

Rapport d'une Étude Statistique sur les Ordinateurs Portables

Ce rapport présente les résultats d'une étude statistique approfondie sur les ordinateurs portables, en analysant les caractéristiques clés telles que la RAM, le poids, le prix, la densité de pixels (PPI), la capacité du SSD, la marque et le processeur. L'étude vise à explorer les relations entre ces variables, identifier les tendances et les corrélations, et à proposer des conclusions et des recommandations basées sur les données analysées.

Sous la guidance compétente de : Mme H. FERJOUCHIAI

Realisee par :

EZ-ZAHERY Ahmed Amine

AIT SIDI AHMED Walid

ATILA Hammam

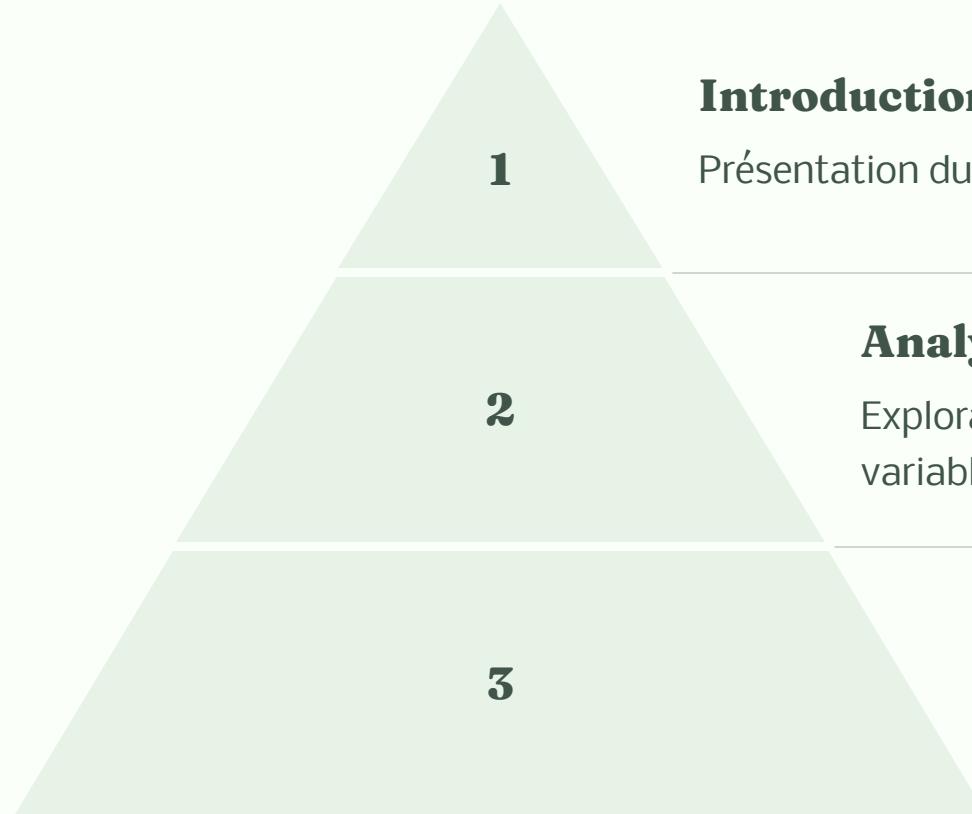
BELAZRI Hamza

HAMMOU Abderrahmane

EL MEHDAOUI Chaimaa

KARIM Ayoub

Sommaire du Rapport



Introduction

Présentation du contexte et des objectifs de l'étude.

Analyse

Exploration approfondie des données et des relations entre les variables.

Résultats et Conclusion

Présentation des résultats clés et des conclusions de l'étude.

Ce rapport présente une étude statistique approfondie sur les ordinateurs portables. Il explore les relations entre les variables, identifie les tendances et les corrélations, et propose des conclusions et des recommandations basées sur les données analysées.

Collecte et Préparation des Données

Sources de Données

L'étude a utilisé des données provenant de plusieurs sources, y compris des sites Web de vente au détail, des plateformes de comparaison de prix et des fabricants.

