

Czy profil użytkownika wpływa na aktywność w aplikacji?

Aleksandra Zachajska, Konstanty Andrzejczak

Table of contents

1	Wprowadzenie	2
1.1	1.1. Dane	2
1.2	1.2. Opis zmiennych	2
2	Duolingo według badanego zbioru danych	4
2.1	Popularność aplikacji	4
2.2	Dlaczego Duolingo?	5
2.3	Geograficzne rozłożenie użytkowników	6
3	Profil odbiorców aplikacji	7
3.1	Wiek i płeć	7
4	Dynamika rozwoju użytkownika	7
4.2	Kraje anglojęzyczne	7
5	Źródła	7
6	A tibble: 15 x 5	7
7	Analiza danych	8
8	Śmietnik	17
8.1	Wprowadzenie	17
8.2	Kto sięga po aplikację (ogólna charakterystyka użytkowników)?	17

1 Wprowadzenie

Rozwój technologii mobilnych i powszechna dostępność urządzeń elektronicznych znacząco wpłynęły na sposób, w jaki ludzie zdobywają wiedzę i rozwijają swoje umiejętności. Aplikacje mobilne, takie jak Duolingo, stanowią doskonały przykład tej rewolucji edukacyjnej, umożliwiając użytkownikom naukę języków obcych w sposób elastyczny, dostosowany do ich indywidualnych potrzeb i harmonogramu. Od momentu wprowadzenia na rynek w 2011 roku, Duolingo zdobyło ogromną popularność na całym świecie, przyciągając miliony użytkowników, którzy regularnie korzystają z aplikacji, aby doskonalić swoje umiejętności językowe.

Za sukcesem aplikacji stoi nie tylko jej innowacyjna metoda nauki, oparta na gamifikacji, ale także zdolność do dostosowania się do różnorodnych potrzeb użytkowników. Aplikacja stała się szczególnie popularna w Stanach Zjednoczonych, Europie i Azji, a jej użytkownicy reprezentują szeroką gamę krajów, kultur oraz poziomów zaawansowania. Zjawisko to skłania do refleksji nad tym, w jaki sposób profil użytkownika – obejmujący takie czynniki jak wiek, kraj pochodzenia, płeć czy wcześniejsze doświadczenia z językiem – może wpływać na jego zaangażowanie oraz efektywność nauki. Celem niniejszej pracy jest zbadanie, w jaki sposób różne cechy użytkowników aplikacji Duolingo kształtują ich aktywność, motywację oraz postępy w nauce, a także jakie czynniki mogą decydować o sukcesie w tym procesie edukacyjnym.

1.1 1.1. Dane

Aby przeprowadzić tę analizę, posłużymy się danymi pochodzącymi z publicznego repozytorium MatthewWolff/DuolingoAnalysis na GitHubie [1]. Autor nie zamieścił informacji o licencji na jakiej udostępniono dane jednak zostaną one użyte wyłącznie w celach edukacyjnych do wykonania niniejszej analizy funkcjonalnej.

1.2 1.2. Opis zmiennych

Zestawienie zostało wykonane przy pomocy ankiety przeprowadzonej przez Duolingo w okresie od 1 sierpnia do 5 listopada 2018 roku. Ankieta obejmowała pytania dotyczące aktywności użytkowników, korzystania z aplikacji oraz ich profilu demograficznego, takiego jak wiek, płeć czy kraj zamieszkania. Baza danych obejmuje 5488 obserwacji i 29 typów zmiennych. Dane zawierają niewielką liczbę braków, które zostaną odpowiednio obsłużone w kolejnych etapach analizy. Dokładniejsza charakterystyka pól została przedstawiona w Tabeli 1.

