



BestMarket

Analyse des retours clients

Vincent

Audoux

09/2023

I. Analyse de la satisfaction client

1	Quel est le nombre de retours clients sur la livraison ?	<pre>SELECT COUNT(cle_retour_client) AS nb_retour_client FROM retour_client WHERE libelle_categorie = "livraison"</pre>	<table><tr><th>nb_retour_client</th></tr><tr><td>639</td></tr></table>	nb_retour_client	639			
nb_retour_client								
639								
2	Quelle est la liste des notes des clients sur les réseaux sociaux sur les TV ?	<pre>SELECT r.note as Notes FROM retour_client r JOIN produit p ON r.cle_produit = p.cle_produit WHERE p.titre_produit = 'TV' AND r.libelle_source = 'réseaux sociaux';</pre>	<table><tr><th>Notes</th></tr><tr><td>8</td></tr><tr><td>9</td></tr><tr><td>10</td></tr><tr><td>10</td></tr></table>	Notes	8	9	10	10
Notes								
8								
9								
10								
10								

I. Analyse de la satisfaction client

3	Quelle est la note moyenne pour chaque catégorie de produit ? (Classé de la meilleure à la moins bonne)	<pre>SELECT p.typologie_produit, ROUND(AVG(r.note), 2) AS note_moyenne FROM retour_client r JOIN produit p ON r.cle_produit = p.cle_produit GROUP BY p.typologie_produit ORDER BY note_moyenne DESC;</pre>	<table><tr><th>typologie_produit</th><th>note_moyenne</th></tr><tr><td>High-Tech</td><td>8.16</td></tr><tr><td>Loisirs</td><td>8.09</td></tr><tr><td>Alimentaire</td><td>8.04</td></tr><tr><td>Maison</td><td>7.85</td></tr></table>	typologie_produit	note_moyenne	High-Tech	8.16	Loisirs	8.09	Alimentaire	8.04	Maison	7.85								
typologie_produit	note_moyenne																				
High-Tech	8.16																				
Loisirs	8.09																				
Alimentaire	8.04																				
Maison	7.85																				
4	Quels sont les 5 magasins avec les meilleures notes moyennes ?	<pre>SELECT p.ref_magasin, ROUND(AVG(r.note), 2) AS note_moyenne, p.libelle_de_commune FROM retour_client r JOIN ref_magasin p ON r.ref_magasin = p.ref_magasin GROUP BY p.ref_magasin ORDER BY note_moyenne DESC LIMIT 5;</pre>	<table><tr><th>ref_magasin</th><th>note_moyenne</th><th>libelle_de_commune</th></tr><tr><td>75</td><td>8.73</td><td>Paris 14^e</td></tr><tr><td>78</td><td>8.55</td><td>Saint-Pierre-du-Perray</td></tr><tr><td>62</td><td>8.5</td><td>Paris 19^e</td></tr><tr><td>23</td><td>8.48</td><td>Paris 11^e</td></tr><tr><td>19</td><td>8.45</td><td>Coulommiers</td></tr></table>	ref_magasin	note_moyenne	libelle_de_commune	75	8.73	Paris 14 ^e	78	8.55	Saint-Pierre-du-Perray	62	8.5	Paris 19 ^e	23	8.48	Paris 11 ^e	19	8.45	Coulommiers
ref_magasin	note_moyenne	libelle_de_commune																			
75	8.73	Paris 14 ^e																			
78	8.55	Saint-Pierre-du-Perray																			
62	8.5	Paris 19 ^e																			
23	8.48	Paris 11 ^e																			
19	8.45	Coulommiers																			

