



Master 2 Informatique

Filière : Systèmes d'Information et Aide à la Décision

Mise en place d'un système décisionnel pour l'étude du marché des véhicules électriques

Réalisé par :

Anas SAAOUDI

Othmane BENHAKKI

Yahya EL FARCI

Yacine IFRADEN

Sous l'encadrement de :

Pr. Engelbert Mephu Nguifo

Année universitaire : 2023-2024

Résumé

Le rapport présent englobe l'ensemble de notre travail réalisé sur une période de deux mois. Le défi posé par ce projet consistait à mettre en place un système d'information d'aide à la décision (BI) pour l'analyse du marché des véhicules électriques.

L'objectif principal de cette étude est double : d'une part, fournir à l'entreprise des insights précis sur l'évolution des ventes sur le marché et des statistiques concernant la répartition des stations de recharge, et d'autre part, établir des prédictions sur les ventes dans les années à venir ainsi que des suggestions quant aux emplacements potentiels pour de nouvelles stations de recharge.

Ce rapport offre une vue d'ensemble exhaustive du projet, organisée en cinq chapitres qui couvrent les différentes étapes méthodologiques de notre démarche, depuis la contextualisation générale du projet jusqu'à la mise en œuvre de la solution.

Mots-clés :

- Système décisionnel ;
- Véhicule électrique ;
- Ventes ;
- Évolution ;
- Prévision

Abstract

The following report encompasses all of our work carried out over a period of 2 months. The challenge behind this project was to establish a Business Intelligence (BI) system for the analysis of the electric vehicle market.

The main objective of this study is twofold : firstly, to provide the company with precise insights into market sales trends and statistics on the distribution of charging stations, and secondly, to make predictions about future sales and provide suggestions on potential locations for charging stations.

This report provides a comprehensive overview of the project, structured around five chapters covering the methodological aspects of our approach, from establishing the general project context to implementing the solution.

Keywords :

- Decision Support System ;
- Electrical vehicle ;
- Sales ;
- Evolution ;
- Forecasting ;

Liste d'abréviations

Abréviation	Signification
BI	Business Intelligence
CRISP-DM	Cross-Industry Standard Process for Data Mining
DWH	Data Warehouse
ETL	Extract Transform Load
FDA	File Deposit Area
MCD	Modèle Conceptuel de Données
SQL	Structured Query Language

Table des figures

2.1	Tableau Kanban	15
2.2	Répartition des tâches	16
2.3	Diagramme de Gantt	16
2.4	Liste des activités	17
4.1	Architecture golbal du système	27
4.2	Modélisation du DWH	30
4.3	Différentes étapes du framework CRISP-DM	33
5.1	Les différents attributs des fichiers de ventes	36
5.2	Les différents attributs des fichiers d'immatriculations	36
5.3	Quelques attributs des fichiers de stations de recharge	36
5.4	Outils utilisé pour faire l'ETL	37
5.5	Exemple : fusion de la table d'immatriculations au niveau de Montana avec l'index des zip code.	38
5.6	Logo de l'outil utilisé pour élaborer les dashboards	38
5.7	Premier Dashboard : ventes et immatriculations (1)	39
5.8	Premier Dashboard : ventes et immatriculations (2)	39
5.9	Deuxième Dashboard part 1 : distribution de station de recharge	40
5.10	Deuxième Dashboard (2))	40
5.11	Troisième Dashboard : Emission des EVs	41
5.12	Courbe de prédiction des ventes potentielles des véhicules électriques dans les années 2024 et 2025	42
5.13	Résultat depuis VS Code du Prédiction géographique des futures installa- tions des ports de charges EVSE	43
5.14	Prédiction géographique des futures installations des ports de charges EVSE selon le ratio du ventes VE et le nombre des EVSE	43

Liste des tableaux

2.1	Exemples des principaux risques identifiés	18
3.1	Matrice dimensionnelle	24
4.1	Fiche de conception du premier dashboard (Ventes)	31
4.2	Fiche de conception du deuxième dashboard (stations de recharge)	32
4.3	Fiche de conception du troisieme dashboard (Emission CO2)	32

Table des matières

Introduction	10
1 Contexte général du projet	11
1.1 Cadre du Travail	12
1.2 Problématique et motivation du projet	12
1.3 Objectifs du projet	13
2 Conduite du Projet	14
2.1 Méthodologie de gestion du projet	15
2.2 Planification du projet	15
2.3 Diagramme de Gantt	16
2.4 Management des délais du projet	17
2.5 Management des risques	18
3 Analyse et Spécification	20
3.1 Périmètre du projet	21
3.2 Les axes d'analyse	21
3.3 Les indicateurs de performance	22
3.4 Matrice dimensionnelle	24
4 Conception de la solution	26
4.1 Architecture générale du système	27
4.2 Modèle conceptuel de données	27
4.3 Conception du Tableau de bord	30
4.3.1 Fiche de conception	30
4.4 Fouille de données : CRISP-DM	33
5 Mise en oeuvre de la solution	35
5.1 Source de données	36
5.2 Réalisation ETL	36
5.2.1 Choix d'outils	37
5.2.2 Étapes de transformation	37

TABLE DES MATIÈRES

5.3	Création des Tableaux de bord	38
5.3.1	Choix d'outil	38
5.3.2	Dashboards	39
5.4	Data Mining	42
5.4.1	Prévision temporelle des ventes	42
5.4.2	Suggestion géographique	43
	Conclusion Générale	45
	Bibliographie	46

Introduction

Le marché des véhicules électriques connaît une croissance fulgurante ces dernières années. Cette croissance est alimentée par un certain nombre de facteurs, notamment l'augmentation des préoccupations environnementales, la baisse du prix des batteries et l'expansion des infrastructures de recharge. Dans ce contexte dynamique, les entreprises du secteur automobile doivent s'adapter rapidement pour saisir les opportunités et relever les défis de cette nouvelle ère de la mobilité.

C'est dans ce contexte que notre équipe d'analystes a été mandatée par une entreprise spécialisée dans les véhicules électriques pour réaliser une étude de marché approfondie. L'objectif de cette étude est de fournir à l'entreprise des insights précis et exploitables sur les tendances du marché.

Ce rapport présente une vue d'ensemble complète du projet, articulée autour de cinq chapitres qui couvrent différentes facettes de notre démarche méthodologique. Nous commençons par établir le "Contexte général du projet", où nous exposons les défis et les opportunités du marché des véhicules électriques, ainsi que les objectifs et la motivation qui sous-tendent notre projet. Ensuite, nous décrivons la "Conduite du Projet", mettant en lumière la méthodologie, la planification et la gestion des risques qui ont guidé notre démarche.

Le troisième chapitre, "Analyse et Spécification", nous plonge dans le processus d'identification des axes d'analyse pertinents et des indicateurs de performance clés pour notre étude de marché. Nous détaillons également la création d'une matrice dimensionnelle pour structurer nos analyses. Dans le "Chapitre 4 : Conception de la solution", nous passons en revue l'architecture du système, la modélisation des données et la conception des tableaux de bord, éléments essentiels pour transformer les données en informations exploitables.

Enfin, le cinquième chapitre, "Mise en œuvre de la solution", nous plonge dans l'aspect pratique de notre projet, détaillant la collecte des données, le processus ETL, la création des tableaux de bord et les analyses de Data Mining réalisées.

À travers ce rapport, nous visons à fournir à notre client une compréhension approfondie du marché des véhicules électriques et à l'aider à élaborer des stratégies efficaces pour rester compétitif dans ce secteur en pleine évolution.

CONTEXTE GÉNÉRAL DU PROJET

Ce chapitre vise à établir le contexte général du projet "Mise en place d'un système décisionnel pour l'étude du marché des véhicules électriques". Dans le cadre d'une étude de cas décisionnelle, notre équipe de data analysts se place dans le rôle d'experts chargés d'analyser le marché des véhicules électriques pour une entreprise spécialisée dans ce domaine. Ce chapitre présentera d'abord le cadre de travail du projet, ses domaines d'expertise et ses clients potentiels. Ensuite, nous aborderons la problématique et les objectifs spécifiques qui guideront notre travail dans cette étude de marché dynamique et en évolution constante.

1.1 Cadre du Travail

Ce projet s'inscrit dans le cadre d'une unité d'enseignement intitulée "Étude de Cas Décisionnelle" au sein de notre programme académique. L'objectif principal de cette étude est la Mise en place d'un système décisionnel pour l'étude du marché des véhicules électriques. Notre équipe est composée de quatre étudiants agissant en tant que data analysts, chargés de réaliser ce projet dans le cadre de notre formation. Notre professeur superviseur endosse le rôle du responsable d'une entreprise spécialisée dans les véhicules électriques, qui est le client de ce projet.

Ce cadre de travail définit les rôles et les responsabilités des différents acteurs impliqués dans le projet, en mettant en avant le caractère académique de l'initiative tout en le plaçant dans un contexte professionnel réaliste.

1.2 Problématique et motivation du projet

La mise en place d'un système décisionnel pour l'étude du marché des véhicules électriques se confronte à plusieurs problématiques majeures au sein du marché américain et de son infrastructure de recharge. Tout d'abord, le manque d'infrastructures de recharge représente un obstacle significatif à l'adoption généralisée des véhicules électriques. La lenteur des progrès dans ce domaine est largement attribuée à la difficulté de se recharger en dehors des grands axes routiers, limitant ainsi la portée et l'attrait des voitures électriques pour de nombreux citoyens.

Une autre problématique spécifique réside dans la complexité accrue de l'achat d'un véhicule électrique, notamment en raison de l'autonomie limitée et du réseau de recharge peu développé. Les conducteurs américains sont confrontés à des défis logistiques lorsqu'il s'agit de planifier leurs trajets en raison de la disponibilité inégale des stations de recharge et des temps de recharge relativement longs par rapport à un plein de carburant traditionnel.

