

Rapport sur le traitement des données Dev'Immediat

Processus d'extraction, de préparation et d'anonymisation des données

<i>Version</i>	<i>Auteur</i>	<i>Description</i>	<i>Date</i>
V1	Nathalie Currid	<i>Proposition de tableau anonymisé, destiné à l'équipe de Performance commerciale, contenant les données nécessaires des clients ayant un dossier complet, suite à une demande de devis faite en 2022, incluant les données tarifaires non détaillées.</i>	21/02/2024

Présentation du tableau anonymisé pour l'équipe Performance commerciale et justification des choix

Tableau anonymisé pour l'équipe

Performance commerciale

- **Champs non sélectionnés** car ils permettent d'identifier les clients et/ou ne sont pas nécessaires à la finalité du traitement des données : *metier, employeur, num_ss, groupe_sanguin, id_site_web, nom, email, id_client, valeur_residence_prin, formation, etat_dossier, adresse, lat, lon*. A noter que le traitement de données de santé est en principe interdit (article 9 du RGPD), cependant il existe des dérogations.
- **Index créé** pour remplacer les valeurs du champ *id_client* (A), dans le cas où celles-ci permettraient l'identification des clients
- **Champ converti en âge** puis en **tranches d'âge** (C) pour une anonymisation maximale : *date_naissance* (devenu *age_client*)
- **Champs convertis en intervalles** : *revenus* (D), *tarif_devis* (O), permettant une analyse statistique tout en rendant le traçage impossible
- **Champs modifiés en type de données logiques** (booléen) permettant d'identifier des tendances : *nombre_enfants* (E) (devenu 'enfants') et *enfant_conduite_accompagne* (F)
- **Champ remplacé par oui/non** : *est_rouge* (I), **champ converti en nombre entier** : *age_vehicule* (H)
- **Champ converti en mois de l'année 2022** : *date_demande* (M)

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
ID	sexe	age	revenus_K€	enfants	enfant_conduite_ac	type_vehicule	age_vehicule	est_rouge	usage_vehicule	type_conduite	points_perdus	date_demande	formule	tarif_devis_€
1	F	[15-25]	[125-150]	oui	non	Sports Car	7	non	Commercial	Highly Urban/ Urban	0	février	dev_integral	[300-450]
2	F	> 75	[50-75]	oui	oui	Minivan	1	non	Private	Highly Urban/ Urban	2	mai	dev_express	< 300
3	F	[45-55]	< 25	oui	non	Pickup	4	non	Private	z_Highly Rural/ Rural	0	janvier	dev_express	< 300
4	F	< 15	< 25	non	non	z_SUV	5	non	Private	Highly Urban/ Urban	3	avril	dev_integral	[300-450]
5	M	[65-75]	< 25	non	non	Panel Truck	12	non	Commercial	Highly Urban/ Urban	2	mars	dev_vip	> 600
6	F	[25-35]	[75-100]	oui	non	z_SUV	15	non	Private	z_Highly Rural/ Rural	2	mars	dev_premium	[450-600]
7	F	> 75	[125-150]	oui	oui	Minivan	9	non	Commercial	z_Highly Rural/ Rural	0	mars	dev_express	< 300

Description des étapes techniques

Extraction des données du CRM

A - Requête SQL avec SQLite



Éditeur SQL 1

CRM (1)

Requête Historique

```
1 SELECT sexe, date_naissance, revenus, nombre_enfants, enfant_conduite_accompagne, type_vehicule, age_vehicule, est_rouge, usage_vehicule,
2 type_conduite, points_perdus, date_demande, formule, tarif_devis
3 FROM base_client
4 WHERE etat_dossier = 'complet' AND date_demande LIKE '%2022%';
```