I. Analyse de la satisfaction client

5	Quels sont les magasins qui ont plus de 12 feedbacks sur le drive ?	<pre>SELECT p.ref_magasin, COUNT(r.cle_retour_client) AS nb_feedbacks, p.libelle_de_commune FROM retour_client r JOIN ref_magasin p ON r.ref_magasin = p.ref_magasin WHERE r.libelle_categorie = 'drive' GROUP BY p.ref_magasin HAVING COUNT(*) > 12;</pre>	<table><tr><th>ref_magasin</th><th>nb_feedbacks</th><th>libelle_de_commune</th></tr><tr><td>45</td><td>13</td><td>Paris 12^e</td></tr><tr><td>63</td><td>13</td><td>Ivry-sur-Seine</td></tr><tr><td>67</td><td>14</td><td>Éragny</td></tr></table>	ref_magasin	nb_feedbacks	libelle_de_commune	45	13	Paris 12 ^e	63	13	Ivry-sur-Seine	67	14	Éragny															
ref_magasin	nb_feedbacks	libelle_de_commune																												
45	13	Paris 12 ^e																												
63	13	Ivry-sur-Seine																												
67	14	Éragny																												
6	Quel est le classement des départements par notes ?	<pre>SELECT p.departement, ROUND(AVG(r.note), 2) AS note_moyenne FROM retour_client r JOIN ref_magasin p ON r.ref_magasin = p.ref_magasin GROUP BY p.departement ORDER BY note_moyenne DESC;</pre>	<table><tr><th>#</th><th>departement</th><th>note_moyenne</th></tr><tr><td>1</td><td>95</td><td>8.14</td></tr><tr><td>2</td><td>75</td><td>8.11</td></tr><tr><td>3</td><td>94</td><td>8.06</td></tr><tr><td>4</td><td>91</td><td>8.05</td></tr><tr><td>5</td><td>77</td><td>8.04</td></tr><tr><td>6</td><td>92</td><td>8.03</td></tr><tr><td>7</td><td>78</td><td>8.02</td></tr><tr><td>8</td><td>93</td><td>7.94</td></tr></table>	#	departement	note_moyenne	1	95	8.14	2	75	8.11	3	94	8.06	4	91	8.05	5	77	8.04	6	92	8.03	7	78	8.02	8	93	7.94
#	departement	note_moyenne																												
1	95	8.14																												
2	75	8.11																												
3	94	8.06																												
4	91	8.05																												
5	77	8.04																												
6	92	8.03																												
7	78	8.02																												
8	93	7.94																												

I. Analyse de la satisfaction client

7	Quelle est la typologie de produit qui apporte le meilleur service après-vente ?	<pre>SELECT p.typologie_produit, ROUND(AVG(r.note), 2) AS note_moyenne FROM retour_client r JOIN produit p ON r.cle_produit = p.cle_produit WHERE libelle_categorie = 'service après-vente' GROUP BY p.typologie_produit ORDER BY note_moyenne DESC LIMIT 1;</pre>	<table><tr><th>typologie_p roduit</th><th>note_moye nne</th></tr><tr><td>Loisirs</td><td>8.51</td></tr></table>	typologie_p roduit	note_moye nne	Loisirs	8.51
typologie_p roduit	note_moye nne						
Loisirs	8.51						
8	Quelle est la note moyenne sur l'ensemble des boissons ?	<pre>SELECT ROUND(AVG(r.note), 2) AS note_moyenne FROM retour_client r JOIN produit p ON r.cle_produit = p.cle_produit WHERE p.titre_produit LIKE '%Boissons%';</pre>	<table><tr><th>note_moyenne</th></tr><tr><td>8.21</td></tr></table>	note_moyenne	8.21		
note_moyenne							
8.21							

