

SQL Projet Transverse

MBA ESG - MBDIA

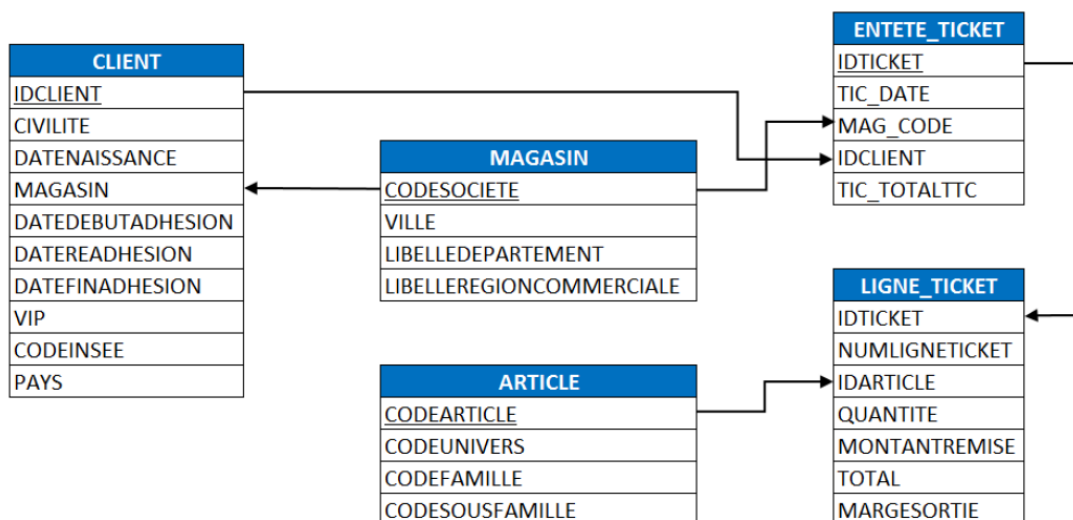
Réalisé par :

- VU Tien Dat
- BUI Thuy Tien
- INES Zouari

Contexte : Une société X a envoyé ces données client ainsi que les achats sur l'année N-2 (2016) et N-1 (2017)

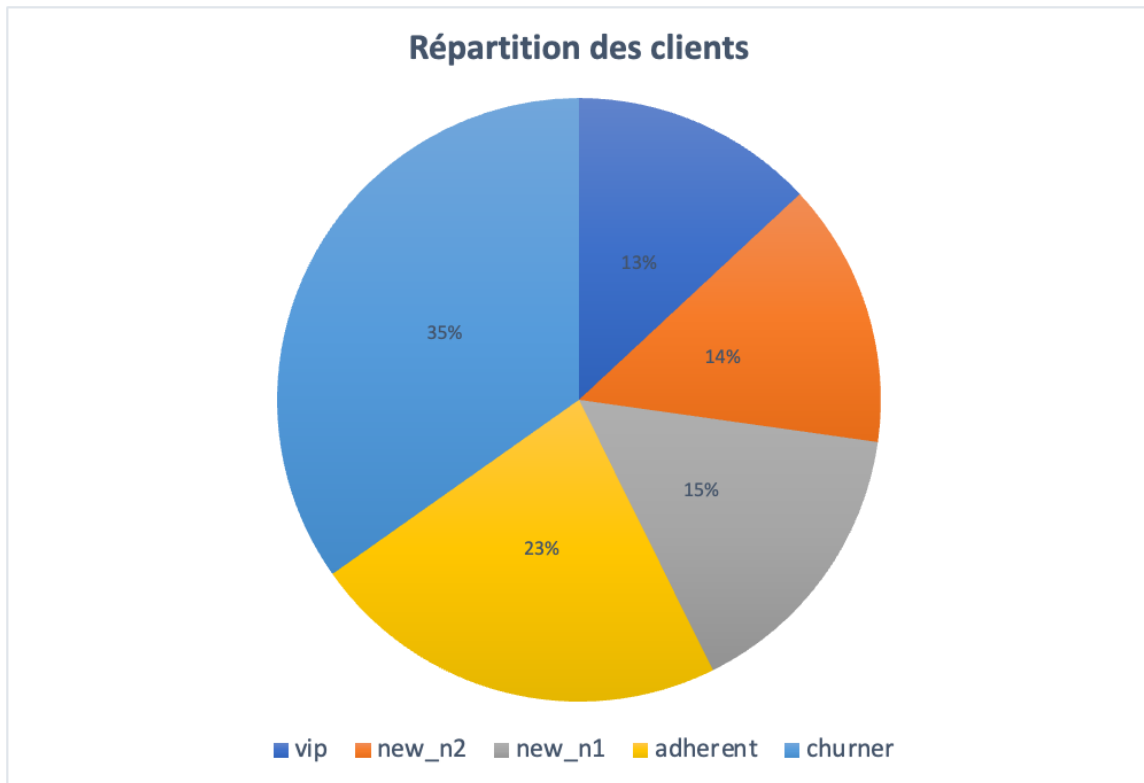
Objectif :