Nettoyage et Transformation

Les données brutes ont été nettoyées, transformées et préparées pour l'analyse, ce qui implique la suppression des valeurs manquantes et la normalisation des données.

Sélection des Variables

Les variables pertinentes pour l'étude ont été soigneusement sélectionnées, telles que la RAM, le poids, le prix, la densité de pixels (PPI), la capacité du SSD, la marque et le processeur.

Introduction à l'Étude Statistique des Ordinateurs Portables

Dans un monde où la technologie évolue rapidement, les ordinateurs portables sont des outils essentiels pour la productivité, l'éducation et le divertissement. Leurs performances, leur portabilité et leur accessibilité financière influencent fortement les choix des consommateurs.

Cette étude vise à analyser et comparer divers ordinateurs portables en fonction de paramètres clés tels que la RAM, le poids, le prix (INR), la densité de pixels (PPI), la capacité SSD, le fabricant (marque) et la marque du processeur (CPU).

Grâce à des techniques statistiques, nous examinerons les relations et les tendances entre ces variables. Par exemple, une densité de pixels élevée est-elle toujours associée à un prix plus élevé ?

Collecte et Préparation des Données

La première étape de l'étude a consisté à collecter un ensemble de données représentatif des ordinateurs portables disponibles sur le marché. Ces données ont été extraites d'une base de données de Kaggle du lien suivant :<https://www.kaggle.com/datasets/abdocan/laptop-prices>. La base de données a été nettoyée et prétraitée pour supprimer les valeurs manquantes, les doublons et les incohérences, assurant la cohérence et la fiabilité des données.

Cette étude vise à analyser et comparer divers ordinateurs portables en fonction de paramètres clés tels que la **RAM**, le **poids**, le **prix (INR)**, la **densité de pixels (PPI)**, la **capacité SSD**, le **fabricant (marque)** et la **marque du processeur (CPU)**.

Analyse Descriptive des Variables

L'analyse descriptive a permis d'obtenir une vue d'ensemble des caractéristiques des ordinateurs portables. Des statistiques descriptives, telles que la moyenne, la médiane, l'écart type, le minimum, le maximum et la variance, ont été calculées pour chaque variable. Cette analyse a révélé des informations précieuses sur la distribution, la tendance centrale et la dispersion des données.

Par exemple, l'analyse du prix des ordinateurs portables a montré que la moyenne des prix était de [51986.2506], avec un écart type de [33842.68855]. Cela indique une certaine variabilité dans les prix des ordinateurs portables. De plus, l'analyse de la RAM a montré que la plupart des ordinateurs portables avaient une capacité de RAM de [8], avec une variation significative en fonction des modèles et des marques.

L'analyse descriptive a également révélé des tendances intéressantes en ce qui concerne les marques et les processeurs. Les marques les plus populaires étaient [DELL], tandis que les processeurs les plus répandus étaient ceux de [Intel Core i5]. Ces informations fournissent un aperçu précieux du marché des ordinateurs portables et de ses principaux acteurs.

Analyse de la Corrélation entre les Variables

L'analyse de la corrélation a permis d'explorer les relations linéaires entre les différentes variables. Le coefficient de corrélation de Pearson a été utilisé pour mesurer la force et la direction de la corrélation entre chaque paire de variables.

Analyse et Résultats de Tableau de la corrélation

coefficient de corrélation entre	la valeur
RAM et Weight	0.106597292
RAM et Price_INR	0.617905481
RAM et ppi	0.272401781
RAM et SSD	0.487535785
Weight et Price_INR	0.028313803
Weight et ppi	-0.337588676
Weight et SSD	-0.068327565
Price_INR et ppi	0.604757087
Price_INR et SSD	0.671715447
ppi et SSD	0.505489681

Une corrélation positive indique que deux variables augmentent ou diminuent ensemble, tandis qu'une corrélation négative indique que les variables varient en sens inverse. Par exemple, une forte corrélation positive entre le prix et la RAM suggère que les ordinateurs portables avec une RAM élevée ont généralement un prix plus élevé. De même, une corrélation négative entre le poids et la densité de pixels suggère que les ordinateurs portables plus légers ont tendance à avoir une densité de pixels plus élevée.

