**Rapport :**

Contexte métier et objectifs   
Méthodologie : CRISP-DM...   
Sources de données : Annual reports, Argus...   
Conception du DWH : Constellation par concurrent   
Processus ETL : Python + Talend + Exemples   
Dashboards avec Power BI   
Conclusion = KPI répondant aux questions   
Défis, acquis, pistes d’amélioration, perspectives

Les livrables à fournir :

§ Un rapport sur le système décisionnel développé renfermant :

• Les outils techniques utilisés dans le projet

• La conception hiérarchique de la Data Warehouse

• Une description des indicateurs et des axes d’analyse

• Des captures d’écran sur les principales interfaces du tableau de bord

Rapport de mise en situation (groupe de travail de Brahim, Charles, Loup, Rédha et Saddem)

Sommaire:

1 Organisation du travail

2 Contexte métier et objectifs

3 Méthodologie adoptée

4 Sources de données utilisées

5 Conception du Datawarehouse

6 Processus ETL

7 Conception des Dashboards

8 Conclusion

9 Défis, acquis, pistes d’améliorations, perspectives

1 Organisation du travail

A Les différentes étapes

* Le choix du sujet: Le Groupe Volkswagen, la voiture du peuple

Après quelques échanges et délibérations, l’équipe de travail s’est rendu compte qu’au niveau des sources, le sujet portant sur la firme Volkswagen semblait plutôt équilibré entre jeu de données commerciales et financières.

L’analyse des ventes du sujet a également semblé plus pertinente et en adéquation avec le métier de Data Analyst.

C’est donc ce choix qui a été fait.

* Identification des concurrents: BMW, Stellantis et Toyota

BMW est l’un des principaux concurrents de Volkswagen en Europe, mais pour apporter plus de points de comparaison au niveau des jeux de données utilisés, deux autres grands concurrents de taille identique ont été choisis, nommément Stellantis et Toyota.

* Création d’un référentiel de travail commun

Au premier jour du démarrage du projet, le groupe a opté pour la solution la plus rapidement disponible pour pouvoir mettre en commun les différentes optiques de création, à savoir Excel.

Dès le lendemain, une solution en ligne lui a été préférée: le système de partage de fichiers Trello, plus efficace dans le mise à jour en temps réel de l’avancement général d’une part, et l’échange et l’envoi de documents de travail d’autre part.

* Conception du calendrier de travail (étapes en nombre de jours avec une marge de manoeuvre pour chacune d’elle)

10 jours de travail ont été prévus pour l’élaboration de la mise en situation. Il a donc fallu construire un calendrier des tâches à prévoir et du délai accordé à chacune d’entre elles. Des étapes ont été prédéfinies à l’avance en nombre de jours en laissant, sur certaines tâches, une marge de manœuvre afin de pouvoir y revenir en cas de délai supplémentaire ou blocage particulier. Le calendrier a été détaillé un peu plus loin.

* Décision des choix techniques (applicatifs ou Python) sur chaque étape

Pour chaque étape à effectuer, il a fallu définir lesquelles d’entre elles nécessitaient de travailler en commun de celles qui exigeaient une division en petits sous-groupes de travail.

Il a fallu également définir pour chacune d’elles l’intégration à travers des lignes de scripts Python ou une orientation applicative.

Le schéma étant clairement défini pour tous, le travail de recherche documentaire et d’études des sources de données allait pouvoir commencer.

2 Contexte métier, cadre d’analyse et objectifs

Contexte métier: marché concurrentiel de l’automobile

Présentation du Groupe Volkswagen AG

Au fil du temps et face à un marché de plus en plus concurrentiel et mondialisé, le Groupe Volkswagen AG, tout comme ses concurrents, a fusionné avec d’autres groupes automobiles pour acquérir en taille et fait aujourd’hui partie des plus grosses firmes automobiles mondiales.

On peut citer parmi ses plus grandes marques l’allemand Audi, rachetée en 1964, puis l’américain Chrysler en 1979, l’espagnol SEAT en 1986 et le tchèque Skoda en 1991. Volkswagen décide ensuite de se diversifier en acquérant des véhicules de luxe: rachat du français Bugatti en 1998, puis de l’italien Lamborghini via sa filiale Audi la même année, pour finir par l’anglais Bentley en 1999.

Puis, la gamme de voitures sportives s’est étendue en 2012 avec la fusion entre Volkswagen et Porsche.

Finalement, toujours dans une stratégie de diversification, le Groupe acquiert les bus et camions MAN en 2011, et les camions du suédois Scania en 2014.

La présence en Chine a été renforcée en 2018 en lançant la création de Sol, la marque chinoise de véhicules électriques et autonomes destinée au marché local, dans le but de développer la voiture connectée du futur.

Présentation succinctes des concurrents:

BMW

La multinationale BMW comporte en plus de la filiale allemande principale les marques anglaises Rolls-Royce et Mini. Toutes trois sont spécialisées dans les véhicules haut de gamme et de luxe.

Stellantis

La mega-firme Stellantis est née de la fusion en 2021 de deux géants européens de la construction automobile: PSA et FCA.

Parmi les grandes marques de PSA: les français Peugeot et Citroën, l’allemand Opel et l’anglais Vauxhall.

Parmi celles de FCA: les italiens Abarth, Alfa Romeo, Fiat, Lancia, Maserati additionnés des américains Chrysler, Dodge, Jeep, Maserati et Ram.

