Combien vaut votre voiture ?

Une estimation du prix de revente de votre voiture

[**1.1. Comment ça marche ?**](#_1beom0ama8cz) **3**

[**1.2. À quoi ça sert ?**](#_r1taqryqnxul) **3**

[**2.1. Les données utilisées**](#_vbhrdh2v0r2o) **4**

[**2.2. Comment utiliser ces données ?**](#_zdis7yeeihp) **4**

[**2.3 Architecture du projet**](#_komg0xg1d2ck) **4**

[**3.1. Sprint 3**](#_q7inv1gi9w7a) **5**

[**3.2. Sprint 4**](#_pokq8cbxrqsw) **6**

Introduction

L'industrie automobile en France est un secteur d'activité majeur qui continue de susciter un vif intérêt et d'occuper une place centrale dans la vie quotidienne des Français. Avec plus de 37,8 millions de voitures particulières en circulation, le paysage automobile français est en constante évolution. Cependant, l'achat d'une voiture neuve demeure souvent un investissement financier considérable, ce qui pousse de nombreuses personnes à se tourner vers le marché des véhicules d'occasion. La question de l'achat et de la revente de voitures d'occasion est un enjeu crucial qui concerne des millions de citoyens. Dans ce contexte, notre équipe s'est engagée dans la création d'une application visant à offrir une estimation du prix de revente des véhicules d'occasion, dans le but de simplifier ce processus et de répondre aux besoins des propriétaires de voitures d'occasion en France.

Dans ce contexte, permettez-moi de vous présenter les membres de cette équipe qui ont

mis leurs compétences au service de cette application :

* **Benjamin Colson (Product Owner)** : Il dirige l'équipe en définissant la vision du produit, priorisant les fonctionnalités, et veillant à ce que l'application réponde aux besoins des utilisateurs. Avec une expertise en gestion de projets, il est le moteur derrière notre application de prédiction sur le prix des voitures.
* **Edgar Saiz (Scrum Master)** : En tant que Scrum Master, il assure l'application des méthodologies agiles, favorise la collaboration au sein de l'équipe et garantit que les objectifs du projet sont atteints dans les délais.
* **Charlotte Papelard (Développeuse VBA)** : Elle se spécialise dans le développement de fonctionnalités et d'outils pour l'analyse des données et la création de modèles de prédiction en utilisant VBA, automatisant ainsi le traitement des données.
* **Isra Friaa (Data Engineer)** : Son rôle essentiel est la collecte, la préparation, et la structuration des données qui alimentent l'application, assurant ainsi la disponibilité de données de qualité pour des prédictions précises en matière de prix du véhicule.
* **Luxon Masseau (Data scientist)** : Il se concentre sur l'élaboration de l'algorithme de prédiction au cœur de l'application, en développant des modèles analytiques et statistiques, contribuant ainsi à faire de notre application un outil fiable pour prédire le prix des voitures des automobilistes.

Notre équipe vous propose un rapport explorant la conception, le développement et le résultat d’une telle application.

1. Présentation du projet

L’objectif de ce nouveau projet est de concevoir un outil permettant aux automobilistes d’estimer rapidement le prix de leur véhicule avec des informations simples (trouvables sur une carte grise, obligatoire pour prouver la possession ou le prêt du véhicule). Pour ce faire, nous allons utiliser une base de données obtenue par *“scraping”* sur le site [www.spoticar.fr](http://www.spoticar.fr), qui recense plein d’offres de voitures d’occasion.

## 

## 1.1. Comment ça marche ?

L'utilisateur pourra accéder à notre outil via un formulaire. Dans celui-ci, il renseignera le nom et les caractéristiques de son véhicule. Une fois ça fait, notre algorithme traitera les données et sortira une estimation du prix du véhicule ainsi qu’une comparaison avec les véhicules du même.

## 1.2. À quoi ça sert ?

Le but est de pouvoir donner une estimation rapide du prix d'un véhicule à partir d'informations simples sans passer plusieurs heures de recherches sur des sites spécialisés. Cet outil n'a pas pour but de remplacer une estimation faite par un professionnel de l'automobile mais de pouvoir donner un ordre d'idée en comparant les différentes voitures entre elles.

1. Comment construire un tel outil ?

## 2.1. Les données utilisées

Afin de mener à bien notre projet, nous utilisons une base de données *“scrappée”* . Les avantages de cette technique contrairement aux bases déjà existantes, c’est dans un premier temps le fait de choisir nous même les données qui nous seront utiles pour notre prédiction, le second est le fait qu’il possible de relancer le “*scraping”* afin de ré-obtenir une nouvelle base à jour suivant ainsi les variations du marché.

