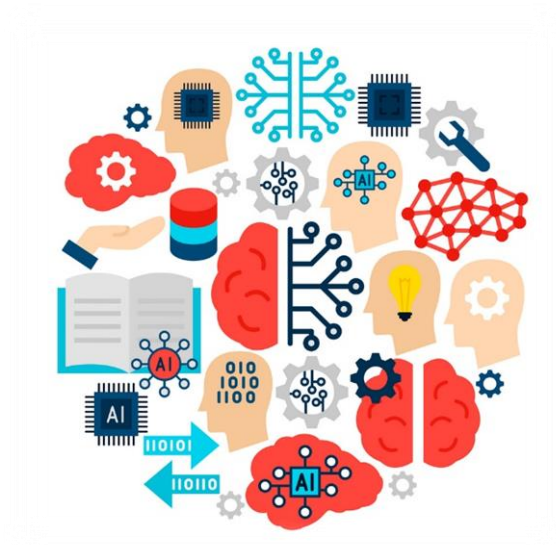


# Lab Data Preparation

## In-Vehicle Coupon Recommendation



*Préparé par :*

*Mohamed Wahchi*

*Malek Hmem*

*Kalthoum Dridi*

*Selim Sahli*

*Nour Hassine*

*Khouloud Zograni*

## Introduction

Les coupons, sous forme de réductions, de bons d'achat ou d'offres spéciales, sont un outil puissant et omniprésent dans le monde du commerce et du marketing. Ils ont une influence significative sur les consommateurs, les entreprises et le marché en général.

Ils favorisent la fidélisation des clients, en influençant la concurrence et les prix, en permettant la collecte de données sur les habitudes d'achat des consommateurs, en impactant le comportement des consommateurs et en permettant le marketing ciblé.

Parmi les différentes approches, les "In-Vehicle Coupons" se distinguent par leur pertinence dans le domaine de la mobilité.

Au vu de la spécificité des "In-Vehicle Coupons" et de leur distribution au sein d'un environnement de mobilité, la question de l'acceptation et de l'utilisation de ces coupons par les conducteurs et les passagers devient essentielle. Ainsi, la problématique centrale de cette introduction pourrait être formulée comme suit :



Comment les facteurs tels que la pertinence, la personnalisation, la confidentialité et la facilité d'accès influencent-ils l'acceptation et l'utilisation des "In-Vehicle Coupons" au sein des véhicules, tout en maximisant leur efficacité dans le domaine de la mobilité ?



## I- Etude du sujet :

### 1. Définition d'un coupon :

Les coupons, sous leurs diverses formes, ont longtemps été un outil incontournable dans le domaine du marketing et du commerce. Ils sont utilisés pour inciter les consommateurs à acheter des produits ou à utiliser des services en offrant des avantages tels que des réductions de prix, des offres spéciales ou des cadeaux. Les coupons ont évolué au fil du temps, passant du papier traditionnel aux versions numériques plus innovantes, tout en maintenant leur rôle crucial dans la stratégie marketing. Dans cette étude, nous allons explorer le concept des coupons et surtout mettre en lumière sur la nouvelle tendance du « In Vehicle Coupon » et leur importance dans le monde du commerce et du marketing.

### 1. Statistiques mondiales sur l'utilisation des coupons :

Les statistiques sur les In-Vehicle Coupons sont encore en développement, car cette forme de marketing est relativement nouvelle. Cependant, certaines tendances émergentes et des données partielles peuvent être relevées à partir de l'utilisation des coupons en général.

