



Projet SuperMarket

Projet SuperMarket.....	1
Contexte du projet.....	1
Les indicateurs de performances.....	2
Tâches à faire.....	3
Livrables.....	4
Annexe.....	6
Dataset:.....	6
Les phases d'un projet de DataViz.....	7
Choisir le bon graphique.....	8
Qualité des données (Data Quality).....	9

Contexte du projet

La croissance des supermarchés dans les villes les plus peuplées augmente et la concurrence sur le marché est également élevée. Le présent cahier des charges définit les spécifications et les objectifs de notre projet d'analyse de données consacré à l'examen des ventes de magasins répartis sur plusieurs sites géographiques au cours d'une période de trois mois.

L'objectif premier de cette analyse est d'acquérir une compréhension approfondie des performances commerciales de nos magasins dans un contexte géographique diversifié et au sein d'une fenêtre temporelle restreinte. Cette démarche vise à mettre en évidence des tendances significatives, des opportunités d'amélioration et à fournir des informations stratégiques pour une gestion plus efficace. Le cahier des charges qui suit détaille les attentes, les exigences et les étapes clés de ce projet d'analyse de données, tout en posant les bases de sa réalisation.

Les indicateurs de performances

Un indicateur de performance (KPI) est une mesure utilisée pour évaluer la réussite, l'efficacité ou la qualité d'un processus, d'une entreprise ou d'un projet. Ils aident à prendre des décisions éclairées et à suivre la progression vers les objectifs.

Voici les mesures que la Direction aimerait suivre dans ce projet.

1. Quel est le chiffre d'affaires global sur la période?
 - a. Répartition globale
 - b. Répartition par branche (A, B, C)
2. Quels sont les bénéfices et pertes ?
 - a. Quel moyen de paiements est le plus utilisé
 - b. La marge bénéficiaire moyenne: Calculer la marge bénéficiaire moyenne pour les produits vendus par le supermarché. Cela peut aider à évaluer la rentabilité globale.
3. Quelles branches comptent le plus de clients membres ?
4. Quelle est la répartition des clients en fonction de leur type ?
5. L'évolution du nombre de clients durant la période
6. Quel est le classement des produits ?
 - a. la proportion du CA
 - b. la quantité (gestion de stocks)
 - c. gestion par (categories de produits)
 - d. plus vendus
 - e. moins vendus
7. Quelle est l'attraction des catégories par genre ?
8. Montrez l'évolution de ventes sur la période
 - a. Mois
 - b. Semaine
 - c. Jour
9. Les clients sont ils satisfaits ? : Calculer le taux de satisfaction des clients
10. Quel est le panier moyen ? : Calculer le montant moyen dépensé par client à chaque visite. Cela peut aider à comprendre le comportement d'achat des clients

Tâches à faire

1. Rédiger un cahier des charges avec les besoins métiers.
Définir les indicateurs de performance. KPI
2. Etudier la disponibilité des données
3. Effectuer un choix technique d'outil de visualisation
 - a. Contrainte (utiliser excel, google sheet)
4. Prendre connaissance du jeu de données
 - a. Lire ou établir le dictionnaire de données **en annexe**
 - b. Noter le nombre de lignes et de colonnes
 - c. Les sources et la fiabilité des données
5. Nettoyer les données si nécessaire (Data Quality)
 - a. les doublons
 - b. les données manquantes
 - c. les incohérences
 - d. les formats inadéquats
 - e. etc ...
6. Répondre à chaque question en choisissant le type de graphe adéquat.
 - a. aide en annexe pour les choix
7. Rassembler les graphiques pour en faire un tableau de board (Dashboard)
 - a. Le dashboard représente la synthèse de votre analyse
8. Quelle conclusion peut-on faire sur cette analyse
9. Quelles recommandations peut-on faire

Livrables

La rédaction d'un rapport d'analyse de données professionnelle et complète est essentielle pour communiquer efficacement vos résultats. Voici une structure suggérée pour votre rapport au format pdf (un ppt en pdf):

Page de Garde :

- Titre du projet.
- Votre nom.
- Date.