- Manipuler et analyser de la Data sous SQL
- Programmer en SQL
- Sortir un rapport d'étude

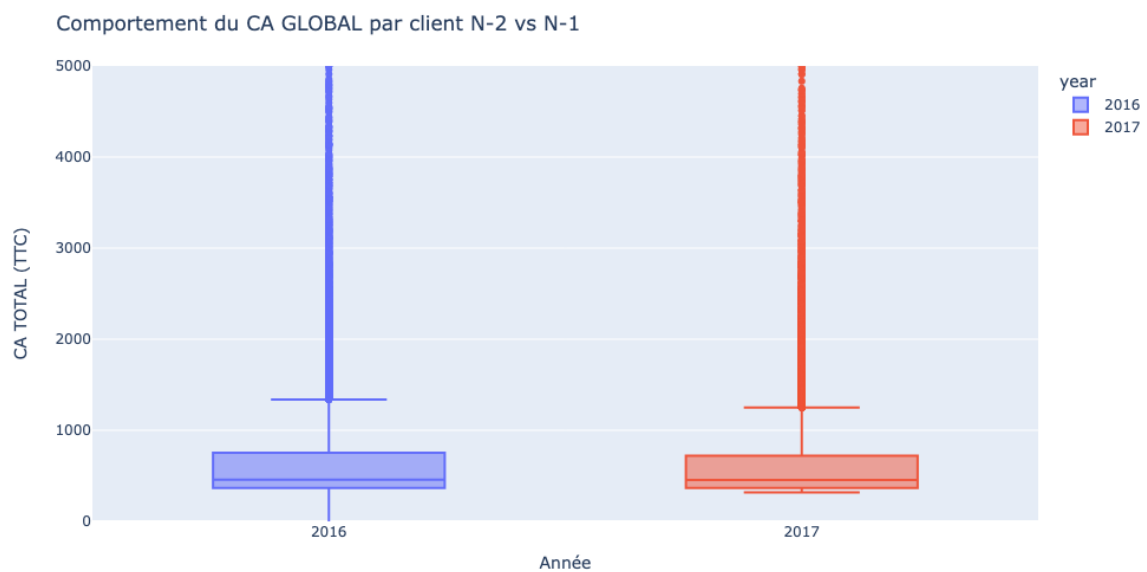


I. Étude globale :

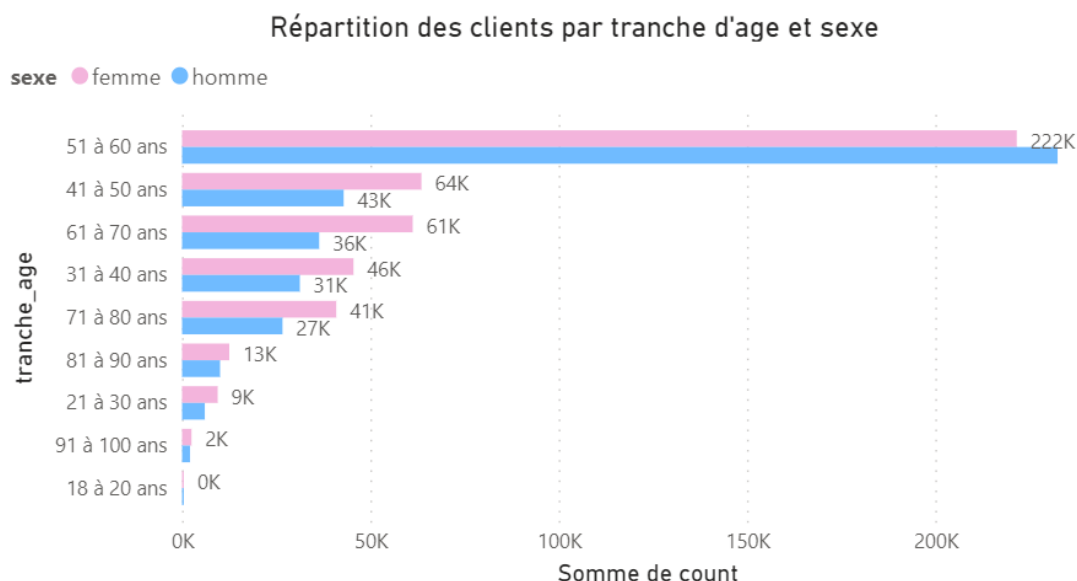
a. *Répartition Adhérant / VIP*



b. *Comportement du CA GLOBAL par client N-2 vs N-1*



c. Répartition par âge x sexe



II. Étude par magasin

a. Résultat par magasin

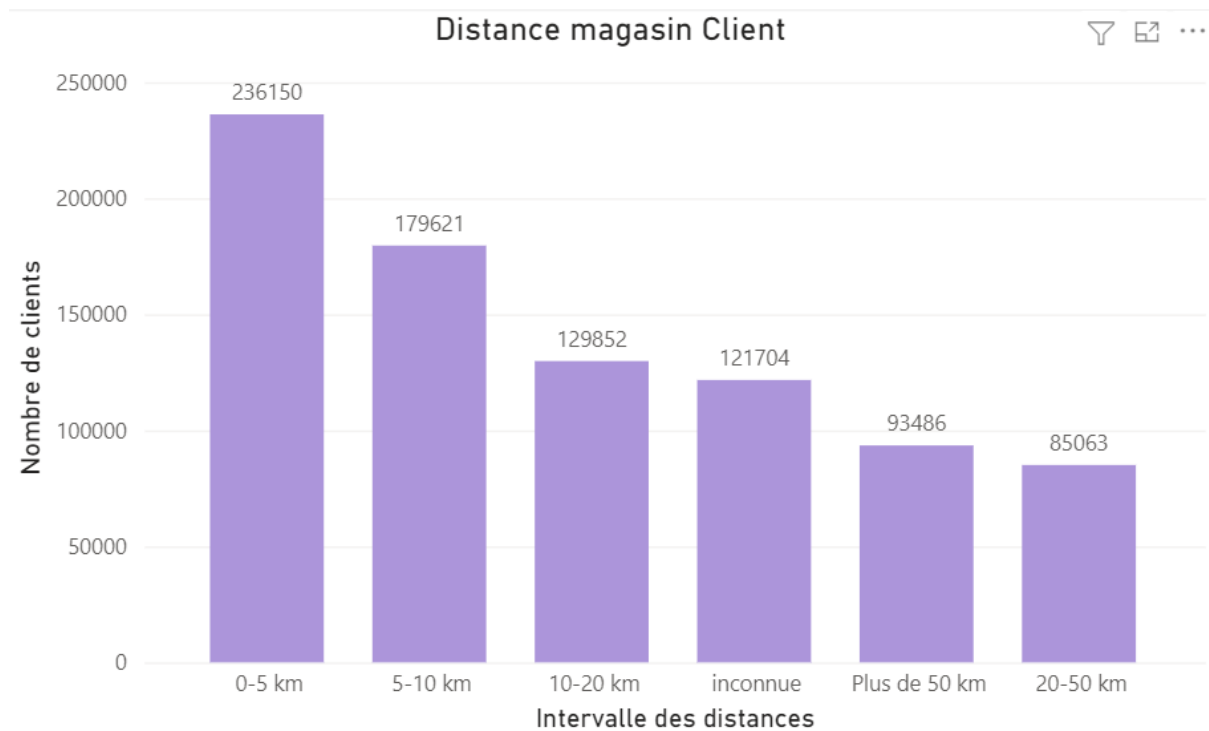
| magasin | ville | nombre_adherent | nombreclientactif_2016 | nombreclientactif_2017 | sales_2016 | sales_2017 | evolution_client (%) | evolution_ca (%) |
|---------|--------------------|-----------------|------------------------|------------------------|------------|------------|----------------------|------------------|
| PRI | SAINT-PRIEST | 26935 | 17157 | 18686 | 5528750,33 | 5557473,28 | 8 | 0,52 |
| SEY | SEYSSINS | 25967 | 17496 | 19242 | 4336406,7 | 4400118,8 | 9 | 1,47 |
| VIB | VILLEURBANNE | 24640 | 15707 | 16975 | 4402349,51 | 4323459,04 | 8 | -1,79 |
| MOU | MOUANS-SARTOUX | 24389 | 16243 | 17401 | 6349088,35 | 6045145,65 | 7 | -4,79 |
| SUR | SURESNES | 23784 | 14783 | 15810 | 4807772,47 | 4966887,74 | 6 | 3,31 |