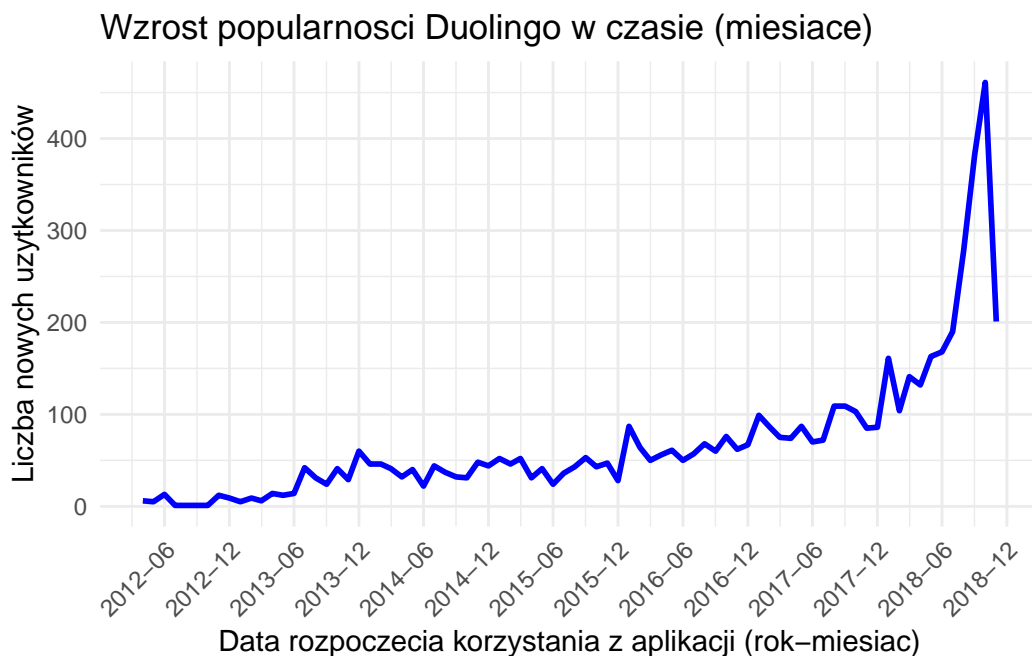
Table 1: Pełny opis zmiennych w zbiorze danych Duolingo.

Zmienna	Opis	Jednostki/Wartości
user_id	Unikalny identyfikator użytkownika	Tekst
age	Grupa wiekowa użytkownika	Kategorie: 'Under 18', '18-34', '35-54', '55-74', '75 or older'
annual_income	Roczny dochód użytkownika	Kategorie: '\$0 - \$25,000', '\$26,000 - \$75,000', '\$76,000 - \$150,000', '> \$150,000'
country	Kraj zamieszkania użytkownika	Kody krajów, np. 'US', 'JP', 'DE'
duolingo_platform	Platforma używana przez użytkownika	Kategorie: 'Android', 'iOS', 'Web'
duolingo_subscriber	Status subskrypcji Duolingo Plus	Kategorie: 'Yes', 'No', 'Previously Paid'
duolingo_usage	Częstotliwość korzystania z aplikacji Duolingo	Kategorie: 'Daily', 'Weekly', 'Monthly'
employment_status	Status zatrudnienia użytkownika	Kategorie: 'Employed full-time', 'Part-time', 'Retired', 'Unemployed'
future_contact	Czy użytkownik zgodził się na kontakt w przyszłości	Prawda/Falsz
gender	Płeć użytkownika	Kategorie: 'Male', 'Female', 'Other'
other_resources	Liczba innych zasobów używanych do nauki języków	Liczba całkowita
primary_language_commitment	Deklaracja zaangażowania w naukę języka głównego	Kategorie: 'Low', 'Medium', 'High'
primary_language_review	Przegląd postępów w nauce języka głównego	Kategorie: 'Poor', 'Average', 'Good'
primary_language_motivation	Motywacja użytkownika do nauki języka głównego	Tekst (deklaracje motywacji)
primary_language_motivation_followup	Dodatkowe informacje o motywacji do nauki głównego języka	Tekst (dodatkowe informacje o motywacji)
primary_language_proficiency	Samodzielnie oceniony poziom znajomości języka głównego	Kategorie: 'Beginner', 'Intermediate', 'Advanced'
student	Czy użytkownik jest studentem	Prawda/Falsz
survey_complete	Czy użytkownik ukończył ankietę	Prawda/Falsz
time_spent_seconds	Czas spędzony na wypełnianiu ankiety (w sekundach)	Liczba całkowita (sekundy)
duolingo_start_date	Data rozpoczęcia korzystania z aplikacji	Data (YYYY-MM-DD)
highest_course_progress	Najwyższy osiągnięty postęp w kursie (w procentach)	Procent (0-100)
took_placement_test	Czy użytkownik wykonał test poziomujący	Prawda/Falsz
purchased_subscription	Czy użytkownik wykupił subskrypcję Duolingo Plus	Prawda/Falsz
highest_crown_count	Najwyższa liczba zdobytych koron	Liczba całkowita
n_active_days	Liczba dni, w których użytkownik był aktywny	Liczba całkowita
n_lessons_started	Liczba lekcji rozpoczętych przez użytkownika	Liczba całkowita
n_lessons_completed	Liczba lekcji ukończonych przez użytkownika	Liczba całkowita
longest_streak	Najdłuższa seria dni korzystania z aplikacji	Liczba całkowita
n_days_on_platform	Łączna liczba dni od rejestracji użytkownika w aplikacji	Liczba całkowita