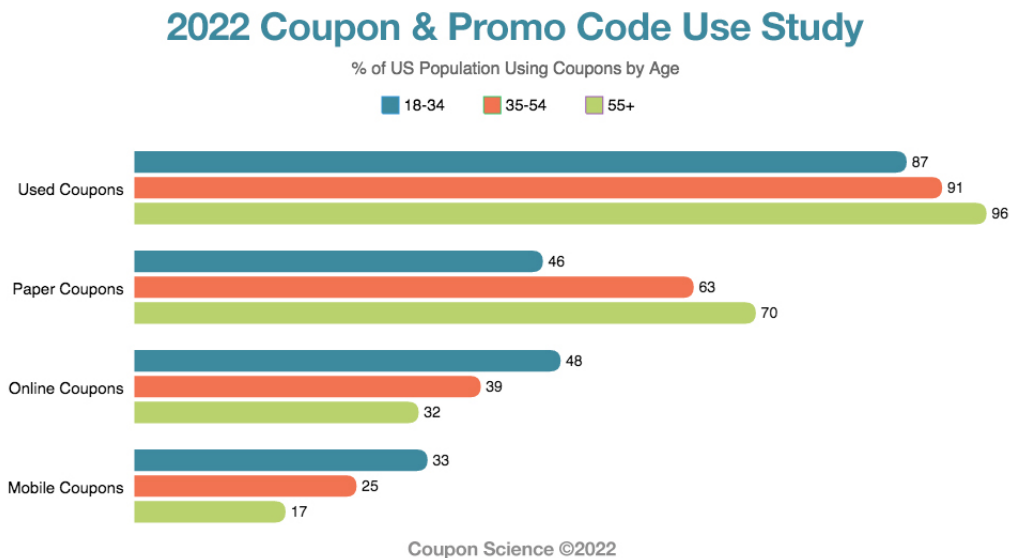
Le graphe ci-dessous représente l'importance de l'utilisation des coupons au Etats Unis :



Ces statistiques effectués lors des années 2017-2020 sur 1000 clients vivant au Etats Unis montrent que les coupons jouent un rôle essentiel dans le monde du commerce et du marketing, et ont été toujours majoritairement utilisés lors de leur shopping.

Les clients ont souvent recours aux coupons lors de leurs achats puisqu'ils représentent un moyen efficace pour les clients d'économiser de l'argent, d'accéder à des offres spéciales, de découvrir de nouveaux produits, et d'améliorer leur expérience de shopping.

Les coupons pourraient aussi prendre différentes formes, et être distribués à partir de divers canaux de communication. En effet, le graphique ci-dessous illustre la diversité des formes de coupons utilisées par toutes les tranches d'âge en Etats Unis lors de l'année 2022.