I. Analyse de la satisfaction client

9	Quel est le classement des jours de la semaine où l'expérience client est la meilleure expérience en magasin ?	<pre>SELECT CASE strftime('%w', date_achat) WHEN '0' THEN 'Dimanche' WHEN '1' THEN 'Lundi' WHEN '2' THEN 'Mardi' WHEN '3' THEN 'Mercredi' WHEN '4' THEN 'Jeudi' WHEN '5' THEN 'Vendredi' WHEN '6' THEN 'Samedi' END AS jour, ROUND(AVG(note), 2) AS moyenne_note FROM retour_client WHERE libelle_categorie = 'expérience en magasin' GROUP BY jour ORDER BY moyenne_note DESC;</pre>	<table><tr><th>#</th><th>jour</th><th>note_moyenne</th></tr><tr><td>1</td><td>Samedi</td><td>8.34</td></tr><tr><td>2</td><td>Dimanche</td><td>8.18</td></tr><tr><td>3</td><td>Vendredi</td><td>8.07</td></tr><tr><td>4</td><td>Jeudi</td><td>8.04</td></tr><tr><td>5</td><td>Mercredi</td><td>7.99</td></tr><tr><td>6</td><td>Mardi</td><td>7.95</td></tr><tr><td>7</td><td>Lundi</td><td>7.74</td></tr></table>	#	jour	note_moyenne	1	Samedi	8.34	2	Dimanche	8.18	3	Vendredi	8.07	4	Jeudi	8.04	5	Mercredi	7.99	6	Mardi	7.95	7	Lundi	7.74
#	jour	note_moyenne																									
1	Samedi	8.34																									
2	Dimanche	8.18																									
3	Vendredi	8.07																									
4	Jeudi	8.04																									
5	Mercredi	7.99																									
6	Mardi	7.95																									
7	Lundi	7.74																									
10	Sur quel mois a-t-on le plus de retour sur le service après-vente ?	<pre>SELECT strftime('%m', date_achat) AS mois, COUNT(*) AS nombre_retours FROM retour_client WHERE libelle_categorie = 'service après-vente' GROUP BY mois ORDER BY nombre_retours DESC LIMIT 1;</pre>	<table><tr><th>mois</th><th>nombre_retours</th></tr><tr><td>10</td><td>55</td></tr></table>	mois	nombre_retours	10	55																				
mois	nombre_retours																										
10	55																										

I. Analyse de la satisfaction client

11	Quel est le pourcentage de recommandations client ?	<pre>SELECT ROUND((SUM(recommandation) * 100) / COUNT(*), 2) AS pourcentage_recommandations, COUNT(*) AS nombre_total_retours FROM retour_client WHERE recommandation IN (0, 1);</pre>	<table><tr><th>pourcentage_rec ommandations</th><th>Nombre_total_re tours</th></tr><tr><td>90</td><td>2326</td></tr></table>	pourcentage_rec ommandations	Nombre_total_re tours	90	2326														
pourcentage_rec ommandations	Nombre_total_re tours																				
90	2326																				
12	Quels sont les magasins qui ont une note inférieure à la moyenne ?	<pre>SELECT p.ref_magasin, ROUND(AVG(r.note), 2) AS moyenne_magasin, ROUND(mg.moyenne_globale, 2) AS moyenne_globale FROM retour_client r JOIN ref_magasin p ON r.ref_magasin = p.ref_magasin CROSS JOIN (SELECT AVG(note) AS moyenne_globale FROM retour_client) mg GROUP BY p.ref_magasin, mg.moyenne_globale HAVING AVG(r.note) < mg.moyenne_globale ORDER BY moyenne_magasin ASC;</pre>	<table><tr><th>ref_maga sin</th><th>moyenne _magasin</th><th>moyenne _globale</th></tr><tr><td>60</td><td>7.38</td><td>8.05</td></tr><tr><td>81</td><td>7.44</td><td>8.05</td></tr><tr><td>82</td><td>7.53</td><td>8.05</td></tr><tr><td>46</td><td>7.56</td><td>8.05</td></tr><tr><td>55</td><td>7.59</td><td>8.05</td></tr></table> <p>Et 34 de plus...</p>	ref_maga sin	moyenne _magasin	moyenne _globale	60	7.38	8.05	81	7.44	8.05	82	7.53	8.05	46	7.56	8.05	55	7.59	8.05
ref_maga sin	moyenne _magasin	moyenne _globale																			
60	7.38	8.05																			
81	7.44	8.05																			
82	7.53	8.05																			
46	7.56	8.05																			
55	7.59	8.05																			