Table Formulaire

Nombre de lignes chargées : 1158

	sexe	date_naissan	revenus	nombre_	enfant_c	type_vehicu	age_veh	est_roug	usage_vehic	type_conduite	points_p	date_demande	formule
1	F	1999-09-25	125301	1	0	Sports Car	7.0	no	Commercial	Highly Urban/ Urban	0	2022-02-25 16:13:38	dev_integr
2	F	1932-01-28	50815	1	1	Minivan	1.0	no	Private	Highly Urban/ Urban	2	2022-05-24 21:07:34	dev_expre
3	F	1978-09-15	21204	2	0	Pickup	4.0	no	Private	z_Highly Rural/ Rural	0	2022-01-24 08:17:21	dev_expre
4	F	2021-02-17	0	0	0	z_SUV	5.0	no	Private	Highly Urban/ Urban	3	2022-04-20 03:46:00	dev_integr
5	M	1956-06-14	NULL	0	0	Panel Truck	12.0	no	Commercial	Highly Urban/ Urban	2	2022-03-23 01:32:40	dev_vip
6	F	1992-04-11	79606	1	0	z_SUV	15.0	no	Private	z_Highly Rural/ Rural	2	2022-03-23 22:20:50	dev_premi
7	F	1941-03-18	130540	2	2	Minivan	9.0	no	Commercial	z_Highly Rural/ Rural	0	2022-03-13 01:10:59	dev_expre
8	M	1956-03-06	51278	0	0	z_SUV	10.0	no	Private	Highly Urban/ Urban	1	2022-03-09 03:43:53	dev_premi
9	M	1956-02-18	115044	0	0	Minivan	14.0	no	Private	Highly Urban/ Urban	0	2022-01-29 10:17:49	dev_premi
10	F	1974-03-07	NULL	0	0	Panel Truck	10.0	yes	Commercial	Highly Urban/ Urban	1	2022-03-26 07:37:33	dev_integr
11	F	1967-02-13	24060	2	1	Minivan	10.0	no	Private	Highly Urban/ Urban	0	2022-05-26 05:51:37	dev_expre

Barre d'état

[12:37:01] Requête terminée en 0.001 seconde(s).