Objectifs de la mise en situation:

* comprendre la place du groupe Volkswagen sur le marché ultra-concurrentiel et mondialisé de l’automobile
* saisir le positionnement en termes de gamme de voiture et la stratégie mise en place sur les différents marchés en fonction de leur zone géographique
* analyser la répartition des ventes des différents modèles par type d’énergie, et apprécier l’évolution de l’hybride et de l’électrique sur les années récentes
* observer le développement du groupe sur un nouveau marché émergent de grande ampleur: la Chine

3 Méthodologie adoptées:

10 jours de travail ont été prévus pour l’élaboration de la mise en situation. Il a donc fallu construire un calendrier des tâches à prévoir et du délai accordé à chacune d’entre elles.

La méthodologie choisie pour effectuer le projet est le CRISP-DM, avec la possibilité de travailler en agilité sur les étapes grâce à l’outil de gestion de projet en ligne Trello.

Étapes du CRISP-DM et nombre de jours associés:

1 Business Understanding (1 journée)

2 Data Understanding (2 à 3 jours)

3 Staging Area (1 journée): Dû au faible nombre de données, il nous a semblé judicieux de passer directement à l’étape suivante.

4 ETL/Datawarehouse (1 journée)

5 Modelling (2 à 3 jours)

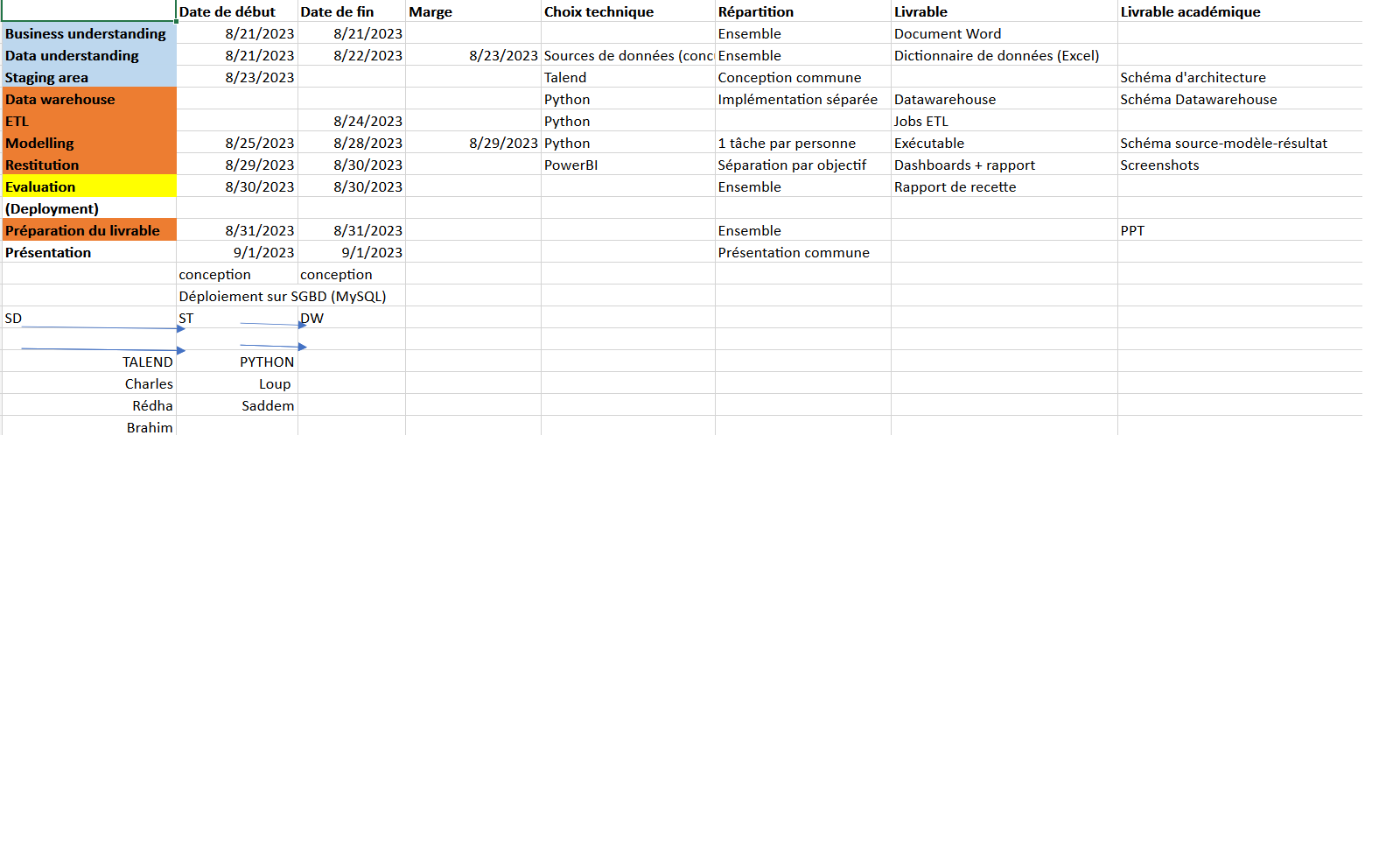
6 Restitution et dashboards (2 jours)

7 (Évaluation et déploiement)

8 Préparation des livrables: rapport et présentation Powerpoint (1 journée)

10 Présentation (dernier jour)

Visuel du calendrier de travail:



