Rapport de recherche

ÉTUDE OBSERVATIONNEL, ET ANALYTIQUE, DE LA FRÉQUENCE D'UTILISATION DE DUOLINGO AU COURS DE LA JOURNÉE EN FONCTIONS DU PAYS.

Asurix · **Github** · **Courriel** · **Site**

Abstract
Cette étude analyse la fréquence d'utilisation de la plateforme Duolingo dans plusieurs pays au cours d'une journée, dans le contexte particulier du défi mondial lancé le 11 février 2025 pour « ressusciter » la mascotte Duo. À l'aide d'une méthodologie de collecte de données, l'étude examine l'évolution des scores globaux et par pays. Les résultats mettent en évidence une augmentation linéaire de l'expérience totale, avec une domination marquée des États-Unis, suivis par la Chine et l'Allemagne. Mais un engagement beaucoup plus marqué par la population tchèque. L'analyse temporelle révèle une utilisation similaire de l'application dans tout les pays. Ces similarités suggèrent que les comportements d'apprentissage des utilisateurs sont liés aux mécaniques misent en œuvres par l'application, mécanismes surpassant les routines personnelles et la culture propre de chaque pays.

Keywords
Duolingo, Apprentissage, Langues, Analyse de données, Comportement des utilisateurs.

Table des matières

1 Introduction **3**

2 Méthodologie commune. **4**

3 Résultats **5**

3.1 Analyse des tendances globales 5

3.2 Analyse des différences par horaire au niveau local 7

Résumé

4 Discussion **11**

4.1 Différence d'horaire entre l'utilisation de Duolingo et du sommeil. 11

4.2 Domination des États-unis dans le score 12

4.3 L'engagement remarquable de la Tchéquie 12

4.4 Limites de l'étude 12

5 Conclusion **13**

6 Remerciements **14**

7 Bibliographie **15**

1. Introduction

Duolingo est un site web et une application mobile d'apprentissage par le jeu, fonctionnant sur le principe de la répétition espacée. Il est principalement utilisé pour apprendre une langue étrangère, mais il propose aussi des activités de mathématiques, de musique et d'échecs. Le service compte 100 millions d'utilisateurs actifs mensuels en août 2024 , et 8 millions d'abonnés payants.

Le 11 février 2025, Duolingo annonçait la mort de sa mascotte : Duo. Pour le ressusciter, il fallait que tous les utilisateurs de l'application atteignent ensemble 50 milliards d'expérience. Sur la page du défi, il était possible de savoir le score de chaque pays dans le top du classement. C'était l'occasion d'analyser la façon dont le score évolue pour chaque pays.

Nous pouvons formuler la problématique suivante : à l'air du numérique nos habitudes s'uniformisent t-elles ? Ou, la culture local reste assez influente ?

Cette étude démontrera que les mécanismes ludiques mise en place par Duolingo, surpassent les habitudes curturelles, dans notre utilisation de l'application.

2. Méthodologie commune.

Du 18 février 2025 à 23h55 (UTC) au 20 févier 2025 à 12h05 (UTC). Toutes les 10 minutes, j'ai collecté les données sur la page officielle du défi de Duolingo. Au total, 36h de données de 15 pays à travers le monde ont été collectées, donc plus d'un cycle de journée.

Les 15 pays concernés :

- *Europe*
  - Allemagne
  - France
  - Royaume-Uni
  - Espagne
  - Tchéquie
  - Pologne
  - Italie
- *Asie*
  - Vietnam
  - Chine
  - Inde
  - Corée du sud
- *Amérique*
  - États-Unis
  - Brésil
  - Mexique
  - Colombie

Les données ont été rangées dans un tableur. Ayant une colone pour l'horodatage, et des autres colonnes pour chaque pays. Chaque ligne représente une collecte de données.

3. Résultats

Ci-dessous, vous trouverez les résultats sous forme de graphiques et leurs commentaires.