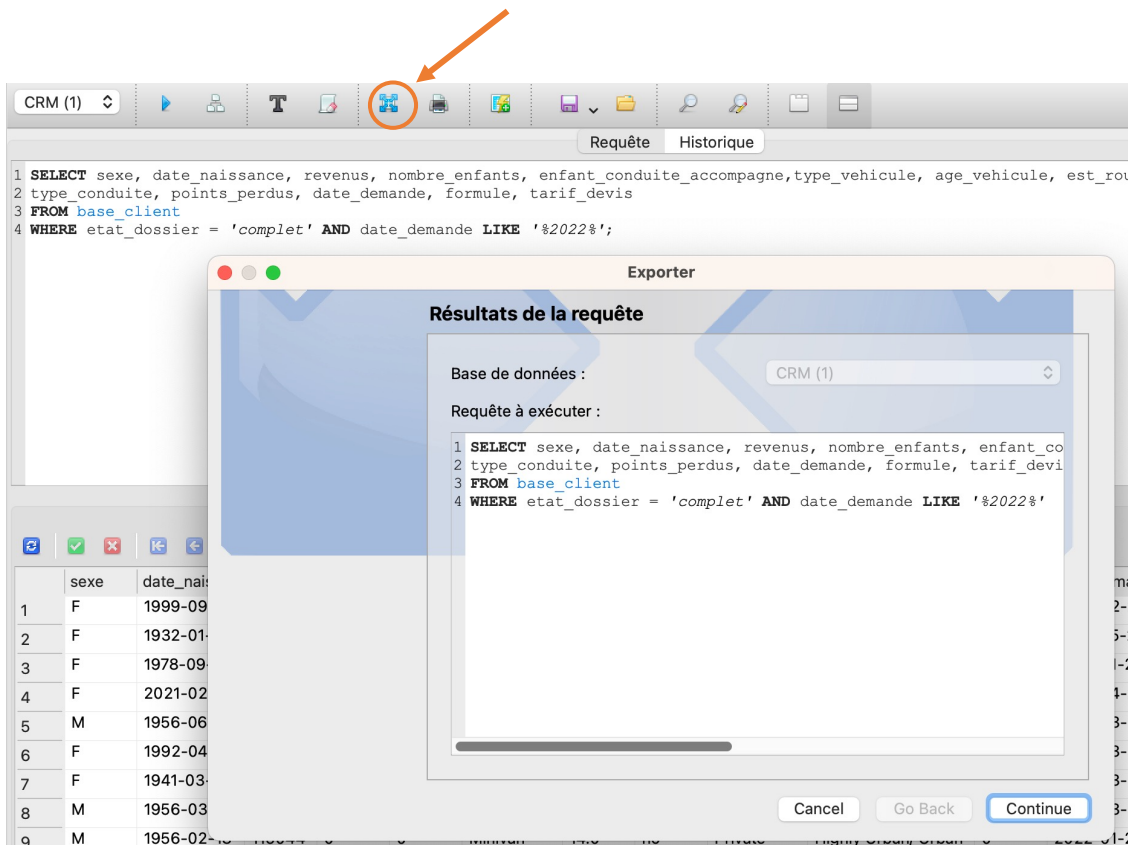
Base de données

Filtre par nom

- DB_P5 (SQLite 3)
- CRM (1) (SQLite 3)
- Tableaux (1)
- base_client
 - Colonnes (28)
 - metier
 - employeur
 - num_ss
 - groupe_sanguin
 - id_site_web
 - nom
 - sexe
 - email
 - date_naissance
 - id_client
 - enfant_conduite_accompagne
 - nombre_enfants
 - revenus
 - valeur_residence_prin
 - formation
 - usage_vehicule
 - type_vehicule
 - est_rouge
 - points_perdus
 - age_vehicule
 - type_conduite
 - date_demande
 - etat_dossier
 - formule
 - tarif_devis
 - adresse
 - lat
 - lon
 - Index
 - Déclencheurs
- Vues

Extraction des données du CRM

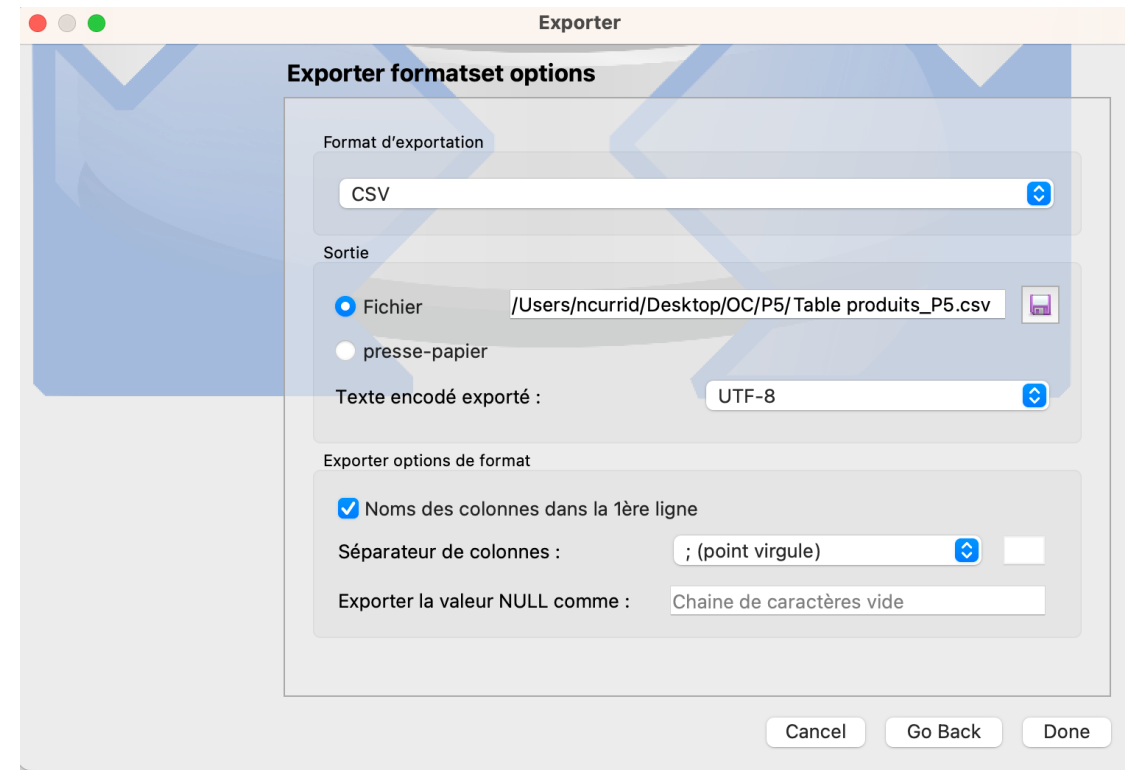
B – Export des résultats de la requête au format CSV