La motivation principale de ce projet est de mettre en place un système décisionnel efficace capable de formuler des plans d'action visant à résoudre ces problèmes spécifiques. En optimisant l'infrastructure de recharge et en proposant des solutions innovantes pour augmenter l'attrait des véhicules électriques.

1.3 Objectifs du projet

Dans notre projet, nous avons identifié plusieurs objectifs clés afin de répondre aux défis et aux besoins de l'entreprise dans le secteur des véhicules électriques.

L'objectif principal de cette étude est d'analyser les tendances annuelles des ventes de véhicules électriques, en examinant les variations saisonnières et les tendances temporelles au fil du temps. De plus, nous nous efforçons d'identifier les zones à forte densité de stations de recharge et leur corrélation avec les ventes de véhicules électriques, afin d'évaluer leur influence sur le marché. Enfin, notre objectif est de proposer des emplacements stratégiques pour de nouvelles installations de stations de recharge, en optimisant l'accessibilité et la couverture du réseau de recharge pour répondre aux besoins croissants des utilisateurs de véhicules électriques.

Conclusion

En conclusion, la mise en place d'un système décisionnel pour l'étude du marché des véhicules électriques s'inscrit dans un cadre de travail académique visant à répondre à des défis concrets rencontrés par les entreprises du secteur. La problématique identifiée, notamment le manque d'infrastructures de recharge et les obstacles à l'adoption des véhicules électriques, ainsi que les motivations sous-jacentes, ont guidé la définition des objectifs stratégiques de ce projet. Ces objectifs visent à analyser les tendances du marché, évaluer l'impact des stations de recharge et proposer des solutions pour optimiser le réseau de recharge, contribuant ainsi à la prise de décision éclairée dans le domaine de la mobilité électrique.

CONDUITE DU PROJET

Dans ce chapitre, nous aborderons la conduite de notre projet en détaillant les différentes étapes et processus clés qui ont été mis en œuvre pour garantir son succès. Nous commencerons par examiner notre méthodologie de gestion du projet, suivie de la planification stratégique qui explique les bases de notre progression. Ensuite, nous allons discuter la gestion des délais du projet, mettant en évidence l'importance de respecter les échéances fixées pour assurer une progression efficace. Enfin, nous aborderons la gestion des risques, en identifiant les défis et en présentant les stratégies mises en place pour les anticiper.

2.1 Méthodologie de gestion du projet

Pour notre projet nous avons adopté la méthode agile pour sa flexibilité et son efficacité, en structurant notre processus de développement autour des principes de visualisation du flux de travail, de limitation du travail en cours et d'amélioration continue, nous avons pu maintenir un flux de travail fluide.

Le framework Kanban nous a permis de structurer notre processus de développement de manière efficace en visualisant le flux de travail et en limitant le nombre de tâches en cours. L'intégration de ce framework dans notre méthodologie nous a offert une plateforme centralisée pour la gestion de nos projets, des tâches et des équipes.

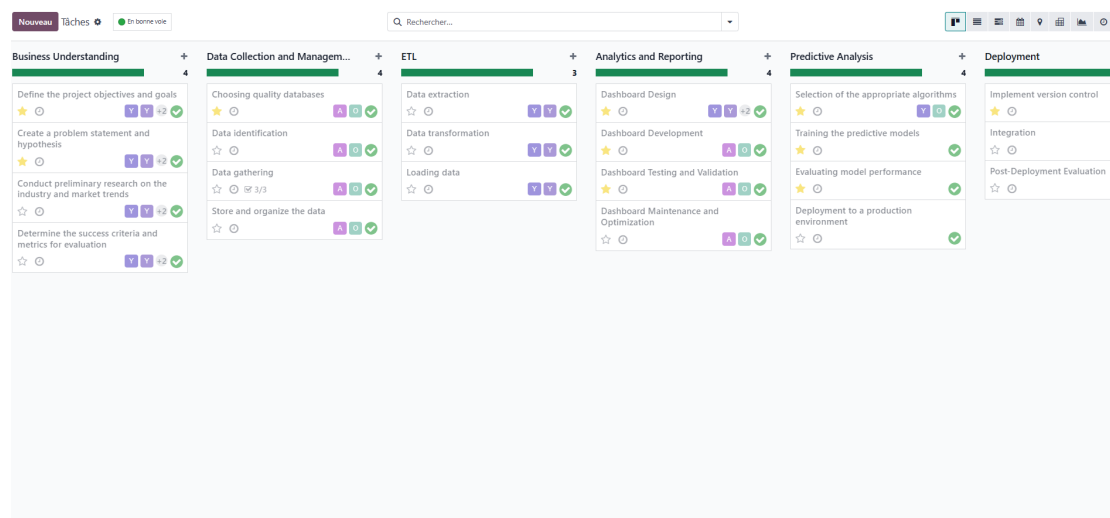


FIGURE 2.1 – Tableau Kanban

Grâce à l'outil Odoo, nous avons pu créer des tableaux Kanban personnalisés pour chaque phase du projet, permettant ainsi une visualisation claire de l'état d'avancement des tâches et des priorités à chaque étape.

2.2 Planification du projet

La planification du projet a été essentiel pour atteindre nos objectifs. Cette phase a impliqué une série d'étapes stratégiques visant à définir clairement les objectifs, à identifier les ressources nécessaires et à élaborer un plan de travail détaillé pour guider l'exécution du projet.

Une fois nos objectifs établis, nous avons procédé à une évaluation des ressources

nécessaires pour les atteindre. Cela comprenait l'identification des membres de l'équipe dotés des compétences requises, l'accès aux données et aux technologies et outils nécessaires, ainsi que l'estimation des besoins en termes de temps.

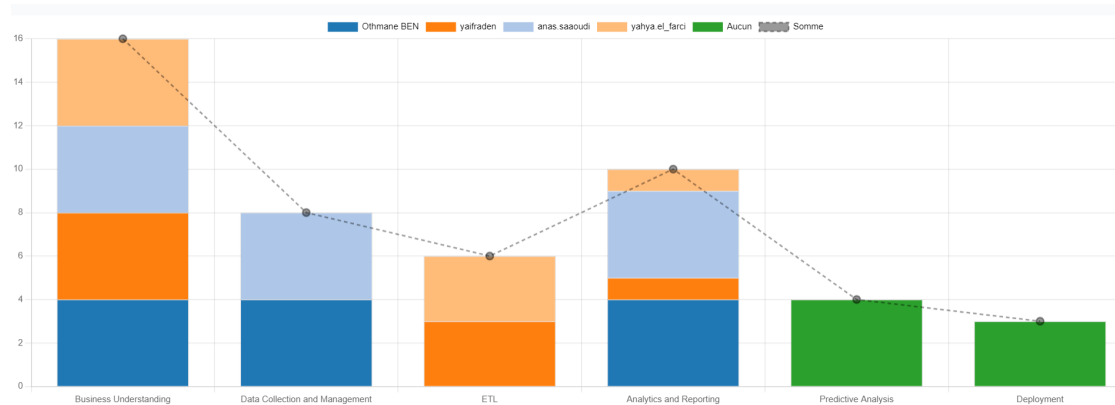


FIGURE 2.2 – Répartition des tâches

Pour assurer une répartition efficace des responsabilités et des tâches au sein de l'équipe, nous avons tenu compte des compétences individuelles de chaque membre. Cette approche a favorisé une répartition équilibrée de la charge de travail et a permis à chaque membre de l'équipe de contribuer de manière significative à la réalisation des objectifs du projet.

2.3 Diagramme de Gantt

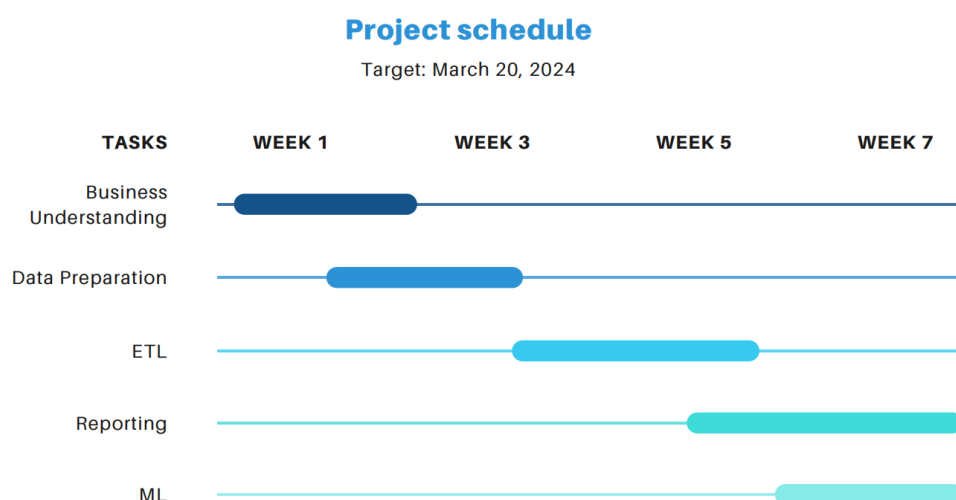


FIGURE 2.3 – Diagramme de Gantt

2.4 Management des délais du projet

Pour garantir que notre projet reste sur la bonne voie et gérer l'achèvement dans le temps voulu, nous avons adopté une approche méthodique pour anticiper, surveiller et gérer les délais potentiels.

Définir les Activités : Ce processus initial consiste à identifier de manière précise et détaillée les actions spécifiques nécessaires pour produire les livrables du projet. En analysant les exigences et les objectifs du projet, nous avons pu décomposer le travail en activités gérables et mesurables, chacune contribuant à l'atteinte des objectifs globaux.

Suivi Régulier de l'Avancement : Pour assurer le suivi continu des délais, nous avons mis en place des mécanismes de suivi réguliers. Des réunions d'avancement ont été organisées pour examiner l'état d'avancement des tâches, identifier les éventuels retards et prendre des mesures correctives si nécessaire.

