablement des ralentissements et des retards dans les livraisons, comme en témoignent les données présentées dans le graphique. Pour faire face à ces défis, les applications et les sites de livraisons ont dû intégrer les prévisions météorologiques dans leurs outils de planification. Cette approche proactive permet d'anticiper les perturbations, d'optimiser les itinéraires et d'informer les clients en

temps réel de l'état de leur commande.

Après avoir étudié l'influence des conditions météorologiques, intéressons-nous à un autre facteur clé : le niveau de trafic. Le graphique suivant met en évidence l'influence de l'intensité du trafic sur la rapidité des livraisons.

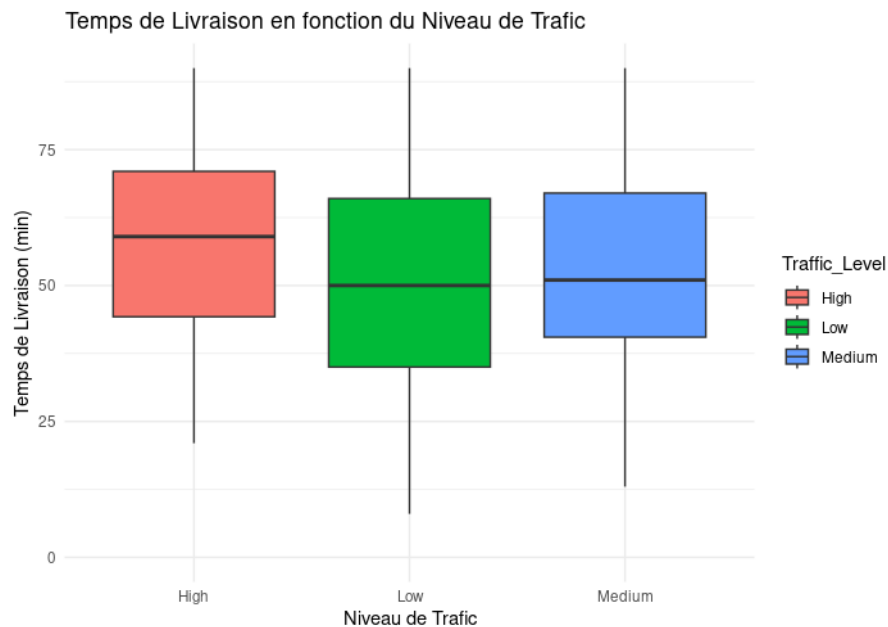


FIGURE 2 – Boxplot pour explorer les temps de livraison en fonction du niveau de trafic

Le trafic, un véritable bouchon pour nos livraisons !

Comme attendu, le niveau de trafic exerce une influence directe sur les délais de livraison. Le boxplot le confirme : en période de forte congestion, le minimum, le premier quartile, la médiane, le troisième quartile et le max sont respectivement supérieurs aux autres niveaux de trafic. À l'inverse, lorsque le trafic est fluide, les livraisons sont plus rapides mais plus variables.

4 Impact de la distance et de l'environnement sur les délais de livraison

Alors que les conditions météorologiques et le trafic échappent en grande partie au contrôle des livreurs et des clients, la distance, elle, est un paramètre directement lié au choix du consommateur. En effet, en sélectionnant un restaurant ou une enseigne plus ou moins éloignée de son domicile, le client exerce une influence directe sur le délai de livraison. Cette étude vise à quantifier l'impact de cette variable sur les temps de livraison.

La distance parcourue par les livreurs constitue un facteur déterminant dans les délais de livraison. Plus la distance à parcourir est longue, plus le temps de livraison tend à s'allonger. Cette relation positive est visualisée par la tendance générale des points à se répartir le long d'une droite ascendante.