| CLA | CLAPIERS | 23744 | 15625 | 16800 | 5847741,89 | 5639065,55 | 7 | -3,57 |
| SSM | LA SEYNE-SUR-MER | 23601 | 15639 | 17564 | 6413411,93 | 6842402,95 | 12 | 6,69 |
| HEI | HEILLECOURT | 23500 | 16224 | 17722 | 5834749,58 | 5796841,23 | 9 | -0,65 |
| VLG | VILLE-LA-GRAND | 22541 | 15679 | 15890 | 6212516,5 | 5653657,78 | 1 | -9 |
| BEA | BEAUMONT | 20714 | 14523 | 15742 | 5739924,27 | 5566069,91 | 8 | -3,03 |
| GEX | ST GENIS POUILLY | 19278 | 14281 | 14563 | 7388908,23 | 6957685,71 | 1 | -5,84 |
| MOB | MONTBONNOT-SAINT | 18929 | 13647 | 15236 | 5977004,79 | 6384418,04 | 11 | 6,82 |
| MET | METZ-TESSY | 17997 | 13029 | 13498 | 4826410,13 | 4633844,77 | 3 | -3,99 |
| IAB | L'ISLE-D'ABEAU | 16885 | 10786 | 12404 | 3222142,17 | 3239040,89 | 15 | 0,52 |
| VAL | VALENCE | 16685 | 11345 | 11099 | 3459987,91 | 2986765,69 | -2 | -13,68 |
| ALM | LES MILLES | 16499 | 11503 | 12651 | 4036788,61 | 4113500,34 | 9 | 1,9 |
| GAI | GAILLARD | 16162 | 11657 | 11789 | 6143169,42 | 5666500,73 | 1 | -7,76 |
| AVI | VILLENEUVE-LES-AV | 15435 | 10474 | 11433 | 4065939,61 | 3951027,58 | 9 | -2,83 |
| ECU | ECULLY | 15410 | 10496 | 11627 | 4112227,44 | 4274237,6 | 10 | 3,94 |