I. Analyse des tendances globales

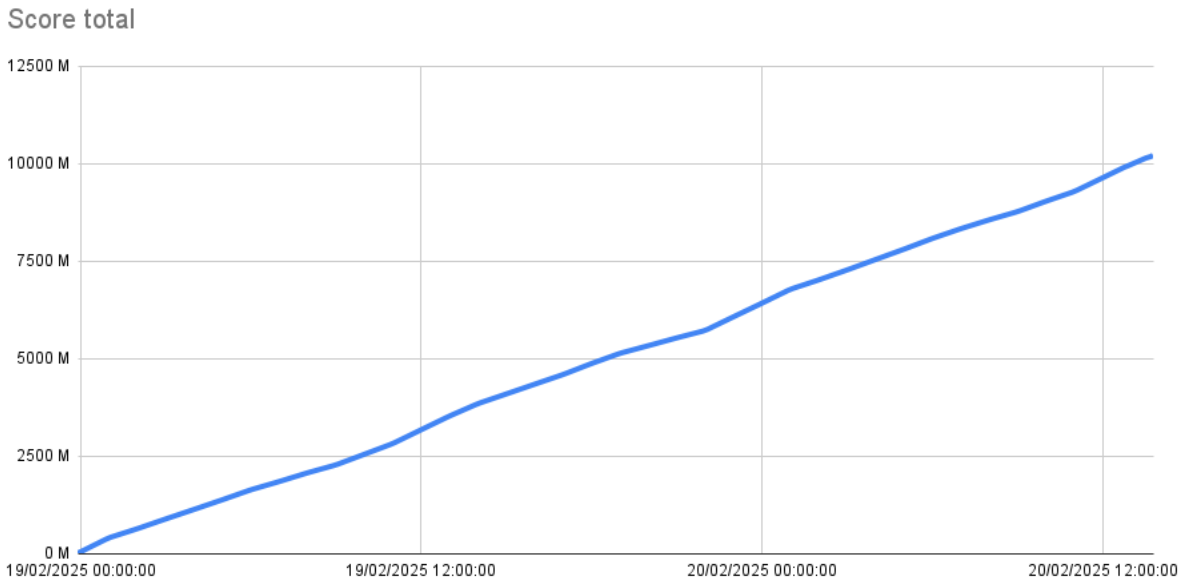


Figure 1. Graph showing the increase in overall score over time.

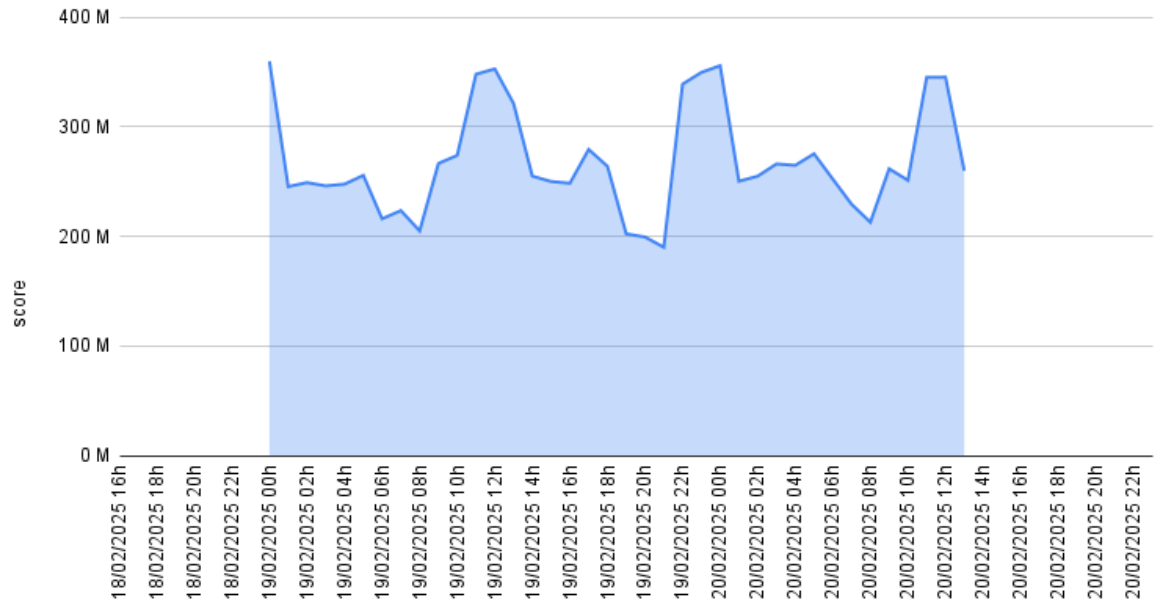


Figure 2. Graph showing overall score gain per hour.

Nous pouvons remarquer dans la Figure 1 que le score posséde une linéarité. Mais si nous analysons le gain de score par heure dans la Figure 2, il subit des variations entre des gains de 200 millions et 350 millions d'expériences.