Méthodes de Calculs de Corrélation

La corrélation permet d'analyser la relation entre deux variables, en mesurant à quel point elles sont liées. Pour chaque couple de variables étudiées, la fonction utilisée est `=CORREL(Variable1; Variable2)`, qui calcule le coefficient de corrélation linéaire de Pearson. Ce coefficient varie entre -1 et 1, indiquant une relation négative, nulle ou positive entre les variables. Les résultats obtenus offrent une meilleure compréhension des dépendances ou indépendances entre les caractéristiques étudiées, comme la RAM, le poids, le prix, le PPI et la capacité SSD, et d'autres variables associées.

Correlation	Fonction Utilisée
RAM/*	"=CORREL(B:B, *:*)"
Weight/*	"=CORREL(C:C, *:*)"
Price_INR/*	"=CORREL(D:D, *:*)"
ppi/*	"=CORREL(E:E, *:*)"
SSD/*	"=CORREL(F:F, *:*)"

L'analyse de la corrélation a également révélé des relations intéressantes entre les variables catégorielles, telles que la marque et le processeur. Par exemple, il est possible de trouver des corrélations significatives entre certaines marques et la capacité du SSD ou le type de processeur utilisé.

Tableaux et Visualisations des Résultats

Les résultats de l'analyse statistique ont été présentés dans des tableaux et des graphiques clairs et concis. Les tableaux ont fourni des informations détaillées sur les statistiques descriptives, les coefficients de corrélation et les résultats des tests statistiques. Les graphiques, tels que des histogrammes, des diagrammes en boîte et des nuages de points, ont permis de visualiser les relations entre les variables et de mettre en évidence les tendances importantes.

Les tableaux et les graphiques ont été soigneusement conçus pour être facilement compréhensibles et interprétables. Ils ont permis de présenter les résultats de l'analyse de manière claire et concise, facilitant la compréhension des conclusions de l'étude.

Laptop Price



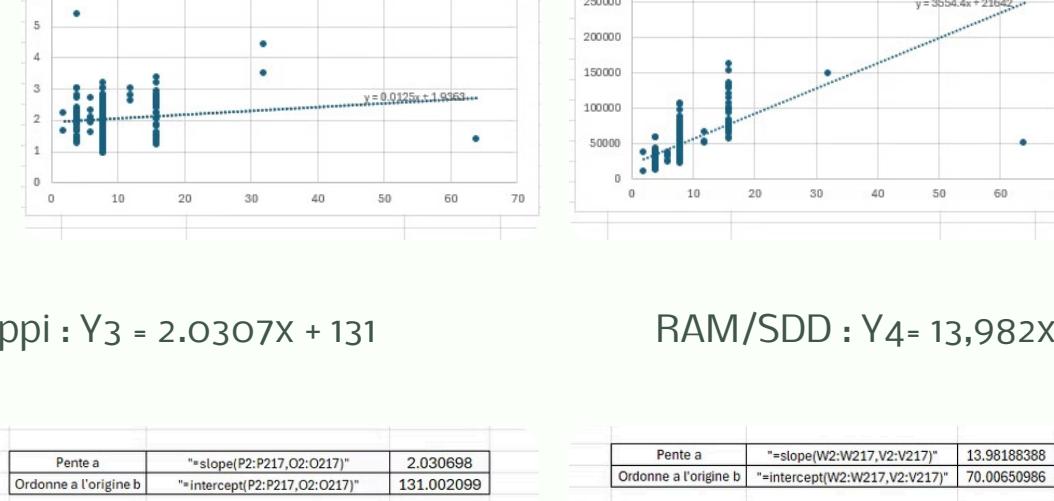
Laptop PAM	\$7. MIT	\$3. WB	\$1AMZ	\$2. AM	\$1. MS
plen	\$1,901	\$7430	\$0100	2503	\$1,950
par	\$1,550	\$1,000	\$2550	\$100	\$1,950
ptam	\$2570	\$1,000	\$4100	3,000	\$1,203
part	\$1,970	\$750	\$2550	\$200	\$1,990

