

Analysez les ventes d'une librairie avec R ou Python

Mission

Mis à jour le mercredi 31 mai 2023



Vous êtes consultant Data Analyst chez Lapage, une grande librairie généraliste en ligne très réputée. Vous êtes directement rattaché au service Marketing.



Lapage était originellement une librairie physique avec plusieurs points de vente. Mais devant le succès de certains de ses produits et l'engouement de ses clients, elle a décidé depuis 2 ans d'ouvrir un site de vente en ligne. Vous intervenez car l'entreprise souhaite faire le

point après deux ans d'exercice, et pouvoir analyser ses points forts, ses points faibles, les comportements clients, etc.

L'équipe Marketing que vous intégrez se compose de :

- Annabelle, la responsable Marketing ;
- Julie, chargée d'études en marketing ;
- Antoine, chargé de produit marketing.

Voici le mail de résumé envoyé par Annabelle, que vous trouvez dès votre arrivée :

Objet : Bienvenue dans l'équipe Marketing

De : Annabelle

À : Moi

Hello,

Bienvenue chez Lapage ! J'espère que tu te plairas parmi nous.

Comme tu l'as certainement compris, nous avons ouvert il y a deux ans de cela notre librairie en ligne, et nous aimerions faire le point à présent. Cela nous permettra de décider de la marche à suivre, par exemple décider si nous devons créer certaines offres, adapter certains prix, etc.

Cette analyse va être découpée en deux parties :

- Une analyse des différents indicateurs de vente, qui a été demandée directement par Antoine. Je t'ai mis en pièce jointe l'e-mail qu'il m'a envoyé résumant ses demandes.
- Une analyse plus ciblée sur les clients : l'objectif serait cette fois-ci de comprendre le comportement de nos clients en ligne, pour pouvoir ensuite comparer avec la connaissance acquise via nos librairies physiques. Pour cela, je t'invite à voir

avec Julie, qui te précisera directement ses demandes à ce niveau.

La base de données est en pièce jointe. N'hésite pas si tu as des questions, je me tiens à ta disposition.

Excellente journée à toi !

Annabelle

Pièces jointes :

- [Mail forwardé d'Antoine](#)
- [Fichier ZIP contenant les bases](#)

Après avoir consulté le mail d'Antoine, vous prenez directement contact avec Julie.

Hello Julie, je suis le nouveau Data Analyst chargé d'analyser les données de ventes en ligne. Annabelle m'a dit que tu avais des demandes concernant l'analyse des comportements clients.

Salut ! Bienvenue dans l'équipe ! J'allais justement prendre contact avec toi. C'est exactement cela ! Je vais avoir besoin que tu regardes certaines corrélations, pour voir s'il n'y a pas certains points que nous pourrions analyser.

Bien sûr. Tu as déjà une idée précise de ce que tu veux que je regarde ?

Oui ! J'aimerais que tu regardes en particulier :

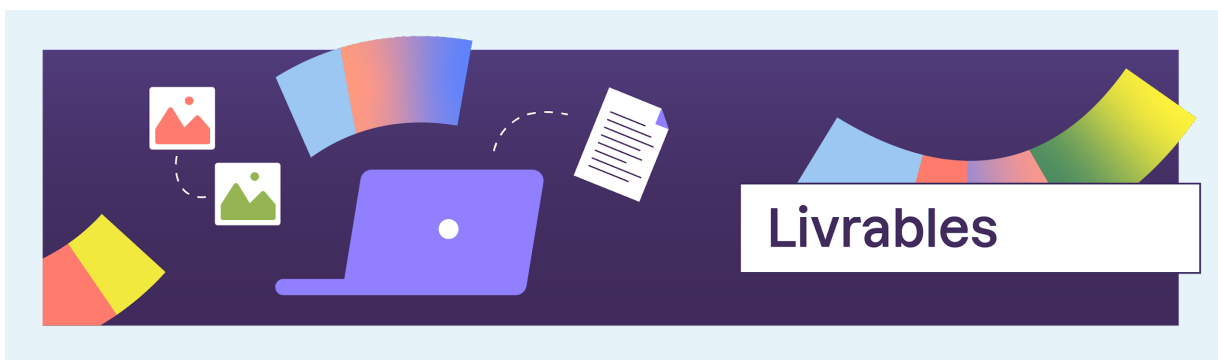
- **le lien entre le genre d'un client et les catégories des livres achetés ;**
- **et ensuite le lien entre l'âge des clients et le montant total des achats, la fréquence d'achat, la taille du panier moyen et les catégories des livres achetés.**

