

Table des matières

1) Introduction	2
2) Diagramme des modèles	3
3) Table de fait	3
4) Justification	4
4.1) Clés primaires du Data Warehouse :	4
4.2) Modélisation des dimensions et faits :	4
5) Limite rencontrée	4
6) Gestion des erreurs	5

1) Introduction

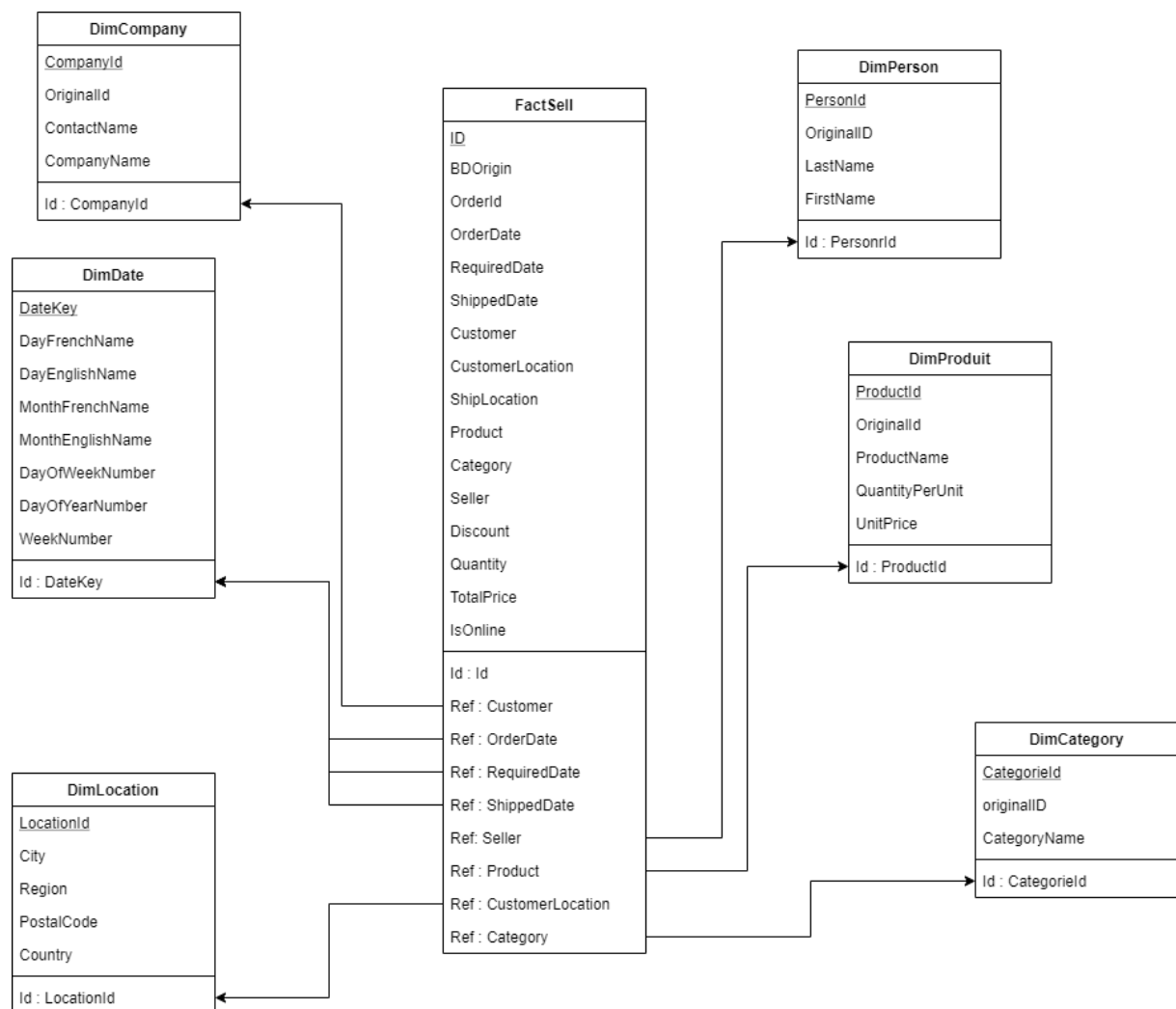
Pour répondre à la demande de BlablaCo d'analyse de ses ventes. Cette entreprise enregistre tout ses commandes dans un base de données et nous a demandé une analyse de ses ventes par critère temporel, par produit et catégorie et par les caractéristiques des clients qui ont acheté des produits.

Pour se faire nous proposons une utilisant la SSIS Toolbox de Visual Studio. Pour se faire nous avons eu recours a différent composant tels que :

- OLE DB Source : permet d'obtenir des données d'une base de données
- OLE DB Destination : permet le stockage des données dans une base de données
- Derived Column : permet d'effectuer des modifications sur les données
- Lookup : permet de vérifier si une donnée existe dans une base de données
- Script component : permet d'effectuer des modifications sur les données à l'aide de C#
- Union All : permet de fusionner les données reçues par d'autres composants.

Notre solution se décompose en deux temps. Dans un premier temps la solution va remplir les différentes tables de dimensions qui sont liées à la table de faits. Une fois les tables de dimension remplies avec les informations de la bd d'origine, la solution remplira la table de faits.

2) Diagramme des modèles



3) Table de fait

La table de fait contient toutes les informations sur une ventes ainsi que toutes les clés étrangères vers les différentes dimensions.

4) Justification

4.1) Clés primaires du Data Warehouse :

Pour les clés des dimensions nous avons concaténé le nom de la BD d'origine avec l'id d'origine de l'endroit où on prend la donnée. Cette manière de faire permet d'aller chercher dans les données d'origines au besoin.

Il y a des exceptions pour les clés de deux dimensions. La première est là pour la DimDate qui a sa clé au format <YYYYMMDD>. La seconde est la DimLocation qui a sa clé composée des 5 premières lettres du pays + les 5 premières lettres du code postal + les 5 premières lettres de la ville + les 5 premières lettres de la région.

La table des faits quand a-t-elle sa clé est un entier auto-incrémenté.

Nous avons préféré faire des concaténations plutôt que des auto-incrémentations afin de pouvoir créer les clés étrangères à partir des données dont on dispose sans avoir à effectuer des recherches dans les dimensions.

4.2) Modélisation des dimensions et faits :

Nous avons gardé les informations nécessaires pour répondre aux différentes questions que BlablaCo nous a posé. De ce fait par exemple il ne nous était pas nécessaire de garder dans le Data Warehouse les informations relatives à l'adresse/numéro de téléphone des employés.

5) Limite rencontrée

Avec le temps donné pour la réalisation du projet il nous a été impossible d'implémenter une traduction des données dans les 2 langues demandées. Avec plus de temps nous pensons qu'il aurait été possible d'implémenter ce traducteur. Cependant notre programme permet tout de même de traduire les dates en français et en anglais.

6) Gestion des erreurs

Dans la base de données d'origine certaine donnée était a NULL nous avons donc du géré ces valeurs.

- Si les données son stockées dans une table de dimension on lui attribue une valeur par défaut. Par exemple pour la dimension personne la valeur est du « Dummy Person »
- Si ces donné doivent être stockées dans une table de fait on leur assigne une valeur par défaut tels que « none » pour les strings si c'est un chiffre la valeur est mise à 0 et pour les booléens sont mis à False.

En cas d'arrête de l'ETL en cours de traitement du a des raisons externe il suffit de relancer l'exécution depuis le début.