Régressions Linéaires et Interprétation des Coefficients

Les droites de regressions :

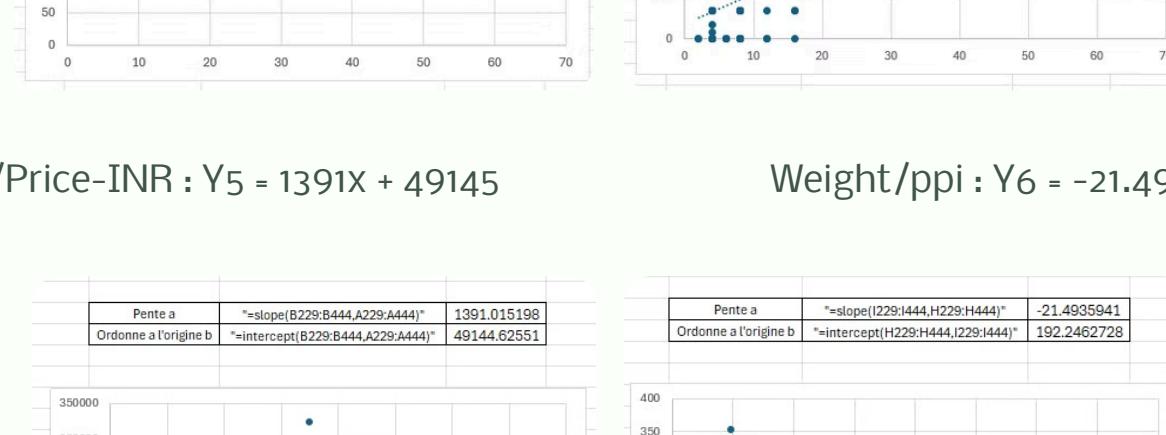
$$\text{RAM/Weight : } Y_1 = 0.0125X + 1.9363$$

$$\text{RAM/Price_INF : } Y_2 = 3554.4X + 21642$$



$$\text{RAM/ppi : } Y_3 = 2.0307X + 131$$

$$\text{RAM/SDD : } Y_4 = 13.982X + 70.007$$



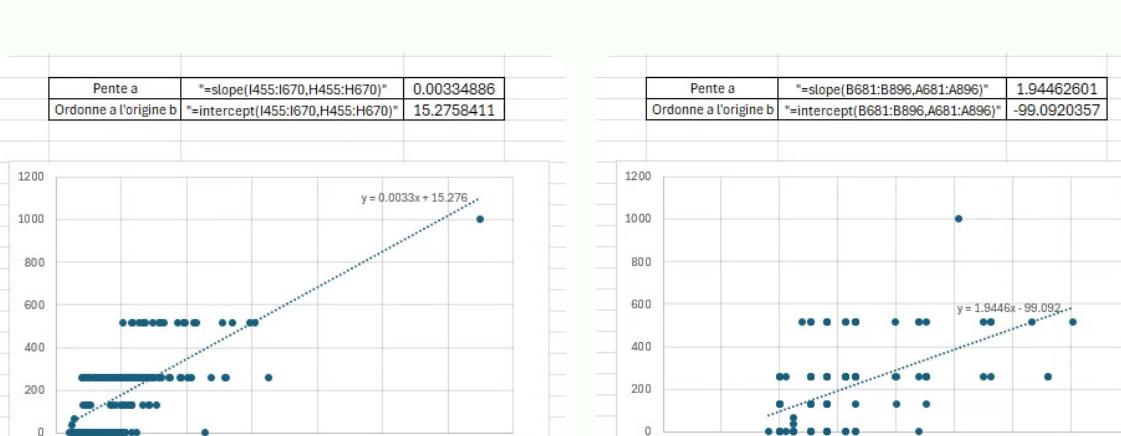
$$\text{Weight/Price-INR : } Y_5 = 1391X + 49145$$