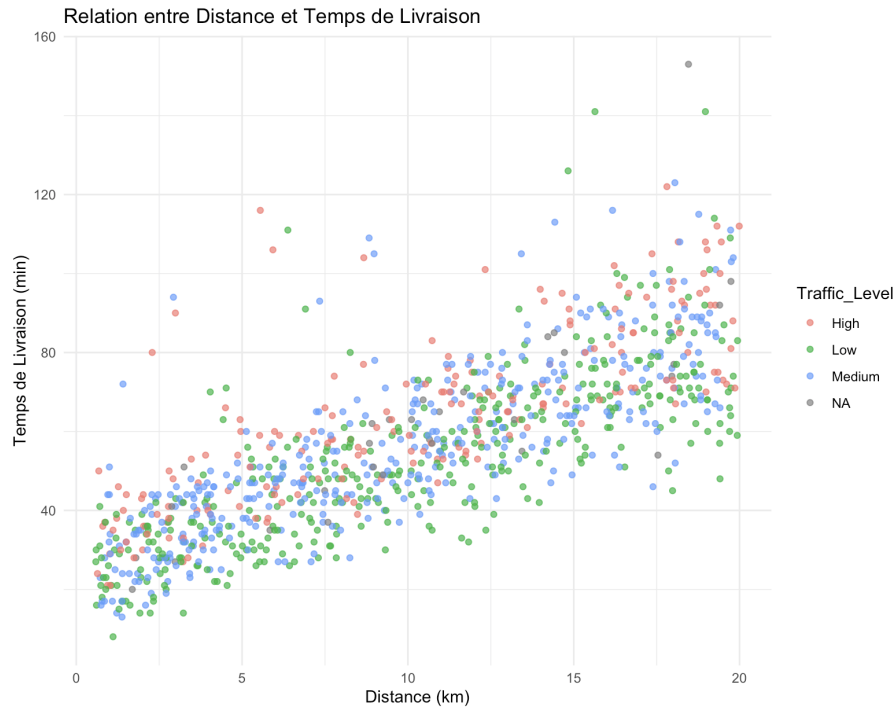


FIGURE 3 – Nuage de points de la relation entre la distance et le temps de livraison

Le graphique met en évidence l'influence combinée de la distance et du niveau de trafic sur les délais de livraison. Si la distance est un facteur prédictif évident des temps de livraison, le niveau de congestion vient nuancer cette relation. En effet, on observe une dispersion plus importante des points pour les trajets effectués en période de fort trafic, suggérant une variabilité accrue des temps de livraison. Cette dispersion est particulièrement marquée pour les longues distances, où les embouteillages peuvent allonger considérablement la durée du trajet. À l'inverse, pour les courtes distances, l'impact du trafic est moins prononcé.

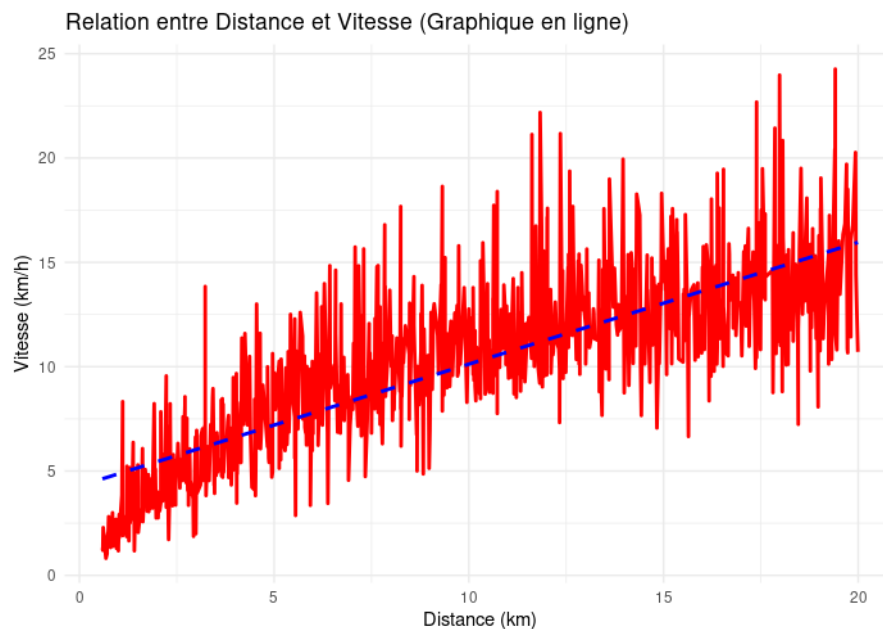


FIGURE 4 – Graphique en ligne pour la relation entre distance et vitesse de livraison