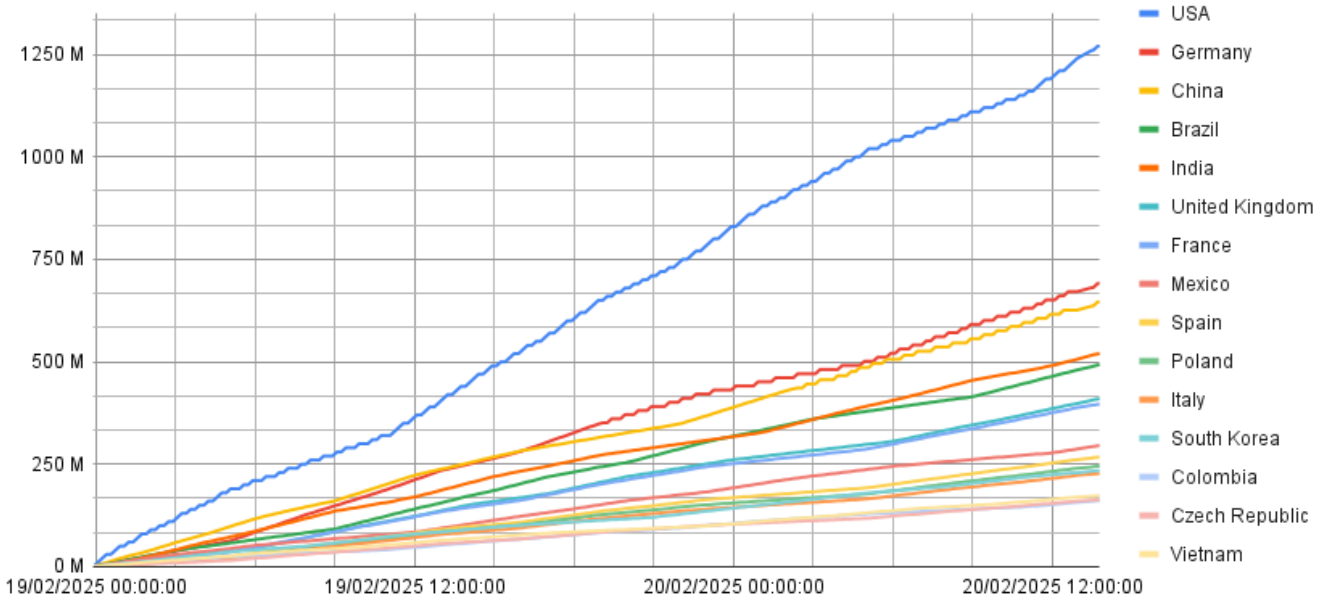


Figure 3. Graph illustrating the increase in each country's score.