▼ Business Understanding (4)					
<input type="checkbox"/>	☆	✓	Determine the success criteria and metrics for evaluation	○ A Y Y	○ Business Understanding
<input type="checkbox"/>	★	✓	Define the project objectives and goals	○ A Y Y	○ Business Understanding
<input type="checkbox"/>	★	✓	Create a problem statement and hypothesis	○ A Y Y	○ Business Understanding
<input type="checkbox"/>	☆	✓	Conduct preliminary research on the industry and market trends	○ A Y Y	○ Business Understanding
▼ Data Collection and Management (4)					
<input type="checkbox"/>	☆	✓	Store and organize the data	○ A	○ Data Collection and Management
<input type="checkbox"/>	☆	✓	Data identification	○ A	○ Data Collection and Management
<input type="checkbox"/>	☆	✓	Data gathering (3/3 sous-tâches)	○ A	○ Data Collection and Management
<input type="checkbox"/>	★	✓	Choosing quality databases	○ A	○ Data Collection and Management
▼ ETL (3)					
<input type="checkbox"/>	☆	✓	Loading data	Y Y	○ ETL
<input type="checkbox"/>	☆	✓	Data transformation	Y Y	○ ETL
<input type="checkbox"/>	☆	✓	Data extraction	Y Y	○ ETL
▼ Analytics and Reporting (4)					
<input type="checkbox"/>	★	✓	Dashboard Testing and Validation	○ A	○ Analytics and Reporting
<input type="checkbox"/>	☆	✓	Dashboard Maintenance and Optimization	○ A	○ Analytics and Reporting
<input type="checkbox"/>	★	✓	Dashboard Development	○ A	○ Analytics and Reporting
<input type="checkbox"/>	★	✓	Dashboard Design	○ A Y Y	○ Analytics and Reporting
▼ Predictive Analysis (4)					
<input type="checkbox"/>	☆	✓	Training the predictive models		○ Predictive Analysis
<input type="checkbox"/>	★	✓	Selection of the appropriate algorithms		○ Predictive Analysis
<input type="checkbox"/>	★	✓	Evaluating model performance		○ Predictive Analysis
<input type="checkbox"/>	☆	✓	Deployment to a production environment		○ Predictive Analysis
▼ Deployment (3)					
<input type="checkbox"/>	☆	✓	Post-Deployment Evaluation		○ Deployment
<input type="checkbox"/>	☆	✓	Integration		○ Deployment
<input type="checkbox"/>	★	✓	Implement version control		○ Deployment

FIGURE 2.4 – Liste des activités

2.5 Management des risques

La gestion des risques est un aspect visant à identifier, évaluer et atténuer les menaces potentielles qui pourraient compromettre le succès du projet.

- **Identification des Risques** : Nous avons débuté par une analyse approfondie des risques potentiels qui pourraient affecter notre projet. Nous avons identifié un éventail de risques, allant des retards dans la livraison des données à des imprévus techniques liés à la mise en œuvre du système.
- **Évaluation des Risques** : Chaque risque identifié a été évalué en termes de probabilité d'occurrence et d'impact potentiel sur le projet.
- **Planification des Réponses aux Risques** : Une fois les risques évalués, nous avons élaboré des plans de réponse adaptés à chaque risque identifié. Ces plans comprenaient des stratégies d'atténuation pour réduire la probabilité d'occurrence des risques,

Risque	Probabilité	Impact	Stratégie d'atténuation
Retard dans la livraison de données	Moyenne	Élevé	Renégociations de délais, Utilisation de données alternatives
Problèmes Techniques Imprévus	Faible	Très Élevée	Réalisation de tests approfondis, Trouver des solutions internes
Changement dans les exigences	Élevée	Moyen	Communication régulière avec les parties prenantes, flexibilité dans la planification

TABLE 2.1 – Exemples des principaux risques identifiés

Conclusion

En conclusion, notre projet a été mené à bien grâce à une approche méthodique, une collaboration efficace et une gestion pro-active. Nous avons pu atteindre nos objectifs en utilisant des méthodologies telles que la méthode Agile et l'outil Odoo pour la gestion du projet, ce qui a permis une planification précise, une allocation efficace des ressources et une gestion optimale des délais et des risques.

ANALYSE ET SPÉCIFICATION

Dans ce chapitre, nous entamons l'analyse et la spécification de notre projet d'étude de marché des véhicules électriques. Nous commencerons par définir le périmètre du projet, identifier les axes d'analyses pertinents, sélectionner les indicateurs de performance essentiels, et créer une matrice dimensionnelle pour guider notre analyse. Ces étapes sont cruciales pour établir une base solide et assurer le succès de notre projet, en fournissant des insights approfondis et des recommandations stratégiques à notre client.

3.1 Périmètre du projet

Dans le cadre de notre projet d'étude de marché des véhicules électriques, nous entreprendrons plusieurs activités essentielles pour atteindre nos objectifs. Tout d'abord, nous procéderons à la collecte de données en identifiant les sources pertinentes telles que les ventes de véhicules électriques et les statistiques sur les infrastructures de recharge. Ensuite, nous analyserons ces données pour identifier les tendances, les modèles et les corrélations qui pourraient influencer le marché des véhicules électriques.

Nous évaluerons également la demande actuelle et future pour les véhicules électriques, en tenant compte des différents facteurs qui pourraient affecter cette demande. De plus, nous étudierons les infrastructures de recharge disponibles et leur impact sur l'adoption des véhicules électriques.

Notre étude sera limitée géographiquement aux États-Unis, en raison de la disponibilité des données sur le marché américain et de l'hypothèse selon laquelle l'entreprise cliente est située sur ce marché. Le projet sera réalisé dans un délai de 7 semaines, comme indiqué dans notre diagramme de Gantt, et aucun budget spécifique n'a été alloué, étant donné que le projet est de nature académique.

En termes de livrables, nous fournirons une analyse approfondie du marché, comprenant les ventes de véhicules électriques et la distribution des stations de recharge, ainsi que des recommandations stratégiques visant à augmenter le taux de vente des véhicules électriques.

3.2 Les axes d'analyse

Nous avons identifié trois axes d'analyse clés. Ces axes nous permettront de segmenter nos données et indicateurs afin de fournir aux utilisateurs les informations nécessaires à la prise de décision.

Le premier axe d'analyse que nous utilisons est celui du temps. Cet axe nous permet d'analyser nos données et indicateurs pour des périodes spécifiques, ce qui nous aide à observer les tendances temporelles et saisonnières des ventes de véhicules électriques, ainsi que l'évolution de la demande au fil du temps.

Ensuite, nous utilisons l'axe de la géographie. Cela nous permet de réaliser des analyses géospatiales et de cartographier nos données par région, département ou autres

zones géographiques pertinentes. Cela nous aide à identifier les zones à forte demande de véhicules électriques, ainsi que les régions où les infrastructures de recharge sont les plus nécessaires.

Enfin, nous utilisons l'axe de la technologie. Cet axe nous permet de suivre l'évolution des technologies liées aux véhicules électriques, en mettant particulièrement l'accent sur les infrastructures de recharge et leur accessibilité technologique. Cela nous aide à évaluer l'impact des avancées technologiques sur le marché des véhicules électriques et à identifier les opportunités stratégiques pour notre client.

En utilisant ces axes d'analyse, nous pouvons créer des tableaux de bord efficaces qui permettent à notre client de visualiser et d'interpréter facilement les données, afin de prendre des décisions éclairées pour son entreprise.

3.3 Les indicateurs de performance

Dans cette partie, nous allons décrire les principaux indicateurs de performance (KPIs) utilisés pour évaluer différents aspects du marché des véhicules électriques.

1. Indicateurs en rapport avec l'évolution annuelle des ventes des véhicules électriques (EV) :

- **Nombre d'immatriculations de véhicules électriques** : Il s'agit du nombre total de voitures électriques vendues sur une période donnée, permettant de mesurer la popularité des EV.
- **Taux de BEV (Battery Electric Vehicle)** : Ce KPI indique la proportion de véhicules électriques entièrement alimentés par batterie par rapport au total des ventes de véhicules électriques.
- **Taux de PHEV (Plug-in Hybrid Electric Vehicle)** : Il représente la proportion de véhicules électriques hybrides rechargeables par rapport au total des ventes de véhicules électriques.
- **Véhicules par marque et modèle** : Cet indicateur analyse la part de marché de chaque constructeur en termes de ventes de véhicules électriques, en identifiant les modèles les plus populaires.

2. Indicateurs en rapport avec la distribution des stations de charge électriques :

- **Nombre de stations de recharge** : Il s'agit du nombre total de stations de recharge électrique disponibles, permettant d'évaluer l'accessibilité des infrastructures de recharge.
- **Nombre de sockets** : Ce KPI représente le nombre total de prises de recharge disponibles dans toutes les stations de charge.
- **Nombre de stations par type** : Il analyse la répartition des stations de recharge en fonction de leur type (rapide, normale, etc.), fournissant des informations sur la diversité des options de recharge.
- **Type d'emplacement d'installation** : Cet indicateur identifie les zones où les stations de recharge sont concentrées, ce qui peut aider à cibler les régions nécessitant davantage d'infrastructures de recharge.

3. Indicateurs en rapport avec l'émission des EVs :

- **Émissions de CO₂ (Kg)** : Ce KPI évalue la quantité totale de dioxyde de carbone émis par les véhicules électriques, permettant de mesurer leur contribution à la réduction des émissions de gaz à effet de serre.
- **Émissions de Nox (g/km)** : Il mesure les émissions de dioxyde d'azote par kilomètre parcouru, fournissant des informations sur l'impact environnemental des véhicules électriques.
- **Utilisation des gaz fluorés** : Cet indicateur analyse le taux d'utilisation des gaz fluorés dans les véhicules électriques, ce qui est important pour évaluer leur empreinte environnementale globale.

Ces indicateurs de performance nous fournissent des données essentielles pour évaluer et surveiller différents aspects du marché des véhicules électriques, ce qui nous permettra de formuler des recommandations stratégiques pour notre client.