4 Sources de données utilisées

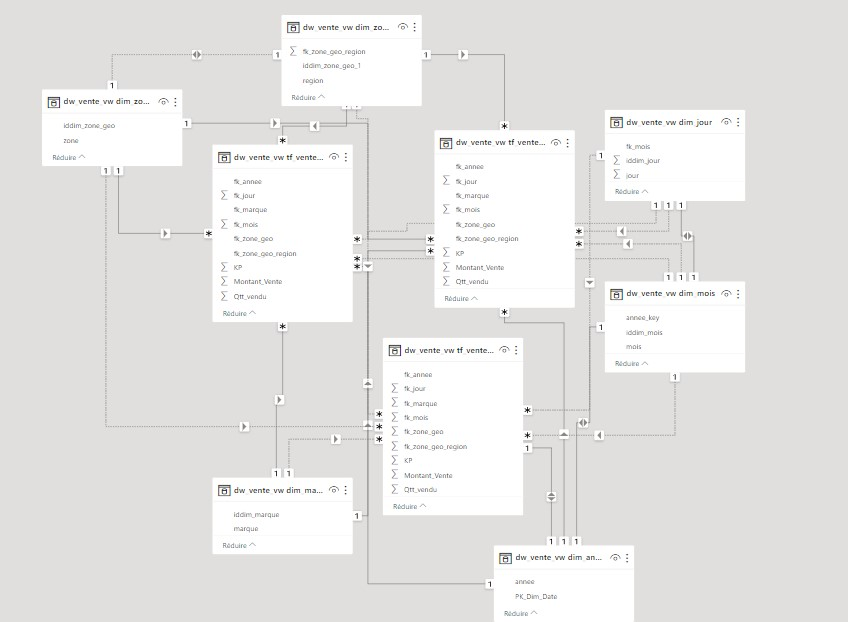
Rapport annuels de Volkswagen et également des différents concurrents (BMW, Stellantis et Toyota) sur l’année 2022.

Le site Internet Largus.fr pour aller chercher les gammes de prix des différents modèles pour chacune des marques.

Voir en annexe le dictionnaire de données.

5 Conception du Data Warehouse

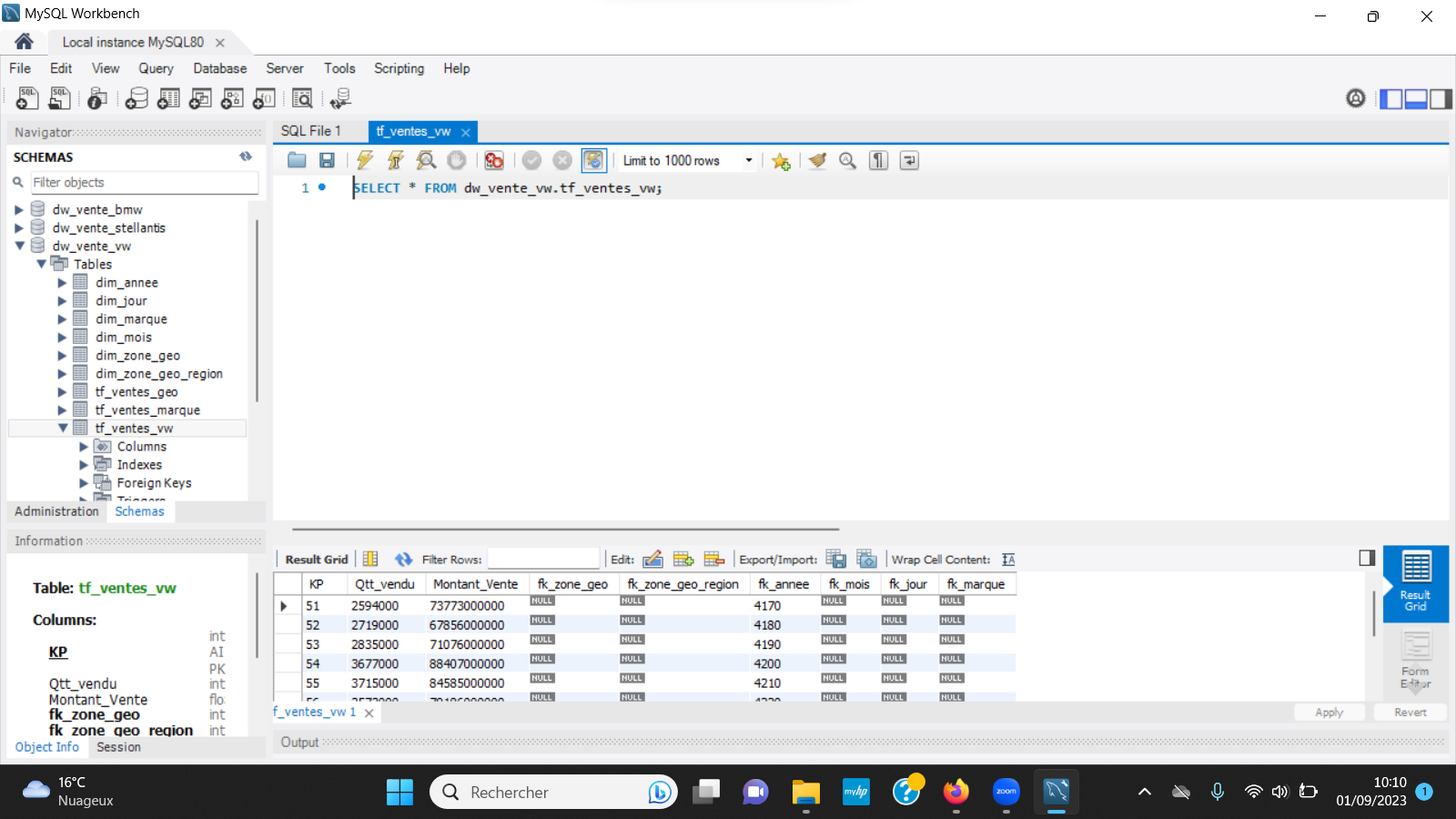
Schéma en forme de constellation par concurrent