$$\text{Weight/ppi : } Y_6 = -21.494X + 192.25$$



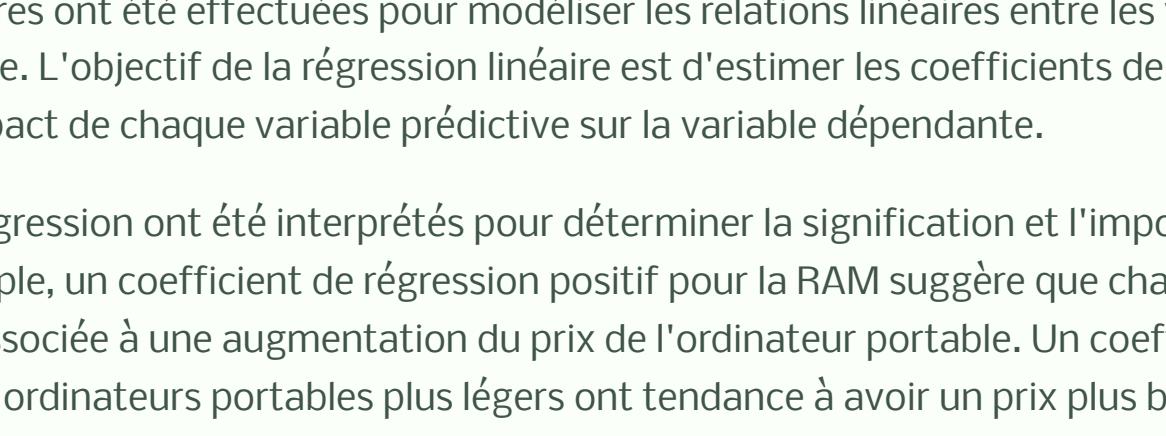
$$\text{Weight/SDD : } Y_7 = -16.736X + 223.56 + 107.59$$

$$\text{Price-INR/ppi : } Y_7 = 0.0008X + 107.59$$



$$\text{ppi/SDD : } Y_{10} = 0.0033X + 15.276$$

$$\text{Price-INR/SSD : } Y_9 = 1.9446X + 99.092$$



Des régressions linéaires ont été effectuées pour modéliser les relations linéaires entre les variables prédictives et la variable dépendante. L'objectif de la régression linéaire est d'estimer les coefficients de la droite de régression, qui représentent l'impact de chaque variable prédictive sur la variable dépendante.

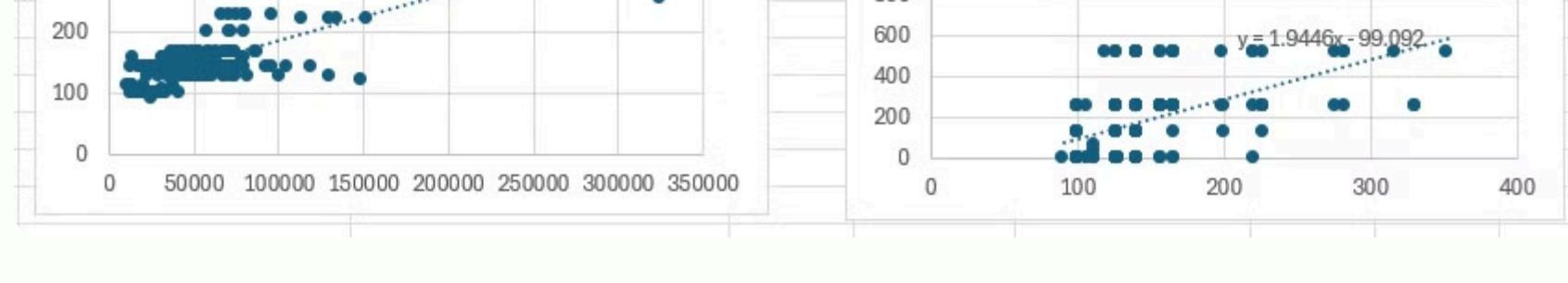
Les coefficients de régression ont été interprétés pour déterminer la signification et l'importance des variables prédictives. Par exemple, un coefficient de régression positif pour la RAM suggère que chaque augmentation d'unité de RAM est associée à une augmentation du prix de l'ordinateur portable. Un coefficient négatif pour le poids suggère que les ordinateurs portables plus légers ont tendance à avoir un prix plus bas.