3.4 Matrice dimensionnelle

La matrice dimensionnelle est un outil essentiel pour organiser et synthétiser les différents indicateurs et axes d'analyse pertinents dans le cadre de notre étude de marché sur les véhicules électriques. Elle permet de visualiser clairement comment chaque indicateur est lié à chaque dimension ou axe d'analyse. Voici la matrice dimensionnelle pour notre projet.

Indicateur	Temps	Géographie	Technologies
Nombre d'immatriculations de VE	Oui	Oui	Oui
Taux de BEV	Oui	Oui	Non
Taux de PHEV	Oui	Oui	Non
Véhicules par marque et modèle	Oui	Non	Oui
Nombre de stations de recharge	Oui	Oui	Non
Nombre de sockets	Oui	Oui	Non
Nombre de stations par type	Oui	Oui	Oui
Type d'emplacement d'installation	Non	Oui	Non
Émissions de CO2 (Kg)	Oui	Non	Non
Émissions de Nox (g/km)	Oui	Non	Non
Taux d'utilisation des gaz fluorés	Oui	Non	Non

TABLE 3.1 – Matrice dimensionnelle

Conclusion

En conclusion, dans ce chapitre d'Analyse et Spécification, nous avons détaillé les aspects clés de notre projet. Nous avons commencé par délimiter le périmètre du projet, définissant ainsi les contours de notre travail, ses objectifs et ses limites. Ensuite, nous avons exploré les axes d'analyse, identifiant les dimensions essentielles selon lesquelles nous structurerons nos analyses et interpréterons nos données. Nous avons ensuite présenté une gamme d'indicateurs de performance spécifiques, offrant un aperçu des mesures que nous utiliserons pour évaluer le succès et les tendances du marché des véhicules électriques. Enfin, nous avons établi une matrice dimensionnelle, consolidant nos axes d'analyse et nos indicateurs dans un cadre synthétique pour orienter notre travail analytique.

CONCEPTION DE LA SOLUTION

Une fois les besoins définis, la phase de conception est l'étape suivante dans le cycle décisionnel. Nous explorerons les éléments fondamentaux de la conception de notre système. Tout d'abord, nous examinerons l'architecture générale du système. Ensuite, nous présenterons le modèle conceptuel de données. Nous aborderons également la conception du tableau de bord, en décrivant le processus de création et en fournissant une fiche de conception détaillée pour guider le développement. Enfin, nous plongerons dans la méthodologie CRISP-DM pour la fouille de données.

4.1 Architecture générale du système

Notre architecture BI s'articule autour d'une file de dépôt (FDA) où sont stockés les différents fichiers Excel et CSV collectés qui contiennent des données sur le marché des véhicules électriques, avant d'être soumis à un processus ETL pour la construction du data warehouse. Cette approche garantit une collecte efficace et ordonnée des données sources, prêtes à être transformées et chargées pour l'analyse ultérieure. À partir du data warehouse, nous passerons ensuite à la création des tableaux de bord et à l'exploration des données, ouvrant ainsi la voie à des analyses approfondies et à des insights stratégiques. Chaque étape de cette architecture est conçue pour garantir la pertinence, la fiabilité et la disponibilité des informations nécessaires à la prise de décision.

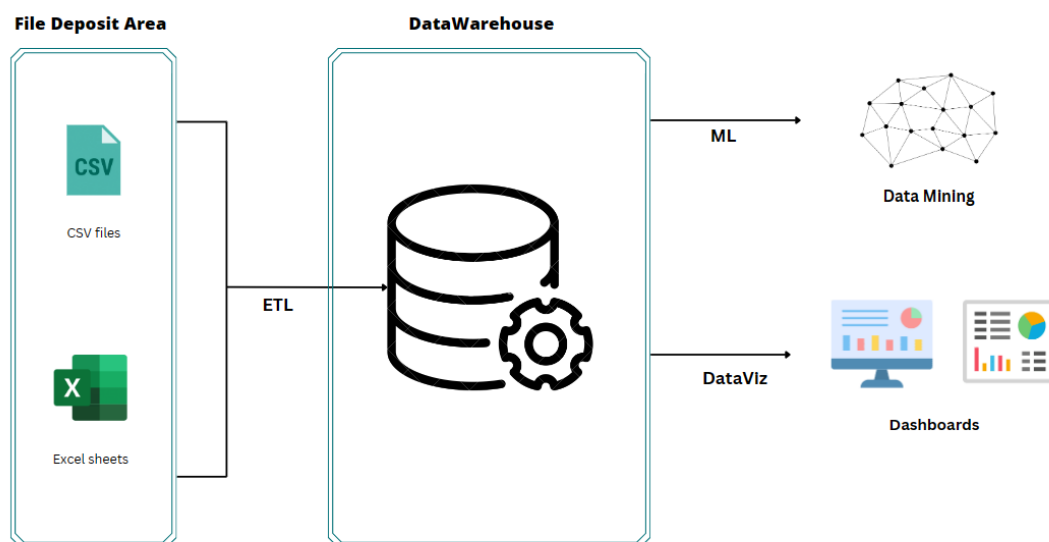


FIGURE 4.1 – Architecture globale du système

4.2 Modèle conceptuel de données

La conception des données est cruciale pour notre projet BI. Notre modèle conceptuel de données (MCD) offre une représentation logique de la structure des données, résultant d'une analyse approfondie des besoins métier et de l'application de meilleures pratiques. Il identifie les entités principales et leurs relations, fournissant ainsi une base solide pour notre système BI.

Entités :

- **FACT_immatriculations_véhicules** : Cette table contient des faits sur les immatriculations de véhicules, avec des attributs comme l'identifiant d'immatriculation, l'identifiant d'état, l'identifiant de véhicule, l'identifiant de localisation, l'identifiant de date d'immatriculation, l'identifiant de type de transmission, les frais d'immatriculation, le coût de possession d'un véhicule électrique, la part de marché des véhicules électriques et le taux d'adoption des véhicules électriques de haute technologie.
- **FACT_station_recharge** : Cette table contient des faits sur les stations de recharge, avec des attributs comme l'identifiant de la station de recharge, l'identifiant de la station-service (si applicable), l'identifiant de la localisation, l'identifiant de l'état et un type de données géographique représentant la densité de la station.
- **DIM_Date** : Cette table de dimension stocke des informations sur les dates, incluant l'identifiant de date et la date elle-même.
- **Dim_type_transmission** : Cette table de dimension capture les détails sur les types de transmission, pouvant inclure des attributs comme l'identifiant du type de transmission et la description du type (par exemple, électrique, essence, hybride).
- **Dim_état** : Cette table de dimension stocke des informations sur les états, contenant probablement l'identifiant d'état et le nom de l'état.
- **Dim_véhicule** : Cette table de dimension capture les détails sur les véhicules, incluant potentiellement l'identifiant du véhicule, la marque, le modèle, l'année, etc.
- **Dim_localisation** : Cette table de dimension stocke des informations sur les localisations, contenant probablement l'identifiant de localisation, la ville, le code postal, etc.
- **Dim_station_service** : Cette table de dimension stocke des informations sur les stations-service (si applicable), pouvant inclure des attributs comme l'identifiant de la station-service, le nom de la station, la marque, etc.

Relations :

FACT_immatriculations_véhicules se connecte à plusieurs tables de dimension avec des relations un-à-plusieurs :

- **DIM_date** : Une seule immatriculation a lieu à une date précise (une immatriculation - une seule date).
- **Dim_type_transmission** : Une immatriculation s'applique à un véhicule avec un type de transmission spécifique (une immatriculation - un type de transmission).
- **Dim_état** : Un véhicule est immatriculé dans un état particulier (une immatriculation - un état).
- **Dim_véhicule** : Un enregistrement d'immatriculation représente un seul véhicule (une immatriculation - un véhicule).
- **Dim_localisation** : Une immatriculation de véhicule doit être liée à une localisation spécifique (une immatriculation - une localisation).

FACT_station_recharge se connecte à plusieurs tables de dimension avec des relations un-à-plusieurs :

- **dim_station_service** : Une station de recharge peut être co-localisée avec une station-service (une station - une ou plusieurs stations-service).
- **Dim_état** : Une station de recharge fonctionne dans un état (une station - plusieurs états).
- **Dim_localisation** : Une station de recharge est située à un endroit précis (une station - plusieurs localisations).