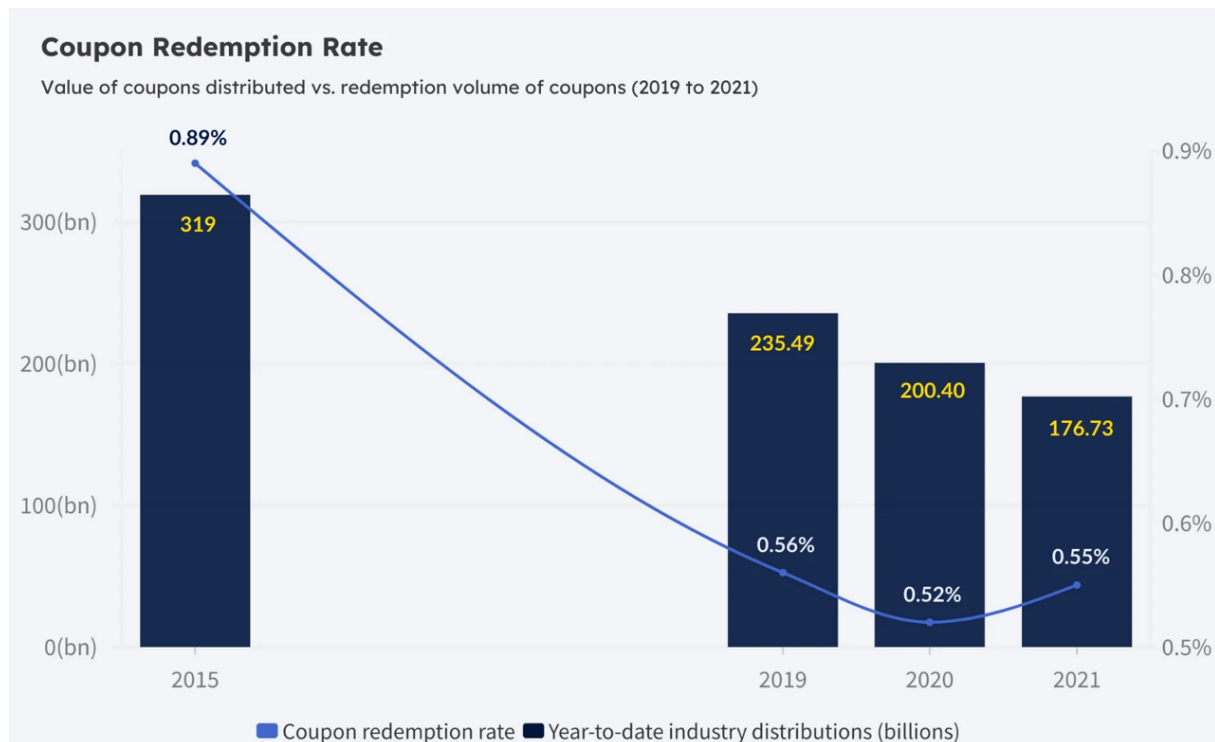


On peut constater que toutes les tranches d'âge ont souvent recours aux coupons. Les coupons traditionnels en papier continuent d'être populaires, avec une utilisation constante chez les consommateurs de toutes générations. Cependant, une tendance claire montre l'adoption croissante des coupons mobiles et en ligne, avec une augmentation significative de l'utilisation parmi les plus jeunes générations.

Cette diversité de canaux de distribution de coupons reflète la manière dont les entreprises qui distribuent ce genre de coupons s'adaptent aux préférences de leur clientèle, offrant une gamme variée d'options pour répondre aux besoins des consommateurs de tous âges.

Cependant, malgré de nombreuses recherches antérieures, la problématique centrale qui demeure concerne la validation et l'acceptation des enquêtes. Pour mieux comprendre ce processus crucial, des enquêtes approfondies sont nécessaires afin d'optimiser ces coupons et d'aider les entreprises à perfectionner leurs techniques de fidélisation.

Pour ce faire, nous allons examiner de près un graphique mettant en lumière les taux d'acceptation de ces coupons lors des années 2015-2021, ce qui nous permettra de dégager des insights essentiels pour notre analyse.



Ce graphe montre que, bien que ces coupons aient le potentiel d'offrir des avantages significatifs aux utilisateurs, leur acceptation varie en fonction de l'année et dépendent sur plusieurs facteurs clés.

Cette variation soulève une question centrale pour notre sujet : Comment pouvons-nous comprendre et influencer les niveaux d'acceptation des "In-Vehicle Coupons" afin d'optimiser leur impact dans le domaine de la mobilité tout en respectant les besoins des utilisateurs et en maximisant les résultats pour les entreprises ?

Cette question sera explorée en analysant les données provenant des enquêtes faites sur le sujet, afin de dégager les facteurs de personnalisation, de pertinence et de convivialité qui façonnent l'acceptation de ces coupons parmi les conducteurs et les passagers.

## 2. Concept des In-Vehicle Coupons :

Le concept des coupons à l'intérieur des véhicules, souvent désignés sous le nom "In-Vehicle Coupons," est un sujet pertinent dans le domaine du marketing, de la publicité et du commerce en constante évolution. Ces coupons sont une forme de promotion qui profite des véhicules comme canal de distribution pour atteindre les consommateurs tout en étant dans leur véhicule.

**Définition :** Les In-Vehicle Coupons sont des offres, des réductions ou des promotions distribuées aux utilisateurs de véhicules, le plus souvent via des canaux numériques, tels que des applications mobiles, des systèmes de navigation intégrés ou des codes QR imprimés.

**Canal de Distribution :** Ils tirent parti des technologies de communication à l'intérieur des véhicules pour diffuser des offres ciblées.

**Objectif :** Stimuler les ventes, la fidélisation des clients, la notoriété de la marque et améliorer l'expérience de conduite.

### 3. Utilisation des In-Vehicle Coupons :

**Distribution ciblée :** Les coupons peuvent être envoyés aux conducteurs en fonction de leur emplacement, de leurs préférences et de leur historique d'achats.

**Personnalisation :** Les offres peuvent être adaptées en fonction des données des utilisateurs, telles que les habitudes de conduite, les destinations fréquentes et les préférences.

**Intégration dans les systèmes de navigation :** Les coupons peuvent être affichés sur l'écran du tableau de bord lorsque le véhicule est à l'arrêt ou pendant le trajet.