I. Analyse de la satisfaction client

13	Quelles sont les typologies produits qui ont amélioré leur moyenne entre le 1er et le 2ème trimestre 2021 ?	<pre>SELECT trimestre1.typologie_produit, trimestrel.moyenne_typologie_produit as moyenne_t1, trimestre2.moyenne_typologie_produit as moyenne_t2 FROM (SELECT typologie_produit, ROUND(AVG(note), 2) as moyenne_typologie_produit FROM retour_client rc LEFT JOIN produit p ON rc.cle_produit = p.cle_produit WHERE strftime('%m', date_achat) IN ('01', '02', '03') GROUP BY typologie_produit) trimestre1, (SELECT typologie_produit, ROUND(AVG(note), 2) as moyenne_typologie_produit FROM retour_client rc LEFT JOIN produit p ON rc.cle_produit = p.cle_produit WHERE strftime('%m', date_achat) IN ('04', '05', '06') GROUP BY typologie_produit) trimestre2 WHERE trimestre2.typologie_produit = trimestre1.typologie_produit AND trimestre2.moyenne_typologie_produit > trimestrel.moyenne_typologie_produit;</pre>	<table><tr><th>typologie_p roduit</th><th>moyenn e_t1</th><th>moyenne _t2</th></tr><tr><td>Alimentaire</td><td>7.99</td><td>8.06</td></tr><tr><td>Loisirs</td><td>8</td><td>8.34</td></tr></table>	typologie_p roduit	moyenn e_t1	moyenne _t2	Alimentaire	7.99	8.06	Loisirs	8	8.34									
typologie_p roduit	moyenn e_t1	moyenne _t2																			
Alimentaire	7.99	8.06																			
Loisirs	8	8.34																			
14	Quels sont les 5 produits ayant reçu les meilleures notes moyennes ?	<pre>SELECT p.titre_produit AS nom_produit, p.typologie_produit, ROUND(AVG(rc.note), 2) AS note_moyenne FROM produit p LEFT JOIN retour_client rc ON p.cle_produit = rc.cle_produit GROUP BY p.cle_produit, p.titre_produit, p.typologie_produit HAVING COUNT(rc.cle_retour_client) >= 10 ORDER BY note_moyenne DESC LIMIT 5;</pre>	<table><tr><th>nom_produit</th><th>typologie_prod uit</th><th>note_ moye nne</th></tr><tr><td>Plantes aromatiques surgelées</td><td>Alimentaire</td><td>9.2</td></tr><tr><td>Sodas</td><td>Alimentaire</td><td>9</td></tr><tr><td>Boissons alcoolisées</td><td>Alimentaire</td><td>8.79</td></tr><tr><td>Sauces au soja</td><td>Alimentaire</td><td>8.77</td></tr><tr><td>Aliments à base de plantes frais</td><td>Alimentaire</td><td>8.71</td></tr></table>	nom_produit	typologie_prod uit	note_ moye nne	Plantes aromatiques surgelées	Alimentaire	9.2	Sodas	Alimentaire	9	Boissons alcoolisées	Alimentaire	8.79	Sauces au soja	Alimentaire	8.77	Aliments à base de plantes frais	Alimentaire	8.71
nom_produit	typologie_prod uit	note_ moye nne																			
Plantes aromatiques surgelées	Alimentaire	9.2																			
Sodas	Alimentaire	9																			
Boissons alcoolisées	Alimentaire	8.79																			
Sauces au soja	Alimentaire	8.77																			
Aliments à base de plantes frais	Alimentaire	8.71																			

