

PROJET PYTHON

# ANALYSE DE PACK-SKI-FIX SUR LE SITE DE GLISSHOP

---

ABICHOU Nour Elhouda  
2022/2023



[Lien vers la présentation animée](#)

[https://www.canva.com/design/DAFSBn49-PY/1430HdJyhkuu4pq0RYI8GA/edit?utm\\_content=DAFSBn49-PY&utm\\_campaign=designshare&utm\\_medium=link2&utm\\_source=sharebutton](https://www.canva.com/design/DAFSBn49-PY/1430HdJyhkuu4pq0RYI8GA/edit?utm_content=DAFSBn49-PY&utm_campaign=designshare&utm_medium=link2&utm_source=sharebutton)



# Plan



- 1 Introduction et objectif
- 2 Stratégie de travail
- 3 Présentation de Dataset
- 4 Analyse des graphiques
- 5 Conclusion générale



1

## Introduction et objectif

Le e-commerce connaît actuellement et depuis la crise sanitaire de Covid-19, une période de prospérité. Ce qui fait, les entreprises de e-commerce doivent attacher une importance particulière aux avis de leurs clients sur ses produits. En effet, Un avis peut avoir un impact sur les ventes de l'entreprise.

Ce projet traite une partie de ce phénomène des avis dans le secteur de vente en ligne, ce qui reflète précisément notre objectif qui est le suivant :

**Déterminer l'impact de la notation d'un produit dans une catégorie donnée sur son attractivité ainsi que l'impact des autres caractéristiques sur sa notation.**

Dans ce contexte et vue le rapprochement de la saison de ski en France, on a décidé de travailler sur une catégorie de produit formé par des de packs ski incluant skis alpins et fixations.



2

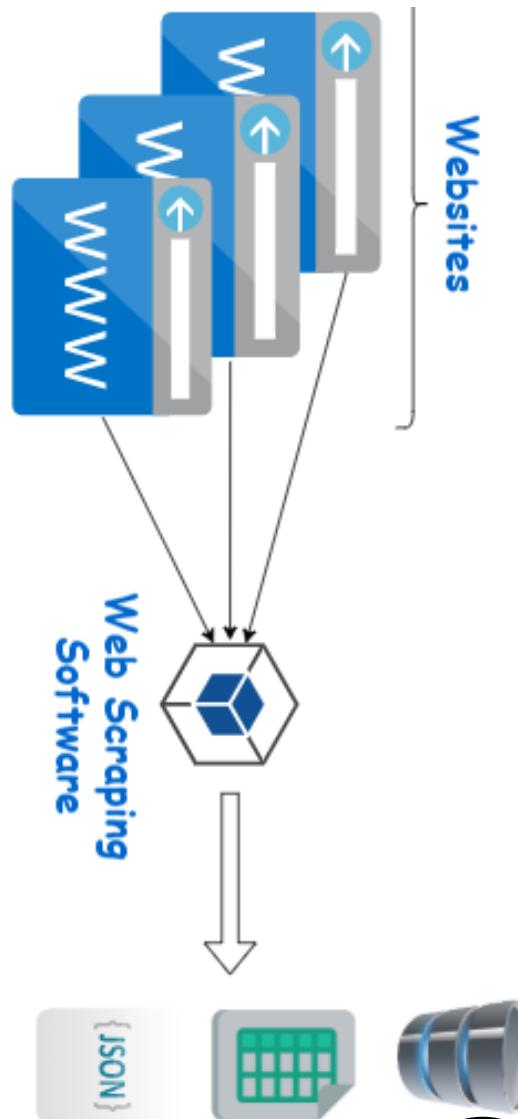


2

## Stratégie de travail

Notre stratégie de travail est basée sur trois étapes :

- Le choix de site et la sélection de son URL : on va scraper le site Glisshop, un site qui dispose d'une grande sélection de produits de ski (de références en ski alpin, snowboard...)
- Inspection de la page : détecter les différents balises à extraire pour effectuer l'analyse du téléchargement, du code HTML de la page à scraper, et son parsing
- Récupération de contenu du site web : il faut simplement envoyer une requête HTTP et recevoir la réponse. Une réponse sous la forme du code HTML de la page via son parseur.
- Coder le script de scraping : notre script est basé sur deux bibliothèques de la langue Python, BeautifulSoup pour parser les documents HTML ainsi que pour les hiérarchiser afin de faciliter la recherche et l'extraction des données. Pour ce faire, ce package travaille avec le package requests en lui transmettant l'URL à consulter. On obtient en retour un objet Réponse, possédant de nombreux attributs.
- Stocker les données sous le format requis : on va stocker les données initialement au format CSV après le nettoyage et les stocker en format Excel.
- Traitements et manipulation des données en utilisant les packages Pandas et numpy, réalisation de graphiques via les packages de matplotlib et seaborn et l'analyse de données.