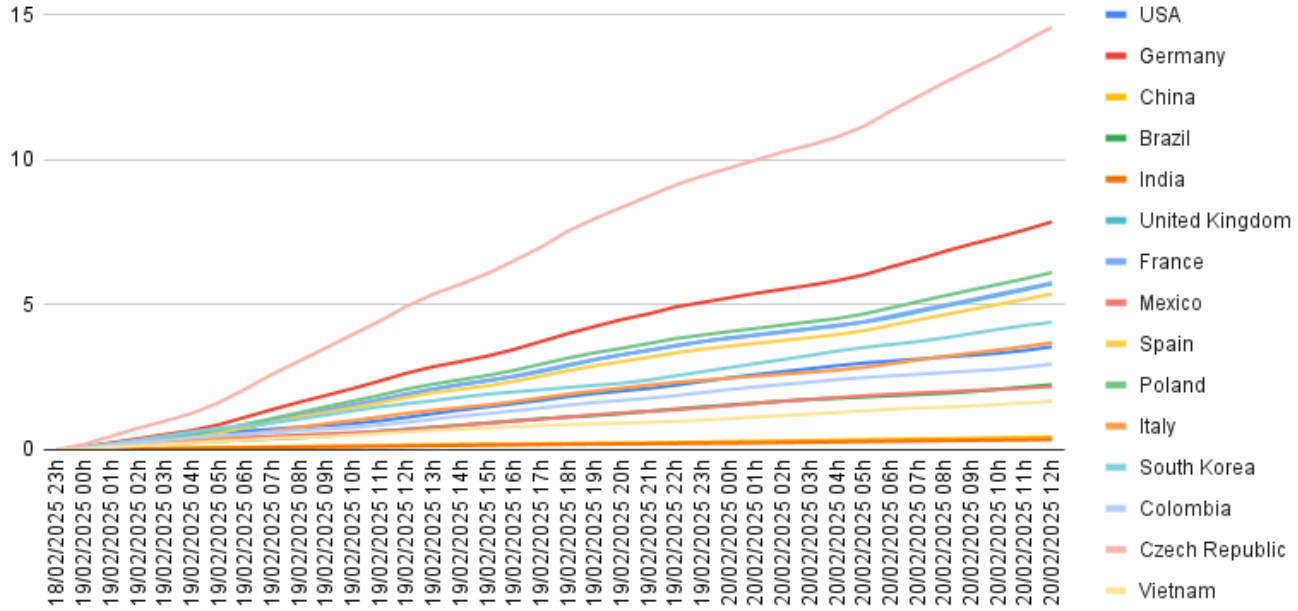


Figure 4. Graph illustrating the increase in each country's score relative to its population.

Nous pouvons remarquer dans la Figure 3 que le score des États-Unis augmente le plus rapidement. Presque deux fois plus que celui de l'Allemagne et de la Chine, qui occupent la deuxième et la troisième position.

Cependant, si nous calculons le score en fonction de la population du pays, le classement change complètement. Comme nous pouvons le voir dans la Figure 4, la Tchéquie arrive de loin en première position, et à apporté cumulé sur toute la durée, presque 90 points d'expérience par habitant. Contrairement au pays en deuxième position, l'Allemagne, qui a apporté cumulé sur toute la durée, presque 50 point d'expérience par habitant.

II. Analyse des différences par horaire au niveau local

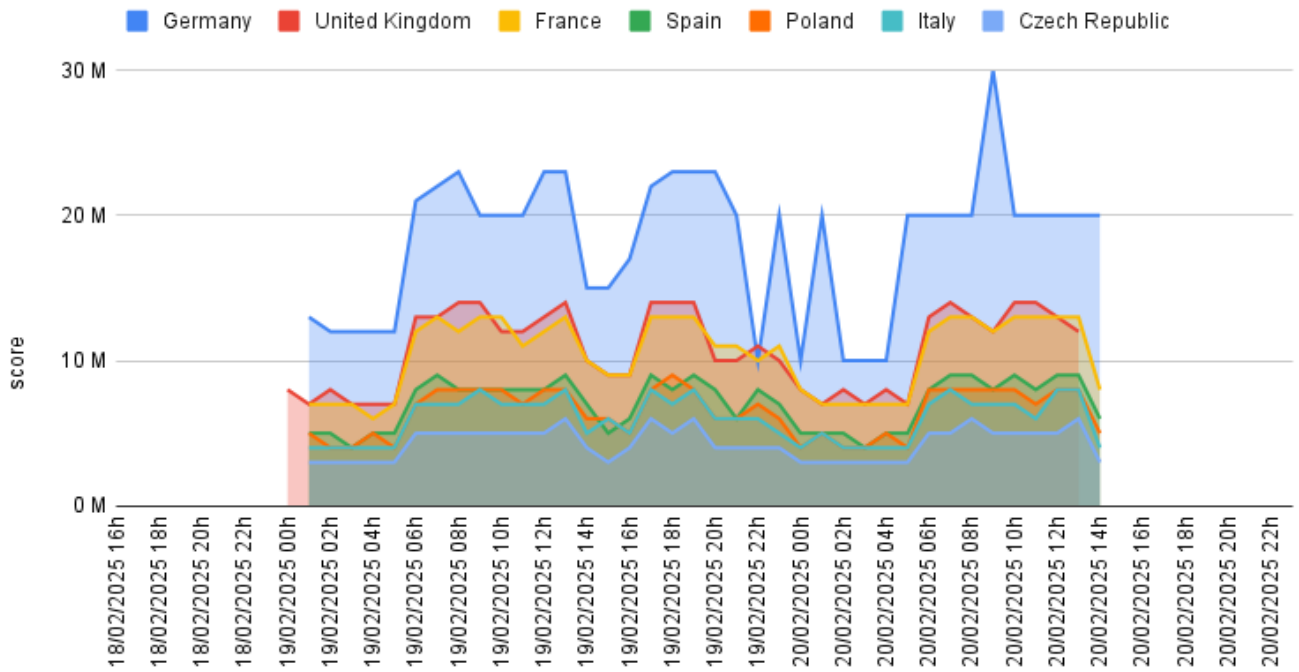


Figure 5. Graph showing the experience points earned by date and local time in certain European countries.