On peut remarquer sur la figure 4 une augmentation progressive de la vitesse moyenne en fonction du temps de livraison. En effet, on peut penser que le temps passé à faible allure (récupération de la commande, chargement dans le moyen de transport, recherche de l'adresse exact, appel au client pour la remise de la course) sont de plus en plus négligeables au fur et à mesure que la durée de la livraison augmente. De plus, on peut imaginer qu'un livreur va être amené à prendre de plus gros axes de circulation au plus la distance de la course sera grande. Le graphique révèle également un aspect particulièrement intéressant : l'existence d'un plafond apparent dans les

vitesse observées. Cette limitation de vitesse maximale, qui semble s'imposer indépendamment de la distance parcourue, suggère l'existence de contraintes externes influençant le comportement des livreurs. Malheureusement, l'absence de données géographiques détaillées dans notre jeu de données ne nous permet pas d'affirmer avec certitude les raisons de ce plafonnement. Cependant, plusieurs hypothèses peuvent être formulées :

- **Limitations réglementaires** : Les livreurs sont soumis à des limitations de vitesse qui varient en fonction des zones géographiques et des types de voies. Ce plafond pourrait ainsi refléter le respect de ces réglementations.
- **Conditions de circulation** : Les zones urbaines, caractérisées par un trafic dense et des infrastructures complexes, peuvent imposer des vitesses moyennes plus faibles, quel que soit la distance à parcourir.
- **Nature des livraisons** : Les livraisons en centre-ville, qui impliquent souvent des manœuvres complexes et des arrêts fréquents, peuvent conduire à des vitesses moyennes plus faibles que les livraisons en zone périurbaine ou rurale.

En l'absence de données plus granulaires, il est difficile de trancher entre ces différentes hypothèses. Néanmoins, l'observation de ce plafonnement dans les vitesses suggère que des facteurs externes, liés à l'environnement urbain et aux contraintes réglementaires, jouent un rôle important dans la détermination de la vitesse des livreurs.

5 L'expérience des Livreurs

L'ancienneté d'un livreur joue un rôle déterminant dans la rapidité et l'efficacité des livraisons. Comme dans tout métier, une expertise acquise au fil du temps permet d'optimiser les trajets, de mieux gérer les imprévus et ainsi de réduire les délais.

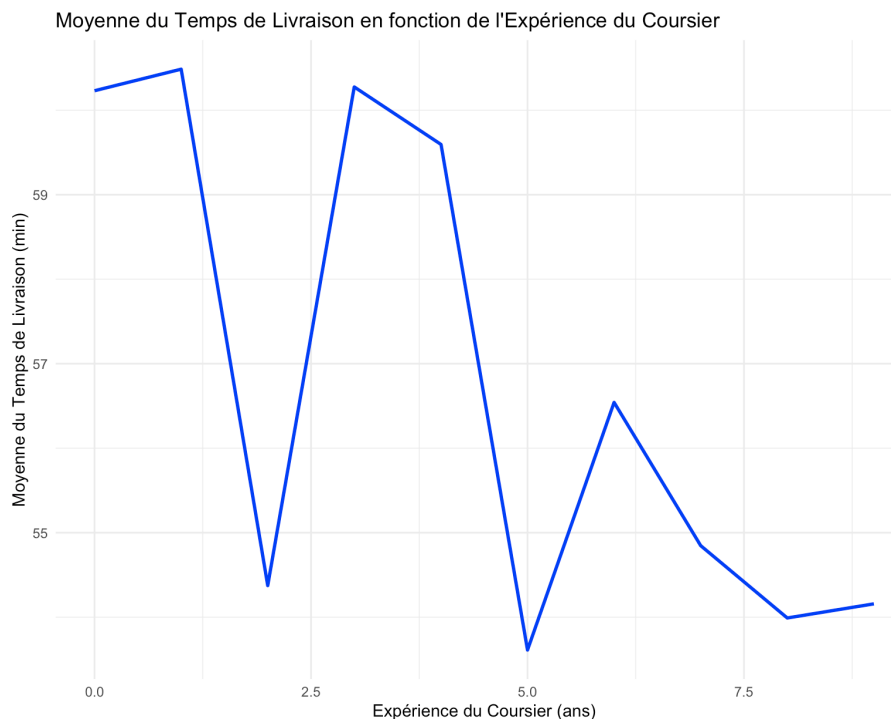


FIGURE 5 – Moyenne du temps de livraison en fonction de l'expérience du livreur

L'évolution du temps de livraison en fonction de l'expérience des livreurs présente des fluctuations inattendues. Après une phase d'adaptation initiale où les temps de livraison augmentent, on observe une amélioration significative liée à l'acquisition d'expérience. Cependant, cette tendance s'inverse par la suite, suggérant que d'autres facteurs, tels que la complexité des livraisons, la motivation des livreurs ou des variations aléatoires dans les données, peuvent entrer en jeu. Ces résultats soulignent la complexité de la relation entre l'expérience et la performance des livreurs et appellent à une analyse plus approfondie en considérant d'autres variables.

