






RAPPORT

Analyse multidimensionnelle d'un site E-commerce



Réalisé par :

 BENALLAL Youssef
 LAZREK Youssef
 ECHCHAIR Zakaria

Encadrant :



 Mme. ELFIRDOUSSI Selwa
 M. HERIMCH Hamid

Table des matières

Introduction :.....	3
Fonctionnalités à implémenter :	3
Nettoyage des données :	4
Implémentation du Dashboard :	4
• Fonctions de Zoom/Resize :	4
• Fonctions de filtrage :	4
• Map Chart :	5
• Chiffre d'affaires :	6
• Sélection des meilleurs produits :	7
• Sélection des meilleurs Clients :	8
• Taux de livraison :	9
• Le design du formulaire :	10
Segmentation :	13
Conclusion :	15

Introduction :

Après la démocratisation de l'internet, le secteur de l'e-commerce ne cesse de réaliser des avancées majeures en enregistrant des chiffres de croissance impressionnants. Ces avancées s'accompagnent d'une concurrence rude entre les stores, cette concurrence motive encore plus la nécessité d'avoir des outils décisionnels permettant d'attirer plus de clientèle, d'augmenter le taux de conversion pour accroître les performances du store. Notre projet a pour objectif de réaliser un Dashboard dynamique et ergonomique qui présentera les métriques d'aide à la décision jugées importantes pour la gestion d'un store, ainsi qu'une partie qui consiste à la gestion générale de la base de données des transactions en conformité avec la base de données initiale.

Fonctionnalités à implémenter :

Nous synthétisons les fonctions à implémenter en deux blocs, la partie Dashboard puis la partie gestion des transactions (CRUD).

Pour la partie Dashboard dynamique les fonctionnalités sont listées comme suit :

- Filtrer les informations du tableau de bord suivant le pays, le produit ainsi que les données temporelles (Année, mois)
- Visualisation de la carte d'exploration des transactions par pays (apparition d'une carte choroplèthe des ventes, du nombre de transactions, du taux de livraison, du taux d'abondant et de la quantité livrée)
- Visualisation de la tendance du chiffre d'affaires par mois avec une granularité atteignant l'évolution du chiffre d'affaires suivant les jours.
- Visualisation des nombres de transactions suivant les heures pour l'observation des heures à forte et faible activité.
- Classification des meilleurs produits de manière dynamique
- Classification des meilleurs clients suivant les achats et le nombre de transactions effectuées
- Segmentation des achats suivant le prix moyen d'un panier (Basket Price)
- Résumer la part de clients VIP à cibler dans les opérations marketing suivant une analyse RF (récence et fréquence)
- Afficher la meilleure transaction

Pour la partie gestion des transactions, les fonctionnalités sont listées comme suit :

- Ajout d'une transaction au niveau de la base de données
- Suppression d'une transaction
- Affichage de la BDD, en particulier les transactions les plus récentes au niveau de la BDD
- Générer un rapport PDF du Dashboard

- Envoi du Dashboard par mail

Nettoyage des données :

En explorant les données initiales on a identifié des irrégularités et des patterns qu'on a supprimé pour le cas des irrégularités et qu'on a plus explicité pour le cas des patterns.

25% des entrées ne sont pas affectées à un client particulier. Avec les données disponibles, il est impossible d'imputer des valeurs pour l'utilisateur et ces entrées sont donc inutiles pour l'exercice en cours. Nous les supprimons donc du data frame.

Si le code InvoiceNo commence par la lettre 'c', cela indique une annulation. On ajoute une variable état de commande qui prend la valeur 'abondant' dans le cas d'un InvoiceNo qui commence avec C.

De plus, on opère la suppression des observations dupliquées.

Implémentation du Dashboard :



Fonctions de Zoom/Resize :

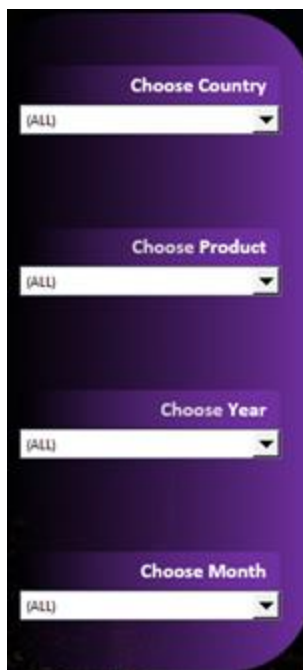
Nous avons créé plusieurs fonctions permettant d'effectuer des zooms sur un graphique au moment où son en-tête est sélectionné. A chaque fois que le zoom est établi, le bouton apparait. Il permet évidemment de retourner à l'état de zoom initial.