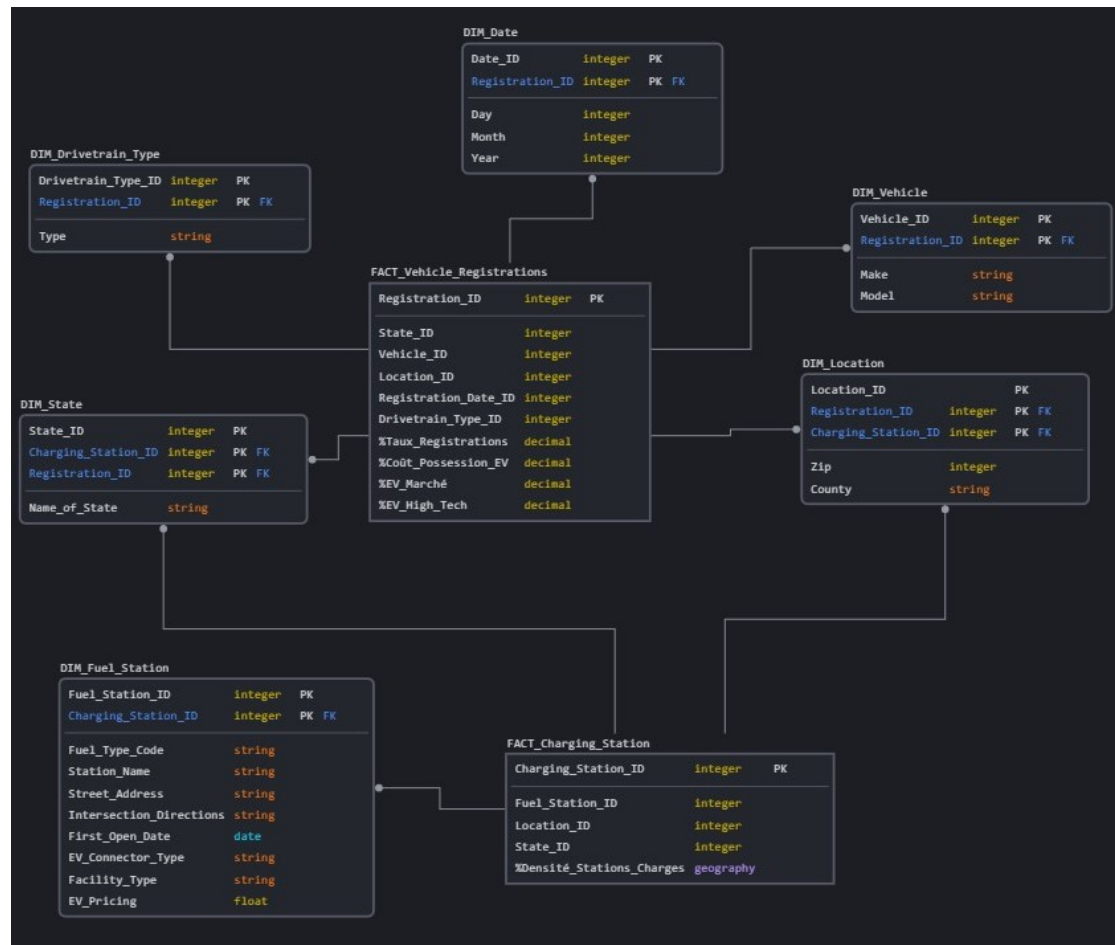


FIGURE 4.2 – Modélisation du DWH

4.3 Conception du Tableau de bord

La modélisation du datamart établie n'est pas suffisante pour achever la phase de conception et passer à la mise en oeuvre. Pour cela, nous devons bien choisir les indicateurs et les présenter de la meilleure façon. Cette présentation doit être aussi simple et compréhensible. Pour arriver à ces objectifs, nous allons présenter nos données qui sont liées aux indicateurs afin de les présenter sous forme des schémas graphiques.

4.3.1 Fiche de conception

La fiche de conception du tableau de bord représente des informations sur une proposition de tableau de bord pour l'utilisateur final, telles que les questions commerciales auxquelles il faut répondre avec ce tableau de bord, les indicateurs de performance à utiliser, les filtres, l'audience visée et les axes d'analyse.

Audience : Direction Générale, Département Ventes et Marketing, Département de Recherche et développement			
Filtre : Géographie, Temps, Technologie			
Approche : Drill down			
Dashboard	Questions métiers	KPIs	Dimensions
Evolution annuelle de vente des EVs	Combien de voiture électrique est vendue ?	Nombre d'immatriculations de véhicules électriques	Temps, Géographie, Technologie
	Combien sont de type BEV ?	taux de BEV	Temps, Géographie
	Combien sont de type PHEV ?	taux de PHEV	Temps, Géographie
	Quel est la part de marché de chaque constructeur	vehicule par marque et modèle	Temps, Technologie

TABLE 4.1 – Fiche de conception du premier dashboard (Ventes)

Audience : Département Ventes et Marketing, Département de R&D			
Filtre : Géographie, Temps, Technologie			
Approche : Drill down			
Dashboard	Questions métiers	KPIs	Dimensions
Distribution de stations de recharge électriques	Quel est le nombre de station de recharge disponible ?	nombre de stations de recharge	Temps, Géographie
	Combien de sockets sont en circulation ?	Nombre de sockets	Temps, Technologie Géographie
	Combien de sockets disponible par type ?	nombre de stations de par type	Temps, Géographie, Technologie
	Les stations de recharge sont concentrée où ?	Type de emplacement d'installation	Geographie

TABLE 4.2 – Fiche de conception du deuxième dashboard (stations de recharge)

Audience : Département de Recherche et développement, Chargé d'environnement			
Filtre : Temps, Technologie			
Approche : Drill down			
Dashboard	Questions métiers	KPIs	Dimensions
Emission des EVs	Les véhicules électriques progressent-elles dans la réduction des émissions ?	Émissions de CO2 (Kg) Émissions de Nox (g/km)	Temps
	Quel est le taux d'utilisation du gaz fluoré ?	Utilisation des gaz fluorés	Temps

TABLE 4.3 – Fiche de conception du troisième dashboard (Emission CO2)

4.4 Fouille de données : CRISP-DM

Nous avons choisi d'adopter l'approche CRISP-DM (Cross-Industry Standard Process for Data Mining) comme cadre méthodologique pour nos activités de datamining. Cette approche est largement reconnue dans l'industrie pour sa structure robuste et sa flexibilité, ce qui en fait un choix idéal pour notre projet.

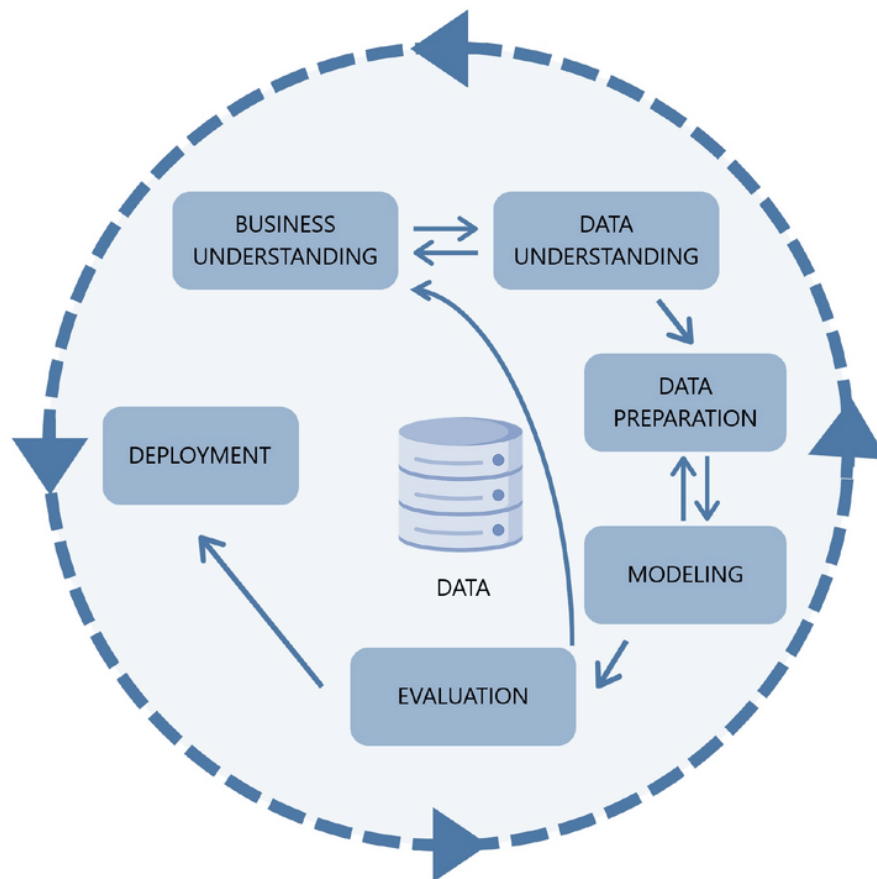


FIGURE 4.3 – Différentes étapes du framework CRISP-DM

Dans le cadre de notre projet, nous pouvons parfaitement intégrer notre expérience préalable en reporting avec l'approche CRISP-DM pour le datamining. Voici comment nous pouvons collaborer entre ces deux éléments :

Compréhension du métier et des données : Nous pouvons tirer parti de notre expérience en reporting pour approfondir notre compréhension des besoins métier et des données disponibles. Cette étape nous permettra de mieux cibler nos efforts de datamining en identifiant les questions clés à explorer et en sélectionnant les données pertinentes

pour l'analyse.

Préparation des données : Notre expérience en reporting nous a probablement déjà confrontés à des défis liés à la qualité et à la structure des données. En intégrant cette expérience dans la phase de préparation des données du CRISP-DM, nous pouvons mettre en place des processus efficaces pour nettoyer, transformer et intégrer les données sources en vue de l'analyse.

Modélisation : L'expérience préalable en reporting nous a familiarisés avec les types de modèles et les techniques d'analyse les plus pertinents pour notre organisation. Nous pouvons utiliser cette connaissance pour guider nos choix de modélisation dans le cadre du datamining, en sélectionnant les méthodes les mieux adaptées pour répondre à nos objectifs métier spécifiques.

Évaluation et déploiement : Nous pouvons utiliser nos compétences en reporting pour évaluer et interpréter les résultats de nos modèles de datamining, en les présentant de manière claire et concise aux parties prenantes de l'organisation. De plus, nous pouvons mettre en place des mécanismes de suivi et de surveillance pour assurer la pérennité des solutions déployées.

Conclusion

En conclusion, ce chapitre a jeté les bases essentielles de notre système. De l'architecture générale au modèle conceptuel de données, en passant par la conception détaillée du tableau de bord et l'utilisation de CRISP-DM pour la fouille de données, nous sommes maintenant prêts à passer à l'implémentation. Ces éléments fournissent une base solide pour notre projet, nous permettant de progresser vers notre objectif final avec confiance et clarté.

MISE EN OEUVRE DE LA SOLUTION

Ce chapitre est consacré à la phase de mise en oeuvre du projet qui concrétise le travail d'analyse et de conception présenté dans les deux chapitres précédents. Dans ce chapitre, nous commencerons par la collection de données, puis la partie ETL et nous passerons à la construction des tableaux de bord et finalement à la partie Machine Learning.

