

iraset.org/qbadou/

<https://scholar.google.com/citations?user=9-Wf9YkAAAAJ&hl=en>

<https://www.researchgate.net/profile/Mohammed-Qbadou>

<https://orcid.org/0000-0001-7822-1135>

<https://www.scopus.com/authid/detail.uri?authorId=56705618500>

Lab. M2S2I, ENSET Mohammedia
mohammed.qbadou@univh2c.ma
qbadou@enset-media.ac.ma
mohammed.qbadou@gmail.com



DATA SCIENCE SOUS PYTHON

APPLICATION DE SYNTHÈSE

ANALYSE & VISUALISATION

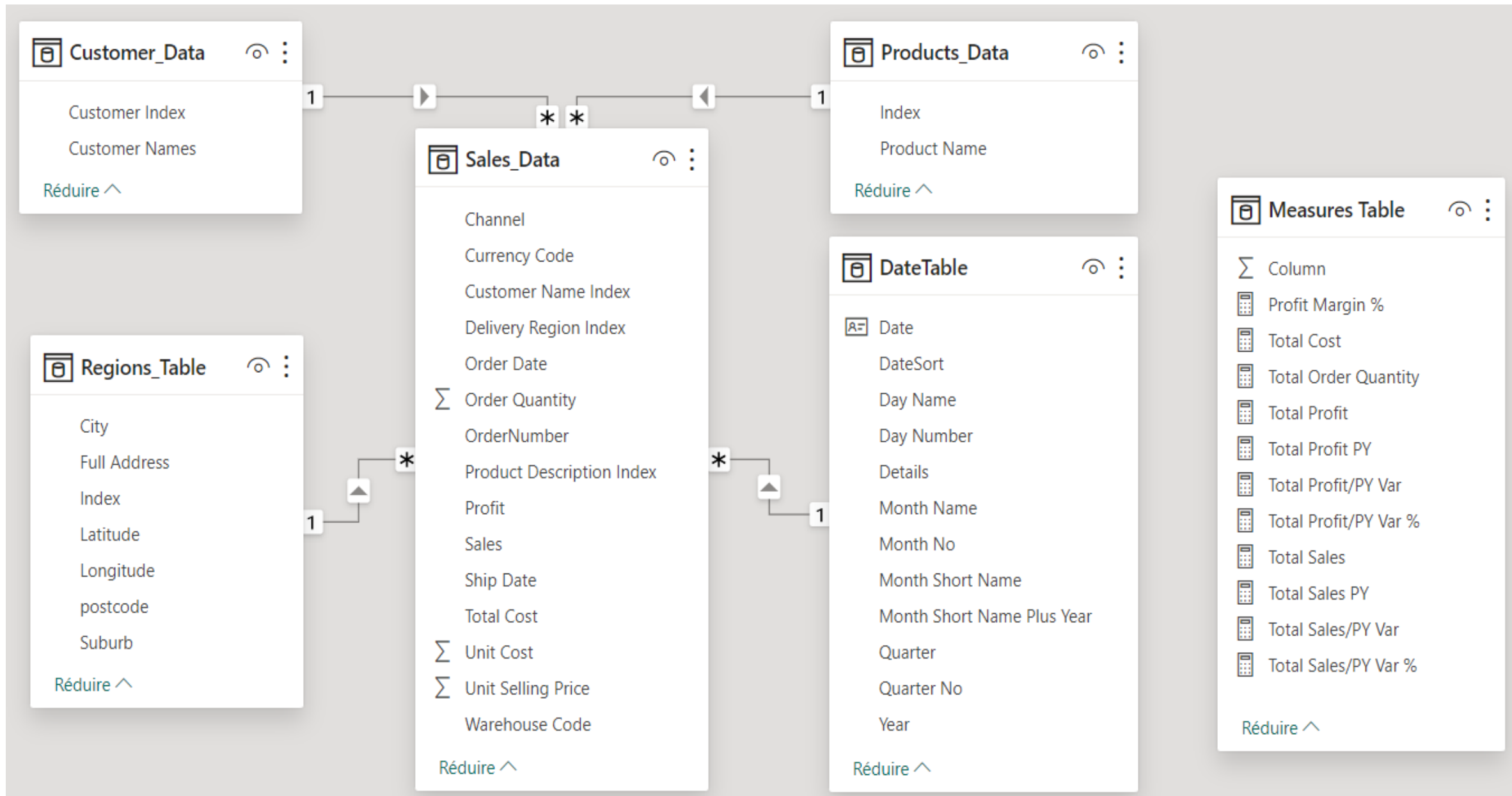
M. Mohammed QBADOU

Département Mathématiques et Informatique

Objectifs

- ❑ Effectuer le nettoyage et la transformation de données sur la vente
- ❑ Créer des visuels , des tableaux de bord et des rapports interactifs, dynamiques et attrayants pour l'étude de la force de vente.
- ❑ Analyser les résultats de la vente fournis dans le fichier Excel « Sales.xlsx » et dont le résultat de l'analyse multidimensionnelle est représentée par le modèle de la page suivante.

Le modèle d'anlyse multidimensionnelle :



Besoins métiers

- **Analyser et de présenter des informations** sur les **ventes**, les **bénéfices**, les **commandes**, la **marge bénéficiaire** et **diverses comparaisons**.
- Fournir une compréhension claire des **indicateurs de performance clés (KPI)** et des **tendances**.

Les besoins métiers auxquels le rapport doit répondre sont résumés comme suit :

1. **Calculer les ventes totales** : calculez et affichez la valeur totale des ventes pour la **période sélectionnée**, ce qui permet aux utilisateurs de comprendre **le chiffre d'affaires global généré**.
2. **Calculer le profit** : Calculez et visualisez le **bénéfice total** réalisé en fonction des données de vente, en fournissant des informations sur les **performances financières**.
3. **Analyser les commandes** : Analysez le **nombre de commandes** passées au cours de la période sélectionnée, ce qui permet d'identifier les **modèles de vente et les tendances des commandes**.
4. **Calculer la marge bénéficiaire** : Calculez et visualisez le **pourcentage de la marge bénéficiaire**, ce qui permet aux utilisateurs d'évaluer la **rentabilité des produits ou des services**.