Ça marche, je m'occupe de ça !

Vous gardez en tête toutes ces informations et commencez à regarder les données.

Dans ce projet, vous êtes libre d'utiliser le langage de programmation de votre choix entre R et Python pour mener à bien votre analyse.

Vous pouvez utiliser Jupyter (pour R et pour Python) ou R Markdown (pour R).



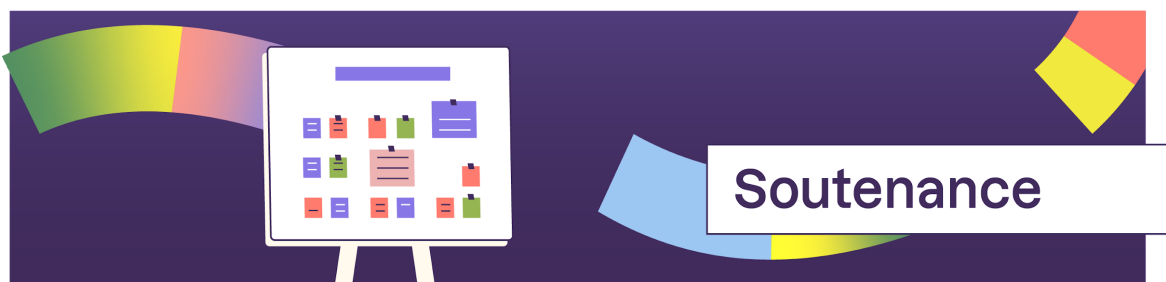
1. Le ou les notebooks R Markdown ou Jupyter contenant les traitements et les différentes analyses effectués pour répondre aux demandes d'Antoine et Julie.

Pour faciliter votre passage devant le jury, déposez sur la plateforme, dans un dossier zip nommé “**Titre_du_projet_nom_prénom**”, votre livrable nommé comme suit : **Nom_Prénom_n° du livrable_nom du livrable_date de démarrage du projet**. Cela donnera :

- *Nom_Prénom_1_notebook_mmaaaa*

Par exemple, votre livrable peut être nommé comme suit :

Dupont_Jean_1_notebook_01203.



Durant la présentation orale, votre évaluateur jouera le rôle d'un des membres de votre équipe (Annabelle, Julie ou Antoine) à qui vous présentez vos résultats :

- **Présentation des livrables (15 minutes)**
 - Détail du nettoyage des données : vous présenterez ici l'ensemble des choix et actions que vous avez réalisés afin d'obtenir un jeu de données exploitable.
 - Restitution de l'analyse : les chiffres clés, KPI, graphiques générés, corrélations calculées, etc.
- **Discussion (5 minutes)**
 - L'évaluateur vous challengera sur vos choix.
- **Debrief (5 minutes)**
 - Votre évaluateur reprend son rôle de mentor pour que vous puissiez débriefer ensemble.

Votre présentation devrait durer 15 minutes (+/- 5 minutes). Puisque le respect des durées des présentations est important en milieu professionnel, les présentations en dessous de 10 minutes ou au-dessus de 20 minutes peuvent être refusées.

Compétences évaluées

- Réaliser un test statistique
- Réaliser une analyse bivariée pour interpréter des données
- Analyser des séries temporelles