Les fonctions de zoom et la fonction Resize se trouvent dans : **Visual Basic > Modules > Zooms**



Fonctions de filtrage :

Il s'agit des fonctions que nous avons établies et qui permettent de filtrer les données de tous les graphs du Dashboard suivant le pays, la description du produit, l'année et mois choisis par l'utilisateur. La section contenant les différents filtres se trouve à gauche du Dashboard :



Ces filtres agissent sur tous les graphes à l'exception de la section identifiant la meilleure transaction et le graphe des meilleurs produits qui ne change pas lorsqu'on utilise le filtre de la description des produits.

La seule partie du Dashboard qui ne dépend de aucun filtre est la phrase qui résume l'analyse RF.

Les fonctions de ces filtres se trouvent dans
Visual Basic > Modules > Filters

Map Chart :

Dans cette partie nous avons réparti les transactions, quantités vendues, nombre de commandes livrées et annulées et la somme des ventes par pays pour ensuite les visualisés dans une carte :



Chaque bouton parmi ceux représentés au-dessus nous transfère vers un métrique précis. Le changement de métriques est assuré par les fonctions établies dans **Visual Basic > Modules > Boutons**

Nous sommes arrivés à réaliser cette représentation à travers le tableau croisé ci-dessous :



Dans la partie inférieure droite du graphique, il y a un bouton responsable du changement de l'échelle de temps, c'est-à-dire qu'il nous permet de voir l'évolution journalière ou par heure du chiffre d'affaires au lieu de l'évolution mensuelle représentée ci-dessus. Voici un exemple :



Sélection des meilleurs produits :

Vu que le choix de maintenir la vente d'un certain produit est crucial pour le succès d'un online store, nous avons songé à représenter les trois meilleurs produits qui ont enregistré les plus grands prix de ventes. Veuillez noter que les produits peuvent changer suivant le pays et la durée précisé.



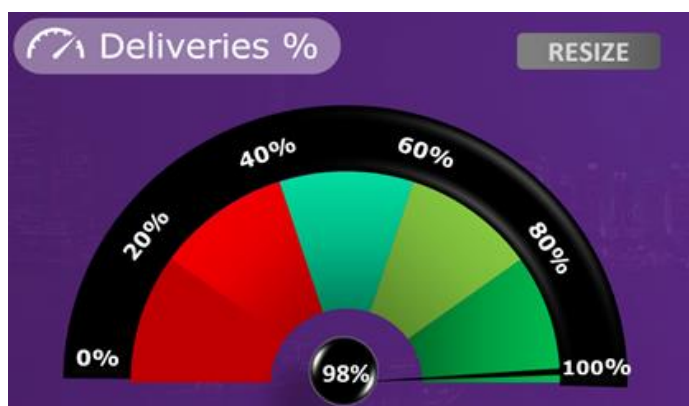
Voici le tableau croisé correspondant :

Country (All)			
Description (All)			
Years (All)			
Months (All)			
Row Labels			
	Count of InvoiceNo	Sum of Quantity	Sum of TotalPrice
€ 77,183.80 12346	2	76215	77183.8
€ 4,320.00 12347	182	2458	4320
€ 1,797.24 12348	70	2342	1797.24
€ 1,757.55 12349	79	431	1757.55
€ 334.40 12350	17	397	334.4
€ 1,665.76 12352	89	593	1665.76
€ 89.00 12353	4	20	89
€ 1,079.80 12354	58	530	1079.8
€ 409.40 12355	13	340	409.4
€ 2,811.43 12356	59	1291	2811.43
€ 4,207.67 12357	131	2708	4207.67
€ 1,168.06 12358	19	248	1168.06
€ 6,110.03 12359	252	1809	6110.03
€ 2,662.06 12360	129	1345	2662.06
€ 189.90 12362	10	91	189.9
€ 5,226.23 12362	274	2229	5226.23
€ 552.00 12363	23	408	552
€ 1,113.30 12364	85	1206	1113.3
€ 320.89 12365	21	173	320.89
€ 368.90 12367	11	173	368.9
€ 5,541.94 12370	394	2352	5541.94
€ 1,887.96 12371	40	591	1887.96
€ 1,298.04 12372	52	794	1298.04
€ 364.60 12373	14	197	364.6
€ 742.93 12374	33	342	742.93
€ 457.50 12375	18	178	457.5
€ 1,628.12 12377	77	944	1628.12
€ 4,008.62 12378	219	2529	4008.62
€ 852.24 12379	41	407	852.24
€ 2,724.81 12380	105	1127	2724.81
€ 1,845.31 12381	95	808	1845.31
€ 1,895.56 12383	100	9107	1895.56



