

Analyse BI pour **US Candy Distributor** et **Churn des Clients d'une Banque Européenne**

Expert BI :
Spero TESSY
Stephene WANTCHEKON

October 11, 2024

1 Introduction

Cet exercice est conçu pour simuler un contexte d'entreprise, permettant d'appliquer des compétences en analyse de données et en Business Intelligence (BI) sur des ensembles de données réels. Les données utilisées proviennent de Maven Analytics, une plateforme reconnue pour fournir des ensembles de données à des fins d'apprentissage et d'analyse. Les ensembles de données, y compris ceux concernant le churn des clients d'une banque européenne et les ventes d'un distributeur de bonbons américain, sont disponibles pour une utilisation gratuite, tant que les utilisateurs respectent les conditions d'utilisation énoncées par Maven.

Vous pouvez utiliser les données fournies par Maven Analytics sans crainte de poursuites tant que vous respectez leurs conditions d'utilisation. Assurez-vous que votre utilisation ne viole pas les droits d'auteur ou les droits d'autres personnes, et évitez de partager des informations sensibles. Si vous envisagez de publier vos résultats ou analyses, assurez-vous également d'avoir les droits nécessaires pour le faire.

Il est impératif de noter que toutes les utilisations de ces données doivent se conformer aux lois et règlements applicables, y compris le respect de la vie privée et des droits d'auteur. Les utilisateurs doivent avoir le droit de publier tout contenu qu'ils soumettent et ne doivent pas publier d'informations qui appartiennent à d'autres sans autorisation. En respectant ces directives, les utilisateurs peuvent explorer les ensembles de données fournis par Maven Analytics sans crainte de poursuites, tout en appliquant des techniques d'analyse pertinentes pour répondre aux problématiques d'entreprise.

De plus, Maven encourage les utilisateurs à partager leurs visualisations (et toute table croisée ou code applicable) sur LinkedIn en mentionnant @Maven Analytics. Cela permet non seulement de montrer votre travail, mais aussi de recevoir des retours constructifs de la part de la communauté. En respectant ces directives, les utilisateurs peuvent explorer les ensembles de données fournis par Maven Analytics sans crainte de poursuites, tout en appliquant des techniques d'analyse pertinentes pour répondre aux problématiques d'entreprise.

2 Contexte et Problématiques : US Candy Distributor

L'entreprise *US Candy Distributor*, un distributeur national de bonbons aux États-Unis, rencontre plusieurs défis à surmonter. Malgré des ventes stables, le chiffre d'affaires stagne, et les coûts de transport élevés affectent la rentabilité. Face à ces défis, l'entreprise a décidé d'engager un expert en BI pour analyser les données de ventes et d'expédition afin d'identifier des leviers d'optimisation.

Les principaux objectifs de l'analyse sont :

- Optimiser les routes d'expédition entre les usines et les clients.
- Identifier les produits les plus rentables.
- Redistribuer certains produits entre différentes usines pour minimiser les coûts de transport.

Les données disponibles couvrent les informations suivantes : commandes, objectifs de vente, clients, produits, ainsi que les localisations géospatiales des usines et des clients.

2.1 Problématique 1: Optimisation des Routes d'Expédition

Problème : Les coûts logistiques sont un point sensible pour l'entreprise. Il est crucial d'identifier les routes d'expédition les plus et les moins efficaces entre les usines et les clients.

Objectif : Trouver des moyens d'optimiser les trajets de livraison pour réduire les coûts et améliorer la rapidité des livraisons.

Variables impliquées :

- Localisation des clients (`customer_location`).
- Localisation des usines (`factory_location`).
- Distance usine-client (`shipping_distance`).
- Temps de livraison (`delivery_time`).
- Coût d'expédition (`shipping_cost`).

2.2 Problématique 2: Analyse des Marges par Produit

Problème : Certains produits ont une marge bénéficiaire inférieure, ce qui impacte la rentabilité globale.

Objectif : Identifier les produits les plus profitables et ajuster la stratégie pour les produits moins rentables.

Variables impliquées :

- Produit (`product_name`).
- Coût de production (`production_cost`).
- Prix de vente (`sales_price`).
- Quantité vendue (`quantity_sold`).

2.3 Problématique 3: Relocalisation des Produits pour Optimiser les Routes

Problème : Certaines lignes de produits sont produites dans des usines éloignées des clients, entraînant des coûts de transport élevés.

Objectif : Minimiser les coûts en redistribuant les produits entre les usines pour rapprocher la production des clients.