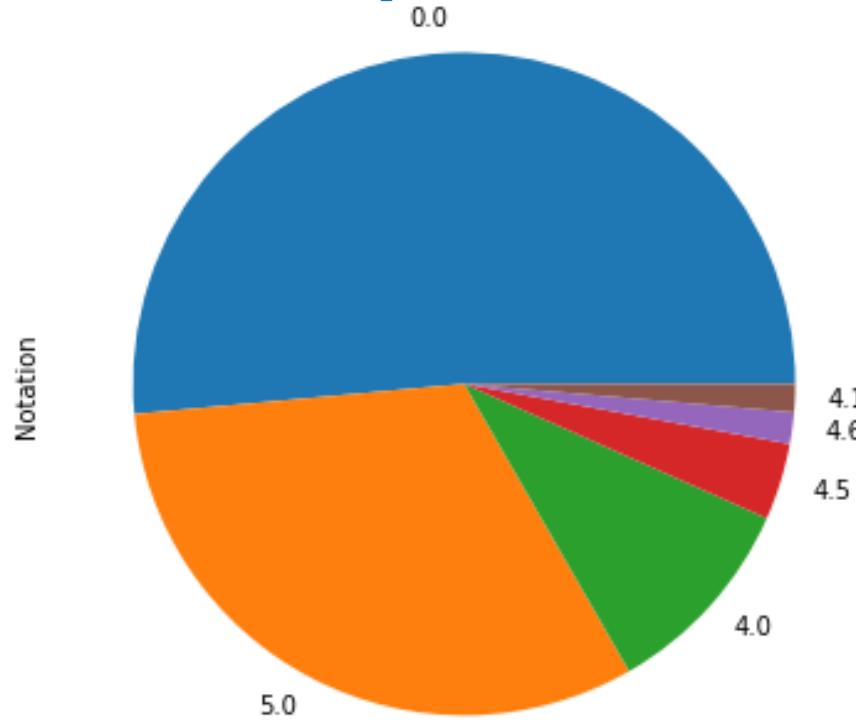


	Produit	La marque	Description produit	Tailles disponibles	Nombre de tailles	Prix initial	Remise	Prix final	Notation	Nombre avis
0	Pack Ski Alpin Experience 80 Carbon Xpress + Fix	Rossignol	Le ski Rossignol Experience 80 Carbon Xpress e...	150 158 166 174 182	5	489.9	0	489.90	5.0	5
1	Pack Ski Alpin Atris + Fix	Black Crows	Modèle phare chez Black Crows l'Atris revient ...	178 184 189	3	819.90	-17	679.90	5.0	5
2	Pack Ski Alpin Captis + Fix	Black Crows	Le Black Crows Captis est un ski all mountain ...	172 184	2	679.90	-13	584.90	5.0	1
3	Pack Ski Alpin Disruption 78 Ti Quikclik Hp +...	K2	Le K2 Disruption 78Ti Quikclik Hp est un petit...	156 163 170 177	4	819.90	-39	499.90	5.0	5

Notre base de données est composé de dix variables :

- Produit : le nom de chaque produit qui appartient a la catégorie de Pack ski +fix
- La marque : la marque de chaque produit de ski
- Description produit : une brève description de chaque produit
- Tailles disponibles : les différents types proposés pour chaque produit de ski
- Nombre de tailles : les nombres de tailles disponibles par produit
- Prix initial : le prix initial avant l'application de la réduction si elle existe
- Remise : Le pourcentage de la réduction proposé par produit
- Prix final : le prix de produit après la déduction des remises octroyées si elles existent
- Notation : la note attribuée à chaque produit
- Nombre avis : le nombre d'avis publié sur chaque produit

**Graphique 1: Camembert de fréquence de chaque notation**



**Interprétation :**

Le graphique montre que notre base de données est composé majoritairement de plus que 52%.

la deuxième note la plus fréquente est la note 5 sur 5 avec 32% suivi par la note 4 avec 10% et finalement les fréquences des autres notation sont presque similaire et ils présentent les pourcentage respectifs : 4%, 1% et 1%.