5.1 Source de données

Après avoir définis le modèle conceptuel de données, maintenant vient l'étape de collection de données afin d'alimenter notre datawarehouse. Pour ce faire nous avons pu acquérir plusieurs fichiers qui contiennent les différentes données sur les ventes et immatriculations des véhicules électriques dans les états unis durant les dernières années et d'autres fichiers avec des données sur la distribution des stations de charge électriques dans les différentes états de l'Amérique.

	A	B	C	D	E	F	G	H	I	J
1	State	County	Registration Date	Vehicle Make	Vehicle Model	Vehicle Model Year	Drivetrain Type	Vehicle GVWR Class	Vehicle Category	Vehicle Count
2	TN	Wilson County	3/1/2022	CHEVROLET	VOLT	2013	PHEV	1	Light-Duty (Class 1	1
3	TN	Blount County	3/1/2022	CHEVROLET	VOLT	2013	PHEV	1	Light-Duty (Class 1	1
4	TN	Bedford County	3/1/2022	CHEVROLET	VOLT	2013	PHEV	1	Light-Duty (Class 1	1
5	TN	Knox County	3/1/2022	CHEVROLET	VOLT	2013	PHEV	1	Light-Duty (Class 1	1
6	TN	Lawrence County	3/1/2022	CHEVROLET	VOLT	2013	PHEV	1	Light-Duty (Class 1	1
7	TN	Shelby County	3/1/2022	CHEVROLET	VOLT	2013	PHEV	1	Light-Duty (Class 1	1
8	TN	Hamilton County	3/1/2022	CHEVROLET	VOLT	2013	PHEV	1	Light-Duty (Class 1	1
9	TN	Rhea County	3/1/2022	CHEVROLET	VOLT	2013	PHEV	1	Light-Duty (Class 1	1

FIGURE 5.1 – Les différents attributs des fichiers de ventes

	A	B	C	D	E	F	G
1	DMV ID	DMV Snapshot (Date)	County	Vehicle Name	Registration Valid Date	Registration	Technology
2	1	Registration Data from FPL (6/30/2018)	Dade	Tesla Model X	6/30/2018		BEV
3	1	Registration Data from FPL (6/30/2018)	Dade	Tesla Model X	6/30/2018		BEV
4	1	Registration Data from FPL (6/30/2018)	Dade	Tesla Model X	6/30/2018		BEV
5	1	Registration Data from FPL (6/30/2018)	Dade	Tesla Model X	6/30/2018		BEV
6	1	Registration Data from FPL (6/30/2018)	Dade	Tesla Model X	6/30/2018		BEV
7	1	Registration Data from FPL (6/30/2018)	Dade	Tesla Model X	6/30/2018		BEV
8	1	Registration Data from FPL (6/30/2018)	Dade	Tesla Model X	6/30/2018		BEV
9	1	Registration Data from FPL (6/30/2018)	Dade	Tesla Model X	6/30/2018		BEV

FIGURE 5.2 – Les différents attributs des fichiers d'immatriculations

	A	B	C	E	F	G	X	Y	Z	AA	AB	AC
1	Fuel Type	Station Name	Street Address	City	State	ZIP	Geocode	Latitude	Longitude	Date Last	ID	Updated At
2	ELEC	Los Angeles Cor	1201 S Figueroa St	Los Angeles	CA	90015	GPS	34.040539	-118.2713	#####	1523	2023-02-1
3	ELEC	California Air Re	9530 Telstar Ave	El Monte	CA	91731	GPS	34.06872	-118.064	#####	1583	2023-02-1
4	ELEC	Scripps Green H	10666 N Torrey Pines R	Patient Parking Structure								
5	ELEC	Galpin Motors	15421 Roscoe Blvd	Sepulveda	CA	91343	200-9	34.221665	-118.4683	#####	6405	2024-01-3
6	ELEC	Galleria at Tyler	1299 Galleria at Tyler	Riverside	CA	92503	GPS	33.909914	-117.4590	#####	6425	2024-01-3
7	ELEC	City of Pasadena	150 E Holly St	Pasadena	CA	91103	200-8	34.14762	-118.1471	#####	6505	2024-01-3
8	ELEC	City of Pasadena	45 De Lacey Ave	Pasadena	CA	91105	GPS	34.145138	-118.1526	#####	6506	2023-02-1
9	ELEC	City of Pasadena	33 E Green St	Pasadena	CA	91105	GPS	34.145119	-118.1501	#####	6507	2023-02-1

FIGURE 5.3 – Quelques attributs des fichiers de stations de recharge

5.2 Réalisation ETL

Après avoir collecter les fichiers de données, il faut extraire ces données et faire les transformations nécessaires afin de les nettoyer et les préparer pour les charger dans le

datawarehouse dans le but de calculer les KPIs et faire les visualisations. Pour ce faire, nous avons réalisé les transformations de données à l'aide de Power Query dans Power BI.

5.2.1 Choix d'outils

Power Query est un outil puissant et intuitif pour la préparation de données. Il permet de se connecter à différentes sources de données, de les transformer et de les charger dans un modèle de données Power BI.



FIGURE 5.4 – Outils utilisé pour faire l'ETL

5.2.2 Étapes de transformation

Les étapes de transformation effectuées dans Power Query sont les suivantes :

- **Extraction des données :**

Connexion à la source de données (fichiers Excel, CSV, etc.).

Sélection des tables et des colonnes nécessaires.

- **Transformation des données :**

Nettoyage des données (suppression des doublons, correction des erreurs, etc.).

Transformation des formats de données (date, zipcode, etc.).

Ajout de nouvelles colonnes calculées.

Fusion de tables (données des différents états, etc.)

- **Chargement des données :** Chargement des données transformées dans le modèle de données Power BI.

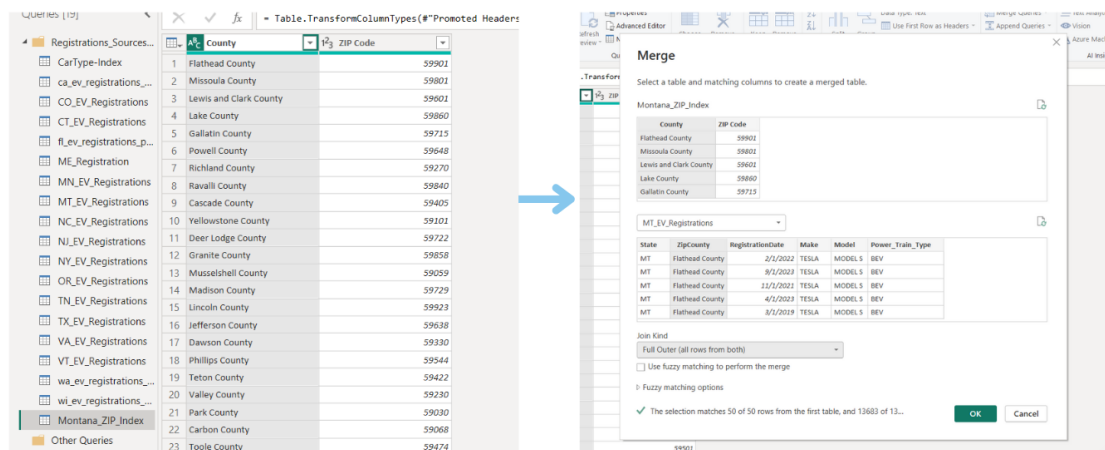


FIGURE 5.5 – Exemple : fusion de la table d'immatriculations au niveau de Montana avec l'index des zip code.

5.3 Création des Tableaux de bord

5.3.1 Choix d'outil

Nous avons opté pour Power BI qui est une solution de Business Intelligence (BI) complète et intuitive qui permet aux entreprises et aux organisations de visualiser et d'analyser leurs données de manière efficace.



FIGURE 5.6 – Logo de l'outil utilisé pour élaborer les dashboards

5.3.2 Dashboards

Le premier tableau de bord représente l'évolution des ventes de véhicules électriques (VE). Dans celui-ci, il est possible de filtrer les données de manière temporelle, géographique, ainsi que par marque et modèle, afin d'extraire la courbe d'évolution des ventes sur un graphique linéaire multiple.

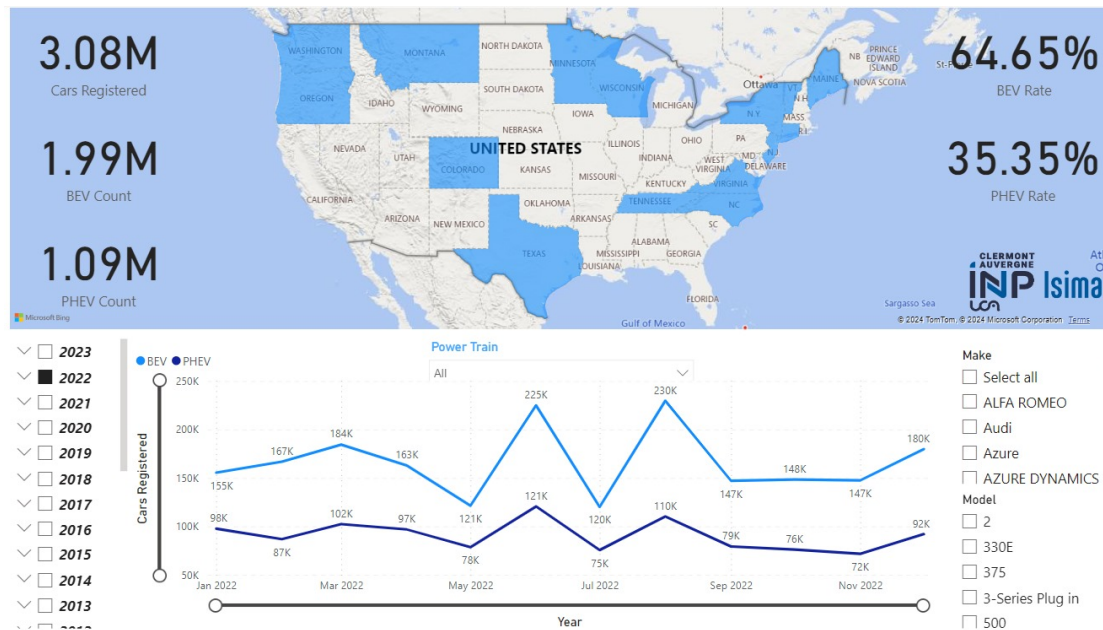


FIGURE 5.7 – Premier Dashboard : ventes et immatriculations (1)