The screenshot shows the CRM interface with a toolbar at the top. An orange circle highlights the export icon (a document with a downward arrow) in the toolbar. Below the toolbar, a SQL query is visible in the 'Requête' tab:

```
1 SELECT sexe, date_naissance, revenus, nombre_enfants, enfant_conduite_accompagne, type_vehicule, age_vehicule, est_ro  
2 type_conduite, points_perdus, date_demande, formule, tarif_devis  
3 FROM base_client  
4 WHERE etat_dossier = 'complet' AND date_demande LIKE '%2022%';
```

An 'Exporter' dialog box is open, titled 'Résultats de la requête'. It shows the database 'CRM (1)' and the same SQL query. At the bottom of the dialog are 'Cancel', 'Go Back', and 'Continue' buttons.



The 'Exporter' dialog box is shown with the 'Exporter formatset options' section. The 'Format d'exportation' is set to 'CSV'. The 'Sortie' section has 'Fichier' selected, with the path '/Users/ncurrid/Desktop/OC/P5/ Table produits_P5.csv'. The 'Texte encodé exporté' is set to 'UTF-8'. The 'Exporter options de format' section has 'Noms des colonnes dans la 1ère ligne' checked. The 'Séparateur de colonnes' is set to '; (point virgule)'. The 'Exporter la valeur NULL comme' is set to 'Chaine de caractères vide'. At the bottom are 'Cancel', 'Go Back', and 'Done' buttons.

Etapes

« Obtenir des données (Power Query) »

« Obtenir les données » / Texte/CSV

Sélectionner le fichier CSV importé depuis SQLite

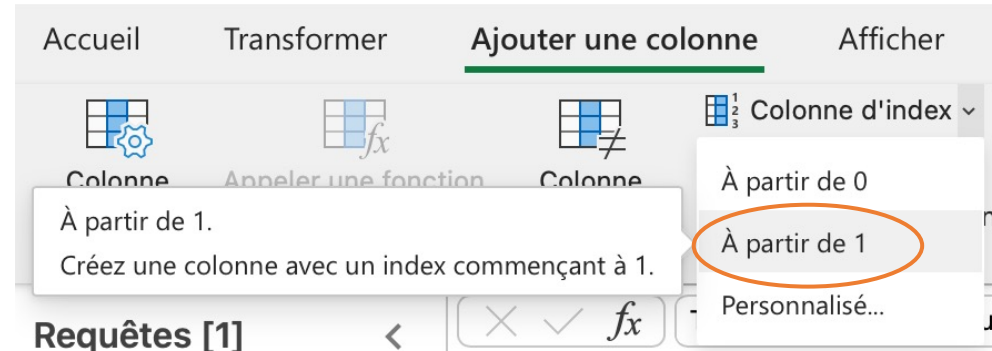
7

Traitement des données sur Power Query

B – Création d'une colonne d'index

Etapes

Cliquer sur « Transformer »
 « Ajouter une colonne »
 « Colonne d'index »
 « A partir de 1 »