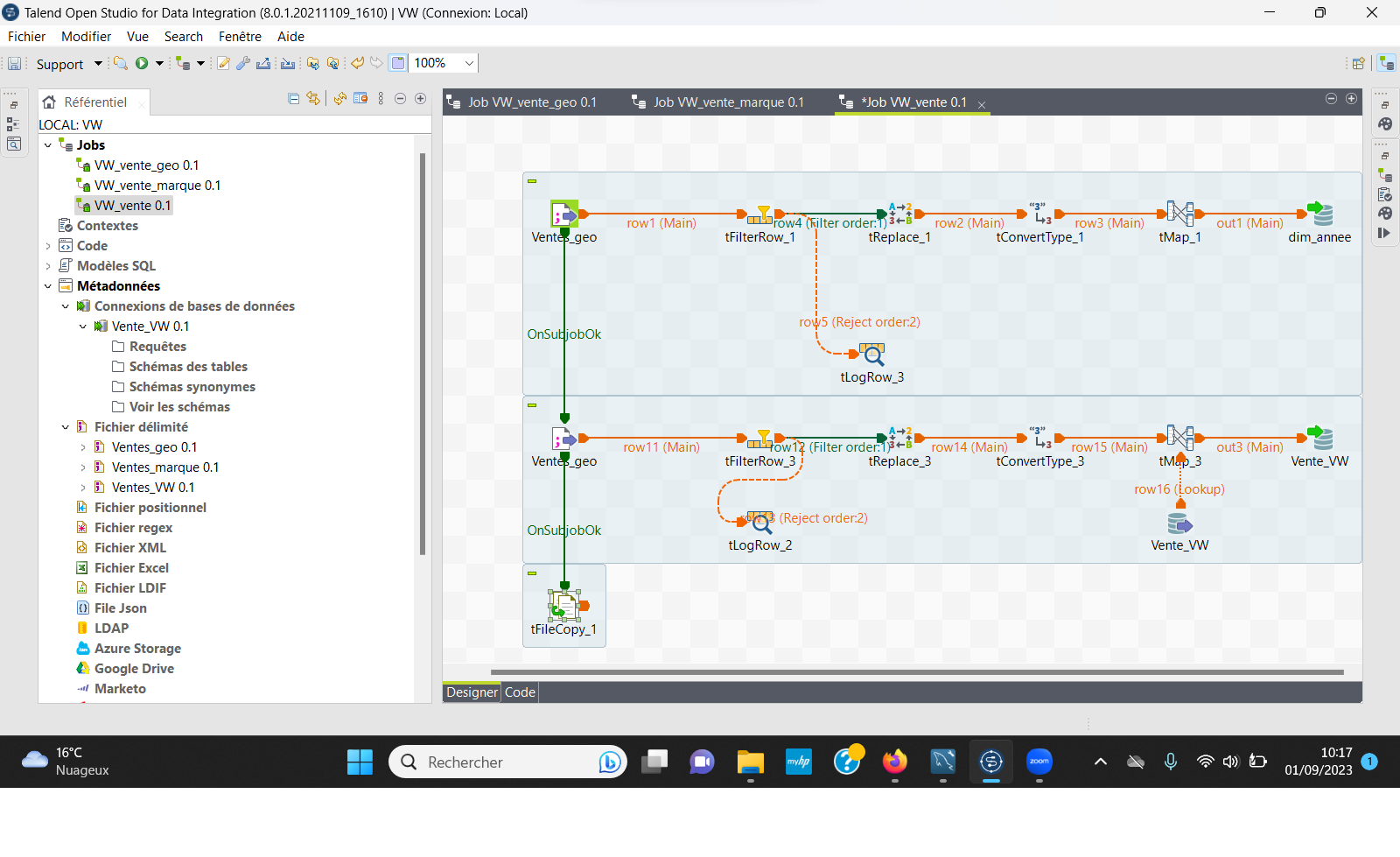
6 Processus ETL

Les outils techniques : MySQL Workbench 8.0, Talend 8.0.1 et Python 3.11.4

Bases de travail et tables sous MySQL Workbench



3 Jobs créés sous Talend:



Le choix de l’ETL s’est porté à la fois sur Python et Talend pour atteindre les résultats escomptés en variant les technologies utilisées.

Les outils MySQL Benchwork et Talend en version d’essai n’offrant pas la possibilité de centraliser la base de données, il a semblé judicieux de segmenter les tâches en deux sous-groupes et fusionner par la suite .

- l’équipe Talend, responsable des données principales de Volkswagen.

- et l’équipe Python, chargé des données des concurrents BMW et Stellantis.

Par contrainte de temps, le jeu de données de Toyota n’a pas été chargé. Une fois le Datawarehouse conçu et développé à l’aide MySQL Workbench, les scripts SQL de création des tables ont été exportés et échangés afin que tous puissent reprendre une base de travail identique pour effectuer la modélisation et la création des dashboards.