5. **Comparer les ventes par produit avec celles de l'année précédente** : Comparez les **performances des ventes de chaque produit** entre la période sélectionnée et l'année précédente, en mettant en évidence **la croissance ou la baisse des ventes**.

6. **Comparer les ventes par mois avec l'année précédente** : comparez les performances des ventes sur différents mois entre la période sélectionnée et l'année précédente, en identifiant les régions présentant des changements significatifs.
7. **Afficher les 5 principales villes** : présentez une visualisation présentant les 5 premières villes en fonction des ventes, ce qui permet aux utilisateurs **d'identifier rapidement les emplacements les plus lucratifs**.
8. **Comparez les bénéfices par chaîne de vente (channel) avec ceux de l'année précédente** : comparez les bénéfices générés par chaque canal entre la période sélectionnée et l'année précédente, en indiquant les améliorations ou les défis en matière de rentabilité.
9. **Analysez les top 5 des ventes par client et comparez-les avec celles de l'année précédente** : analysez les données de vente par client, en mettant en évidence les performances de chaque client et en les comparant à celles de l'année précédente.
10. **Analysez les last 5 des ventes par client et comparez-les avec celles de l'année précédente** : analysez les données de vente par client, en mettant en évidence les performances de chaque client et en les comparant à celles de l'année précédente.

Etapes à suivre pour le projet

1. Recueillir les données

Collecter les données nécessaires au projet. Il peut s'agir de données provenant de diverses sources telles que des bases de données, des feuilles de calcul ou des services Web. Assurez-vous que les données sont exactes et pertinentes par rapport aux besoins et objectifs.

2. Extraction, transformation et chargement des données pour le nettoyage et la transformation des données. Nous l'utiliserons pour nettoyer et transformer les données pour les rendre aptes à l'analyse. Il peut s'agir de **supprimer les doublons**, de **gérer les valeurs manquantes**, de **fusionner des jeux de données** ou de **créer des colonnes calculées**.