2 Duolingo według badanego zbioru danych

2.1 Popularność aplikacji

Jak wspomniano wcześniej, aplikacja Duolingo z biegiem czasu zyskiwała coraz większą popularność na rynku międzynarodowym. Aby przeanalizować przyrost liczby użytkowników, posłużono się danymi ankietowymi, które przedstawiają liczbę nowych użytkowników w poszczególnych miesiącach. Poniższy wykres ilustruje tę dynamikę.

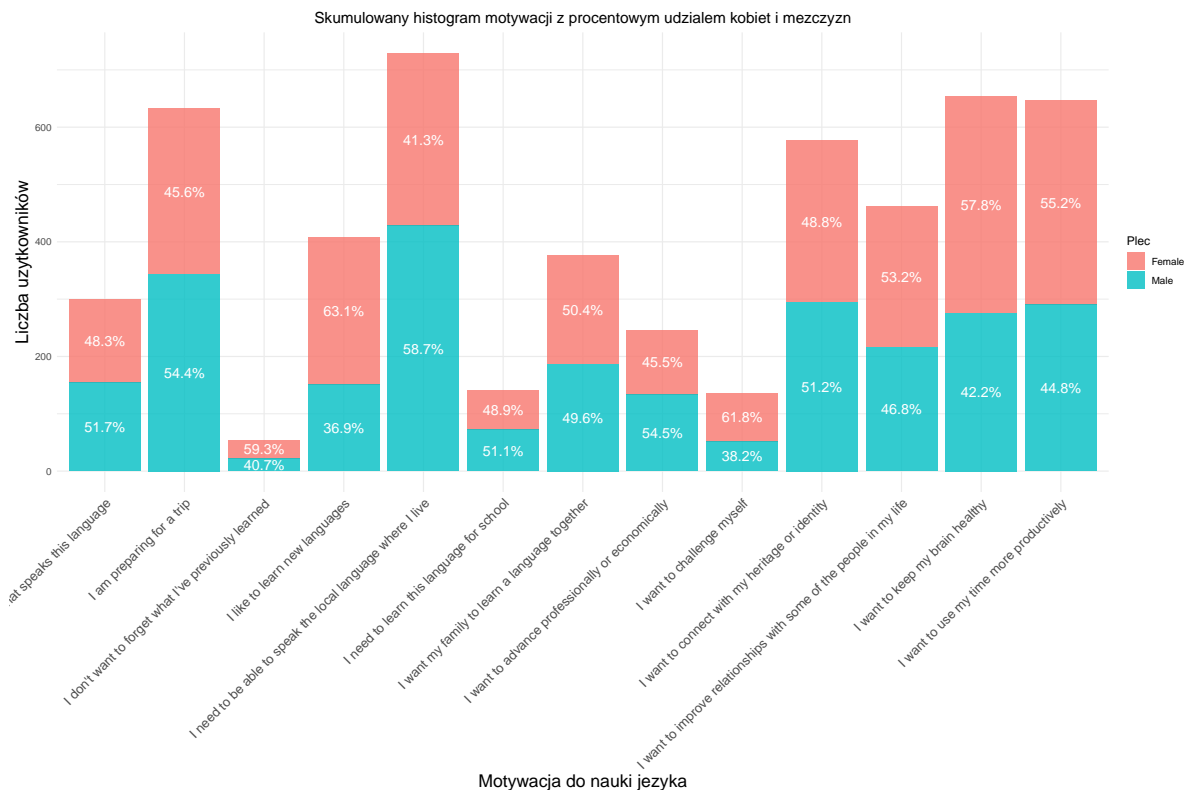


Wykres ten przedstawia liczbę nowych użytkowników aplikacji w kolejnych miesiącach. Można zauważyć wyraźną tendencję wzrostu, przy czym największy przyrost liczby użytkowników miał miejsce w trzecim kwartale 2018 roku. Ta gwałtowna ekspansja może być efektem zintensyfikowanych działań marketingowych w tym okresie, a także wyróżnienia aplikacji w sklepach App Store i Google Play.