Ci-après un tableau récapitulatif des jobs Talend :

| Source de données | Principaux traitements | Destination |
| --- | --- | --- |
| Fichier des ventes par années et par zone géographique de VW | * remplacement des valeur vide en NA pour les identifier * conversion de type numériques (quantité , montant des ventes en) en type numérique et entier * corriger les virgules * filtrer les données inutiles (Edge on ..) | Table de dimensions : année , zone géographique  Table de fait : vente par zone |
| Fichier des ventes par marque et par année | * remplacement des valeur vide en NA pour les identifier * conversion de type numériques (quantité , montant des ventes en) en type numérique et entier * filtrer les données inutile (Other ..) | Table de dimensions : année ,marque  Table de fait : vente par marque |
| Fichier des ventes par année | * remplacement des valeur vide en NA pour les identifier * conversion de type numériques (quantité , montant des ventes en) en type numérique et entier * filtrer les données inutile (Other ..) identifier * conversion de type numériques (quantité , montant des ventes en) en type numérique et entier * filtrer les données inutiles (Edge on ..) | Table de dimensions : année  Table de fait : vente totale |

Tableau récapitulatif des traitements Python :

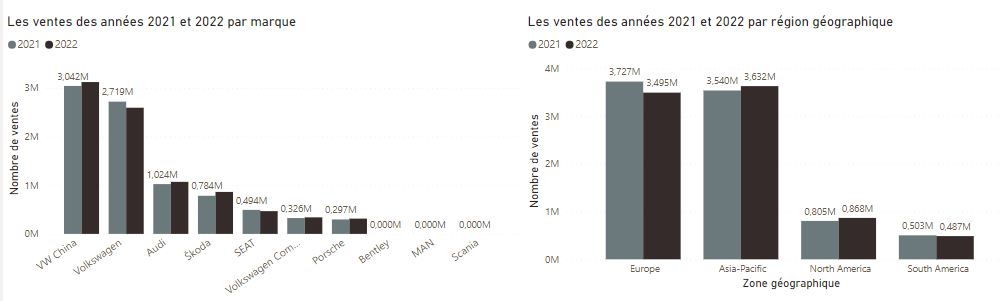
| Source de données | Principaux traitements | Destination |
| --- | --- | --- |
| Fichier des ventes par années, zone géographique et marque de BMW | - Charger les données dans des dataframes  - Transformer les dataframes large en un format long  - Enlever les caractères inutiles et conversion de type | Table de dimensions : année , zone géographique, marque  Table de fait : vente par zone, vente par marque |
| Fichier des ventes par années, zone géographique et marque de Stellantis | - Charger les données dans des dataframes  - Transformer les dataframes large en un format long  - Enlever les caractères inutiles et conversion de type | Table de dimensions : année , zone géographique, marque  Table de fait : vente par zone, vente par marque |
| Fichier des prix des modèles BMW et Stellantis | - Charger les données dans des dataframes  - Transformer les dataframes large en un format long  - Enlever les caractères inutiles et conversion de type | Table référentiel de prix BMW et Stellantis |

7 Conception Dashboards avec Power BI

Grâce à la modélisation, on a pu identifier des KPI, permettant de répondre aux questions suivantes:

1 Quelle est la quantité commercialisée des marques GVW par rapport à la même

période de l'année précédente par zone géographique ?



2 Quels sont les prix des marques par rapport à ceux des concurrents ?

Comparaison des prix moyens par marque et gamme

Au niveau des marques et gamme, BMW ressort avec le prix moyen le plus élevé à 98 200 € suivi du groupe Volkswagen à 79 500 € et enfin Stellantis avec 47 800 € de prix moyen.

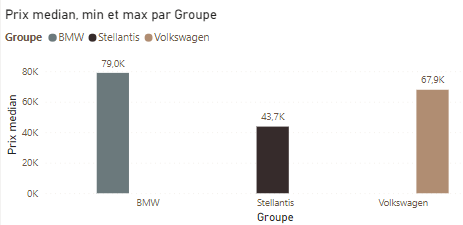
Sur le segment des versions luxe, c’est la gamme BMW qui fait ressortir le prix moyen le plus élevé à près de 238 485 € suivi de Volkswagen à 182 355 € et Sellantis à 171 894 €.

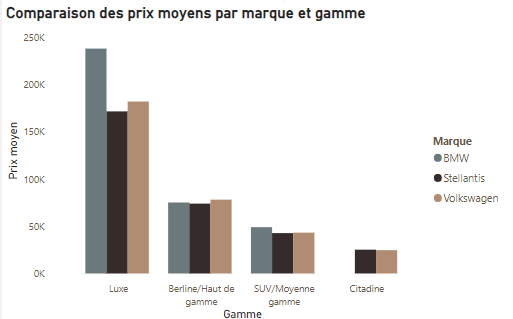
Prix médian, prix min, prix max

Pour les prix min, Stellantis dispose du plus petit prix ressortant à 12 300 €, suivi de Volkswagen avec la Skoda à 16615 € et enfin BMW dont le prix min de la Mini ressort à 30400 €, ce qui en fait le prix min les plus élevé des 3.

Pour les prix Max, la Rolls-Royce du Groupe BMW ressort avec le montant le plus élevé à 558 000 € suivi d’un montant de 432 000 € attribuable à Audi dans le Groupe Volkswagen et enfin veint Stellantis avec un prix Max ressortant à 235200 €.