## 2.2. Comment utiliser ces données ?

Variable intérêt :

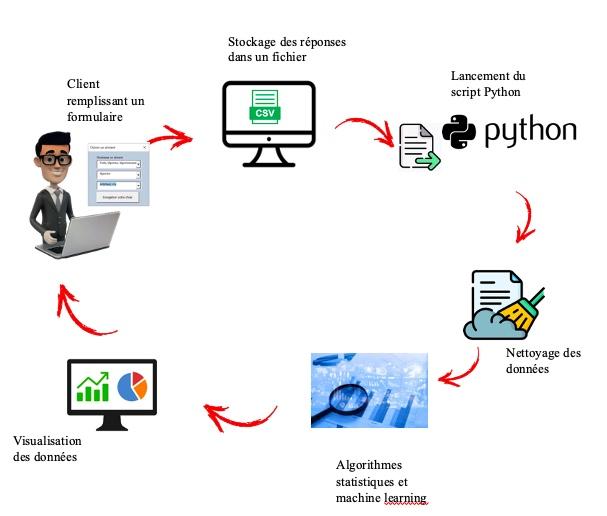
Notre variable d’intérêt est logiquement le prix du véhicule. Toutes les autres variables présentes dans notre base de données servent à définir cette variable. Nous chercherons à créer un graphique qui nous permettrait de comparer le prix prédit avec le prix des véhicules du même modèle afin de pouvoir donner une idée si le véhicule fait partie des meilleurs ou non.

Les variables explicatives :

Voici la liste des variables que nous allons dans un premier temps utilisées et qui peuvent être demandées à l’utilisateur :

* Modèle : Nom de la voiture comprenant la marque de véhicule et son modèle
* Marque : Marque du constructeur automobile ex: Renault, Peugeot …
* Kilométrage : Nombre de kilomètres parcourus par le véhicule
* Carburant : Type de carburant qu’utilise la voiture (Essence, Gasoil, Electrique, …)
* Transmission : Type de boîtier du véhicule (Automatique/Manuel)
* Année : Année sur laquelle le véhicule a été produit

## 2.3 Architecture du projet



3 L’avancée du projet

## 3.1. Sprint 3

**Réalisations :**

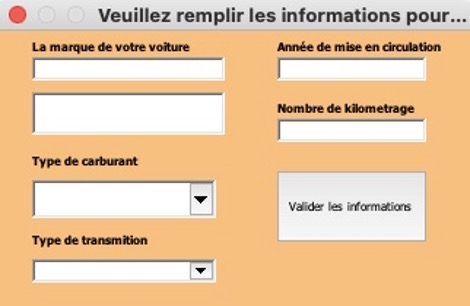
**Product Owner et Scrum Master**

* Mise en place d’une nouvelle application et de leur objectif
* Trouver les différentes options qu l’utilisateur pourra utiliser
* Création de la nouvelle roadmap pour s’adapter aux nouveau projets
* Mise en place des objectifs et DoD
* Listing et explication des variable
* Rédaction du rapport
* Rédaction du ReadMe du git
* Elaboration d’une notice d'utilisation
* Réalisation du Loom

**Dev VBA**

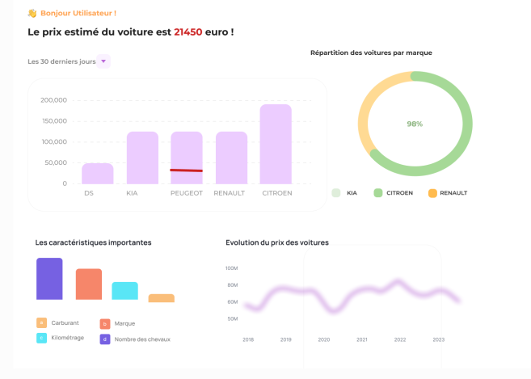
* Scrapping du site Spoticar pour la création de la base de données en temps réel
* Réalisation du formulaire utilisateur en adéquation avec le produit
* Réalisation du PowerPoint en accord avec les autres membres du groupe
* Début de création de graphique pour le dashboard

L’utilisateur remplit le formulaire suivant :



* La **marque** de la voiture doit être entrée à la main et une liste déroulante apparaît avec les premières lettres inscrites. L’objectif pour le prochain sprint c’est d’améliorer cette entrée en créant une sous partie “modèle”, à partir de la marque. L’utilisateur devrait entrer la marque, puis choisir son modèle de voiture.
* Le type de **carburant** et de **transmission** sont renseignés par une liste déroulante. Pas de modifications prévues.
* **L’année** de mise en circulation et le nombre de **kilométrage** sont entrés à la main, et seulement les valeurs numériques sont autorisées. Pour les prochains sprint, on pourrait bloquer à 4 le nombre de chiffres pour l’année de mise en circulation afin d’éviter toute erreur.

Ébauche de l’interface de sortie :



**Dev Data Préparation et Algo de prédiction**

* Familiarisation avec les données
* Préparation de la base avant d'effectuer les traitements
* Réalisation de script de lien entre VBA et Python dans les 2 sens
* Statistiques descriptives uni et bivarié en rapport avec la variable cible
* Réalisation d’une première régression linéaire avec le prix
* Réalisation d’une régression logistique
* Sélection de modèles à l’aide de procédés statistiques
* Mise en place d’un algo de prédiction (reg linéaire) : R^2 = 0,79
* Elaboration d’une ébauche de dashboard

## 3.2. Sprint 4