Dodatkowo, warto zaznaczyć, że Duolingo wprowadziło w tym czasie nowe funkcje, które mogły znacząco wpłynąć na wzrost liczby użytkowników. Przykładem jest system „Crown Levels”, który miał na celu motywowanie użytkowników do kontynuowania nauki poprzez zdobywanie kolejnych poziomów, co nawiązywało do mechanizmów znanych z gier mobilnych. Takie innowacje mogły zachęcić większą liczbę osób do regularnego korzystania z aplikacji, przyczyniając się tym samym do jej popularności.

2.2 Dlaczego Duolingo?

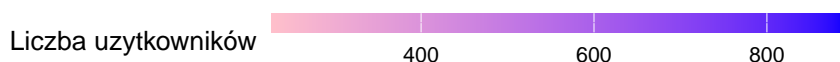
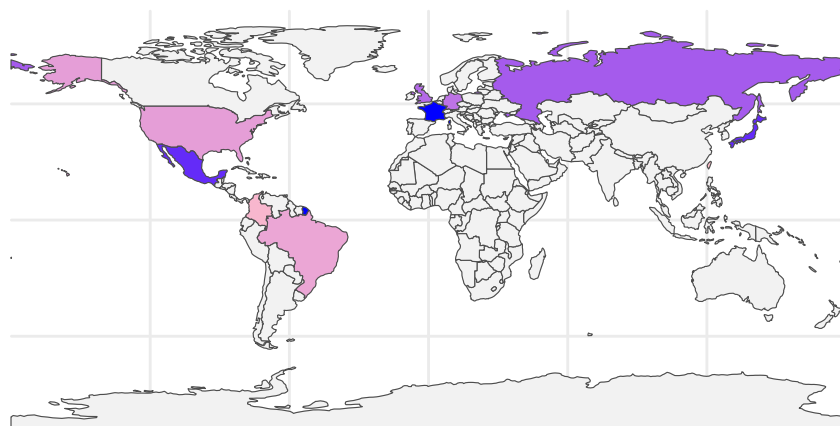
W kontekście oceny sukcesu aplikacji kluczową rolę odgrywa zrozumienie motywacji użytkowników, czyli czynników, które skłaniają ich do wyboru konkretnego produktu. W przypadku Duolingo, istotnym aspektem jest nie tylko sama funkcjonalność aplikacji, ale także jej zdolność do przyciągania szerokiego grona odbiorców. Brak zainteresowania ze strony użytkowników skutkowałby brakiem popytu na aplikację, co mogłoby wpłynąć na jej dalszy rozwój. Aby lepiej zrozumieć, dlaczego użytkownicy wybierają Duolingo, przeprowadzono ankietę, w której respondenci mieli możliwość wskazania czynników, które miały wpływ na ich decyzję.



Wykres powyżej przedstawia wyniki ankiety, ukazując, jakie motywacje były najczęściej wskazywane przez użytkowników. Znacząca większość osób wskazała chęć nauki języka w celu podróżowania, co podkreśla rolę aplikacji jako narzędzia ułatwiającego komunikację w obcych krajach. Kolejnymi ważnymi powodami były chęć rozwoju zawodowego oraz potrzeba nauki nowych umiejętności w sposób wygodny i dostępny. Takie odpowiedzi świadczą o szerokim zakresie zastosowań Duolingo, który odpowiada na różnorodne potrzeby użytkowników.