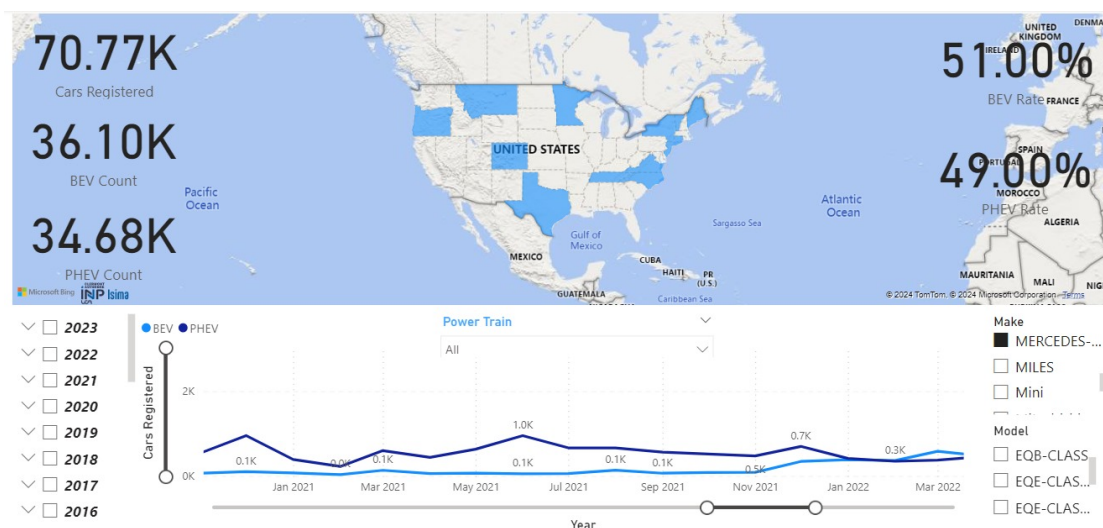


FIGURE 5.8 – Premier Dashboard : ventes et immatriculations (2)

Le deuxième tableau de bord offre une visualisation détaillée de la distribution des stations de recharge à travers les États-Unis. Les indicateurs y affichent le nombre de stations de recharge ainsi que le nombre de ports EVSE pour chaque type. Trois types de filtres sont disponibles : géographique, par type de véhicule, et par infrastructure.

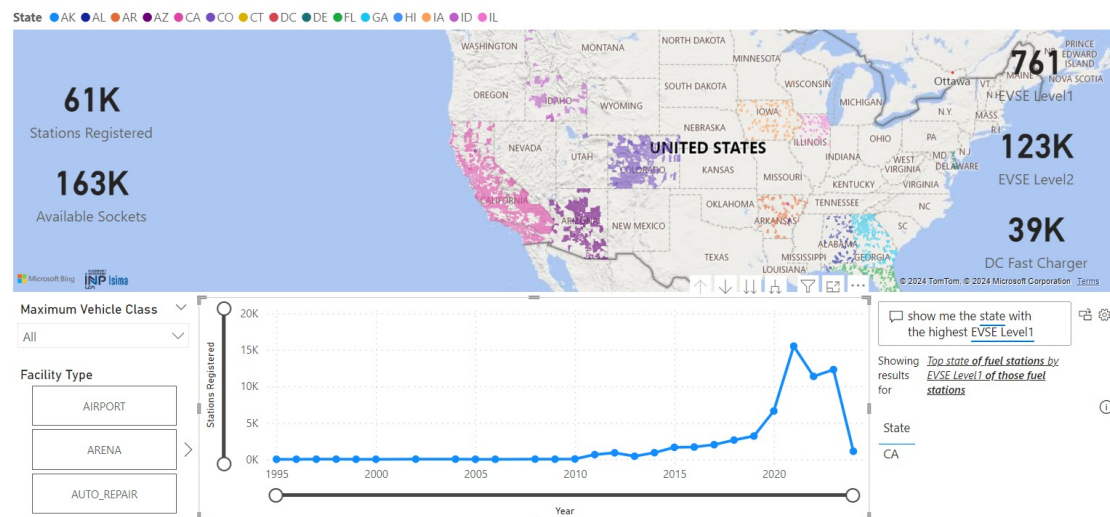


FIGURE 5.9 – Deuxième Dashboard part 1 : distribution de station de recharge

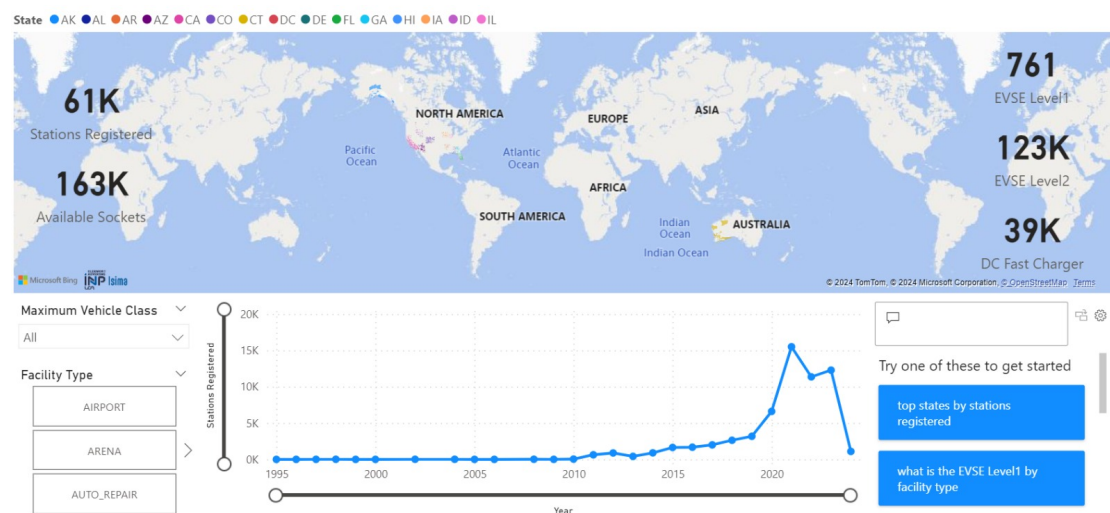


FIGURE 5.10 – Deuxième Dashboard (2))

Ce tableau de bord projette l'évolution des émissions au fil des années par type. En outre, il calcule le pourcentage d'évolution des émissions de CO₂, de Nox, de gaz fluorés et mesure la performance actuelle des gaz à effet de serre (GHG).

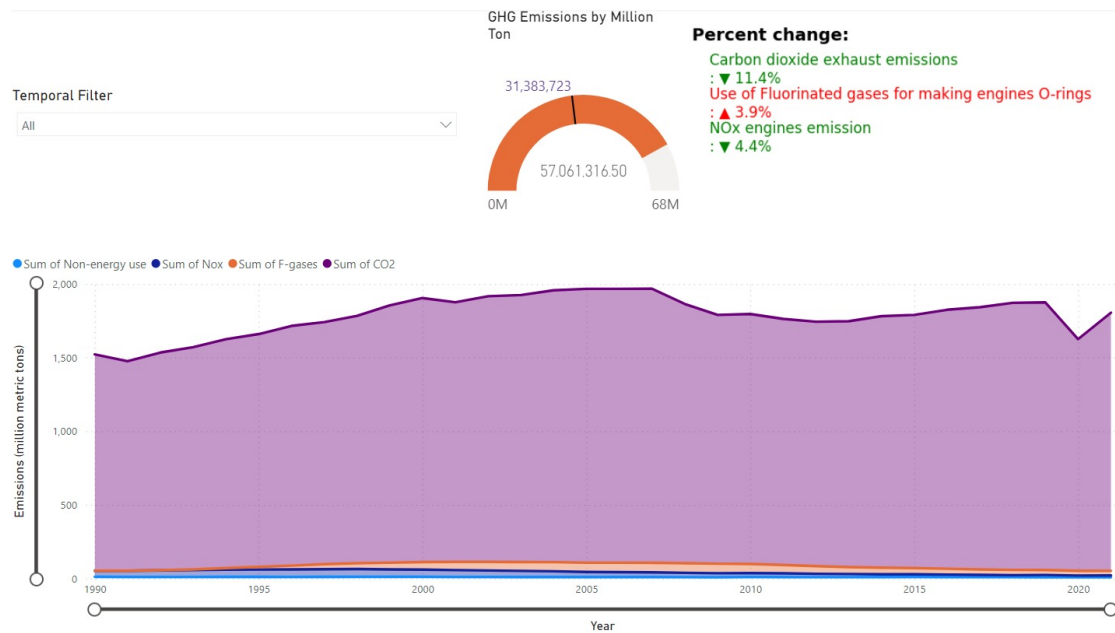


FIGURE 5.11 – Troisième Dashboard : Emission des EVs

5.4 Data Mining

5.4.1 Prévision temporelle des ventes

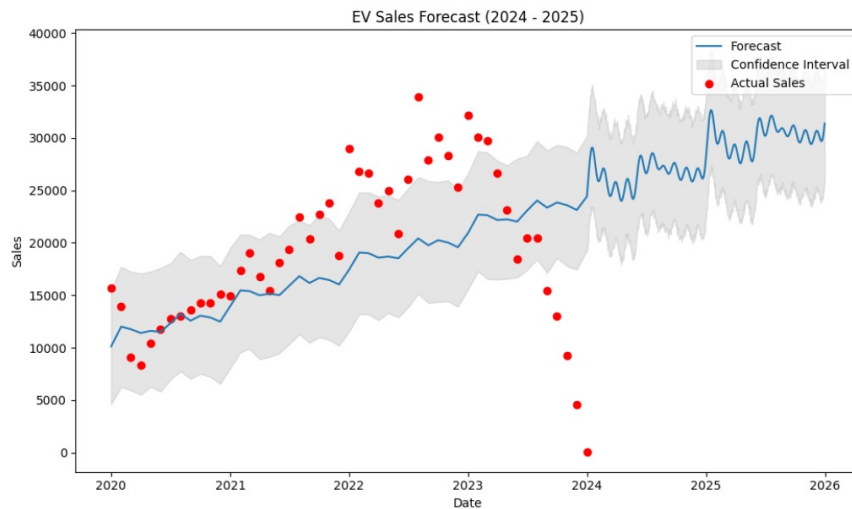


FIGURE 5.12 – Courbe de prédiction des ventes potentielles des véhicules électriques dans les années 2024 et 2025

Pour prédire les ventes de véhicules électriques pour les années 2024 et 2025, le modèle Prophet a été choisi. Ce dernier, une solution développée sur l'API Sklearn de Meta (anciennement Facebook), a été préféré pour plusieurs raisons. L'ETS (Exponentiel Smoothing), bien que disponible sur Power BI, a été écarté pour deux motifs principaux : d'une part, la gestion des séquences de dates dans les données présentait des lacunes, notamment le problème des "ventes nulles", et d'autre part, sa limitation au niveau de la visualisation, puisqu'il ne permet que l'utilisation de graphiques linéaires, restreignant ainsi l'exploitation des informations.