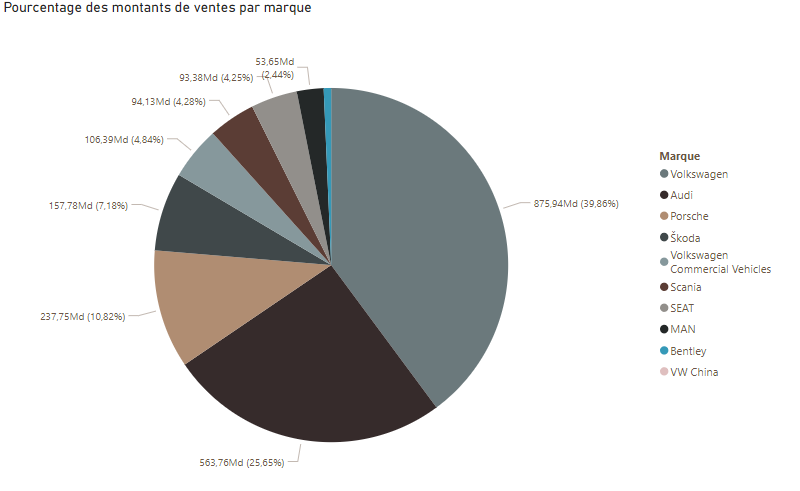
Pour les prix médian, le Groupe BMW est celui faisant ressortir le montant le plus élevé à 98 200 € , vient ensuite le groupe VOLKSWAGEN avec un montant de 79 521 € en enfin Stellantis à 47819 €



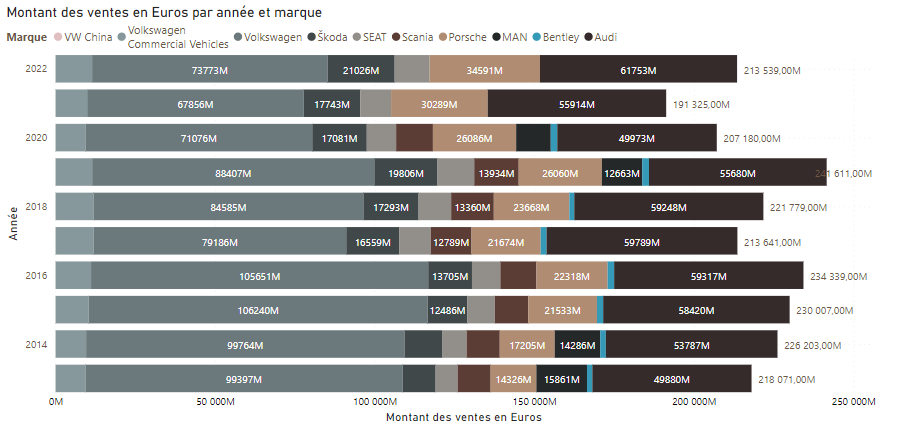


3 Quelle est la marque/modèle qui se vend le mieux ?

La marque qui se vend le mieux parmi les voitures du groupe Volkswagen sont les Volkswagen elles-mêmes (nommées Passenger Cars sur le camembert ci-dessous) pour une grosse part de marché (39.86%) suivi de Audi (25.65%) et de Porsche (10.82%).



et ceci de 2012 à 2022 :



4 Quel est le pourcentage des ventes d'un modèle par rapport au total des ventes pour

l'année actuelle ?

Pourcentage des ventes des marques dans le total vente

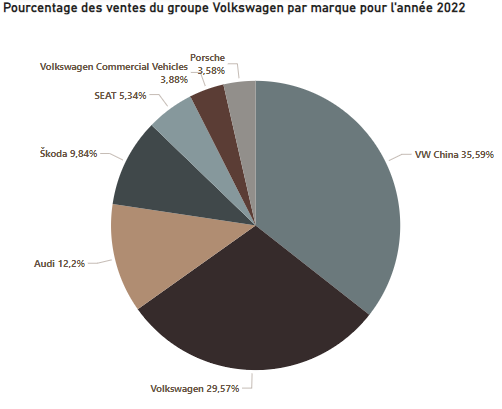
Dans le Groupe Volkswagen en 2022, la marque Volkswagen ressort avec le plus de quantité vendue (2.594 millions sur un total groupe de 5.649 millions de voitures vendues) et il possède également les revenus les plus élevés avec des ventes se chiffrant à 74 milliards sur un total Groupe de 213.5 milliards.

En termes de pourcentage cela représente respectivement 46% et 34.5% du total du Groupe.

Ensuite suit Audi avec des quantités vendues de 1 .070 million de voitures et un chiffre d’affaires de près de 62 milliards pour l’année 2022.

Les ratios sont respectivement de près de 19% et 29% pour ces deux montants.

Il est à noter que Porsche quand bien-même ne représentant que 5.56% des quantités vendues contribue à hauteur de 16.2% au chiffre d’affaires total du groupe.



5 Quel est le niveau des ventes réelles pour cette période par rapport aux ventes

planifiées ?

Écart vente palifiées –ventes réelles

Les ventes 2022 sont ressorties à 8.3 millions de véhicules vendues contre des prévisions 9,57 millions et des prévisions ajustées de 8.9 millions.

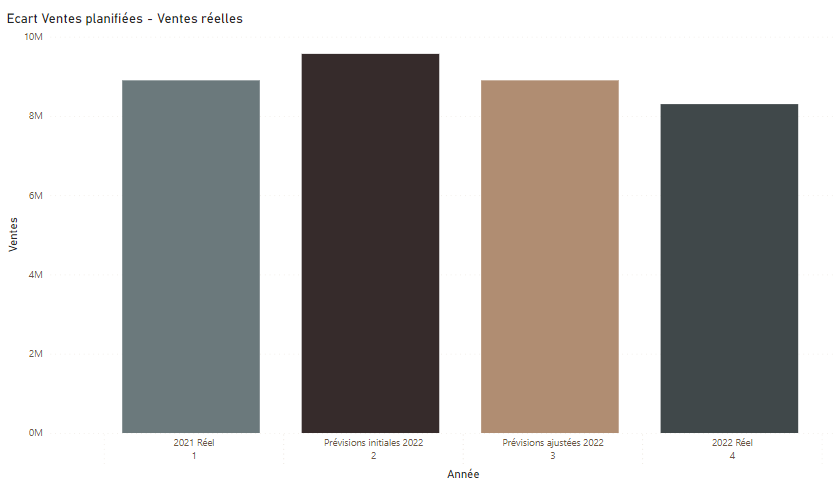
Par rapport aux prévisions ajustées, cela représente une baisse de 6.75%.