2.3 Geograficzne rozłożenie użytkowników

Rozkład użytkowników Duolingo według krajów



Zródło: Dane Duolingo

W danych ankietowych zanotowano aktywność użytkowników z 10. krajów:

	country	country_full
1	JP	Japan
2	RU	Russia
3	CO	Colombia
4	TW	Taiwan
5	GB	United Kingdom
6	BR	Brazil
7	MX	Mexico
8	DE	Germany
9	FR	France
10	US	United States of America

Z danych wynika, że największa liczba użytkowników pochodzi z Francji i Meksyku, co wskazuje na ich silną obecność na rynku aplikacji edukacyjnych. Oba te kraje, ze względu na swoje rozmiary, liczbę ludności oraz popularność języka hiszpańskiego w Meksyku i francuskiego we Francji, stanowią naturalne rynki dla aplikacji uczącej języków obcych. Dalsza analiza rozkładu użytkowników na mapie pokazuje, że Duolingo cieszy się także dużym zainteresowaniem w takich krajach jak Japonia, Rosja czy Niemcy. Zjawisko to może świadczyć o szerokim zasięgu aplikacji na różnych kontynentach, obejmujących zarówno

regiony rozwinięte, jak i wschodzące rynki. Powyższe obserwacje mogą sugerować, że Duolingo dostosowuje swoje funkcje i podejście do specyfiki każdego kraju, co pozwala na utrzymanie dużej bazy użytkowników na całym świecie.

3 Profil odbiorców aplikacji

3.1 Wiek i płeć

Patrząc na ## Status zatrudnienia ##

4 Dynamika rozwoju użytkownika

4.1

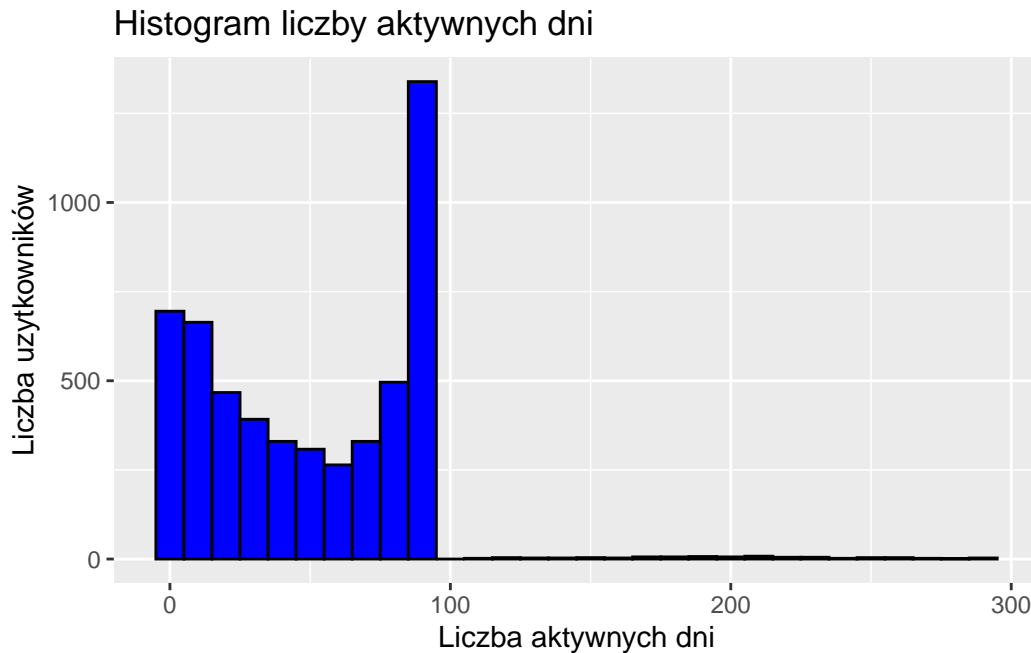
4.2 Kraje anglojęzyczne

5 Źródła

[1] <https://github.com/MatthewWolff/DuolingoAnalysis.git>

6 A tibble: 15 x 5

Variable active lessons streak days 1 n 51.7 257. NA 624. 2 n 50 241. NA 423 3 longest NA
NA 103. NA 4 n 38.8 175 NA 597. 5 n 1 161 NA 2 6 longest NA NA 19 NA 7 n 289 307. NA
2402 8 n NA 290. NA NA 9 longest NA NA 203. NA 10 n NA 1 NA NA 11 n NA 0 NA NA
12 longest NA NA 0 NA 13 n NA 4786 NA NA 14 n NA 4686 NA NA 15 longest NA NA 1820



