

# Informatique décisionnelle 2 – TD 1

## objectifs, besoins, modélisation dimensionnelle

Au cours des trois prochains TD, nous allons nous intéresser à un cas d'utilisation simple et réaliser en partie les étapes nécessaires à la mise en place d'un système d'information décisionnel. Un **rapport court** sera à rendre sur les résultats obtenus sur les trois TD, **par groupe de 3**. Assurez vous donc de bien conserver vos notes sur les différentes tâches par groupe.

**T1:** Organisez vous par groupes de 3. Vous devrez conserver le même groupe au cours des trois TD.

**Description du cas d'étude :** On s'intéresse à une entreprise (on va l'appeler **BuyStuff**) qui vend différents articles de maisons (meubles, fournitures de bureau, technologie) en ligne. **BuyStuff** a donc un ensemble de produits dans son catalogue, des stocks, des commandes, etc. Tous les produits ne sont pas populaires de la même façon, n'apportent pas les mêmes revenus, etc. Les nombres de ventes peuvent aussi varier beaucoup dans le temps généralement, et par produit.

A l'heure actuelle, les stocks sont gérés directement par les trois départements de **BuyStuff** (*furnitures, office supplies, technology*) dans trois entrepôts différents. L'entreprise souhaite optimiser revoir cette stratégie et voir si l'espace de stockage pourrait être mieux utilisé (i.e. pour s'assurer d'avoir du stock pour les produits les plus avantageux proche de la demande).

Le système d'information opérationnel de l'entreprise contient plusieurs sous systèmes qui sont autant de sources de données:

**La base de données centrale de l'entreprise :** Contient des tables les départements (ville, capacité de stockage), les produits (nom, catégorie), etc.

**La base de données client :** Contient des tables sur les clients (nom, catégorie) et leurs adresses.

**La base de données des commandes :** Contient des tables sur les commandes (date, livraison, client) et les produits commandés (quantité, prix, bénéfices).

Des extraits de chacune des tables de ces trois sources de données sont fournis à la fin de ce document et leur contenu se trouve sur Arche au format CSV et aux liens suivants: `addresses.csv`, `customers.csv`, `departments.csv`, `order_details.csv`, `orders.csv` and `products.csv`.

**T2:** Définissez un ou plusieurs objectifs du système d'information décisionnel, avec les indicateurs et les axes d'analyse associés d'après le patron vu en cours (voir ci-dessous).

### Patron de définition d'objectif :

1. **Domaine** : Commencer par découper par domaines, e.g. vente, approvisionnement, gestion, etc.
2. **Objectif** : Le coeur de la question, ce que le système d'information décisionnel doit permettre.
3. **Cible** : La conséquence de l'objectif sur l'entreprise, i.e. les avantages attendus.
4. **Acteurs** : Décideurs, responsables des données, veilleurs, opérateurs, etc.
5. **Indicateurs** : Ce qu'on va vouloir mesurer, suivre et analyser pour répondre à l'objectif.
6. **Axes d'analyse** : Les facteurs à intégrer qui influence les indicateurs.

**T3:** Identifiez, pour chaque indicateur et axe d'analyse, les sources de données nécessaires dans l'entreprise et possiblement externes.

**T4:** Dessinez le ou les modèle(s) de diffusion pour les indicateurs (mesures) et les axes d'analyse (dimensions) considérés. Utilisez un modèle en étoile.

**T5:** Indiquez quelles mesures sont additives, semi-additives et non-additives.

**T6:** Essayez de télécharger Pentaho Data Integration ou Talend Open Studio et de commencez a manipuler les données avec un de ces outils.

Table 1: Table **Departments** de la base centrale

Department	City	Storage Capacity
Furniture	New York City	3003
Office Supplies	San Francisco	4043
Technology	Philadelphia	2257

Table 2: Extrait de la table **Products** de la base centrale

Product ID	Department	Sub-Category	Product Name
FUR-BO-10001798	Furniture	Bookcases	Bush Somerset C...
FUR-CH-10000454	Furniture	Chairs	Hon Deluxe Fabr...
OFF-LA-10000240	Office Supplies	Labels	Self-Adhesive A...
FUR-TA-10000577	Furniture	Tables	Bretford CR4500...
OFF-ST-10000760	Office Supplies	Storage	Eldon Fold 'N R...

Table 3: Extrait de la table <b>Customers</b> de la base client		
Customer ID	Customer Name	Segment
CG-12520	Claire Gute	Consumer
DV-13045	Darrin Van Huff	Corporate
SO-20335	Sean O'Donnell	Consumer
BH-11710	Brosina Hoffman	Consumer
AA-10480	Andrew Allen	Consumer

Table 4: Extrait de la table <b>Addresses</b> de la base client						
Address ID	Customer ID	Country	City	State	Postal Code	Region
CG-12520_42420	CG-12520	United States	Henderson	Kentucky	42420	South
DV-13045_90036	DV-13045	United States	Los Angeles	California	90036	West
SO-20335_33311	SO-20335	United States	Fort Lauderdale	Florida	33311	South
BH-11710_90032	BH-11710	United States	Los Angeles	California	90032	West
AA-10480_28027	AA-10480	United States	Concord	North Carolina	28027	South

Table 5: Extrait de la table <b>Orders</b> de la base commandes				
Order ID	Order Date	Ship Date	Customer ID	Address ID
CA-2016-152156	2021-11-02	2021-11-05	CG-12520	CG-12520_42420
CA-2016-138688	2021-06-06	2021-06-10	DV-13045	DV-13045_90036
US-2015-108966	2020-10-04	2020-10-11	SO-20335	SO-20335_33311
CA-2014-115812	2019-06-03	2019-06-08	BH-11710	BH-11710_90032
CA-2017-114412	2022-04-09	2022-04-14	AA-10480	AA-10480_28027

Table 6: Extrait de la table <b>Order_details</b> de la base commandes				
Order ID	Product ID	Sales	Quantity	Profit
CA-2016-152156	FUR-BO-10001798	261.96	2	41.91
	FUR-CH-10000454	731.94	3	219.58
CA-2016-138688	OFF-LA-10000240	14.62	2	6.87
US-2015-108966	FUR-TA-10000577	957.58	5	-383.03
	OFF-ST-10000760	22.37	2	2.52