Sommaire (Table des Matières) :

- Liste des sections du rapport avec les numéros de page correspondants.

Introduction :

- Contexte du projet : Présentez brièvement le domaine des supermarchés et le but de votre analyse.
- Objectifs : Définissez les objectifs spécifiques de l'analyse.
- Importance du projet : Expliquez pourquoi cette analyse est pertinente.
- Aperçu du rapport : Donnez un aperçu des principales sections du rapport.

1. Collecte de Données :

- Description des sources de données : Où avez-vous obtenu les données sur les supermarchés ? Décrivez les caractéristiques de ces données.
- Processus de collecte : Expliquez comment vous avez rassemblé et préparé les données pour l'analyse.

2. Nettoyage et Prétraitement des Données :

- Identification et gestion des valeurs manquantes.
- Détection et résolution des doublons.
- Transformation des données, le cas échéant.
- Exploration des données pour repérer les incohérences.
- Expliquez comment vous avez évalué et assuré la qualité des données utilisées dans le projet.

3. Analyse des Données :

- Présentation des statistiques descriptives pour les variables clés (moyenne, écart-type, médiane, etc.).
- Visualisations : Créez des graphiques pertinents pour illustrer la distribution des données.
- Calcul des indicateurs de performance (par exemple, la marge bénéficiaire, les ventes par produit, etc.).

4. Tableau de Bord :

- Présentation du tableau de bord interactif, avec des captures d'écran.
- Explication des indicateurs clés inclus dans le tableau de bord.

5. Justification des Choix :

- Expliquez pourquoi vous avez choisi certains indicateurs de performance et des méthodes de visualisation spécifiques.

6. Commentaires des Graphiques :

- Fournissez des commentaires détaillés sur les graphiques présentés.
- Interprétez les tendances et les conclusions tirées de chaque graphique.

7. Problèmes Rencontrés et Solutions :

- Identifiez les défis ou les problèmes que vous avez rencontrés lors de l'analyse des données.
- Présentez les solutions que vous avez mises en œuvre pour résoudre ces problèmes.

8. Axes d'Améliorations :

- Discutez des domaines où le projet pourrait être amélioré à l'avenir, y compris les données supplémentaires, les améliorations techniques, etc.

9. Conclusion : Bilan du Projet :

- Résumez les principales conclusions tirées de l'analyse.
- Revoyez les objectifs initiaux et expliquez si vous les avez atteints.
- Mettez en évidence l'impact de votre analyse.

10. Références :

- Incluez toutes les sources, outils ou méthodes que vous avez utilisés dans votre projet. - Si nécessaire, ajoutez des annexes pour inclure des données supplémentaires, des captures d'écran, des scripts, etc.

Note finale : Assurez-vous que votre rapport est bien structuré, que les informations sont présentées de manière logique et que les graphiques et les tableaux sont lisibles. N'oubliez pas de relire et de réviser votre rapport pour corriger toute erreur grammaticale ou de syntaxe. La clarté et la précision sont essentielles pour un rapport professionnel.



Annexe

Dataset:

https://www.kaggle.com/datasets/aungpyaeap/supermarket-sales?resource=download&select=supermarket_sales+-+Sheet1.csv

Méta data (Dictionnaire de données)

Invoice id: Computer generated sales slip invoice identification number

Branch: Branch of supercenter (3 branches are available identified by A, B and C).

City: Location of supercenters

Customer type: Type of customers, recorded by Members for customers using member card and Normal for without member card.

Gender: Gender type of customer

Product line: General item categorization groups - Electronic accessories, Fashion accessories, Food and beverages, Health and beauty, Home and lifestyle, Sports and travel

Unit price: Price of each product in \$

Quantity: Number of products purchased by customer

Tax: 5% tax fee for customer buying

Total: Total price including tax

Date: Date of purchase (Record available from January 2019 to March 2019)

Time: Purchase time (10am to 9pm)