Concernant l'Europe, nous pouvons remarquer que le Royaume-Unis, la France, l'Expagne, la Polande et l'Italie, leur cycle est similaire. De 00h à 05h l'activité est calme, puis augmente subitement à 6h pour rester haute jusqu'à 13h. Une baisse à 14h, une augmentation à 17h. Puis, une nouvelle baisse à 20h jusqu'à 05h. Puis une réaugmentation à 6h, pour une baisse à 14h.

Horaire	06h	13h	14h	16h	17h	19h	20h	00h	05h	06h
Variation	→	↘	→	↗	→	↘	↘	→	↗	

L'Allemagne est différente, avec une baisse d'activité le soir à 22h au lieu de 19h, et une augmentation le matin à 05h au lieu de 06h.

Horaire	05h	13h	14h	16h	17h	21h	22h	04h	05h
Variation	→	↘	→	↗	→	↘	→	↗	

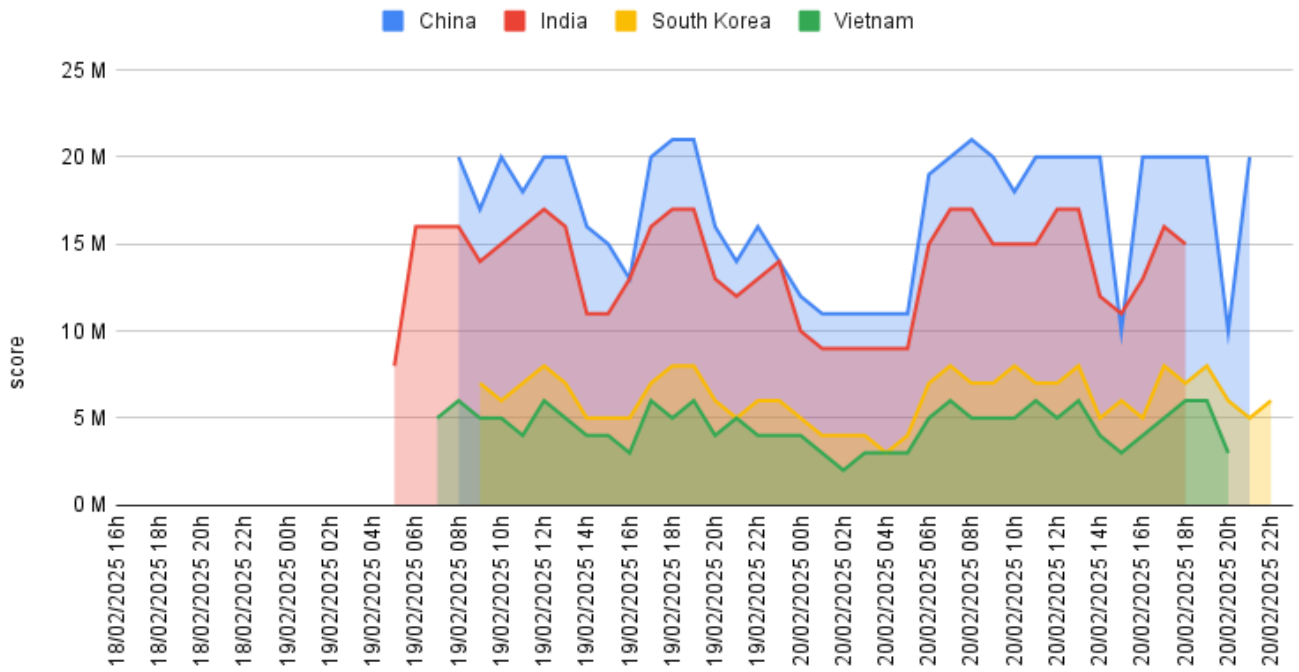


Figure 6. Graph showing the experience points earned by date and local time in certain Asian countries.

Concernant l'Asie, nous pouvons remarquer que pour la Chine, l'Inde, la Corée du sud, et le Vietnam, leur cycle est similaire. De 06h à 13h nous sommes en période très active, puis de 14h jusqu'à 16h nous sommes en période faiblement active. Puis, nous avons une augmentation de l'activité de 17h à 19h. Suivie d'une baisse de l'activité de 20h à 05h. À 06h la période très active revient jusqu'à 13h.

