

Bike Purchase Project:

Introduction:

L'objectif de cette analyse de données est de déterminer les tendances clés dans les achats de vélo, en se basant sur la base de données 'bike_buyers'.

Méthodologie:

La base de données contient 1000 lignes et 13 colonnes, elle étudie les achats de vélos en fonction des caractéristiques suivantes:

- Revenu
- Distance
- Age
- Genre
- Statut familial
- Région
- Possession de voitures
- Possession de maison
- Éducation
- Occupation

Pour parvenir aux résultats on a suivi les étapes suivantes:

- **Extraire et transformer les données:** supprimer les redondants, remplacer quelques valeurs qui se répètent dans plus d'une colonne pour éviter les confusions...
- Faire des **analyses statistiques** exploratoires pour déterminer les facteurs qui influencent l'achat de vélo. Pour ce faire on a étudié la **valeur moyenne** du revenu et son effet sur l'Achat des vélos. En plus, on a utilisé les tableaux croisés dynamiques et les formules d'Excel pour calculer le nombre de personne qui ont acheté ou non des vélos, pour chaque catégorie d'âge et chaque intervalle de distance. Pour les catégories d'âge on a considéré la mesure suivante:
$$\frac{\text{nombre de personne qui ont acheté (ou non)}}{\text{le nombre total dans la catégorie d'âge}}$$

Le choix de cette mesure est dû au fait que la population n'est pas uniformément distribuée par rapport aux catégories d'âge. Une mesure similaire a été considérée dans l'étude en fonction de l'intervalle de distance.

Résultats:

- Le revenu des clients qui ont acheté des vélos sont légèrement plus importants.
- Quand on visualise le pourcentage des clients qui ont acheté des vélos sur le total, en fonction de la distance, on remarque que: quand la distance augmente les clients tendent plus à ne pas acheter des vélos. Chose qui semble logique, puisque pour ces distances l'utilisation d'une voiture semble plus adéquate. En plus, utiliser ce **rapport** comme **mesure** est important dans ce cas car la population n'est pas **uniformément distribuée** en fonction de la distance.
- Les clients dans la catégorie d'âge : [30-60] ont tendance plus à acheter des vélos, que les catégories 30 et moins et 60 et plus. Aussi dans ce cas on a utilisé le **rapport** comme **mesure** car la population n'est pas **uniformément distribuée** en fonction d'âge.
- La clientèle célibataire qui achète des vélos est légèrement plus importante (la voiture peut faciliter un déplacement en famille).
- La clientèle de la région du pacifique tend plus à acheter des vélos (Conditions météorologiques?)
- Les clients avec un bachelor ou un graduate degree tendent plus à acheter des vélos.
- L'achat de vélos diminue quand le nombre de voitures augmente.
- Les personnes qui travaillent dans des emplois de type management ou manual tendent plus à ne pas acheter de vélos.

Le dashboard qui résume tous ces résultats en plus de toutes les traces de notre démarche est mis en pièce jointe.

Recommandations:

En se basant sur les résultats de nos analyses, nous présentons la recommandation suivante:

Les caractéristiques des clients potentiels sont généralement âgées de [30-60] ans, qui ne possèdent pas de voiture, et qui parcourent des distances qui ne dépasse pas 5 miles dans la majorité des cas. En plus, on remarque aussi que le nombre d'achats de vélos est plus important aussi dans la région du Pacifique, entre les gens avec des bachelor ou un graduate degree ou ceux qui travaillent dans des emplois de type management ou manual.

Conclusion:

En conclusion, cette analyse a permis de répondre à notre objectif qui était de trouver les caractéristiques communes entre les clients de vélos, afin de les utiliser pour repérer la clientèle potentielle. Ces points communs sont, en principe l'âge, la distance parcourue, la non possession de voiture et les conditions climatiques.