NA

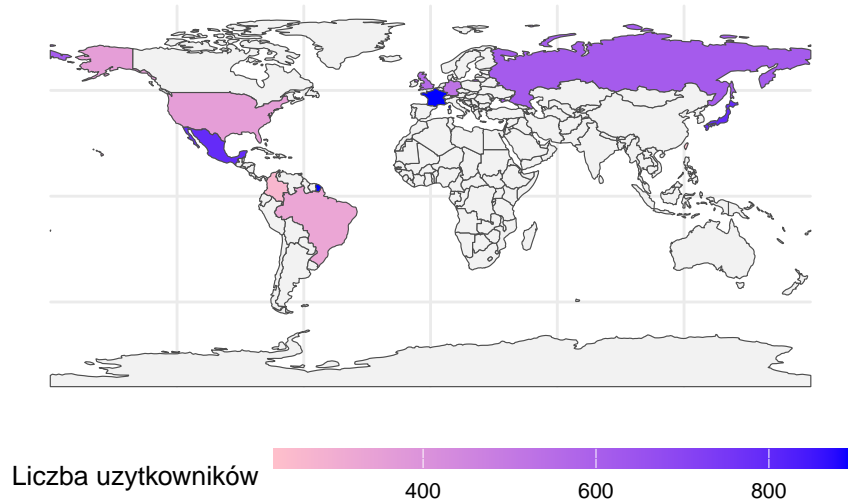
Zbiór danych Duolingo zawiera 29 zmiennych opisujących profile użytkowników oraz ich aktywność w aplikacji. Dane tekstowe, takie jak `age`, `annual_income` czy `duolingo_platform`, zostały sklasyfikowane jako zmienne kategoryczne, co pozwala na ich efektywną analizę. Zmienne logiczne, takie jak `future_contact` (zgoda na kontakt w przyszłości) czy `took_placement_test` (informacja o wykonaniu testu poziomującego), umożliwiają badanie binarnych decyzji użytkowników. Zmienne liczbowe, takie jak `n_active_days` (liczba aktywnych dni) czy `n_lessons_completed` (liczba ukończonych lekcji), zostały zachowane w odpowiednim formacie numerycznym, aby umożliwić analizę ilościową. Dodatkowo, `duolingo_start_date`, reprezentująca datę rozpoczęcia korzystania z aplikacji, została przekonwertowana na typ `date`, co pozwala na analizę danych czasowych.

7 Analiza danych

Wyniki ankiet pochodzą z różnych krajów

Rozkład użytkowników Duolingo według krajów

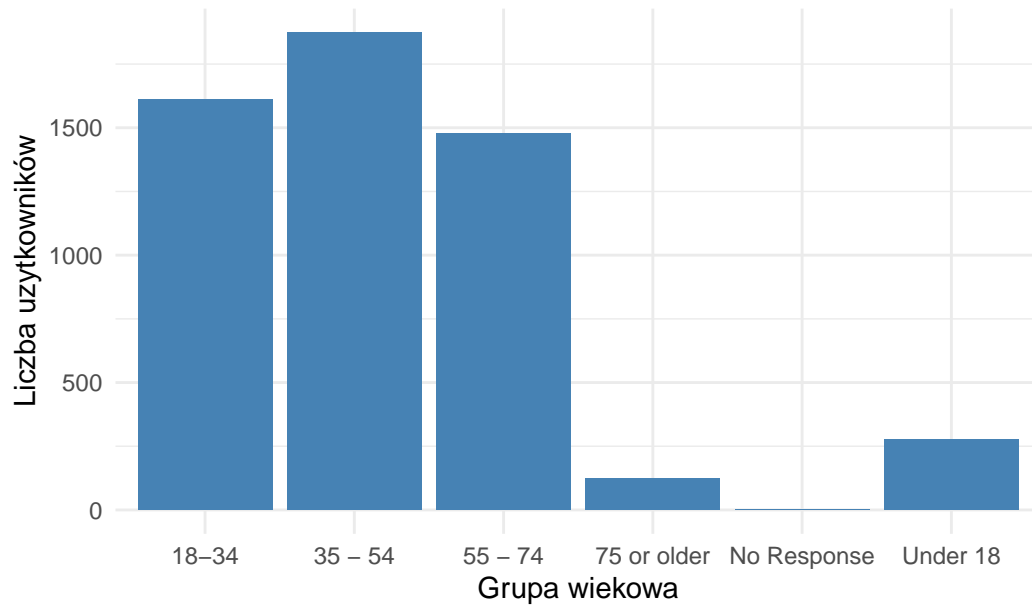
Kolor przedstawia liczbę użytkowników w danym kraju