Horaire	06h	13h	14h	16h	17h	19h	20h	00h	05h	06h
Variation	→	↘	→	↗	→	↘	↘	→	↗	

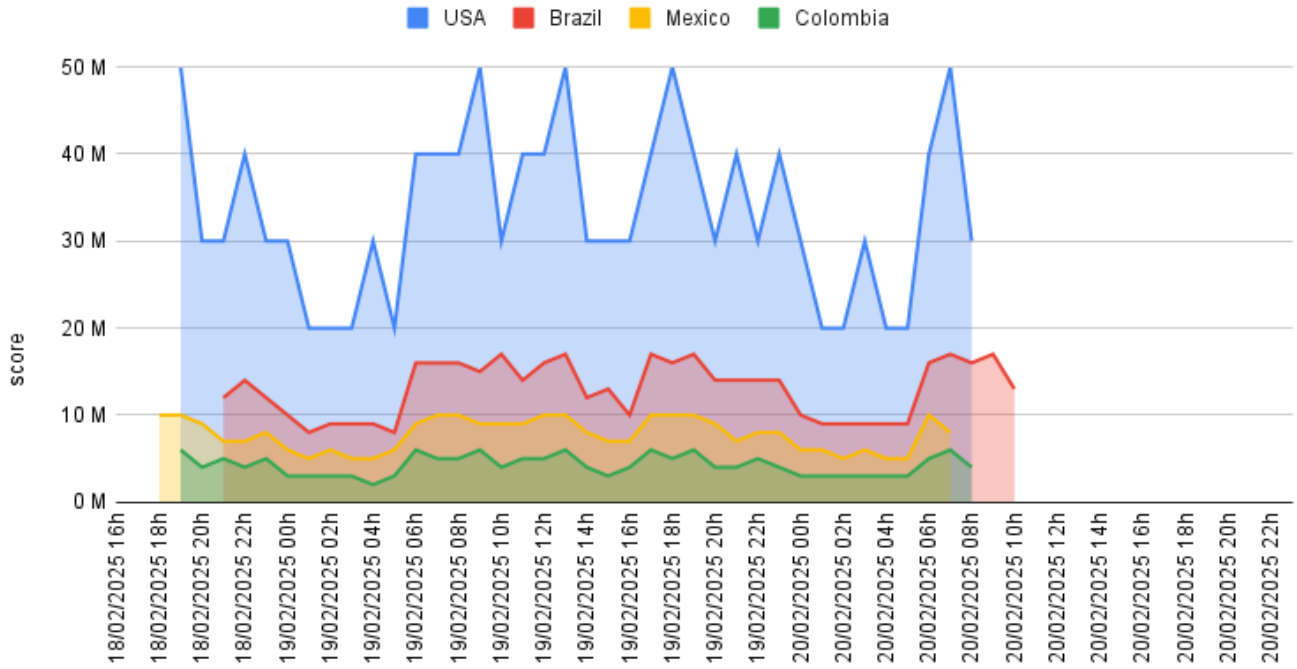


Figure 7. Graph showing the experience points earned by date and local time in certain American countries.

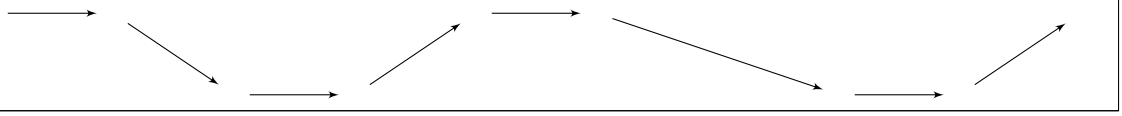
Concernant l'Amérique, les quatre pays auxquels nous avons des données ont le même cycle, même si pour la Colombie c'est plus compliqué à dire au vu de la faible quantité de données. Période haute d'activité de 06h à 13h, puis faible d'activité de 14h à 16h. Augmentation de l'activité à 17h jusqu'à 19h. Puis, une baisse à 20h, une nouvelle baisse à 00h, qui dure jusqu'à 05h.