| FRV | FRANCHEVILLE | 15377 | 10613 | 11261 | 3922133,37 | 3764133,92 | 6 | -4,03 |
| VIT | VITROLLES | 14673 | 9586 | 10665 | 3007212,43 | 3089915,26 | 11 | 2,75 |
| MAC | MACON | 14632 | 9678 | 10730 | 2798360,47 | 2821781,51 | 10 | 0,84 |
| MUL | MULHOUSE | 14240 | 9883 | 10101 | 3332191,16 | 3172041,51 | 2 | -4,81 |
| VIF | VILLEFRANCHE-SUR | 14180 | 9652 | 10038 | 2709178,85 | 2586192,37 | 3 | -4,54 |
| VAR | VARENNES VAUZELL | 12948 | 10134 | 9663 | 3150351,97 | 2701641,48 | -4 | -14,24 |
| ALB | GILLY SUR ISERE | 12828 | 8739 | 9386 | 2960491,65 | 3049735,09 | 7 | 3,01 |
| DIJ | DIJON | 12603 | 8693 | 9299 | 3043839,41 | 2900305,72 | 6 | -4,72 |
| SMR | SAINT-MITRE-LES-RE | 12532 | 8815 | 9225 | 3044483,67 | 3063414,7 | 4 | 0,62 |
| STE | LA FOUILLOUSE | 12001 | 8187 | 8102 | 2526952,24 | 2279578,34 | -1 | -9,79 |
| PON | LE PONTET | 11788 | 7702 | 7517 | 2577362,88 | 2363867,85 | -2 | -8,28 |
| BLA | BLAGNIAC | 11748 | 7135 | 8019 | 2143924,9 | 2255599,64 | 12 | 5,21 |
| QUE | QUETIGNY | 11699 | 7344 | 8411 | 1783723,09 | 1898024,29 | 14 | 6,41 |
| VIC | VILLECHETIF | 11556 | 7481 | 8231 | 2114766,5 | 2133378,76 | 10 | 0,88 |
| RAV | LA RAVOIRE | 11396 | 7882 | 8072 | 2625700,87 | 2302863,76 | 2 | -12,3 |