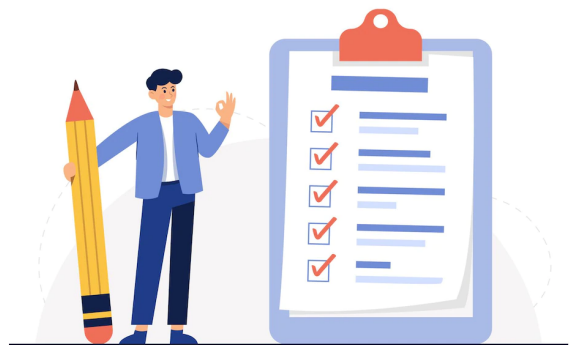
Payment: Payment used by customer for purchase (3 methods are available – Cash, Credit card and Ewallet)

COGS: Cost of goods sold

Gross margin percentage: Gross margin percentage

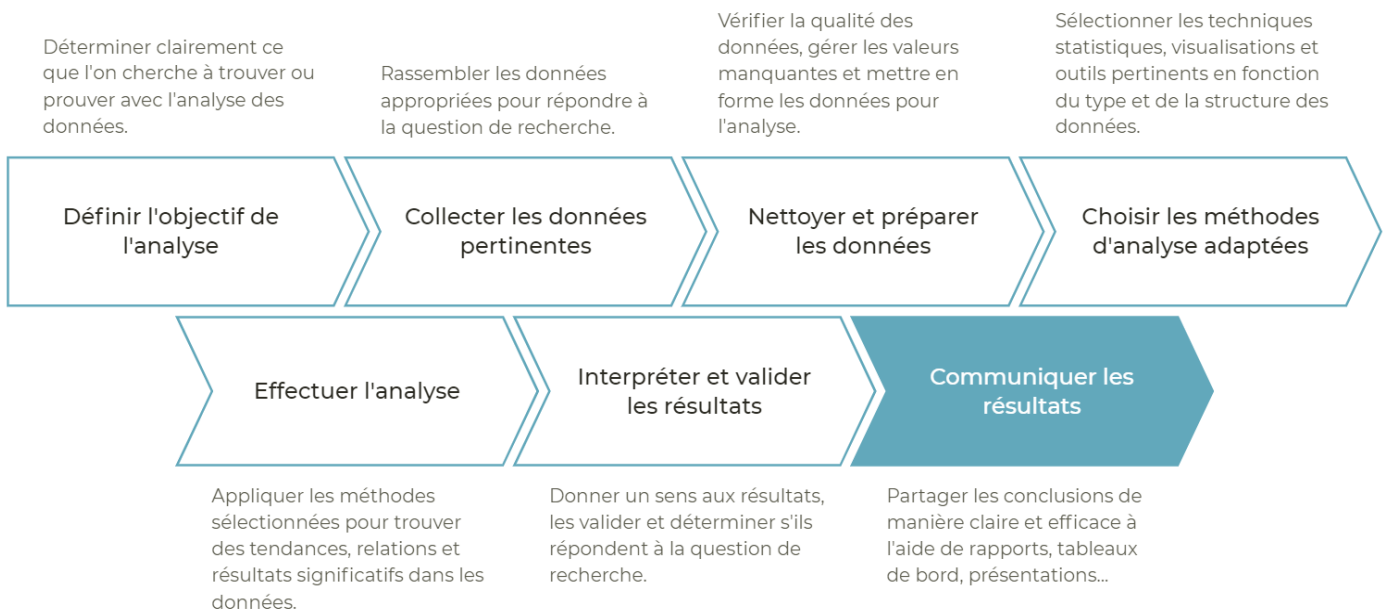
Gross income: Gross income

Rating: Customer stratification rating on their overall shopping experience (On a scale of 1 to 10)