Si le graphique précédent nous a permis d'observer une relation complexe entre l'expérience du livreur et le temps de livraison, l'analyse de la distribution des temps de livraison en fonction de l'expérience et du type de véhicule apporte un éclairage supplémentaire sur les facteurs influençant les performances.

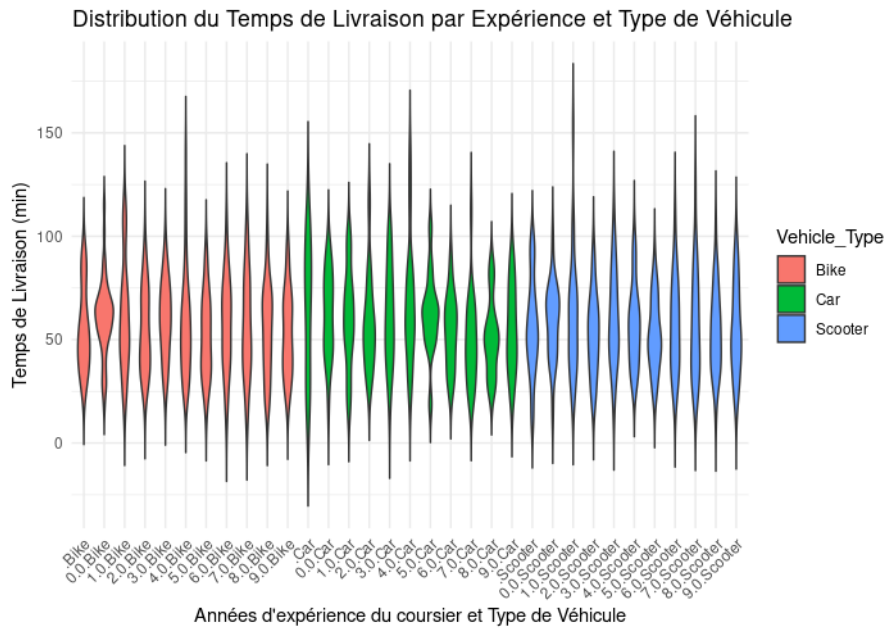


FIGURE 6 – Distribution des Temps de Livraison par Expérience et Type de Véhicule

Le graphique en violon permet de visualiser la distribution des temps de livraison en fonction de l'expérience du livreur et du type de véhicule utilisé. Chaque violon représente une combinaison spécifique de ces deux variables, et la largeur du violon indique la densité des données.

Premièrement, on constate que l'expérience du livreur a un impact notable sur le temps de livraison. En général, plus l'expérience du livreur est grande, plus le temps de livraison diminue, quel que soit le type de véhicule utilisé. Cela suggère que l'expérience permet une meilleure optimisation des itinéraires.

Concernant le type de véhicule, il influence également fortement les délais de livraison. Les livreurs à vélo sont généralement plus rapides pour les courtes distances, tandis que les voitures sont souvent plus adaptées pour les longues distances. Les scooters semblent offrir un compromis entre ces deux options, bien que leur performance puisse varier selon les conditions de trafic et l'infrastructure.

Enfin, il existe des interactions entre l'expérience du livreur et le type de véhicule. Par exemple, un livreur expérimenté en vélo peut être plus rapide qu'un livreur débutant en voiture. La variabilité des temps de livraison est également importante, ce qui suggère que d'autres facteurs, comme les conditions météorologiques ou le trafic, influencent également les délais de livraison.

6 Conclusion

L'enquête sur les délais de livraison révèle la complexité des facteurs influençant les temps de livraison, tels que l'expérience du livreur, le type de véhicule, la distance, le trafic et les conditions météorologiques. Malgré certaines tendances identifiées, quelques questions restent ouvertes. Par exemple, une analyse géographique plus poussée pourrait permettre de repérer des zones où les délais sont systématiquement plus longs et optimiser les itinéraires. En croisant ces données avec des informations locales, comme la densité de population ou le type de voirie, on pourrait affiner l'analyse.