**Etude de cas :**

Un supermarché présent dans plusiseurs villes, gère un grand volume de données sur les ventes. Ci-dessous vous verrez un échantillon de ces données. Imaginez qu’on vous a confiée la mission de créer un processus ETL pour extraire, transformer et charger ces données dans un entrepôt de données afin de faciliter l'analyse des performances des ventes et des tendances.

**Sources de données (deux fichiers):**

Tableau des Commandes :

* ID\_LigneCommande : string
* ID\_Commande : Int
* Nom\_produit : String
* ID\_Client (identifiant unique de chaque client ) : Int
* Prix\_unitaire : String
* Quantité : Int
* Ville : String

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ID\_LigneCommande** | **ID\_Commande** | **Nom\_produit** | **ID\_Client** | **Prix\_unitaire** | **Quantité** | **Ville** |
| 1\_1001 | 1001 | Smart Screen | 1 | 817,36 Dollar | 3 | Casa |
| 2\_1001 | 1001 | iPhone 14 Pro | 1 | 999,09 Dollar | 2 | Casa |
| 1\_1002 | 1002 | iPhone 14 Pro | 2 | 999,09 Dollar | 2 | Rabat |
| 2\_1002 | 1002 | Dell XPS 13 | 2 | 1181 Dollar | 2 | Rabat |
| 1\_1003 | 1003 | Dell XPS 13 | 3 | 1181 Dollar | 1 | Oujda |
| 2\_1003 | 1003 | iPhone 14 Pro | 3 | 999,09 Dollar | 2 | Oujda |
| 1\_1004 | 1004 | iPhone 14 Pro | 2 | 999,09 Dollar | 2 | Fes |
| 1\_1005 | 1005 | HP EliteBook | 2 | 727,18 Dollar | 4 | Casa |
| 1\_1006 | 1006 | Samsung Galaxy | 1 | 1181,73 Dollar | 3 | Casa |
| 1\_1007 | 1007 | MacBook Pro 16 P | 3 | 2181,73 Dollar | 1 | Casa |
| 1\_1008 | 1008 | Smart Screen | 1 | 1409 Dollar | 1 | Rabat |
| 1\_1009 | 1009 | Xiaomi Mi 12 | 2 | 772,64 Dollar | 2 | Casa |
| 1\_1010 | 1010 | TCL 8K Mini-LED TVs | 1 | 2545,36 Dollar | 1 | Casa |
| 1\_1011 | 1011 | Razer Blade 14 | 2 | 899,91 Dollar | 3 | Fes |

*Tab 1. fichier* ***Commandes.csv***

Tableau des clients :

* ID\_Client : Int
* Société : String
* DateCreationSociete : Date

|  |  |  |
| --- | --- | --- |
| **ID\_Client** | **Société** | **DateCreationSociété** |
| 1 | maroc Numeric Cluster | 12/12/2020 |
| 2 | mascir | 12/12/2014 |
| 3 | add | 12/12/2010 |

*Tab 2. fichier* ***Clients.csv***

**Partie 1**

**Phase 1 : Extraction des données**

1. Quel est le composant Talend qui permet d’extraire le fichier **Commandes.csv ?**

TFILEINPUTDELIMITED………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

1. Quels sont les champs à remplir pour configurer l’extraction des deux fichiers **Commandes.csv** et **Clients.csv ?** (Cochez les bonnes réponses) 
   * Chemin du fichier
   * Supprimer les données manquantes
   * Caractère de séparation
   * Définir l’entète
   * Le type de donnée dans chaque colonne
2. Ajouter donc ce composant à Talend pour extraire les données des deux fichiers **Commandes.csv** et **Clients.csv ?**
3. Ajoute les captures d’ écran essentielles qui résument les configurations réalisées sur le composant au niveau de Talend, Ajoute également une liste ordonnée de noms des configurations réalisées

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

**Phase 2 : Transformation des données**

Votre mission est de transformer les données brutes de vente en un format cohérent pour permettre une analyse facile et efficace.

* **Tache 1: Conversion de type**

1. Dans la table **Commandes.csv,** quel est le nom de la colonne nécessitant une conversion de type?

…………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

1. Dans Talen Open Studio, quel est le composant convenable qui permet de convertir cette colonne?

…………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

1. Ajouter donc ce composant à Talend et faites la configuratoin neccessaire ?
2. Colle ici les captures d’ écran essentielles qui résument les configurations réalisées sur le composant au niveau de Talend, Ajoute également une liste ordonnée de noms des configurations réalisées

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

* **Tache 3: Jointure des données**

1. Quel est le nom de la colonne sur laquelle vous pouvez appliquer la jointure des deux tableaux **Clients** et **Commandes ? (1 Pts)**

…………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

1. Dans Talend Open Studio, quel est le nom du composant convenable qui permet d’appliquer la jointure entre les tables d’entrées? **(1 Pts)**

……………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………

1. Aprés la jointure, nous voulons que la table de sortie contiendra seulement les colonnes suivantes:

- ID\_Commande

- Nom\_produit

- Société

- DateCreationSociete

- Prix\_unitaire

- Quantité

- Ville

1. Ajouter donc le composant de jointure et toute la configuration neccessaire pour n’afficher que les colonnes ci-avant cités.
2. Colle ici les captures d’ écran essentielles qui résument les configurations réalisées sur le composant au niveau de Talend, Ajoute également une liste ordonnée de noms des configurations réalisées

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

1. C’est le moment pour tester ton Job Talend et ses composants préalablement ajoutés, quel est le composant qui permet d’afficher les données sur la console sous forme de texte/tables ?