**Graphique 2: Statistique descriptive**

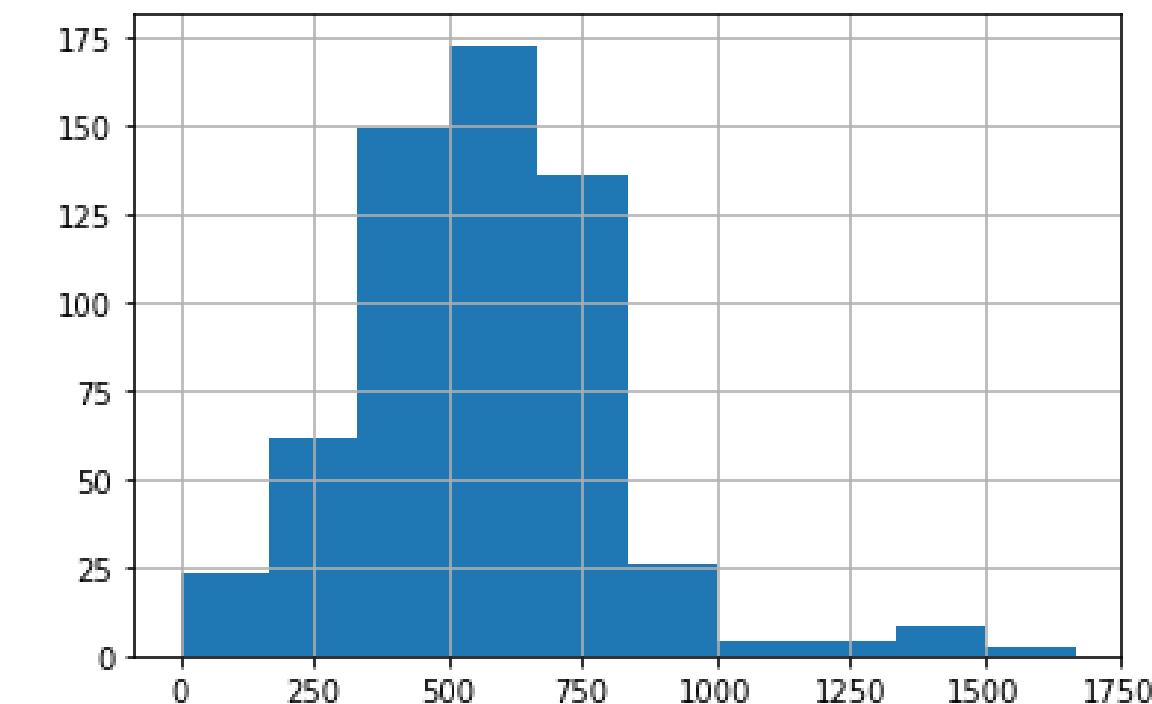
	Nombre de tailles	Remise	Prix final	Notation	Nombre avis
count	587.000000	587.000000	587.000000	587.000000	587.000000
mean	2.759796	-13.480409	569.570256	2.289949	1.577513
std	1.110861	11.569830	233.975995	2.377106	2.438746
min	1.000000	-51.000000	0.000000	0.000000	0.000000
25%	2.000000	-20.000000	431.900000	0.000000	0.000000
50%	3.000000	-13.000000	584.900000	0.000000	0.000000
75%	4.000000	0.000000	685.400000	5.000000	2.000000
max	6.000000	0.000000	1669.900000	5.000000	9.000000

**Interprétation :**

Dans le tableau de statistique descriptive on se focalise sur les valeurs suivantes :

- Chaque produit propose au moyen presque trois tailles disponibles. Cette variable varie entre au minimum un et six tailles par produit.
- Le taux de réduction moyen est de presque 14% Une note moyenne supérieure à 2 attribuée à un produit donné.
- Le nombre des avis moyen est presque égale à deux avis et il varie entre 0 et 9 par produit.
- Le prix final moyen est à peu près égale à 570 euros. On constate que 75% de produits ont un prix inférieur à 685 euros.