Taux de livraison :

Nous avons représenté le taux de livraison sous un graphique jauge :



Country	(All)	
Description	(All)	
Years	(All)	
Months	(All)	
Row Labels	Count of TotalPrice	98%
Delivered	392413	
Cancelled	8620	
Grand Total	401033	
0%	Worst	98%
30%	Bad	98%
50%	Average	98%
75%	Good	98%
95%	Best	98%
	Best	
	98%	
Needle :	1%	
	101%	

0%	Worst	98%
30%	Bad	98%
50%	Average	98%
75%	Good	98%
95%	Best	98%
	Best	
	98%	
Needle :	1%	
	101%	

Needle :	1%
	101%

Add/Delete Order

Invoice Number: 536367

Description : _____

Quantity: _____

Unit Price: _____

Country: _____

Stock Code: _____

Date of Entry: _____

Client N°: _____

State: _____

☒ *Delete*

Data to delete:

Column 1	Column 2	Column 3	Column 4	Column 5	Column 6
Invoice	Stock Code	Description	Quantity	Invoice Date	Unit Price
536367	C000	ASSORTED COLOUR BIRD ORNAMENT	32	01/12/2000 00:34	11.69
536367	D000	PORPHY'S PLAYHOUSE BEDROOM	6	01/12/2000 00:34	22.30
536367	C022	PORPHY'S PLAYHOUSE KITCHEN	6	01/12/2000 00:34	22.30
536367	D007	FELTONCRAFT PRINCESS CHARLOTTE DOL	6	01/12/2000 00:34	23.75
536367	D007	IVORY KNOTTED MUS COIN	6	01/12/2000 00:34	11.65
536367	B005	BOX OF 4 ASSORTED COLOUR TREASUR	6	01/12/2000 00:34	14.25
536367	A102	BOX OF VINTAGE JESSIE BLOCKS	3	01/12/2000 00:34	44.95
536367	C004	BOX OF VINTAGE ALPHABET BLOCKS	2	01/12/2000 00:34	45.95
536367	C006	HOME BUILDING BLOCK WORD	3	01/12/2000 00:34	15.95
536367	A172	LOVE BUILDING BLOCK WORD	3	01/12/2000 00:34	15.95
536367	C006	RECIPE BOX WITH JETTY HEART	4	01/12/2000 00:34	27.95
536367	C003	DOORMAT NEW ENGLAND	4	01/12/2000 00:34	27.95

10

La partie de l'insertion dans la base de données :

Add/Delete Order

Invoice Number: _____

Stock Code: _____

Show BDD

☐ Last Entry

Description : _____

Date of Entry: _____

DASHBOARD

Quantity: _____

Client N°: _____

Cancel

Unit Price: _____

State: _____

☐ Delete

Country: _____

Validate

Normalement par défaut, L'option ajouter à la base de données est activée tant que l'option « Delete » est désactivée.

En sélectionnant un champ dans le formulaire, le nom de ce champ devient bleu ainsi que la ligne qui le souligne. Et en le quittant il reste en bleu si et seulement si le champ n'est pas vide, sinon il revient à son état initial.

De plus une option de suggestions est déjà intégrée dans les champs « Invoice number – Stock Code – Description – Client N° - State – Country ». Cette option permet de donner des suggestions basées sur notre base de données (en prenant en considération les nouvelles modifications) pour remplir le formulaire.

Nous justifions le choix de ces champs, par le fait qu'ils seront très utiles pour remplir rapidement et sans fautes le formulaire, cependant nous permettons aux utilisateurs d'utiliser un historique provisoire dans le reste des champs « Date of Entry, Quantity, Unit Price ». Et qui s'efface en quittant la fenêtre à but de faciliter les saisies répétitives surtout pour la date.