Variables impliquées :

- Produit (`product_name`).
- Usine actuelle (`current_factory`).
- Coût d'expédition (`shipping_cost`).
- Distance usine-client (`shipping_distance`).

3 Contexte et Problématiques : Churn des Clients d'une Banque Européenne

L'une des préoccupations majeures d'une banque européenne est la fidélité de ses clients. Une hausse du taux de churn (attrition ou départ des clients) affecte directement la rentabilité et la croissance de la banque. En effet, chaque départ de client signifie une perte de revenus récurrents, mais également des coûts supplémentaires liés à l'acquisition de nouveaux clients. Le département des services clients a observé un nombre croissant de départs parmi les 10,000 clients dont ils détiennent des informations détaillées, notamment les scores de crédit, les soldes des comptes, les produits utilisés, ainsi que l'indication de si le client a déjà quitté la banque ou non.

Afin de contrer cette tendance, la banque cherche à mieux comprendre les raisons du churn et à utiliser ces informations pour retenir ses clients avant qu'ils ne partent. De plus, la banque souhaite adapter sa stratégie marketing et ses services en fonction des différents segments de clients identifiés.

Les objectifs d'analyse sont les suivants :

1. **Identifier les attributs communs aux clients churners** : Quelles sont les caractéristiques qui rendent certains clients plus susceptibles de quitter la banque ? En comprenant ces facteurs, la banque peut développer des stratégies personnalisées pour fidéliser les clients à risque. Par exemple, si les clients avec un faible score de crédit et un faible solde de compte sont plus enclins à partir, la banque pourrait proposer des produits de crédit plus adaptés ou des services spécifiques à ces segments.
2. **Étudier la possibilité de prédire le churn à partir des variables disponibles** : L'objectif est ici d'anticiper le départ des clients avant qu'il ne se produise. Notre entreprise aimerait être capable de prédire si un client est sur le point de partir afin de prendre des mesures proactives pour le retenir. Grâce à un modèle prédictif basé sur des algorithmes de machine learning, la banque pourrait cibler ces clients avec des offres spéciales ou un suivi personnalisé avant qu'ils ne décident de fermer leur compte.
3. **Analyser les différences démographiques au sein des clients de la banque** : Il est crucial pour la banque de comprendre les caractéristiques démographiques de sa clientèle, telles que l'âge, le sexe, et la localisation géographique. Cette analyse permettra de déterminer s'il existe des tendances spécifiques dans certaines catégories de clients, ce qui permettra à la banque d'ajuster son offre de produits et de services. Par exemple, si les clients plus jeunes ont un taux de churn plus élevé, la banque pourrait développer des offres plus adaptées à cette population.
4. **Étudier le comportement des clients par nationalité (allemands, français, espagnols)** : Étant une banque européenne, il est important d'identifier s'il existe des différences dans les comportements financiers entre les nationalités. Cela permettra à la banque de personnaliser ses services en fonction des besoins spécifiques de chaque pays. Par exemple, les clients allemands pourraient avoir une plus grande aversion au risque et donc une préférence pour des produits d'épargne plus sécurisés, tandis que les clients espagnols pourraient être plus enclins à investir dans des produits à rendement élevé.
5. **Segmenter la clientèle pour mieux comprendre les groupes de clients et adapter les services en conséquence** : En créant des segments de clientèle homogènes, la banque pourra personnaliser son marketing et ses services de manière plus efficace. Par exemple, les clients à haut revenu pourraient bénéficier de services premium, tandis que les clients ayant des besoins plus basiques pourraient se voir proposer des produits bancaires standard. Cette segmentation permettra d'optimiser les efforts de fidélisation et de maximiser la rentabilité par segment.

Les variables disponibles dans la base de données incluent :

- Le score de crédit (**credit_score**) : Cette variable permet d'évaluer la solvabilité du client, ce qui peut influencer à la fois son comportement vis-à-vis des produits bancaires et son risque de churn.
- Le solde des comptes (**balance**) : Le montant détenu sur les comptes peut indiquer la relation financière du client avec la banque et peut être un facteur prédictif important du churn.
- Les produits bancaires détenus (**products_held**) : Le nombre et le type de produits détenus par un client peuvent refléter son engagement auprès de la banque.
- Les informations sur la nationalité des clients (**nationality**) : Cette variable permettra d'explorer des différences de comportement selon les pays.
- L'indicateur de churn (**churn**) : Cette variable indique si le client a quitté la banque ou non. C'est la variable cible dans les analyses de churn.