Horaire	06h	13h	14h	16h	17h	19h	20h	00h	05h	06h
Variation	→	↘	→	↗	→	↘	↘	→	↗	

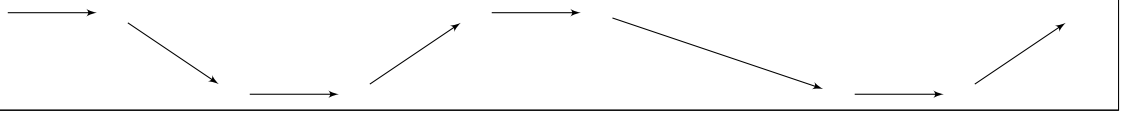
A. Résumé

Ci dessous, les tableaux des variations de l'activité en fonction de l'horaire, pour chaque continent.

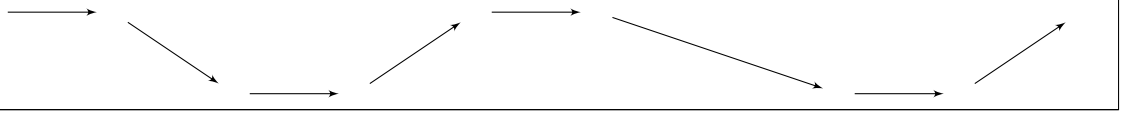
Europe

Horaire	06h	13h	14h	16h	17h	19h	20h	00h	05h	06h
Variation										

Asie

Horaire	06h	13h	14h	16h	17h	19h	20h	00h	05h	06h
Variation										

Amérique

Horaire	06h	13h	14h	16h	17h	19h	20h	00h	05h	06h
Variation										

Nous pouvons remarquer que les utilisateurs de ces différents continents on les mêmes habitudes.

4. Discussion

I. Différence d'horaire entre l'utilisation de Duolingo et du sommeil.

Country	Bedtime	Wake up
Europe		
Allemagne	23h13	07h23
France	23h01	07h44
Royaume-Uni	23h19	07h29
Espagne	00h17	07h54
Tchéquie	23h06	07h32
Pologne	00h11	07h39
Italie	00h10	07h36
Average	*23h38*	*07h38*
Difference	*01h16*	*00h29*
Asia		
Vietnam	00h15	07h30
Chine	00h52	08h15
Inde	00h04	07h08
Corée du sud	00h46	07h52
Average	*00h28*	*07h41*
Difference	*00h48*	*01h07*
America		
États-Unis	23h34	07h06
Brésil	23h12	06h57
Mexique	23h13	07h04
Colombie	23h47	06h27
Average	*23h29*	*06h47*
Difference	*00h35*	*00h39*

Figure 8. Bedtime et wake up par country.

Les données des horaires de sommeils par pays, proviennent de l'étude de [O. J. Walch, A. Cochran and D. B. Forger, 2016 (1)].

Nous pouvons remarquer que les pays ont des horaires de levé et couché différents. Même entre les continents en prenant la moyenne des horaires des pays qui les contiennent, les horaires de levé et de couché sont différent. Pourtant, nous pouvons voir que les horaires d'utilisation de Duolingo sont presque identiques pour tous les pays et continents.

Pour comprendre cette différence, il faut prendre en compte l'existence des coffres, qui sont des bonus permettant d'avoir des amplificateurs d'expérience. Si nous faisons une leçon entre 18h et 00h nous gagnons un coffre utilisable entre 06h et 12h. Si nous faisons une leçon entre 6h et 12h, alors nous gagnons un coffre utilisable entre 18h et 00h. Cela rend du coup moins attractifs les horaires entre 00h et 05h, ainsi qu'entre 12h et 17h.

Couplés à la fonctionnalité des Ligues, les coffres peuvent influer sur l'utilisation de l'application. En effet, pour passer à la ligue supérieure, il faut être dans le top du classement, et votre position dans le classement dépend du nombre d'expérience que vous avez gagné. D'autant plus, que dans un événement où il faut gagner le plus d'expérience possible afin de ressuciter la mascotte Duo, ces coffres deviennent encore plus important.

Ceci expliquerait les baisses et augmentations sur les graphiques.

Nous pouvons donc supposer, que les habitudes d'utilisation des utilisateurs de Duolingo, ne sont pas principalement impactées par la culture de leur pays, mais par les mécaniques de l'application Duolingo.

II. Domination des États-unis dans le score

L'augmentation dominante du score de la part des États-Unis peut s'expliquer par le fait que 25% du trafic du site duolingo.com provient des Etats-Unis. Ainsi que le siége de Duolingo est à Pittsburgh, en Pennsylvanie, aux États-unis.