En compassion annuelle les ventes de 2022 par rapport à 2021 ont également connu une baisse de 6.75% passant de 8.9 à 8.3 millions de vente.

Question 3

Au fil des années 2013 à 2022, le modèle qui se vend le mieux est le Volkswagen Passenger car avec des ventes record de 105.65 milliards en 2016 et 103.24 milliards en 2015.

A partir de 2017, ces chiffres ont commencé à baissé compte tenu de la baisse global des ventes du groupe qui sont passées de 230 milliards en 2015 à 191 milliards en 2021. Ainsi en 2021 et 2022 le Volkswage



8 Conclusion

Question 2 : Comparaison des prix moyens par marque et gamme

Au niveau des marques et gamme, BMW ressort avec le prix moyen le plus élevé à 98 200 € suivi du groupe Volkswagen à 79 500 € et enfin Stellantis avec 47 800 € de prix moyen.

Sur le segment des versions luxe, c’est la gamme BMW qui fait ressortir le prix moyen le plus élevé à près de 238 485 € suivi de Volkswagen à 182 355 € et Sellantis à 171 894 €.

Prix médian, prix min, prix max

Pour les prix min, Stellantis dispose du plus petit prix ressortant à 12 300 €, suivi de Volkswagen avec la Skoda à 16615 € et enfin BMW dont le prix min de la Mini ressort à 30400 €, ce qui en fait le prix min les plus élevé des 3.

Pour les prix Max, la Rolls-Royce du Groupe BMW ressort avec le montant le plus élevé à 558 000 € suivi d’un montant de 432 000 € attribuable à Audi dans le Groupe Volkswagen et enfin vient Stellantis avec un prix Max ressortant à 235200 €.

Pour les prix médian, le Groupe BMW est celui faisant ressortir le montant le plus élevé à 98 200 € , vient ensuite le groupe VOLKSWAGEN avec un montant de 79 521 € en enfin Stellantis à 47819 €.

Question 4 : Pourcentage des ventes des marques dans le total vente

Dans le Groupe Volkswagen en 2022, la marque Volkswagen ressort avec le plus de quantité vendue (2.594 millions sur un total groupe de 5.649 millions de voitures vendues) et il possède également les revenus les plus élevés avec des ventes se chiffrant à 74 milliards sur un total Groupe de 213.5 milliards.

En termes de pourcentage, cela représente respectivement 46% et 34.5% du total du Groupe.

Ensuite suit Audi avec des quantités vendues de 1 .070 million de voitures et un chiffre d’affaires de près de 62 milliards pour l’année 2022.

Les ratios sont respectivement de près de 19% et 29% pour ces deux montants.

Il est à noter que Porsche quand bien même ne représentant que 5.56% des quantités vendues contribue à hauteur de 16.2% au chiffre d’affaires total du groupe.

Question 5

Écart vente planifiées –ventes réelles

Les ventes de 2022 sont ressorties à 8.3 millions de véhicules vendues contre des prévisions 9,57 millions et des prévisions ajustées de 8.9 millions.

Par rapport aux prévisions ajustées, cela représente une baisse de 6.75%.

En compassion annuelle, les ventes de 2022 par rapport à 2021 ont également connu une baisse de 6.75% passant de 8.9 à 8.3 millions de ventes.

Question 3

Au fil des années 2013 à 2022, le modèle qui se vend le mieux est le Volkswagen Passenger car avec des ventes record de 105.65 milliards en 2016 et 103.24 milliards en 2015.

A partir de 2017, ces chiffres ont commencé à baisser compte tenu de la baisse globale des ventes du groupe qui sont passées de 230 milliards en 2015 à 191 milliards en 2021. Ainsi en 2021 et 2022 le Volkswagen Passenger Cars représentait respectivement 68 milliards et 73.77 milliards de revenus.

Question 1

9 Bilan: Défis, acquis, pistes d’améliorations, perspectives

* bilan: respect des délais ?

Grâce aux différents échanges au sein de l’équipe, le défi relevé était de pouvoir:

* trouver sa place au sein de l’équipe et travailler de manière DevOps: ambiance de travail studieuse avec un désir commun d’avancer dans la bonne humeur, échanges réguliers, identification des points de blocages et résolution rapide afin de tous avancer en même temps, respect des pauses et des horaires fixées, point d’avancement de chacun, propositions d’axes d’amélioration dans l’écoute et décision collégiale à la proportionnelle (main levée ou accord tacite), partage des connaissances pour faire avancer l’équipe !
* identifier les difficultés liés au projet

Pistes d’amélioration:

-identification des limites de nos données et de leur adéquation avec l’objectif final: sources manquantes pour certains concurrents, données stratégiques manquantes (données des ventes par marque), granularité insuffisante

Défis:

* gérer au fur et à mesure du projet les lacunes du jeu de données plus de données dans le jeu de données !!!
* Obtenir un Datawarehouse fonctionnel pour la modélisation
* S’organiser à distance à cinq
* Gérer le niveau d’expérience de chacun sur les outils et partager

Acquis:

* S’améliorer sur la technique, les outils, le travail de groupe, le travail d’analyse et de synthèse, le cadre de la construction d’un projet collaboratif

Perspectives:

* Obtenir un niveau d’exigence professionnel
* Être opérationnel pour un futur employeur

9 Documents annexes

| Fichier | Champs | Définition |
| --- | --- | --- |
| Toyota financial summary | Consolidated Vehicle Sales | Ventes consolidées de véhicules en milliers de véhicules |
|  | Yen to U.S. Dollar Rate | Taux de change dollar yen |
|  | Yen to Euro Rate | Taux de change dollar Euro |
|  | Net Revenues / Sales Revenues | Chiffre d'affaire en milliards de Yen |
|  | Net Income (Loss)\*1 | Résultat net en milliards de Yen |
|  | R&D Expenses | Dépenses R&D |
|  | Return on Equity | Résultats nets sur capitaux propres |
|  | Return on Assets ROA | Résultats nets sur total actifs |
|  |  |  |
| Historique-VW/div-vw-pc-key-figures | Deliveries (thousand units) | Livraisons en milliers d'unité |
|  | Vehicle sales | Ventes en milliers d'unité |
|  | Production | Ventes en milliers d'unité |
|  | Sales revenue (€ million) | Ventes en millions d'euros |
|  | Operating return on sales (%) | Résultat d'opération |
|  |  |  |
| Historique-VW/div-figures-market | Vehicule sales | Quantité vendue en milliers d'unité |
|  | SALES REVENUE | Ventes en millions d'Euros |
|  |  |  |
| Historique-VW/figures\_brand\_field | Vehicule sales | Quantité vendue en milliers d'unité |
|  | SALES REVENUE | Ventes en millions d'Euros |
|  | Marque |  |
| Classeur 2(4) / CA Stellantis | Chiffre d'affaires Stellantis (millions d'Euros) |  |
|  | Chiffre d'affaires consolidé PSA (historique) (millions d'Euros) |  |
|  |  |  |
|  | Chiffres d'affaires CNHI, Iveco, Ferarri N.V (millions d'Euros) |  |
|  | CNHI (machines agricoles) |  |
|  | Iveco (utilitaires, poids lourds) |  |
|  | Ferarri N.V |  |
|  |  |  |
| Classeur 2(4) / Unités de vente Stellantis | Unités de ventes par zone géographique (milliers d'unités) |  |
|  | Amérique du Nord |  |
|  | Unités de Maserati vendues |  |
|  |  |  |
| Classeur 2(4) / Parts de marché par zone géographique | Parts de marché par zone géographique |  |
|  |  |  |
|  |  |  |
| KPI | Employees at year-end |  |
|  | Share of women in management positions in the BMW Group (in %) |  |
|  | Deliveries 5 |  |
|  | Share of electrified vehicles in deliveries (in %) |  |
|  | CO2 emissions of the EU new vehicle fleet (in g/km) | Émissions de CO2 de la flotte UE g/km |
|  | CO2 emissions per vehicle produced (in tonnes) | Émissions de CO2 de véhicules produit en tonnes |
|  |  |  |
|  | EBIT margin in % | Résultat avant impôts |
|  | RoCE in % 13 | EBIT / (total actif - total passif) |
|  | Deliveries to customers | Livraisons au clients |
|  |  |  |
|  | RoE in % | Résultat net / capitaux propres |
|  |  |  |
| Ventes par marque |  |  |
|  | AUTOMOTIVE SEGMENT |  |
|  | Deliveries by brand | Livraison par marque en millions |
|  | Production by brand | Productions par marques ( |
|  | MOTORCYCLES SEGMENT |  |
|  | Production |  |
|  | FINANCIAL SERVICES SEGMENT |  |
|  |  |  |
| BMW report 2022 |  |  |
| Ventes par zone géo | Ventes par milliers d'unité |  |
|  |  |  |
| Ventes modèle BMW | in units | Ventes par unité |
| Ventes modèle Mini | in units | Ventes par unité |
| Ventes modèle Rolls | in units | Ventes par unité |
|  |  |  |
| Revenues |  |  |
|  | Sales of products, related goods and revenue of service contracts | Ventes des véhicules |
|  | Revenues related to financial services | Revenu des activités financières |
|  |  |  |
| Ventes Toyota |  |  |
|  | Zone géo / series | unité par marques |
|  | Zone géo / spower train | unités par force motrice |
|  |  |  |
|  |  |  |
| Prix BMW |  |  |
|  |  |  |
|  | Energie | Modèle |
|  |  |  |
|  |  |  |
| CA Stellantis |  |  |
|  | Chiffre d'affaires Stellantis (millions d'Euros) |  |
|  | Chiffre d'affaires consolidé PSA (historique) (millions d'Euros) |  |
|  |  |  |
|  | Chiffres d'affaires CNHI, Iveco, Ferarri N.V (millions d'Euros) |  |
|  | CNHI (machines agricoles) |  |
|  | Iveco (utilitaires, poids lourds) |  |
|  | Ferarri N.V |  |
|  |  |  |
| Unités de vente Stellantis | Unités de ventes par zone géographique (milliers d'unités) |  |
|  | Unités de Maserati vendues |  |
|  |  |  |
| Parts de marché Stellantis | Part de marché par Zone géographique et par année |  |