Zródło: Dane Duolingo

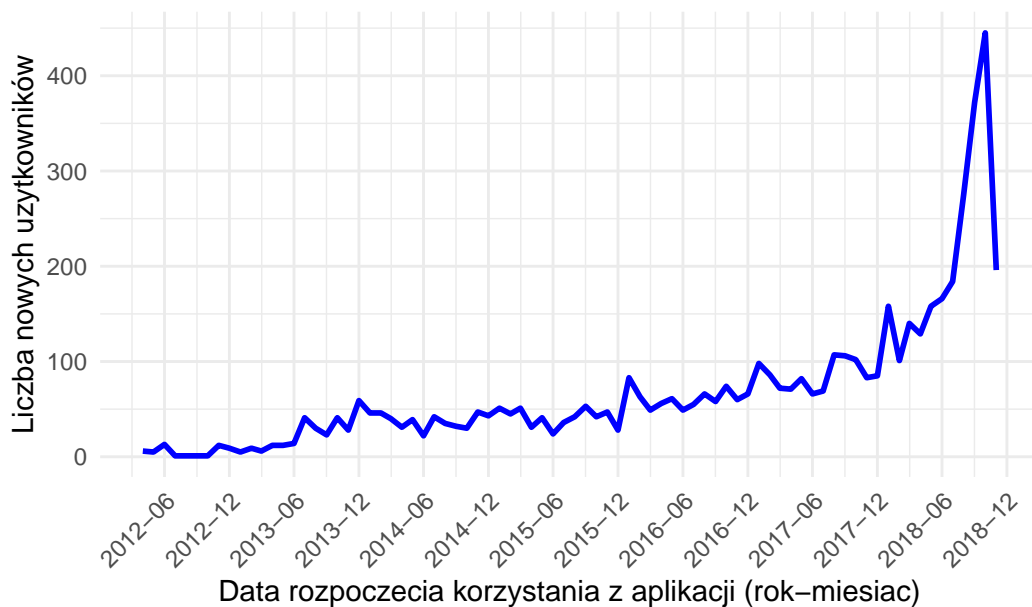
jaki jest rozkład wiekowy użytkowników

Rozkład wiekowy użytkowników

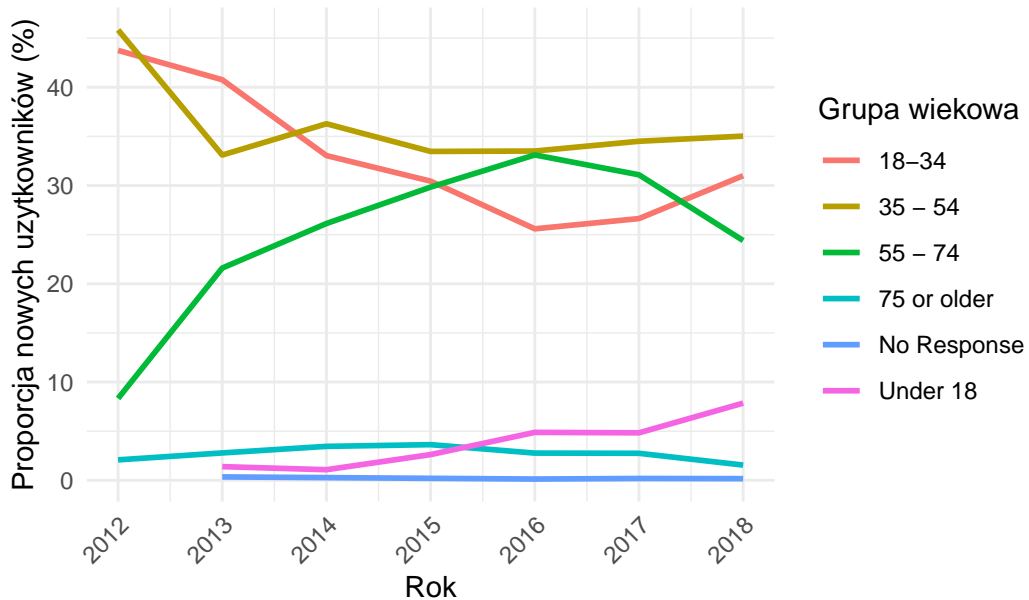


Jak aplikacja Duolingo zyskiwała na popularności?