Résultat

Accueil Transformer **Ajouter une colonne** Afficher Aide

Colonne personnalisée Appeler une fonction personnalisée Colonne conditionnelle Colonne d'index

Table.ReorderColumns("#Index ajouté", {"Index", "sexe", "date_naissance", "revenus", "nombre_enfants",

Requêtes [1]

	1 ² Ind...	AB se...	date_naissa...	1 ² 3 reven...	1 ² 3 nombre_enfa...	1 ² 3 enfant_conduite_accomp...	AB type_vehic...	AB age_vehicule	AB est_rou...	AB usag
1	1	F	25/09/1999	125301	1		0 Sports Car	7.0	no	Comm
2	2	F	28/01/1932	50815	1		1 Minivan	1.0	no	Private
3	3	F	15/09/1978	21204	2		0 Pickup	4.0	no	Private
4	4	F	17/02/2021	0	0		0 z_SUV	5.0	no	Private
5	5	M	14/06/1956	null	0		0 Panel Truck	12.0	no	Comm
6	6	F	11/04/1992	79606	1		0 z_SUV	15.0	no	Private
7	7	F	18/03/1941	130540	2		2 Minivan	9.0	no	Comm
8	8	M	06/03/1956	51278	0		0 z_SUV	10.0	no	Private
9	9	M	18/02/1956	115044	0		0 Minivan	14.0	no	Private
10	10	F	07/03/1974	null	0		0 Panel Truck	10.0	yes	Comm
11	11	F	13/02/1967	24060	2		1 Minivan	10.0	no	Private
12	12	M	14/05/2017	20822	4		0 z_SUV	1.0	no	Private

Paramètres de requête

Propriétés

Nom
CRM_P4

Étapes appliquées

- Source
- En-têtes promus
- Type de colonne changé
- Index ajouté
- Colonnes permutées

Etapes

- Transformer Ajouter une colonne Afficher Aide

Utiliser la première ligne pour les en-têtes

Table.ReorderColumns("#Index ajouté", {"Index ajouté", "date_naissance", "revenu", "type_voiture", "âge"})

Retourne la durée entre l'heure locale actuelle et les valeurs des colonnes sélectionnées.

Currid_CRM...	123 Ind...	123 se...	date_naissance...	123 reven...	123	123	123	123
1	1	F	25/09/1999	125301		1	0	Sports Car
2	2	F	28/01/1932	50815		1	1	Minivan
3	3	F	15/09/1978	21204		2	0	Pickup

Âge

Année

1.2 date_naiss... 123 enfant_conduite_acci

1.2 Nombre décimal

\$ Monnaie

123 Nombre entier

% Pourcentage

Date/heure

Date

Heure

Date/heure/fuseau horaire

Durée

Texte

Vrai/Faux

Binaire

Utilisation des paramètres régionaux...

86,13061882

Ajouter une colonne conditionnelle ?

Ajoutez une colonne conditionnelle calculée en fonction des autres colonnes ou valeurs.

Résultat

9

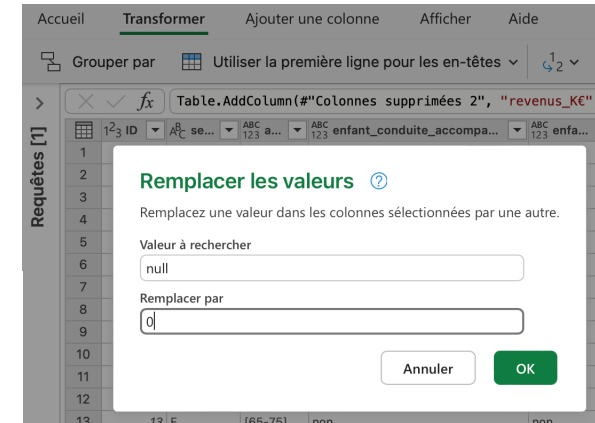
Traitement des données sur Power Query

C – Conversion des revenus en intervalles de revenus

Etapes

- 1 Transformer / Remplacer les valeurs « null » par « 0 »
- 2 Ajouter une colonne / Colonne conditionnelle

1



2 Ajouter une colonne conditionnelle ?

Ajoutez une colonne conditionnelle calculée en fonction des autres colonnes ou valeurs.