**Graphique 3: Histogramme de réparation de prix**

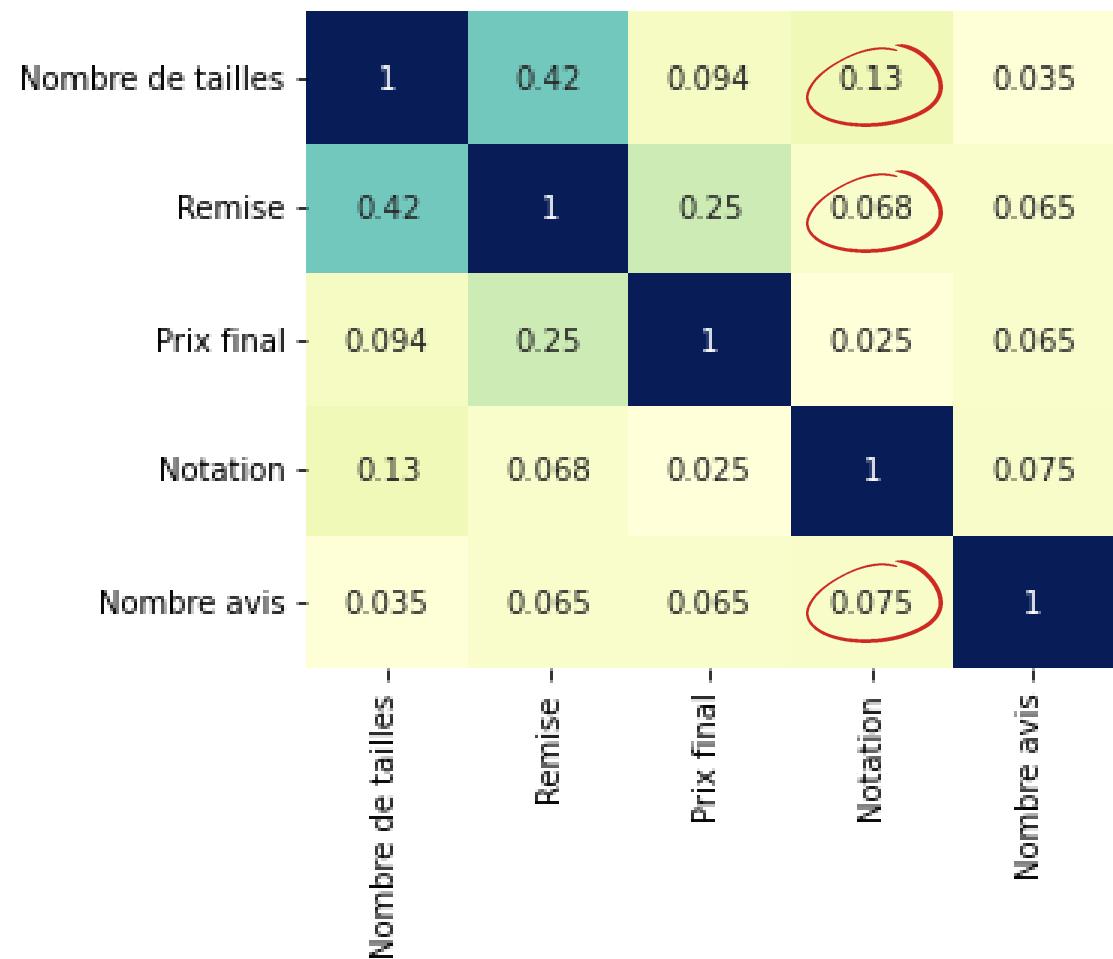


**Interprétation :**

L'histogramme de la réparation de prix montre que ce dernier présente une distribution normale autour de la moyenne.

Le prix varie principalement entre 280 euros et 800 euros. Les valeurs les plus rares de prix varient entre 1000 et 1669 euros.

**Graphique 4: Tableau de corrélation entre les variables quantitatives**



**Interprétation :**

Le tableau de corrélations nous permettant d'estimer des relations entre deux variables quantitatives. On va présenter les corrélations les élèves et qui sont en relation avec la notation :

- Une faible corrélation et positive entre la notation et le nombre de avis, en autre terme on peut dire que l'augmentation de nombre des avis va accroître la note attribue.
- Une corrélation faible et positive entre la notation et la remise accordée de 7%. On peut dire que l'augmentation de remise pourra augmenter la note accordée par le client.
- Une autre corrélation positive entre la notation et les tailles disponibles pour chaque produit, autrement dit, un produit proposant un choix large de taille va avoir comme conséquence une augmentation de sa notation.

