

Paid-plan students		
	minutes_watches_21	minutes_watches_22
Moyenne	33,80	273,02
Mediane	26,33	40,28
Ecartype standard	28,21	854,58
Kurtosis	-0,85	58,48
Skewness	0,63	7,07

Moyenne : Parmi les étudiants ayant regardé entre 1 et 100 minutes en 2021, le temps moyen de visionnage des abonnés a fortement augmenté du T4 2021 au T4 2022, passant d'environ 33,80 à 273,02 minutes. Cela suggère une hausse significative de l'engagement chez ces étudiants initialement peu impliqués.

Médiane : Le temps médian de visionnage de ces abonnés peu engagés a également augmenté du T4 2021 au T4 2022, passant de 26,33 à 40,28 minutes. Bien que moins spectaculaire que l'augmentation de la moyenne, cela indique une hausse généralisée de l'engagement, non limitée à quelques cas particuliers.

Écart-type : L'écart-type de ces abonnés peu engagés a considérablement augmenté, passant de 28,21 minutes au T4 2021 à 854,58 minutes au T4 2022. Cela signifie une plus grande variabilité dans les minutes regardées, reflétant une gamme étendue d'engagements, certains regardant peu et d'autres beaucoup.

En résumé, les résultats suggèrent une augmentation significative de l'engagement des abonnés peu impliqués en 2022. Cependant, la plus grande variabilité indique une diversité d'engagements. Comprendre les raisons derrière cette variabilité pourrait fournir des idées pour stimuler davantage l'engagement, en s'inspirant des facteurs motivants des étudiants ayant considérablement augmenté leur participation.

Free-plan students		
	minutes_watches_21	minutes_watches_22
Mean	25,39	117,64
Median	14,17	11,83
Ecartype standard	26,23	468,93
Kurtosis	0,36	315,76
Skewness	1,17	15,06

Moyenne : Parmi les étudiants ayant regardé entre 1 et 100 minutes en 2021, le temps moyen de visionnage des utilisateurs en plan gratuit est passé d'environ 25,39 minutes au T4 2021 à environ 117,64 minutes au T4 2022. Cela indique une augmentation globale de l'engagement chez ces étudiants initialement peu engagés en plan gratuit, bien que cette augmentation soit moins prononcée que celle observée chez les utilisateurs en plan payant.

Médiane : De manière intéressante, le temps médian de visionnage de ces étudiants en plan gratuit a diminué du T4 2021 au T4 2022, passant de 14,17 minutes à 11,83 minutes. Cela suggère une baisse de l'engagement pour l'étudiant typique de ce groupe. L'augmentation de la moyenne pourrait être due à quelques utilisateurs en plan gratuit ayant fortement augmenté leur engagement au T4 2022, tandis que la majorité n'a pas augmenté ou a même réduit son engagement.

Écart-type : L'écart-type pour les étudiants en plan gratuit peu engagés a augmenté de 26,23 minutes au T4 2021 à 468,93 minutes au T4 2022, indiquant une plus grande variabilité dans les minutes regardées. Le comportement de ces étudiants est devenu plus diversifié au T4 2022, avec certains regardant beaucoup de contenu et d'autres très peu.

En résumé, bien que la moyenne suggère une augmentation de l'engagement chez les utilisateurs en plan gratuit, la diminution de la médiane et l'augmentation de l'écart-type indiquent une complexité. Un petit nombre d'utilisateurs en plan gratuit pourrait avoir fortement augmenté leur engagement, tandis que la majorité n'a pas suivi. Cela pourrait nécessiter des stratégies ciblées pour stimuler l'engagement chez la majorité des utilisateurs en plan gratuit initialement peu engagés.

La dissymétrie

Pour les étudiants en plan payant, la dissymétrie (skewness) est passée de 0,63 au T4 2021 à 7,07 au T4 2022.

La dissymétrie pour les étudiants en plan gratuit est passée de 1,17 au T4 2021 à 15,06 au T4 2022, indiquant une dissymétrie positive.

La moyenne est plus grande que la médiane dans une distribution inclinée vers la droite car la queue de la distribution tire la moyenne vers la droite. Cette observation est confirmée par les valeurs moyennes et médianes des deux années. Une dissymétrie croissante suggère que de plus en plus d'étudiants regardent significativement plus de contenu que la plupart au fil du temps, tirant la moyenne vers le haut.

Dans les deux cas, la moyenne est plus élevée que la médiane ($33,80 > 26,33$ en 2021 et $273,02 > 40,28$ en 2022).