Un message d'erreur « Veuillez remplir tout le formulaire » s'affichera si l'utilisateur n'a pas rempli le formulaire entier.

Pour éviter toutes sortes de confusion entre les descriptions et les StockCode « par exemple la majuscule et la minuscule dans la description » qui peuvent falsifier les statistiques dans le Dashboard. Nous optons pour le remplissage automatique de la description si l'utilisateur a rempli le champ du StockCode, et de même s'il a rempli la description, le StockCode sera automatiquement choisi.

Le bouton « Cancel » permet de vider tous les champs ainsi que la cash « les filtres exécutés en arrière-plan » mais il garde l'historique.

La possibilité d'insérer des nouvelles données est toujours valide. Autrement dit les suggestions ne sont qu'une option pour faciliter la saisie.

La partie de la suppression de la base de données :

Add/Delete Order

Invoice Number: 536367 Stock Code: _____

Description: _____ Date of Entry: _____

Quantity: _____ Client N°: _____

Unit Price: _____ State: _____

Country: _____

☒

Data to delete:

InvoiceNo	StockCode	Description	Quantity	InvoiceDate	UnitPrice
536367	C300	ASSORTED COLOUR BIRD ORNAMENT	31	01/12/2000 08:34	£1.69
536367	D481	PORPHY'S PLAYHOUSE BEDROOM	6	01/12/2000 08:34	£2.10
536367	C132	PORPHY'S PLAYHOUSE KITCHEN	6	01/12/2000 08:34	£2.10
536367	D807	PELTOLIFT PRINCESS CHARLOTTE DOL	6	01/12/2000 08:34	£3.75
536367	D037	IVORY KNETTED MUG COSY	6	01/12/2000 08:34	£1.65
536367	B661	BOX OF 8 ASSORTED COLOUR TEASPOK	6	01/12/2000 08:34	£4.25
536367	A103	BOX OF VINTAGE SAGSAW BLOCKS	3	01/12/2000 08:34	£4.95
536367	CN11	BOX OF VINTAGE ALPHABET BLOCKS	2	01/12/2000 08:34	£9.95
536367	C198	HOME BUILDING BLOCK WORDS	3	01/12/2000 08:34	£5.95
536367	A772	LOVE BUILDING BLOCK WORD	3	01/12/2000 08:34	£5.95
536367	C100	RECIPE BOX WITH METAL HEART	4	01/12/2000 08:34	£7.95
536367	C613	DOORMAT NEW ENGLAND	4	01/12/2000 08:34	£7.95

Criteria:

InvoiceNo	StockCode	Description	Quantity	InvoiceDate	UnitPrice
536367					

Le choix « Delete » permet d'afficher deux listes :

- La première liste filtre la base de données selon les critères insérés dans le formulaire « la confirmation d'un critère se fera par une clique sur le bouton TAB ».
- La deuxième liste présente le tableau des critères sur lequel on applique « Advanced filtre » pour obtenir la 1ère liste.

Si l'utilisateur clique sur le bouton valider avec la case « Delete » cochée, il va supprimer toutes les lignes qui s'affiche dans la 1ère liste. De ce fait, nous recommandant la saisie du maximum des critères.

Cette fois-ci les suggestions dans chaque champ sont limité aux possibilités disponibles. En d'autres termes, ces suggestions sont basées sur la 1ère liste au lieu de la base de données entière.

Nous permettons aux utilisateurs d'auto remplir tous les champs en cliquant doubles cliques sur la ligne qu'ils veulent supprimer dans la 1ère liste. Ensuite le bouton valider permet d'effectuer la suppression.

Segmentation :

Un des objectifs de notre analyse est le clustering des clients. À défaut d'utiliser le K-means présents clustering sur Excel et qui agit sur les données non labelisées et dont le résultat n'offre une bonne interprétabilité, nous utilisons une analyse assez connue dans le domaine du Marketing et de l'E-commerce dont le nom est l'analyse RFM. (Recency, frequency, Monetary value).

L'essence de l'analyse RFM (Récence Fréquence Montant) est de diviser la base client en fonction de la date à laquelle ils ont effectué leur dernier achat, de la fréquence à laquelle ils achètent des articles et de la valeur moyenne de leurs commandes. Pour chacune de ces métriques, nous attribuons les clients à l'un des groupes, auxquels une note est attribuée.