| MAN | MANOSQUE | 11343 | 8244 | 8090 | 2719688,26 | 2461591,49 | -1 | -9,49 |
| SJV | SAINT-JEAN-DE-VED | 11065 | 7157 | 7768 | 2254727,34 | 2414626,02 | 8 | 7,09 |
| LAB | LABEGE CEDEX | 11058 | 7091 | 7767 | 2053101,79 | 2021049,86 | 9 | -1,56 |
| PEG | PERRIGNY | 11012 | 7394 | 7734 | 2118864,43 | 2044229,97 | 4 | -3,52 |
| SNO | SEYNOD | 10828 | 6683 | 8699 | 2071762,57 | 2292897,87 | 30 | 10,67 |
| FEG | FEGRSHEIM | 10561 | 7215 | 7446 | 2347398,26 | 2161775,78 | 3 | -7,91 |
| THO | THONON-LES-BAINS | 10549 | 7281 | 7837 | 2897265,27 | 2979303,98 | 7 | 2,83 |
| VEN | VENELLES | 10544 | 7821 | 8068 | 3354293,2 | 3182815,7 | 3 | -5,11 |
| SEM | FEVES | 10360 | 6720 | 6308 | 1959365,37 | 1781800,21 | -6 | -9,06 |
| BSN | CESSON | 10302 | 6646 | 5730 | 2020922,33 | 1665792,86 | -13 | -17,57 |
| POC | PONTAULT-COMBAU | 10182 | 6685 | 6613 | 2406910,18 | 2235013,54 | -1 | -7,14 |
| BRE | BRETAGNY-SUR-ORG | 10106 | 6261 | 7288 | 2089464,49 | 2121890,22 | 16 | 1,55 |
| ORL | ORLEANS | 9891 | 6785 | 7042 | 2695002,71 | 2467711,98 | 3 | -8,43 |
| ORF | ORFÈVRE | 8760 | 5684 | 7476 | 4789906,91 | 4789906,91 | 0 | 0,0 |