I. Analyse de la satisfaction client

15

Quels sont les 5 magasins qui ont le plus de feedbacks ?

```
SELECT p.ref_magasin, COUNT(r.cle_retour_client) AS feedbacks, p.libelle_de_commune FROM retour_client r JOIN ref_magasin p ON r.ref_magasin = p.ref_magasin GROUP BY p.ref_magasin ORDER BY feedbacks DESC LIMIT 5;
```

ref_magasin	feedbacks	libelle_de_commune
29	55	Mareuil-lès-Meaux
6	49	Osny
80	47	Lognes
5	45	Villecresnes
83	44	Mitry-Mory

16

Calcul du NPS

```
WITH NPSData AS ( SELECT CASE WHEN note >= 9 THEN 'promoteur' WHEN note >= 7 THEN 'passif' ELSE 'detracteur' END AS categorie FROM retour_client ) SELECT COUNT(*) AS total, SUM(CASE WHEN categorie = 'promoteur' THEN 1 ELSE 0 END) AS promoteurs, SUM(CASE WHEN categorie = 'passif' THEN 1 ELSE 0 END) AS passifs, SUM(CASE WHEN categorie = 'detracteur' THEN 1 ELSE 0 END) AS detracteurs, ((SUM(CASE WHEN categorie = 'promoteur' THEN 1 ELSE 0 END) - SUM(CASE WHEN categorie = 'detracteur' THEN 1 ELSE 0 END)) * 100.0 / COUNT(*)) AS nps FROM NPSData;
```

total	promoteurs	passifs	detracteurs	nps
3000	1200	1529	271	30.96

II. Méthodologie suivie



- **MAJ dictionnaire de la base**

Nous avons effectué une révision et une documentation complète de toutes les tables, les colonnes, les contraintes, les types de données ainsi que leur descriptions.

- **Création du schéma de la base**

Nous avons créé le schéma de la base de données, qui définit la structure globale de la base, à savoir : les tables, les relations entre les tables ainsi que les clés primaires et étrangères. Ce schéma a été réalisé à l'aide de MySQL Workbench.

- **Importation des données**

Nous avons importé les données brutes dans notre base de données SQLite à partir d'un fichier SQL externe. Ce fichier SQL contenait les données à importer, et nous avons utilisé une commande SQL pour transférer les informations dans les tables de notre base de données, en respectant la structure définie précédemment

```
sqlite3 feedback.db  
.read CustomerDataFeedback.sql
```

- **MAJ de la base avec un nouveau fichier**

Nous avons reçu un nouveau fichier contenant des données importantes à ajouter à notre base de données existante. Nous avons donc utilisé une commande SQL pour intégrer les nouvelles données du fichier dans une table appropriée au sein de notre base de données.

```
sqlite3 feedback.db  
.mode csv  
.separator ;  
.import ref_magasin.csv ref_magasin  
.exit
```

- **Élaboration des requêtes**

Une fois les données correctement structurées dans la base de données, nous avons élaboré des requêtes SQL afin de mener des analyses et d'extraire les informations pour répondre aux questions spécifiques posées dans le cadre du projet.

III. Cohérence des données



- Lors de l'analyse de données, nous avons effectué une évaluation de la cohérence et de la qualité des données avec lesquelles nous avons travaillé. Cela inclut l'examen des valeurs nulles dans les colonnes, la recherche de doublons dans les tables, la vérification de l'intégrité des clés étrangères, et l'inspection des données pour détecter toute anomalie ou incohérence.
- Pour nous assurer de l'exactitude de nos résultats, nous avons tout d'abord effectué des contrôles de qualité des données pour garantir que les valeurs étaient cohérentes et conformes aux spécifications de notre schéma de base de données. Nous avons ensuite vérifié que les clés étrangères correspondaient correctement aux enregistrements correspondants dans d'autres tables. Enfin, nous avons utilisé des requêtes de validation pour nous assurer que les résultats obtenus correspondaient aux résultats attendus. Par exemple, nous avons vérifié que le NPS était bien compris entre -100 et 100, comme il se doit.