3. Création d'un DataFrame de dates : La table de dates est une table qui contient des colonnes de **dates**, l'année, le trimestre, le mois, le jour de la semaine, etc. Elle est essentielle dans tout modèle qui **implique une analyse en fonction du le temps**.

Cette table doit répondre aux **exigences suivantes** :

- Elle doit avoir une colonne de type de données Date (ou date/heure), appelée **colonne de dates**.
- La colonne de dates doit contenir des **valeurs uniques**.
- La colonne de dates **ne doit pas contenir de valeurs vides**.
- Il ne doit **pas y avoir de dates manquantes** dans la colonne de dates.
- La colonne de dates doit **couvrir des années entières**.
- La table de dates **doit être marquée comme table de dates**

4. Créer un Dataframe des mesures et d'indicateurs

Utiliser des fonctions d'aggregation et de synthèse : SUM, AVERAGE, COUNT, MIN, MEAN, STD, ...

On doit créer les mesures suivantes dont les expressions sont formulées en langage DAX et qu'il faut transformer en python en utilisant pandas, numpy et expressions lambda :

1. **Total Sales = SUM(Sales_Data[Sales])**
2. **Total Sales PY = CALCULATE([Total Sales], SAMEPERIODLASTYEAR(DateTable[Date])) // Previous Year Total Sales**
3. **Total Sales/PY Var = [Total Sales] - [Total Sales PY] //Difference Between Current Year Sales & Previous Year Sales**
4. **Total Sales/PY Var % = DIVIDE([Total Sales/PY Var],[Total Sales],0) //Percentage de ↑↓ in sales year on year (YOY%)**
5. **Total Order Quantity= SUM(Sales_Data[Order Quantity])**
6. **Total Profit = SUM(Sales_Data[Profit])**
7. **Total Profit PY = CALCULATE([Total Profit], SAMEPERIODLASTYEAR(DateTable[Date]))**
8. **Total Profit/PY Var = [Total Profit]- [Total Profit PY]**
9. **Total Profit/PY Var % = DIVIDE([Total Profit/PY Var], [Total Profit], 0)**

10. Profit Margin % = $\text{DIVIDE}([\text{Total Profit}], [\text{Total Sales}], 0)$
11. Total Cost = $\text{SUM}(\text{Sales_Data}[\text{Cost}])$
12. Total Order Quantity/PY = $\text{CALCULATE}([\text{Total Order Quantity}], \text{SAMEPERIODLASTYEAR}(\text{DateTable}[\text{Date}]))$
13. Total Order Quantity/PY Var = $[\text{Total Order Quantity}] - [\text{Total Order Quantity/PY}]$
14. Total Order Quantity/PY Var % = $\text{DIVIDE}([\text{Total Order Quantity/PY Var}], [\text{Total Order Quantity}], 0)$

5. Développer les visuels et rapports

Utiliser Matplotlib et Seaborn pour créer des **graphiques** pour représenter efficacement les données. Appliquer des **filtres**, des **fonctionnalités d'extraction** pour permettre aux utilisateurs d'interagir avec les données. Créer des visuels :

- 1) Ventes par produit et comparaison avec les ventes de l'année dernière.
- 2) Ventes par mois et comparaison avec les ventes de l'année dernière.
- 3) Ventes des 5 premières villes
- 4) Comparez le bénéfice par canal avec le bénéfice de l'année précédente
- 5) Top 5 des Ventes par client et comparaison avec les ventes de l'année dernière
- 6) Last 5 des Ventes par client et comparaison avec les ventes de l'année dernière
- 7) Créez des cartes pour les ventes, les bénéfices, la marge bénéficiaire et les produits vendus

Résultat final à obtenir :



Livrable :

Livrables :

- Les codes Sources Python commentés
- Rapport docx ou pdf qui documente l'application (10 pages max)
- Une présentation Powerpoint du travail (10 diapos max)

Date limite pour rendre le travail :

- Travail (le livrable) à rendre avant le : 20/04/2025

Sur la classRoom :

code : sfujoff

Lien : <https://classroom.google.com/c/NzQzNDU5MzI0Mzgw?cjc=sfujoff>