b. **Distance CLIENT / MAGASIN :**

Nous avons effectué les étapes suivantes pour calculer la distance entre le client et le magasin :

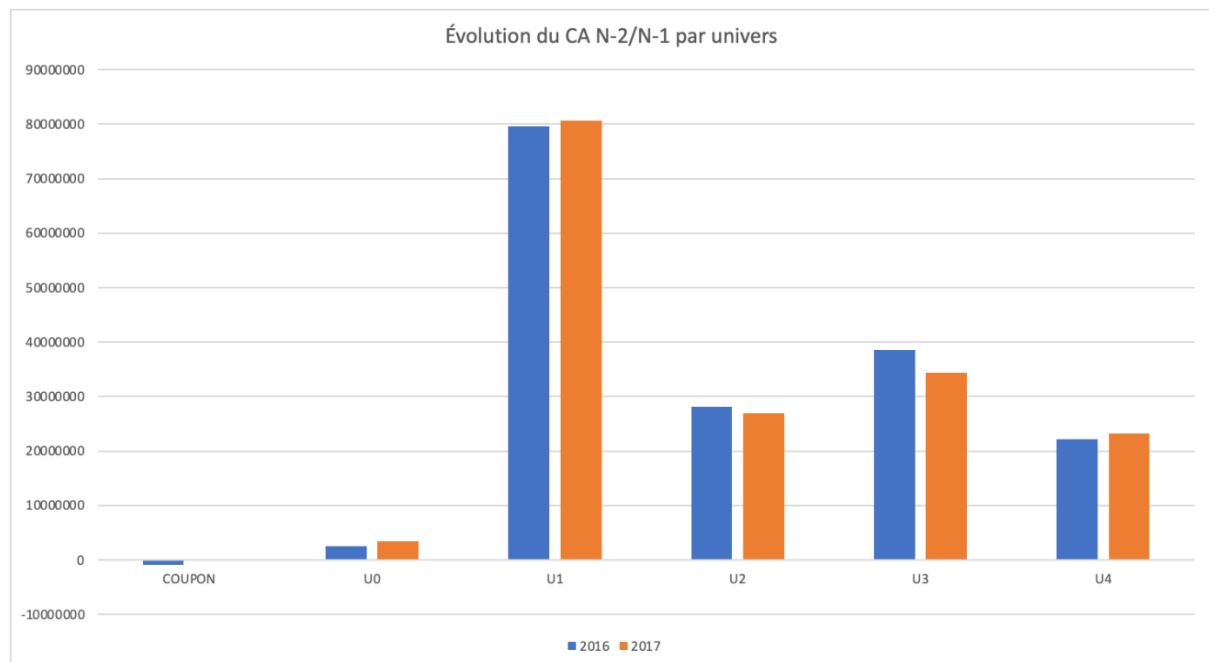
1. Nous avons préparé la table de données GPS en ne conservant que les six colonnes nécessaires, à savoir : code INSEE, code postal, commune, latitude et longitude.
2. Nous avons ajouté deux colonnes de longitude et latitude pour chacune des tables client et magasin en comparant le code INSEE pour la table client. Pour la table de magasin, nous les avons comparées avec la ville et la commune. Nous avons remarqué qu'il était nécessaire de retirer les espaces, les tirets, de remplacer "st" par "saint" et de supprimer le mot "CEDEX".
3. Nous avons créé une nouvelle table de jointure entre la table magasin et client, où nous avons récupéré les colonnes de longitude et latitude, ainsi que l'ID du client pour effectuer le calcul de la distance entre deux points (magasin et client). Nous avons utilisé la fonction "haversine_distance" pour calculer la distance, et nous avons vérifié les calculs avec le site Lexilogos, qui a renvoyé les mêmes valeurs.
4. Nous avons créé une nouvelle table pour représenter le nombre de clients et le pourcentage de clients par distance : 0 à 5km, 5km à 10km, 10km à 20km, 20km à 50km, et plus de 50km.

| Somme de client_count | client_percentage | distance_range |
|-----------------------|-------------------|----------------|
| 236150 | 27.92 | 0-5 km |
| 179621 | 21.23 | 5-10 km |
| 129852 | 15.35 | 10-20 km |
| 121704 | 14.39 | inconnue |
| 93486 | 11.05 | Plus de 50 km |
| 85063 | 10.06 | 20-50 km |



III. Étude par univers :

- a. *Constituer un histogramme N-2 / N-1 évolution du CA par univers*



**b. Afficher le top 5 des familles les plus rentable par univers
(en fonction de la marge obtenue)**