Entre SARIMA et Prophet, le choix s'est porté sur Prophet, en raison de son taux d'erreur plus faible, que ce soit en termes d'erreur quadratique moyenne ou d'erreur absolue moyenne, et de sa meilleure gestion du problème des "ventes nulles".

Une agrégation suivie d'une opération de sommation a été réalisée sur les dates pour constituer un dataframe. Ce dernier comprenait une colonne 'ds', représentant les dates, et une colonne 'y', indiquant le nombre de ventes de véhicules électriques, variable clé pour la prédiction. Les paramètres du modèle ont ensuite été définis : la date maximale,

la période de prédiction, et la configuration de la saisonnalité mensuelle, cette dernière étant activée suite à une analyse préalable de la saisonnalité.

Les résultats obtenus avec ce modèle prédisent une hausse continue des ventes de véhicules électriques, avec un pic attendu en février 2025, où l'on estime que 30 000 véhicules électriques seront vendus dès le début du mois.

5.4.2 Suggestion géographique

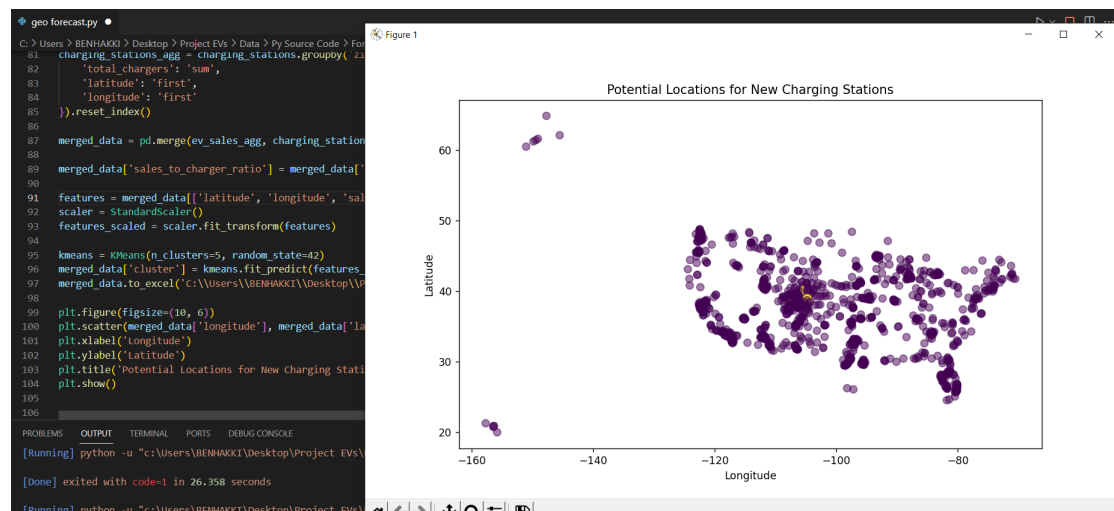


FIGURE 5.13 – Résultat depuis VS Code du Prédiction géographique des futures installations des ports de charges EVSE

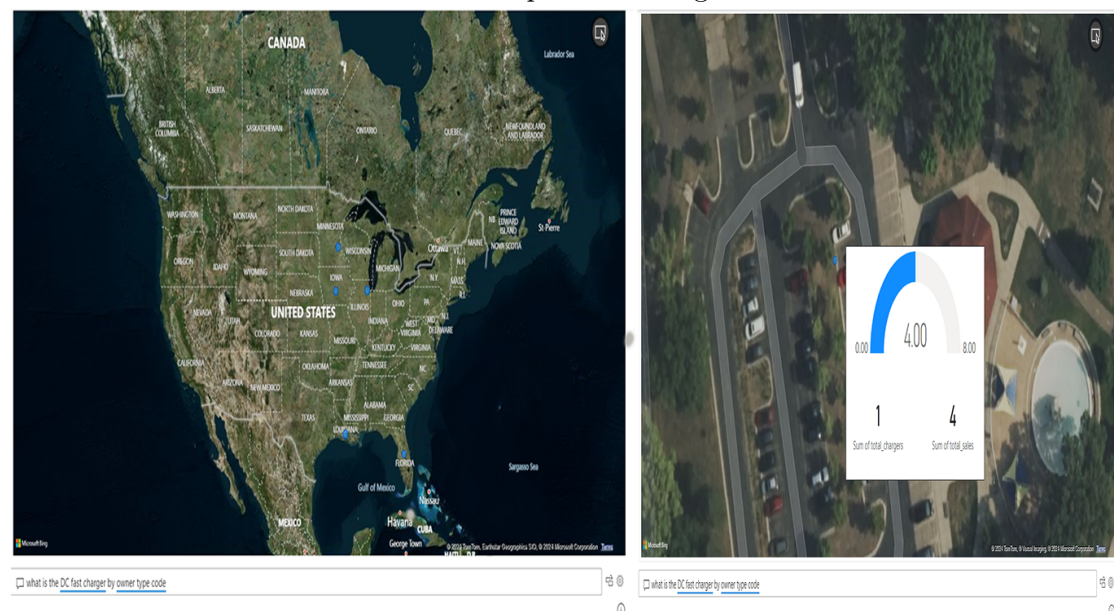


FIGURE 5.14 – Prédiction géographique des futures installations des ports de charges EVSE selon le ratio du ventes VE et le nombre des EVSE

Pour assister nos futurs clients, tels qu'EVgo, Tesla, SemaConnect Network, Electrify America, Volta, etc., dans la réussite de leurs nouvelles installations de stations de recharge ou l'ajout de nouveaux ports EVSE, nous avons créé un ratio (`Sales_To_Charging_Stations`) par code postal, latitude et longitude. Cette méthode nous permet de classer les régions (en utilisant Sklearn pour Kmeans et la bibliothèque `uszipcode`) selon si le nombre de ports disponibles répond adéquatement au nombre de véhicules électriques qui circulent fréquemment dans ces régions. Selon Evadoption, la réglementation américaine considère qu'un ratio égal ou inférieur à 1.89 est satisfaisant et ne nécessite pas de nouvelles installations. Cependant, certaines régions affichent un ratio allant jusqu'à 35, voire plus, indiquant qu'une installation dans ces régions, avec un ratio élevé, représenterait un investissement judicieux pour nos clients.

Conclusion Générale

Dans un monde en perpétuelle évolution, où les préoccupations environnementales et les avancées technologiques façonnent de manière significative les choix des consommateurs, l'industrie automobile se trouve à un point de convergence crucial. Notre projet de "Mise en place d'un système décisionnel pour l'étude du marché des véhicules électriques" s'inscrit dans ce contexte dynamique, où la transition vers la mobilité électrique est devenue une réalité incontournable.

Au terme de notre analyse approfondie et de la mise en œuvre pratique de notre solution, plusieurs conclusions importantes se dégagent. Tout d'abord, nous avons pu identifier les principaux défis auxquels est confronté le marché des véhicules électriques, notamment le manque d'infrastructures de recharge. Cependant, nous avons également identifié des opportunités prometteuses, telles que l'augmentation constante de la demande de véhicules électriques et le développement rapide de l'infrastructure de recharge.

Ensuite, notre démarche méthodologique, allant de l'analyse du contexte à la mise en œuvre pratique, a permis de développer un système décisionnel robuste et pertinent pour étudier le marché des véhicules électriques. Grâce à des outils et des techniques avancés tels que Power BI et le modèle Prophet, nous avons pu transformer des données brutes en informations exploitables, offrant ainsi des perspectives précieuses aux entreprises du secteur.

Enfin, notre projet n'est pas seulement une étude académique, mais aussi une contribution tangible à la transition vers une mobilité plus durable. En proposant des solutions innovantes pour optimiser l'infrastructure de recharge et stimuler l'adoption des véhicules électriques, nous aspirons à jouer un rôle actif dans la construction d'un avenir plus propre et plus durable pour tous.

Bibliographie

- [1] Données des stations de recharge (février 2024)
https://afdc.energy.gov/stations/#/analyze?country=US&fuel=ELEC&ev_levels=all
- [2] PHEV Sales
<https://afdc.energy.gov/data>
- [3] EVs registration data in USA
<https://afdc.energy.gov/data>
- [4] Documentation de Microsoft Power BI
<https://learn.microsoft.com/en-us/power-bi/>
- [5] Emission CO2 des véhicules en USA
<https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions>
- [6] Documentation de la bibliothèque Prophet
https://facebook.github.io/prophet/docs/quick_start.html#python-api