**Notification via application mobile :** Les conducteurs peuvent recevoir des notifications push sur leurs smartphones.

**Rétroaction en temps réel :** Les entreprises peuvent recevoir des données sur la performance des coupons en temps réel.

En conclusion, nous pourrions souligner que les In-Vehicle coupons est une excellente idée pour promouvoir le marketing de plusieurs entreprises, et celles-ci devront continuer à surveiller les tendances, à développer des méthodes d'analyse et à s'adapter aux préférences changeantes des consommateurs pour maximiser l'efficacité des In-Vehicle Coupons et leur acceptation par les clients.

## II- Business Problem :

**Comment les données recueillies à partir des coupons distribués à l'intérieur des véhicules peuvent-elles être exploitées pour optimiser l'efficacité des campagnes marketing et l'acceptation des coupons, tout en améliorant l'expérience des utilisateurs et en maximisant les résultats commerciaux ?**

Cette problématique pose des questions importantes sur l'utilisation des données provenant des "In-Vehicle Coupons" pour comprendre le comportement des consommateurs, personnaliser les offres, cibler les bons publics et mesurer l'impact de ces coupons sur la

croissance de l'entreprise. Elle soulève également des préoccupations quant à la manière dont les entreprises peuvent équilibrer le respect de la vie privée des utilisateurs tout en maximisant le pourcentage d'acceptation de ces coupons pour leurs clients.

Prédire si un client acceptera ou non un coupon représente un défi complexe. Dans cette problématique, notre objectif est de prédire si un client acceptera ou refusera le coupon qui lui est offert, en se basant sur son profil, son comportement et son environnement.

### III- Business Understanding :

Le "business understanding" est une phase essentielle dans le processus de gestion des données et de l'analyse de données, qui vise à bien comprendre les besoins et les objectifs d'une entreprise avant de collecter, traiter et analyser des données

Ces données ont été recueillies à la suite d'une enquête visant à obtenir des informations approfondies sur le comportement et les préférences des individus en ce qui concerne l'utilisation de coupons dans divers contextes de conduite. Cette enquête revêt une importance significative pour plusieurs raisons :

Tout d'abord, elle permet de mieux comprendre comment des facteurs tels que la durée du trajet, les choix de restaurants et l'emplacement géographique influencent la décision d'accepter ou de refuser les coupons. Ces informations sont cruciales pour les entreprises, notamment les restaurants et les bars, car elles leur permettent de cibler plus efficacement leurs efforts promotionnels, d'améliorer l'acquisition de nouveaux clients et de renforcer la fidélisation de la clientèle.

De plus, ces données peuvent être précieuses pour les entreprises cherchant à créer des segments de clientèle en fonction des comportements et des préférences. En segmentant la clientèle de manière plus précise, les entreprises peuvent personnaliser davantage leurs offres pour chaque groupe, ce qui améliore la pertinence des coupons proposés.

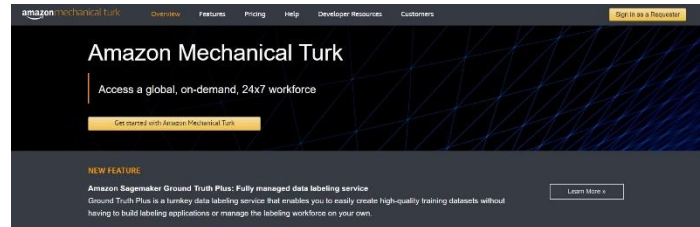
Enfin, cette enquête vise à identifier les types de coupons et d'offres les plus susceptibles d'être acceptés en fonction des préférences des clients. Cette connaissance permet aux entreprises de concevoir des offres plus attirantes, augmentant ainsi les chances que les clients les acceptent.

Dans l'ensemble, l'objectif de cette prédiction est d'anticiper si un client acceptera ou refusera un coupon pour un lieu spécifique, en se basant sur des caractéristiques démographiques et contextuelles. Cette classification binaire ( $Y=1$  pour acceptation et  $Y=0$  pour refus) représente un enjeu clé pour les entreprises cherchant à optimiser leur stratégie marketing et à améliorer leur relation avec les clients.