**Graphique 5: Top 10 de marque**

La marque	Notation	Classement
Rossignol	162.5	1
Salomon	116.5	2
Black Crows	112.1	3
Dynastar	107.7	4
K2	96.7	5
Atomic	88.5	6
Volkl	79.5	7
Elan	73.0	8
Head	68.0	9
Fischer	66.7	10

**Interprétation :**

Après le traitement de notre données scapées, on a décidé de classer les marques existantes en fonction d'une note globale. Le calcul de cette note est basé sur la somme des notes accordés a chaque produit qui appartient a une marque donnée.

Selon les données disponibles on constate que la marque la plus noté est la marque Rossignol avec un score total de 162.5.

Ensuite en second rang la marque Salomon. En dernier rang on trouve la marque Fischer avec une globale de 66.7.

**Graphique 6: Les meilleurs produits classés en fonction de note et de nombre des avis**

		Produit	La marque	Description produit	Tailles disponibles	Nombre de tailles	Prix initial	Remise	Prix final	Notation	Nombre avis
475	Pack Ski Alpin Revo 82 + Fix	Movement		Le Revo 82 est un ski all mountain de la marqu...	153 160 168	3	749.0	-29	525.0	5.0	9
200	Pack Ski Alpin Kore X 80 Lyt-Pr + Fix	Head		Du fait de sa construction à la fois légère et...	163 170 177	3	539.9	0	539.9	5.0	9
220	Pack Ski Alpin Redster Q5 + Fix	Atomic		Le ski all-mountain Atomic Redster Q5 est lége...	154 161 169 177	4	549.9	0	549.9	5.0	9
241	Pack Ski Alpin Stance 96 + Fix	Salomon		Le Salomon Stance 96 est un ski polyvalent et ...	176 182	2	779.9	-35	499.9	5.0	9
263	Pack Ski Alpin Ripstick 86 TW Junior + Fix	Elan		Le modèle pour enfant Elan Ripstick 86 TW Juni...	138 148 158 168	4	459.9	-32	309.9	5.0	9
286	Pack Ski Alpin Belle Dc 72 + Fix	Nordica		Le ski alpin femme Nordica Belle DC 72 est un ...	150 156 162 168	4	669.9	0	669.9	5.0	9
310	Pack Ski Alpin Cochise 106 + Fix	Blizzard		Le Cochise 106 de chez Blizzard est un ski de ...	177 185 192	3	929.9	-12	809.9	5.0	9
361	Pack Ski Alpin Fly 105 + Fix	Movement		Vous êtes un freeskier qui apprécie particuli...	177 184	2	729.9	-13	629.9	5.0	9
388	Pack Ski Alpin Kore 87 + Fix	Head		Le Kore 87 de chez Head est un ski all mountai...	170 177 184 191	4	829.9	-13	719.9	5.0	9

**Interprétation :**

Le tableau ci-dessus est une extraction du tableau récapitulant les 20 meilleurs produits ayant obtenu la note de 5 sur 5 qui ont été classées selon leurs nombres d'avis. Comme de nombreux autres produits possèdent une note de 5 et un nombre d'avis de 9, nous ne pourrons pas classer les différents produits dans un ordre particulier, mais nous pouvons considérer que la table présente les meilleurs articles en les classant dans l'ordre alphabétique.

Nous constatons une relation avec les analyses précédentes. En effet, les produits figurant dans ce tableau ont des prix appartenant à la fourchette le plus fréquemment rencontrée dans notre histogramme de prix. Aussi, ils offrent au minimum deux pointures pour chaque produit, ce qui peut partiellement justifier la corrélation entre le nombre de pointures disponibles et la note accordée.



## 5 Conclusion générale

On conclu que :

- L'importance des avis de clients ne doit pas être ignorer par les entreprises puisqu'ils peuvent construire un manque à gagner pour eux.
- Plusieurs facteurs peuvent impacter soit positivement soit négativement la note accordée à un produit. Selon notre analyse ont trouver comme facteur : la réduction accordée, les taille ainsi que le nombre de avis par produit.
- Notre analyse est limitée par la qualité notre échantillon est déséquilibrer puisqu'il majoritairement (52%) par des produits sans notations ce qui peut erroné les analysés effectués.
- Finalement, le scraping est une méthode utile pour la surveillance des sites de e-commerce. Il permet de recueillir des informations de caractère commercial concernant les produits des compétiteurs dans le but de renforcer considérablement la compréhension et la visibilité du marché.
- MAIS, La question qui se pose toujours par rapport au scraping est : **Est-ce illégal ?**



Merci pour votre attention

ABICHOU NOUR ELHOUDA