La kurtosis

Pour les étudiants en plan payant, la kurtosis est passée de -0,85 au T4 2021 à 58,48 au T4 2022.

La kurtosis a augmenté chez les étudiants en plan gratuit, passant de 0,36 au T4 2021 à 315,76 au T4 2022.

Les valeurs de kurtosis supérieures à 0 indiquent que les données ont des queues plus lourdes et un sommet plus prononcé que la distribution normale (leptokurtique). Une distribution leptokurtique a une kurtosis positive élevée, suggérant qu'elle est très pointue et a un nombre relativement important de valeurs aberrantes. Ce type de distribution a une fréquence plus élevée de valeurs extrêmes ou de valeurs aberrantes. L'augmentation de la kurtosis au fil du

temps suggère plus de cas extrêmes dans les données au T4 2022 par rapport au T4 2021, en particulier pour les étudiants en plan gratuit.

Dans l'ensemble, l'augmentation de la dissymétrie et de la kurtosis pour les deux groupes du T4 2021 au T4 2022 suggère un nombre croissant d'étudiants regardant significativement plus de contenu que la majorité. Cela est particulièrement vrai pour les étudiants en plan gratuit, avec une dissymétrie et une kurtosis plus élevées au T4 2022 que pour les étudiants en plan payant.

Intervalle de confiance

Paid-plan students		
	minute_watches_21	minute_watches_22
Nombre	3433,00	5104,00
Moyenne	332,50	368,35
Écart type standard	485,86	596,41
limite inférieure	316,25	351,99
limite supérieure	348,76	384,72
L'erreur standard	8,29	8,35

Free-plan students		
	minute_watches_21	minute_watches_22
Nombre	32 171,00	120 658,00
Moyenne	133,93	69,15
Écart type standard	367,26	255,62
limite inférieure	129,92	67,71
limite supérieure	137,95	70,59
L'erreur standard	2,05	0,74

Pour les étudiants en plan payant, l'engagement a augmenté du T4 2021 au T4 2022, confirmé par l'intervalle de confiance pour la moyenne des minutes regardées passant de 316,25 à 348,76 minutes (T4 2021) à 351,91 à 384,72 minutes (T4 2022). On peut ainsi être confiant à 95 % que la moyenne réelle des minutes regardées par tous les étudiants en plan payant a augmenté.

Chez les étudiants en plan gratuit, l'intervalle de confiance pour la moyenne des minutes regardées a diminué de 129,92 à 137,95 minutes (T4 2021) à 67,71 à 70,59 minutes (T4 2022). On peut être confiant à 95 % que la moyenne réelle des minutes regardées par tous les étudiants en plan gratuit a diminué.

Comparaison entre étudiants en plan payant et gratuit (T4 2022) :

Les étudiants en plan payant regardent significativement plus que ceux en plan gratuit, selon l'intervalle de confiance pour la moyenne des minutes regardées qui est de 351,99 à 384,72 minutes (payant) et de 61,71 à 70,59 minutes (gratuit). On peut être confiant à 95 % que les étudiants en plan payant ont regardé beaucoup plus de minutes.

Ces conclusions basées sur les intervalles de confiance ne déterminent pas directement la cause de l'augmentation ou de la diminution de l'engagement. Par exemple, l'engagement plus élevé des abonnés en plan payant peut être dû à des fonctionnalités supplémentaires ou à un contenu

plus attractif, plutôt qu'à une simple incitation financière. De même, la diminution de l'engagement chez les utilisateurs en plan gratuit pourrait être influencée par divers facteurs tels que des changements dans la plateforme ou une concurrence accrue.

Test d'hypothèses

Test d'égalité des variances (F-Test)		
Paid-plan students		
	<i>minute_watches_21</i>	<i>minute_watches_22</i>
Moyenne	331,7329371	367,621456
Variance	234098,3641	353024,041
Observations	3432	5103
Degré de liberté	3431	5102
F	0,663123008	
P(F<=f) unilatéral	0	
Valeur critique pour F (unilatéral)	0,949789955	
Free-plan students		
	<i>minute_watches_21</i>	<i>minute_watches_22</i>
Moyenne	133,7908589	69,09569888
Variance	134233,0253	65018,16883
Observations	32170	120657
Degré de liberté	32169	120656
F	2,064546384	
P(F<=f) unilatéral	0	
Valeur critique pour F (unilatéral)	1,014667355	

On a effectué un test F à deux échantillons pour les variances afin de prouver l'hypothèse de variances inégales entre les échantillons des abonnés en plan gratuit et payant.