## IV- Data Understanding :

Ces données ont été obtenues grâce à une enquête publiée en 2017 sur Amazon Mechanical Turk et rendue accessible sur le site de l'UCI.



L'enquête recueille des informations concernant divers scénarios de conduite, notamment la destination de l'utilisateur, l'heure actuelle, les conditions météorologiques, la présence de passagers, les caractéristiques des coupons, les caractéristiques de l'utilisateur, ainsi que les caractéristiques contextuelles. **Ensuite, elle interroge l'utilisateur pour savoir s'il acceptera ou non le coupon.**

Le fichier de jeu de données se nomme "in-vehicle-coupon-recommendation.csv". Il comprend plus de 12 684 points de données et 26 attributs, la plupart étant de nature catégorielle, et quelques-uns sont de type catégoriel numérique.

En ci-dessous, se trouve l'explication de la dataset et ses attributs. En commençant, nous allons explorer les variables d'entrées (les inputs).

### 1) Les données sur l'utilisateur :

Nom de la colonne :	Description :
<b>Gender</b>	<b>Le sexe de l'utilisateur :</b> Male / Femelle (Donnée de type string)
<b>Age</b>	<b>L'âge moyen de l'utilisateur :</b> 21/46/26/31/41/50plus/36/below21 (Donnée de type string)
<b>Marital Status</b>	<b>L'état civil de l'utilisateur:</b> Unmarried partner, Single, Married partner, Divorced, Widowed (Donnée de type string)
<b>Has_Children</b>	<b>Variable indiquant si l'utilisateur a des enfants :</b> 1:has children/0:no children (donnée de type int)
<b>Education</b>	<b>Le niveau d'éducation de l'utilisateur:</b> Some college — no degree, Bachelors degree, Associates degree, High School Graduate, Graduate degree (Masters or Doctorate), Some High School (Donnée de type string)



<b>Occupation</b>	<b>L'occupation que l'utilisateur occupe:</b> Unemployed, Architecture & Engineering, Student, Education & Training & Library, Healthcare Support, Healthcare Practitioners & Technical, Sales & Related, Management, Arts Design Entertainment Sports & Media, Computer & Mathematical, Life Physical Social Science, Personal Care & Service, Community & Social Services, Office & Administrative Support, Construction & Extraction, Legal, Retired, Installation Maintenance & Repair, Transportation & Material Moving, Business & Financial, Protective Service, Food Preparation & Serving Related, Production Occupations, Building & Grounds Cleaning & Maintenance, Farming Fishing & Forestry (Donnée de type string)
<b>Income</b>	<b>Le salaire annuel de l'utilisateur :</b> Less than \$12500, \$12500 — \$24999, \$25000 — \$37499, \$37500 — \$49999, \$50000 — \$62499, \$62500 — \$74999, \$75000 — \$87499, \$87500 — \$99999, \$100000 or More (Donnée de type string)
<b>Car</b>	<b>Le véhicule que l'utilisateur conduit :</b> Scooter and motorcycle, crossover, Mazda5 (Donnée de type string)
<b>Bar</b>	<b>Combien de fois l'utilisateur fréquente le bar chaque mois :</b> never, less1, 1~3, 4~8, gt8, nan (Donnée de type string)
<b>CoffeeHouse</b>	<b>Combien de fois l'utilisateur fréquente un salon de thé chaque mois :</b> never, less1, 1~3, 4~8, gt8, nan (Donnée de type string)
<b>CarryAway</b>	<b>Combien de fois l'utilisateur prend sa commande à emporter chaque mois :</b> never, less1, 1~3, 4~8, gt8, nan (Donnée de type string)
<b>RestaurantLessThan20</b>	<b>Combien de fois l'utilisateur fréquente un restaurant avec un montant inférieur à 20\$ chaque mois :</b> never, less1, 1~3, 4~8, gt8, nan (Donnée de type string)
<b>Restaurant20To50</b>	<b>Combien de fois l'utilisateur fréquente un restaurant avec un montant moyen de 20\$ à 50\$ chaque mois :</b> never, less1, 1~3, 4~8, gt8, nan (Donnée de type string)