……………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………

1. Ajouter donc le composant convenable et tester ton Job Talend.
2. Colle ici les captures d’ écran essentielles qui résument les configurations réalisées sur le composant au niveau de Talend, Ajoute également une liste ordonnée de noms des configurations réalisées

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

1. Nous aimerons perfectionner le Job Talend, au lieu de garder les deux colonnes Prix\_Unitaire et Quantite, nous aimerons avoir une colonne PrixTotal, dont l’expression conditionnelle est PrixTotal = Prix\_Unitaire \* Quantite, Comment pouvons nous résoudre ce problème au niveau de Talend et quel composant faut-il configurer ?

……………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………

1. Ajouter le neccessaire sur le job Talend pour faire cette factorisation de colonnes.
2. Colle ici les captures d’ écran essentielles qui résument les configurations réalisées sur le composant au niveau de Talend, Ajoute également une liste ordonnée de noms des configurations réalisées

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

* **Tache 4: Aggrégation des données**

Dans cette tache, vous devez calculer et afficher la **somme** des **prix totals** par **Société**.

1. Quel est le composant qui permet d’appliquer l’aggrégation**?**

**…………………………………………………………………………………………………………………………………………………………………**

**…………………………………………………………………………………………………………………………………………………………………**

1. A la sortie de ce composant d’aggrégation, vous devez garder combien de colonnes ?

**…………………………………………………………………………………………………………………………………………………………………**

**…………………………………………………………………………………………………………………………………………………………………**

1. Ajouter et configurer ce composant pour calculer et afficher la **somme des ventes** par **Société**.
2. Colle ici les captures d’ écran essentielles qui résument les configurations réalisées sur le composant au niveau de Talend, Ajoute également une liste ordonnée de noms des configurations réalisées

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

* **Tache 5: Trier les données agrégées.**

Dans cette tache, vous devez trier par order **décroissant** la **somme** des **prix totals** par **Société**.

1. Quel est le composant qui permet d’appliquer le tri sur une colonne

……………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………

1. Ajouter et configurer ce composant dans Talend pour trier par ordre décroissant la somme des prix totals par Société. Note: la table de sortie contiendra seulement deux colonne: **Société | SommeVentes**
2. Colle ici les captures d’ écran essentielles qui résument les configurations réalisées sur le composant au niveau de Talend, Ajoute également une liste ordonnée de noms des configurations réalisées

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

**Phase 3 : Chargement des données vers une BD**

Aprés les transformations effectuées, vous devez charger la table résultante dans une base de données MySql.

1. Quel est le composant qui permet de charger les données dans une base de données Mysql

……………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………

1. Ajouter donc le composant sur Talend et toute la configuration neccessaire
2. Colle ici les captures d’ écran essentielles qui résument les configurations réalisées sur le composant au niveau de Talend, Ajoute également une liste ordonnée de noms des configurations réalisées

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

**Partie 2**

Dans ce qui suit nous aimerons faire sortir la somme des ventes par rapport à plusieurs dimensions à savoir :

1. La somme des ventes par Ville
2. La somme des ventes par Produit
3. La somme des ventes par Commande et Société
4. La somme des ventes par Ville et Produit

**La somme des ventes par Ville :**

Pour faire cela, suivez ces étapes :

1. Dupliquer le job de la partie 1 de l’étude de cas et nommer le job **VentesParVille**
2. Faites les adaptions neccessaire sur le Job
3. Tester le job
4. Colle les captures d’écrans neccessaire et lister les configurations faites

**La somme des ventes par Produit :**

Pour faire cela, suivez ces étapes :

1. Dupliquer le job de la partie 1 de l’étude de cas et nommer le job **VentesParProduit**
2. Faites les adaptions neccessaire sur le Job
3. Tester le job
4. Colle les captures d’écrans neccessaire et lister les configurations faites

**La somme des ventes par Commande et Société :**

Pour faire cela, suivez ces étapes :

1. Dupliquer le job de la partie 1 de l’étude de cas et nommer le job **VentesParCommandeSociete**
2. Faites les adaptions neccessaire sur le Job
3. Tester le job
4. Colle les captures d’écrans neccessaire et lister les configurations faites

**La somme des ventes par Ville et Produit :**

Pour faire cela, suivez ces étapes :

1. Dupliquer le job de la partie 1 de l’étude de cas et nommer le job **VentesParVilleProduit**
2. Faites les adaptions neccessaire sur le Job
3. Tester le job
4. Colle les captures d’écrans neccessaire et lister les configurations faites