La p-value indique la probabilité d'obtenir la valeur f observée si l'hypothèse nulle (variances égales) était vraie. Les variances d'échantillon ne sont pas identiques car la p-value dans les deux cas est de 0.

Ensuite, nous utilisons un test t à une queue en supposant des variances inégales pour les étudiants payants et gratuits.

Paid-plan students

Test d'égalité des espérances: deux observations de variances différentes

	2973,67	4110,17
Moyenne	331,7329371	367,621456
Variance	234098,3641	353024,041
Observations	3432	5103
Différence hypothétique des moyennes	0	
Degré de liberté	8228	
Statistique t	-3,061806083	
P(T<=t) unilatéral	0,001103579	
Valeur critique de t (unilatéral)	1,645038841	
P(T<=t) bilatéral	0,002207159	
Valeur critique de t (bilatéral)	1,960252343	

	minutes watched 21	minutes watched 22
Moyenne	332,50	368,35
Ecart type standard	485,86	596,41
Nombre de valeur	3433	5104

Free-plan students

Test d'égalité des espérances: deux observations de variances différentes

	4716,68	6338,07
Moyenne	133,7908589	69,09569888
Variance	134233,0253	65018,16883
Observations	32170	120657
Différence hypothétique des moyennes	0	
Degré de liberté	40833	
Statistique t	29,80526492	
P(T<=t) unilatéral	1,9754E-193	
Valeur critique de t (unilatéral)	1,644890945	
P(T<=t) bilatéral	3,9508E-193	
Valeur critique de t (bilatéral)	1,960022083	

	minutes watched 21	minutes watched 22
Moyenne	133,93	69,15
Ecart type standard	367,26	255,62
Nombre de valeur	32171	120658

Suit au résultat, pour les étudiants en plan payant, nous rejetons l'hypothèse nulle car le t-statistique (-3,06) calculé est inférieur à la valeur critique (1,64). Le t-statistique négatif indique que la moyenne des minutes regardées par les étudiants au T4 2021 est significativement inférieure à celle des étudiants au T4 2022.

Chez les étudiants en plan gratuit, nous ne rejetons pas l'hypothèse nulle car le t-statistique (29,78) calculé est supérieur à la valeur critique (1,64). Cela signifie qu'il n'y a pas suffisamment de preuves pour conclure que μ_1 est inférieur à μ_2 . Ainsi, les données soutiennent l'hypothèse nulle selon laquelle μ_1 est plus grand ou égal à μ_2 .

Comparaison entre étudiants US et IN

	minute_watches_22_US	minute_watches_22_IN
Moyenne	73,07	78,42
Ecartype standard	308,56	319,34
Nombre	6 459,00	21 210,00

Test d'égalité des espérances: deux observations de variances différentes

	US	IN
Moyenne	73,07631465	78,42450469
Variance	95223,17116	101980,237
Observations	6458	21209
Différence hypothétique des moyennes	0	
Degré de liberté	10999	
Statistique t	-1,209473456	
P(T<=t) unilatéral	0,113253499	
Valeur critique de t (unilatéral)	1,644992176	
P(T<=t) bilatéral	0,226506998	
Valeur critique de t (bilatéral)	1,960179689	

Test d'égalité des variances (F-Test)

	US	IN
Moyenne	73,07631465	78,42450469
Variance	95223,17116	101980,237
Observations	6458	21209
Degré de liberté	6457	21208
F	0,933741419	
P(F<=f) unilatéral	0,000354528	
Valeur critique pour F (unilatéral)	0,967312251	

La valeur p indique la probabilité d'obtenir la valeur f observée si l'hypothèse nulle (variances égales) était vraie. Les variances d'échantillon ne sont pas identiques car la valeur p est inférieure à 0. Nous devons effectuer un test t unilatéral en supposant des variances inégales.

Le t-statistique calculé est supérieur à la valeur critique et la valeur p est supérieure au niveau de signification spécifié α (0,05), donc nous rejettons l'hypothèse nulle. Si l'hypothèse selon laquelle les étudiants américains regardent autant ou moins de contenu que les étudiants indiens est rejetée, cela suggère que les étudiants américains regardent en moyenne moins de contenu que les étudiants en Inde.

En conclusion, nous trouvons facilement la différence de marché entre US et IN. Ces détails pourraient indiquer que la plateforme est plus engageante ou pertinente pour les étudiants en Inde que pour les étudiants américains. De plus, ces résultats sont utiles lors de la décision d'allouer des ressources et opportunités de croissance.