L'analyse de la régression linéaire a permis d'identifier les facteurs les plus importants qui influencent le prix des ordinateurs portables, ainsi que de prévoir le prix d'un ordinateur portable en fonction de ses caractéristiques.

Ressources Excel :

Méthodes de Calcul des droites de regressions :

Pente a	Ordonne à l'origine b
"=SLOPE(plage des données)"	"=INTERCEPT(plage des données)"



Analyse de la Variance, Tests Statistiques, Moyenne, Médiane, Mode, Variance, Écart Type, Aplatissement, Asymétrie, Min, Max et Total

Variable	Moyenne	Médiane	Mode	Variance	Écart Type	Aplatissement	Asymétrie	Min	Max	Total
RAM (Go)	[8.53 7037 037]	[8]	[8]	[34.4 52331 96]	[5.88 32453 09]	[37.6 81835 78]	[4.70 32737 43]	[2]	[64]	[1844]
Weight(Kg)	[2.04 28425 93]	[2.02]	[2.2]	[0.47 23316 7]	[0.68 8860 342]	[16.2 02342 34]	[2.66 93881 85]	[0.92]	[7.2]	[441.2 54]
Prix (INR)	[5198 6.250 6]	[4632 6.96]	[3724 2.72]	[1140 02512 16]	[3384 2.688 55]	[19.5 9430 212]	[3.137 5454 06]	[1060 2.72]	[3249 54.72]	[1122 9030. 13]
PPI	[148. 33824 34]	[141.2 11998 1]	[141.2 11998 1]	[1914. 6501]	[43.8 58356 66]	[5.971 8884 36]	[2.16 84577 5]	[90.5 83401 72]	[352. 46514 72]	[3204 1.060 58]
SSD (Go)	[189. 37037 04]	[256]	[256]	[2833 5.899 86]	[168. 72372 31]	[1.526 4002 09]	[0.82 9098 868]	[0]	[1000]	[409 04]

Des tests statistiques, tels que les tests t et les analyses de variance, ont été effectués pour comparer les moyennes des groupes et pour déterminer si les différences entre les groupes étaient significatives. Ces tests ont permis de valider les conclusions tirées de l'analyse descriptive et de la régression linéaire.

Pour la **RAM (Go)**, la moyenne est de 8,54 Go, ce qui indique une taille de mémoire vive typique proche de cette valeur. La médiane et le mode, tous deux égaux à 8 Go, montrent que cette valeur est également la plus fréquente et se situe au milieu de la distribution. L'écart type (5,88) et la variance (34,45) reflètent une certaine variabilité, indiquant des différences significatives entre les capacités de RAM des appareils, allant de 2 Go à un maximum de 64 Go. L'asymétrie positive (4,70) et un aplatissement élevé (37,68) indiquent une distribution fortement concentrée sur les valeurs faibles avec quelques valeurs très élevées.

Pour le **poids (Kg)**, la moyenne est de 2,04 kg, proche de la médiane (2,02 kg) et du mode (2,2 kg), ce qui signifie que la plupart des appareils pèsent environ 2 kg. L'écart type (0,69) et la variance (0,47) montrent une faible variabilité, tandis que l'asymétrie (2,67) et l'aplatissement (16,20) suggèrent la présence de quelques appareils très légers ou très lourds, bien que les poids varient globalement entre 0,92 kg et 7,2 kg.

Concernant le **prix (INR)**, la moyenne est de 51 986,25 INR, avec une médiane plus basse (46 326,96 INR) et un mode encore plus bas (37 242,72 INR). Cela indique une distribution asymétrique positive (3,14), où la majorité des appareils coûtent moins que la moyenne, bien que certains modèles haut de gamme (jusqu'à 324 954,72 INR) tirent cette moyenne vers le haut. L'écart type (33 842,69) et la variance élevée (114 002 512,66) traduisent une très grande dispersion des prix.