Nouveau nom de colonne *

revenus_K€

	Nom de la colonne *	Opérateur *	Valeur ⓘ	Alors	Sortie ⓘ
Si	1.2 revenus	est inférieur à	25000		< 25
Autrement si	1.2 revenus	est inférieur à	50000		[25-50]
Autrement si	1.2 revenus	est inférieur à	75000		[50-75]
Autrement si	1.2 revenus	est inférieur à	100000		[75-100]
Autrement si	1.2 revenus	est inférieur à	125000		[100-125]
Autrement si	1.2 revenus	est inférieur à	150000		[125-150]

Ajouter une clause

Else ⓘ

> 150

Annuler

OK

Traitement des données sur Power Query

D - Modification de champs en types de données logiques

Etapes

Ajouter une colonne / Colonne personnalisée

Résultat

Table.RemoveColumns(#"Colonnes permutées 4", {"enfant_conduite_accompagne..."})

	Ind...	se...	age	revenus...	enfa...	enfant_conduite_accompa...	type_vehic...	age_vehicule	est_rou...
1	1	F	[15-25]	[125-150]	oui	non	Sports Car	7.0	no
2	2	F	> 75	[50-75]	oui	oui	Minivan	1.0	no
3	3	F	[45-55]	< 25	oui	non	Pickup	4.0	no
4	4	F	< 15	< 25	non	non	z_SUV	5.0	no
5	5	M	[65-75]	< 25	non	non	Panel Truck	12.0	no
6	6	F	[25-35]	[75-100]	oui	non	z_SUV	15.0	no
7	7	F	> 75	[125-150]	oui	oui	Minivan	9.0	no
8	8	M	[65-75]	[50-75]	non	non	z_SUV	10.0	no
9	9	M	[65-75]	[100-125]	non	non	Minivan	14.0	no
10	10	F	[45-55]	< 25	non	non	Panel Truck	10.0	yes
11	11	F	[55-65]	< 25	oui	oui	Minivan	10.0	no
12	12	M	< 15	< 25	oui	non	z_SUV	1.0	no
13	13	F	[65-75]	[50-75]	non	non	z_SUV	18.0	no
14	14	M	< 15	< 25	non	non	Van	6.0	no
15	15	M	> 75	[25-50]	oui	non	Minivan	1.0	no
16	16	M	> 75	[25-50]	non	non	Minivan	13.0	no
17	17	F	[45-55]	[100-125]	non	non	z_SUV	15.0	no
18	18	F	> 75	[50-75]	non	non	Van	8.0	yes
19	19	F	[65-75]	[75-100]	non	non	Minivan	11.0	yes
20	20	F	> 75	[25-50]	oui	non	z_SUV	10.0	no
21	21	M	[35-45]	[50-75]	oui	non	Minivan	16.0	no
22	22	M	> 75	[25-50]	oui	oui	Minivan		no
23	23	F	[15-25]	< 25	non	non	Van	7.0	yes
24	24	F	> 75	[25-50]	oui	non	Sports Car	1.0	no
25	25	M	[15-25]	< 25	non	non	Pickup	1.0	no

Colonne personnalisée ?

Ajoutez une colonne calculée en fonction d'autres colonnes ou valeurs.

Nouveau nom de colonne *

enfants

Type de données

Formule de colonne personnalisée *

```
= if [nombre_enfants] = 0 then "non"
else "oui"
```

Colonne(s) disponible(s)

ID
sexe
age
enfant_conduite_accompagn...
nombre_enfants
revenus
valeur_residence_prin...