En utilisant un système de notation de 1 à 4 cela donne le tableau suivant :

Score / Caractéristique	Récence	Fréquence	Valeur monétaire
1	Clients de longue date	Achète rarement	Faible valeur des achats
2	Clients relativement récents	Achète rarement	Valeur moyenne des achats
3	Clients récents	Achète souvent	Valeur élevée des achats
4	Clients très récents	Achète très souvent	Values très élevées

- Procédure du scoring :
 1. Afin de déterminer la date du dernier achat pour chaque client, combien de fois le client a acheté au total. On utilise un tableau croisé dynamique.
 2. On combine toutes leurs transactions en une seule ligne. Pour ce faire, nous ajoutez la valeur Client dans les lignes.
 3. Pour avoir la fréquence d'achat on passe le nombre (count) de la colonne date commande.
 4. Dans cette étape, on doit maintenant calculer la date du dernier achat pour chaque client. Pour ce faire, ajoutez la date de commande aux valeurs (Valeur Max.).

C'est tout pour la préparation de l'analyse.

Dans une 2ème partie il s'agira de renseigner les variables Récence, Fréquence.

La variable récence sera renseignée En utilisant la formule = DATE de l'extraction – Dernier Achat vous calculez depuis combien de jours un client a effectué son dernier achat.

La formule de la fréquence sera spécifiée après la présentation du tableau de critère du scoring.

	Récence	Fréquence
25 %	Score4 x = CENTILE.INCLUDE(Range récence ;0,25)	Score1 = CENTILE.INCLUDE(Range fréquence ;0,25)
50 %	Score3 y = CENTILE.INCLUDE(Range récence;0,5)	Score2 = CENTILE.INCLUDE(Range fréquence ;0,5)
75 %	Score2 z = CENTILE.INCLUDE(Range récence;0,75)	Score3 = CENTILE.INCLUDE(Range fréquence ;0,75)

Vous savez maintenant que 25 % des clients ont acheté vos produits il y a moins de x jours, 50 % ont acheté il y a moins de y jours, 75 % ont acheté il y a moins de z jours.

En conséquence, vous attribuez aux clients qui ont acheté il y a moins de x jours la valeur la plus élevée de 4. Ceux qui ont passé une commande entre x à y jours se voient attribuer une valeur de 3. Ceux qui ont passé une commande entre y à z jours se voient attribuer une valeur de 2. Le reste, qui a acheté il y a plus de z jours, se voit attribuer la valeur de 1.

On suit le même principe pour le scoring de la fréquence cependant cette fois-ci ce sont les valeurs élevées de fréquence qui reçoivent un scoring élevé.

Suite au scoring on pourra réaliser une matrice RF qui nous renseignera sur la part de clients dans différents segments ainsi que la part de chiffre d'affaires qu'ils génèrent.

Les segments RF sont résumés dans la matrice suivante :

	Lost	Leaving VIP	Leaving VIP	VIP
	Lost	Leaving Clients	Regular clients	Regular clients
	Lost	Leaving Clients	Regular clients	Regular clients
	Lost	One-time clients	Regular clients	

Ce qui nous intéresse en particulier ce sont les clients de la classe 4 4 qui sont des clients VIP

ils génèrent la plus grande partie du chiffre d'affaires.

Une métrique utile est de voir dans le Dashboard de manière dynamique la part de ces clients et la part de chiffre d'affaires qu'ils génèrent.

Cette analyse est donc particulièrement intéressante pour définir des actions marketing ciblant un échantillon précis de clients.

Conclusion :

Dans ce projet ayant pour problématique d'assurer une visibilité des métriques clés pour un store e-commerce, nous avons à la fois pu mettre en pratique et faire évoluer nos acquis en VBA à travers la réalisation de la partie visualisation et gestion du store. En parallèle, ce projet nous a permis d'explorer davantage le choix des métriques importants suivant la data e-commerce fournie.

Dans la partie analyse purement e-commerce, nous avons voulu explorer l'analyse d'association c'est à dire les produits qui ont tendance à être vendu ensemble mais le module permettant cette analyse n'était pas présent dans les versions antérieures à 2013. La segmentation à travers le K-means n'est pas facile à interpréter.

On a donc opté pour une analyse RF pour la segmentation et on s'est intéressé particulièrement à la part des clients VIP et la part du CA qu'ils génèrent.