## 2) Les données contextuels :

Nom de la colonne :	Description :
<b>Destination</b>	<b>La destination de l'utilisateur :</b> No Urgent Place, Home, Work (donnée de type string)
<b>Passenger</b>	<b>Les passagers qui se trouvent avec l'utilisateur dans la voiture:</b> Alone, Friend(s), Kid(s), Partner (donnée de type string)
<b>Weather</b>	<b>L'état météorologique lorsque l'utilisateur est en train de conduire :</b> Sunny, Rainy, Snowy (donnée de type string)
<b>Temperature</b>	<b>La temperature en Fahrenheit lorsque l'utilisateur est en train de conduire :</b> 55,80,30 (donnée de type int)
<b>Time</b>	<b>L'heure lorsque l'utilisateur est en train de conduire :</b> 2PM, 10AM, 6PM, 7AM, 10PM (donnée de type string)
<b>toCoupon_GEQ5min</b>	<b>Variable qui indique si la distance en voiture jusqu'au restaurant/bar pour utiliser le coupon est supérieure à 5 minutes :</b> 0,1 (donnée de type int)
<b>toCoupon_GEQ15min</b>	<b>Variable qui indique si la distance en voiture jusqu'au restaurant/bar pour utiliser le coupon est supérieure à 15 minutes :</b> 0,1 (donnée de type int)
<b>toCoupon_GEQ25min</b>	<b>Variable qui indique si la distance en voiture jusqu'au restaurant/bar pour utiliser le coupon est supérieure à 25 minutes :</b> 0,1 (donnée de type int)
<b>Direction_same</b>	<b>Variable qui indique si le restaurant/bar se trouve dans la même direction que la destination actuelle de l'individu :</b> 0,1 (donnée de type int)
<b>Direction_opp</b>	<b>Variable qui indique si le restaurant/bar se trouve dans la même direction que la destination opposée de l'individu :</b> 0,1 (donnée de type int)

## 3) Les données sur le coupon :

Nom de la colonne :	Description :
<b>Coupon</b>	<b>Type du coupon offert par l'entreprise :</b> Restaurant(<\$20), Coffee House, Carry out & Take away, Bar, Restaurant\$20-\$50 (donnée de type string)
<b>Expiration</b>	<b>Délais d'expiration du coupon:</b> 1d,2h (donnée de type string)

#### 4) L'attribut cible (output) :

L'attribut cible, représenté par la colonne Y, est d'une importance capitale pour notre analyse. Il s'agit d'une variable binaire qui indique si le coupon a été accepté (représenté par la valeur 1) ou refusé (représenté par la valeur 0).

Cette variable nous permet de mesurer directement le succès ou l'acceptation des "In-Vehicle Coupons" par les utilisateurs, ce qui est au cœur de notre enquête. La compréhension de cette variable cible sera essentielle pour notre analyse ultérieure, alors que nous chercherons à identifier les facteurs qui influencent les décisions d'acceptation ou de refus des coupons.

## Conclusion

La réalisation de notre recherche a été marquée par un processus rigoureux, commençant par la formulation de notre problématique centrée sur l'acceptation des "In-Vehicle Coupons." Notre démarche s'est ensuite déployée en plusieurs étapes, notamment la phase de "Business Problem" et "Business Understanding" qui ont souligné l'importance de comprendre la problématique et les besoins des entreprises et des utilisateurs dans ce contexte spécifique.

Nous avons également exploré la phase de "Data Understanding," au cours de laquelle nous avons examiné les données relatives aux taux d'acceptation des coupons, éclairant ainsi notre analyse. Ce processus a jeté les bases pour notre démarche ultérieure, qui consistera en une phase cruciale de "Data Cleaning" et de "Data Preparation."

Cette phase est essentielle pour garantir que nos données sont fiables, complètes et prêtes à être analysées. Nous veillerons à éliminer les valeurs aberrantes, à gérer les données manquantes et à les structurer de manière à ce qu'elles soient prêtes pour l'analyse.

En conclusion, cette recherche nous a permis de poser des questions fondamentales sur l'acceptation des "In-Vehicle Coupons" dans le domaine de la mobilité. Notre démarche méthodique et structurée nous a préparés à entreprendre la phase d'analyse, au cours de laquelle nous explorerons les facteurs qui influencent l'acceptation de ces coupons.