III. L'engagement remarquable de la Tchéquie

Comme nous pouvons le voir dans la Figure 4, la Tchéquie est de loin première en fonction du score par habitant.

La Tchéquie est un pays enclavée dans l'Europe, entourée de plusieurs pays : l'Allemagne (Allemand), la Pologne (Polonais), la Slovaquie (Slovaque), l'Autriche (Allemand). Donc deux pays qui parle Allemand, ce qui pourrait expliqué pour la deuxième langue apprise par les Tchéques soit l'Allemand, d'après le rapport de Duolingo de 2024 [Duolingo, 2024 (2)]. La première langue apprise par les Tchéques est l'anglais.

L'hypotèse d'un engagement aussi fort, est la prise de décision du gouvernement un mois avant le défi. En effet, le ministère de l'éducation de la Tchéquie en janvier 2025, a supprimé le russe dans les école et a rendu l'anglais obligatoire dés la première classe, d'après l'article de [Radio Prague International (3)]. Ce qui a pu augmentée l'utilisation de Duolingo.

IV. Limites de l'étude

- Le fait que nous n'avons pas le nombre d'utilisateurs de l'application Duolingo par pays.
- La population de certains pays pourraient se concentrer sur une seule langue, et donc passer moins de temps sur l'application, donnant moins d'expérience, ce qui influence le score.

5. Conclusion

Le rapport a démontré trois niveaux d'influences.

En premier : que les mécaniques dans l'application influent sur l'utilisation de l'application, indépendamment de la zone géographique de l'utilisateur.

En deuxième : l'influence du marché. Les États-Unis, la Chine et l'Allemagne sont les leaders en termes de points d'expérience totaux. Cela est dû au fort trafic d'utilisateur venant de ses pays et de la popularité de l'application dans ces pays, par exemple grâce à la publicité.

En troisième : la situation socio-économique. Avec la Tchéquie, malgré le peu d'expérience apporté dans le score global, la population dans ce pays a eu un engagement très fort par apport au nombre d'habitants. L'hypothèse de cet engagement est sûrement due aux décisions du gouvernement.

Nous pourrions effectuer une étude complémentaire, dans un contexte de non-événement, pour savoir si les statistiques récoltées sont spécifiques à cet événement, et changent par apport à une situation d'utilisation classique. Et donc, en même temps, quantifier l'influence d'un tel événement.

6. Remerciements

Merci à **Duolingo** d'avoir mis en place ce défi, permettant la collecte des données.

Merci à **vous** d'avoir pris le temps de lire mon document.

7. Bibliographie

Références

[1] O. J. Walch, A. Cochran, and D. B. Forger, "A global quantification of "normal" sleep schedules using smartphone data." *Science Advences*, 2016. [Online]. Available : https://www.science.org/doi/10.1126/sciadv.1501705

[2] C. Blanco, "Le rapport duolingo sur les langues 2024." *Duolingo*, 2024. [Online]. Available : https://blog.duolingo.com/fr/rapport-duolingo-langues-2024/

[3] G. Narguet, "Anglais obligatoire dès la 1ère classe, suppression du russe : les nouveautés du ministère tchèque de l'Éducation," *Radio Prague International*, 2025. [Online]. Available : https://francais.radio.cz/anglais-obligatoire-des-la-1ere-classe-suppression-du-russe-les-nouveautes-du-8840150

[4] M. Claudel, "La chouette de duolingo va-t-elle être ressuscitée ?" *Numerama*, 2025. [Online]. Available : https://www.numerama.com/tech/1906963-la-chouette-de-duolingo-va-t-elle-etre-ressuscitee.html

[5] N. Elhajjioui, "La mascotte de duolingo réssuscite à paris !" *J'ai un pote dans la com*, 2025. [Online]. Available : https://jai-un-pote-dans-la.com/duo-revint-a-vie/

[6] A. Boero, "Elle a ressuscité ! duo, la mascotte de duolingo, est revenue à la vie grâce à ses utilisateurs," *Clubic*, 2025. [Online]. Available : https://www.clubic.com/actualite-554808-elle-a-ressuscite-duo-la-mascotte-de-duolingo-est-revenue-a-la-vie-grace-a-ses-utilisateurs.html

[7] Google, "Évolution de l'intérêt pour la recherche duolingo dans google." *Google Trends*, 2025. [Online]. Available : https://trends.google.com/trends/explore?date=2025-01-02%202025-02-28&q=%2Fm%2F0gxzb2p