Insérer une colonne

Annuler

OK

Paramètres de requête

Propriétés

Nom

Nathalie_Currid_CRM_anonymisation

Étapes appliquées

- Source
- En-têtes promus
- Type de colonne changé
- Index ajouté
- Colonnes permutées
- Âge calculé
- Total calculé des années
- Type de colonne changé 1
- Colonne conditionnelle insérée
- Colonnes permutées 1
- Colonnes supprimées
- Valeur remplacée
- Colonne conditionnelle insérée 1
- Colonnes permutées 2
- Colonnes supprimées 1
- Colonne personnalisée ajoutée
- Colonnes permutées 3
- Colonnes supprimées 2
- Colonne personnalisée ajoutée 1
- Colonnes permutées 4
- Colonnes supprimées 3

Traitement des données sur Power Query

Conversion des tarifs des devis en intervalles de tarifs

Etapes

Ajouter une colonne / Colonne conditionnelle

Résultat

Ajouter une colonne conditionnelle

Ajoutez une colonne conditionnelle calculée en fonction des autres colonnes ou valeurs.

Nouveau nom de colonne *

tarif_devis_€

Si 1,2 tarif_devis est inférieur à 300 Alors < 300

Autrement si 1,2 tarif_devis est inférieur à 450 Alors [300-450]

Autrement si 1,2 tarif_devis est inférieur à 600 Alors [450-600] ...

Ajouter une clause

Else > 600

Annuler OK

Accueil

Transformer

Ajouter une colonne

Afficher

Aide

Fermer et charger

Obtenir les données

Entrer des données

Options

Gérer les paramètres

Actualiser

Propriétés

Éditeur avancé

Gérer

Choisir des colonnes

Supprimer les colonnes

Conservier les lignes

Supprimer des lignes

Filter les lignes

Trier

Transformer

Combiner

Exporter un modèle

Fermer

Nouvelle requête

Options

Paramètres

Requête

Gérer les colonnes

Réduire les lignes

Trier

Partager

Table.RemoveColumns(#"Colonne conditionnelle insérée 2", {"tarif_devis"})

Requêtes [1]