Les phases d'un projet de DataViz

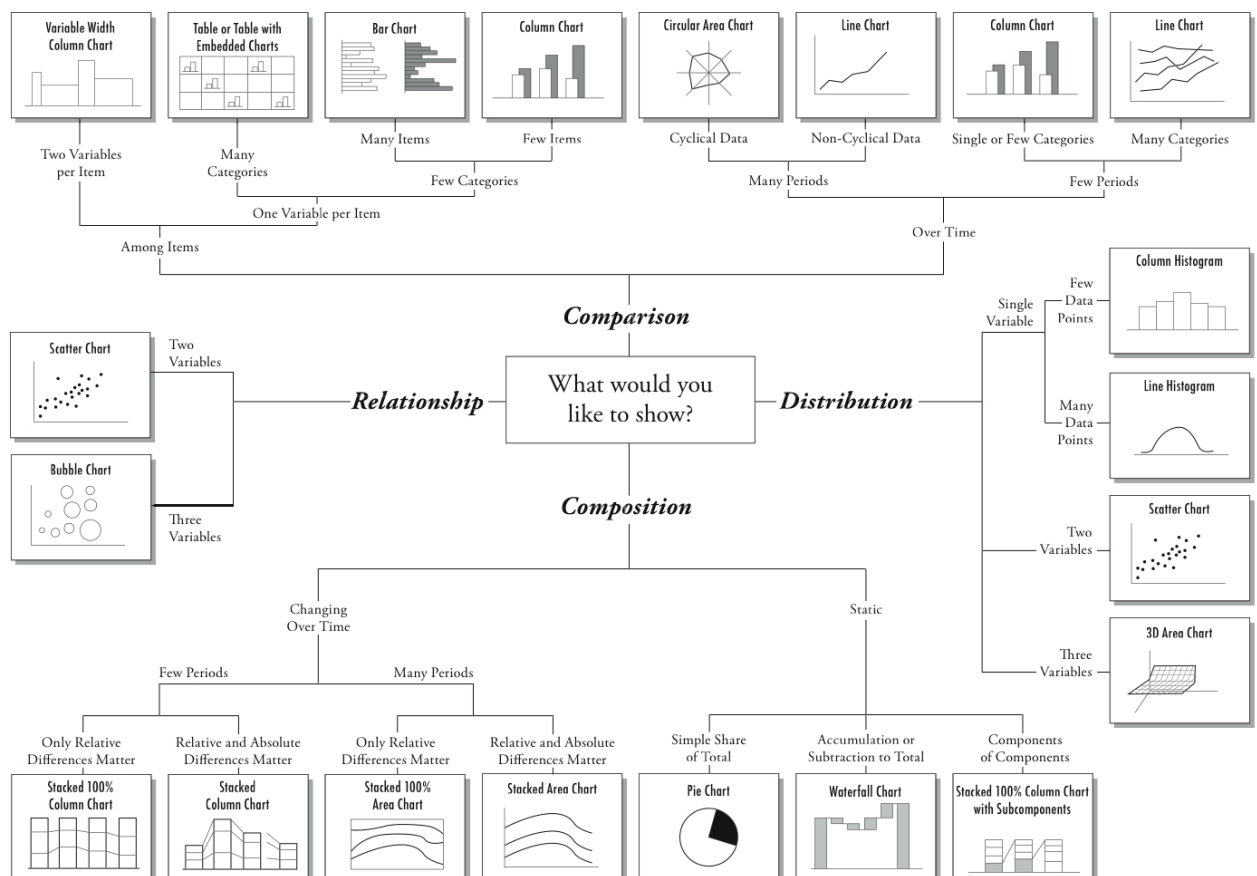
Analyse des données



Choisir le bon graphique

- <https://datavizcatalogue.com/FR/methodes/brainstorming.html>
- <https://extremepresentation.typepad.com/files/chart-chooser-2020.pdf>
- <https://datavizproject.com/>

Chart Suggestions—A Thought-Starter



www.ExtremePresentation.com
© 2009 A. Abela — a.v.abela@gmail.com

Qualité des données (Data Quality)

Qualité des Données



Importance de la qualité des données

La qualité des données est cruciale pour obtenir des visualisations précises et fiables qui aident à la prise de décision.



Sources d'erreurs

Les erreurs peuvent provenir de données manquantes, de valeurs aberrantes, d'incohérences, etc.



Détection des erreurs

Utiliser des statistiques descriptives, des graphiques et des méthodes automatisées pour identifier les problèmes de qualité.



Correction des erreurs

Supprimer ou remplacer les valeurs erronées, normaliser les données pour améliorer la cohérence.