Pour le **PPI (pixels par pouce)**, la moyenne est de 148,34 avec une médiane et un mode identiques à 141,21, ce qui montre une répartition symétrique autour de ces valeurs. La variance (1914,65) et l'écart type (43,86) indiquent une variabilité modérée. L'asymétrie (2,17) et l'aplatissement (5,97) montrent une distribution légèrement biaisée avec quelques valeurs très élevées (jusqu'à 352,47 PPI).

Enfin, pour le **SSD (Go)**, la moyenne est de 189,37 Go, avec une médiane et un mode de 256 Go, ce qui montre que cette capacité est la plus courante. La variance (28 335,90) et l'écart type (168,72) révèlent une forte variabilité, allant de 0 Go (sans SSD) jusqu'à 1000 Go. L'asymétrie (0,83) et l'aplatissement (1,53) indiquent une distribution globalement équilibrée avec une légère concentration sur les valeurs faibles.

Méthodes de Calcul des Indicateurs

Les différentes statistiques présentées dans le tableau ont été calculées à l'aide de fonctions spécifiques. La **moyenne** a été obtenue avec la fonction =AVERAGE(plage des données), qui calcule la valeur moyenne des données. La **médiane**, correspondant à la valeur centrale, a été calculée avec =MEDIAN(plage des données), tandis que le **mode**, représentant la valeur la plus fréquente, a été déterminé avec =MODE(plage des données). La **variance** et l'**écart type**, mesurant respectivement la dispersion et l'écart moyen autour de la moyenne, ont été calculés à l'aide des fonctions =VAR.P(plage des données) et =STDEV.P(plage des données).

Pour les mesures de forme de la distribution, l'**aplatissement** (kurtosis) a été calculé avec =KURTOSIS(plage des données), et l'**asymétrie** (skewness) avec =SKEW(plage des données). Les valeurs **minimale** et **maximale** ont été déterminées grâce aux fonctions =MIN(plage des données) et =MAX(plage des données), tandis que la somme totale des données a été obtenue avec =SUM(plage des données). Ces fonctions permettent de fournir une analyse descriptive complète et précise des données.

K	Fonction Utilisée
"=AVERAGE(plage des données)"	
"=MEDIAN(plage des données)"	
"=MODE(plage des données)"	
"=VAR.P(plage des données)"	
"=STDEV.P(plage des données)"	
"=KURTOSIS(plage des données)"	
"=SKEW(plage des données)"	
"=MAX(plage des données)"	
"=MIN(plage des données)"	
"=SUM(plage des données)"	

Globalement, les données montrent une grande diversité dans les caractéristiques étudiées, en particulier pour les prix et les capacités SSD, tandis que le poids et le PPI restent relativement constants avec une moindre variabilité.

Conclusion et Recommandations

L'étude statistique sur les ordinateurs portables a révélé des relations intéressantes entre les variables analysées. Le prix des ordinateurs portables est fortement influencé par la RAM, la capacité du SSD, la marque et le processeur. Les ordinateurs portables plus légers ont tendance à avoir une densité de pixels plus élevée. Les marques les plus populaires sont [Insérer les marques], tandis que les processeurs les plus répandus sont ceux de [Insérer les marques de processeurs].

Sur la base des conclusions de l'étude, voici quelques recommandations pour les fabricants et les consommateurs d'ordinateurs portables :

Les fabricants devraient continuer à investir dans le développement de technologies de pointe pour améliorer la RAM, la capacité du SSD et les processeurs, tout en minimisant le poids des ordinateurs portables.

Les consommateurs devraient tenir compte des caractéristiques clés, telles que la RAM, la capacité du SSD, la marque et le processeur, lorsqu'ils choisissent un ordinateur portable.

Les fabricants et les consommateurs devraient se tenir au courant des dernières tendances du marché et des technologies émergentes, afin de faire des choix éclairés.