	ABC 123 enfant_c...	type_v...	123 age_v...	est_rou...	usage_v...	type_condu...	123 points...	date_dema...	form...	ABC 123 tarif_devis...
1	non	Sports Car	7	non	Commercial	Highly Urban/ Urban	0	février	dev_integral	[300-450]
2	oui	Minivan	1	non	Private	Highly Urban/ Urban	2	mai	dev_express	< 300
3	non	Pickup	4	non	Private	z_Highly Rural/ Rural	0	janvier	dev_express	< 300
4	non	z_SUV	5	non	Private	Highly Urban/ Urban	3	avril	dev_integral	[300-450]
5	non	Panel Truck	12	non	Commercial	Highly Urban/ Urban	2	mars	dev_vip	> 600
6	non	z_SUV	15	non	Private	z_Highly Rural/ Rural	2	mars	dev_premium	[450-600]
7	oui	Minivan	9	non	Commercial	z_Highly Rural/ Rural	0	mars	dev_express	< 300
8	non	z_SUV	10	non	Private	Highly Urban/ Urban	1	mars	dev_premium	[300-450]
9	non	Minivan	14	non	Private	Highly Urban/ Urban	0	janvier	dev_premium	[300-450]
10	non	Panel Truck	10	oui	Commercial	Highly Urban/ Urban	1	mars	dev_integral	[300-450]
11	oui	Minivan	10	non	Private	Highly Urban/ Urban	0	mai	dev_express	< 300
12	non	z_SUV	1	non	Private	Highly Urban/ Urban	0	janvier	dev_premium	[300-450]
13	non	z_SUV	18	non	Private	Highly Urban/ Urban	0	mai	dev_express	< 300
14	non	Van	6	non	Commercial	z_Highly Rural/ Rural	5	mai	dev_integral	[300-450]
15	non	Minivan	1	non	Private	Highly Urban/ Urban	0	avril	dev_integral	[300-450]
16	non	Minivan	13	non	Private	Highly Urban/ Urban	0	mai	dev_express	< 300
17	non	z_SUV	15	non	Private	Highly Urban/ Urban	0	avril	dev_integral	[300-450]
18	non	Van	8	oui	Commercial	Highly Urban/ Urban	1	mars	dev_express	< 300
19	non	Minivan	11	oui	Private	Highly Urban/ Urban	0	février	dev_express	< 300
20	non	z_SUV	10	non	Private	Highly Urban/ Urban	0	avril	dev_integral	< 300
21	non	Minivan	16	non	Private	z_Highly Rural/ Rural	0	mars	dev_express	< 300
22	oui	Minivan	null	non	Private	Highly Urban/ Urban	1	mars	dev_premium	[300-450]
23	non	Van	7	oui	Commercial	Highly Urban/ Urban	0	mai	dev_premium	[300-450]
24	non	Sports Car	1	non	Commercial	Highly Urban/ Urban	5	février	dev_premium	[450-600]
25	non	Pickup	1	non	Commercial	Highly Urban/ Urban	1	avril	dev_express	< 300
26	non	Minivan	null	oui	Private	z_Highly Rural/ Rural	1	avril	dev_express	< 300
27	non	Sports Car	1	non	Private	z_Highly Rural/ Rural	0	mars	dev_integral	< 300
28	non	z_SUV	null	non	Private	Highly Urban/ Urban	5	avril	dev_integral	< 300
29	non	z_SUV	1	non	Private	Highly Urban/ Urban	0	mars	dev_vip	> 600
30	non	z_SUV	17	non	Private	Highly Urban/ Urban	0	février	dev_premium	[300-450]
31	non	z_SUV	1	non	Private	Highly Urban/ Urban	0	janvier	dev_express	< 300

Paramètres de requête

Propriétés

Nom

Nathalie_Currid_CRM_anonymisation

Étapes appliquées

Type de colonne changé

Index ajouté

Colonnes permutees

Âge calculé

Total calculé des années

Type de colonne changé 1

Colonne conditionnelle insér...

Colonnes permutees 1

Colonnes supprimées

Valeur remplacée

Colonne conditionnelle insér...

Colonnes permutees 2

Colonnes supprimées 1

Colonne personnalisée ajout...

Colonnes permutees 3

Colonnes supprimées 2

Colonne personnalisée ajout...

Colonnes permutees 4

Colonnes supprimées 3

Valeur remplacée 2

Type de colonne changé 2

Valeur remplacée 2

Valeur remplacée 3

Nom du mois extrait

Valeur remplacée 4

Type de colonne changé 3

Colonne conditionnelle insér...

Colonnes supprimées 4

Paramètres de requête

Propriétés

Nom

Nathalie_Currid_CRM_anonymisation

Étapes appliquées

- Type de colonne changé
- Index ajouté
- Colonnes permutées
- Âge calculé
- Total calculé des années
- Type de colonne changé 1
- Colonne conditionnelle insérée...
- Colonnes permutées 1
- Colonnes supprimées
- Valeur remplacée
- Colonne conditionnelle insérée...
- Colonnes permutées 2
- Colonnes supprimées 1
- Colonne personnalisée ajoutée...
- Colonnes permutées 3
- Colonnes supprimées 2
- Colonne personnalisée ajoutée...
- Colonnes permutées 4
- Colonnes supprimées 3
- Valeur remplacée 1
- Type de colonne changé 2
- Valeur remplacée 2
- Valeur remplacée 3
- Nom du mois extrait
- Valeur remplacée 4
- Type de colonne changé 3
- Colonne conditionnelle insérée...
- Colonnes supprimées 4