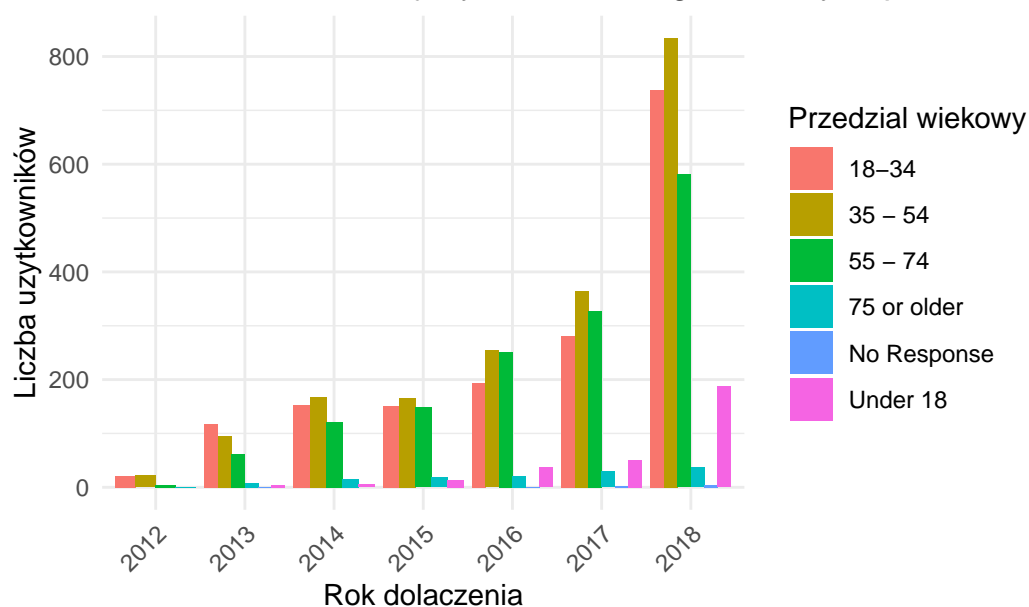
Wzrost popularności Duolingo w czasie (miesiące)



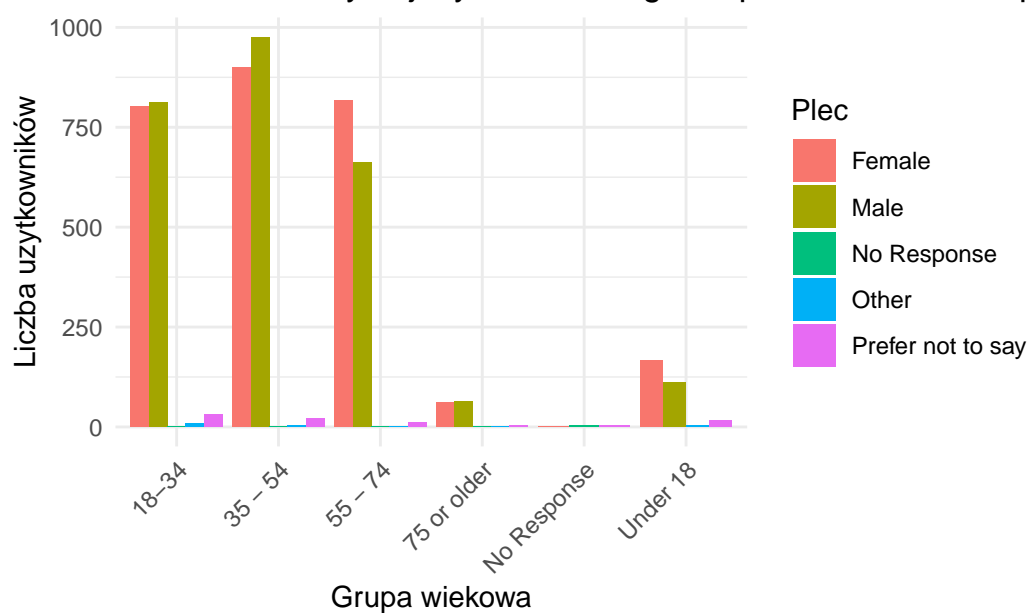
Proporcjonalny wzrost użytkowników w grupach wiekowych (rok)

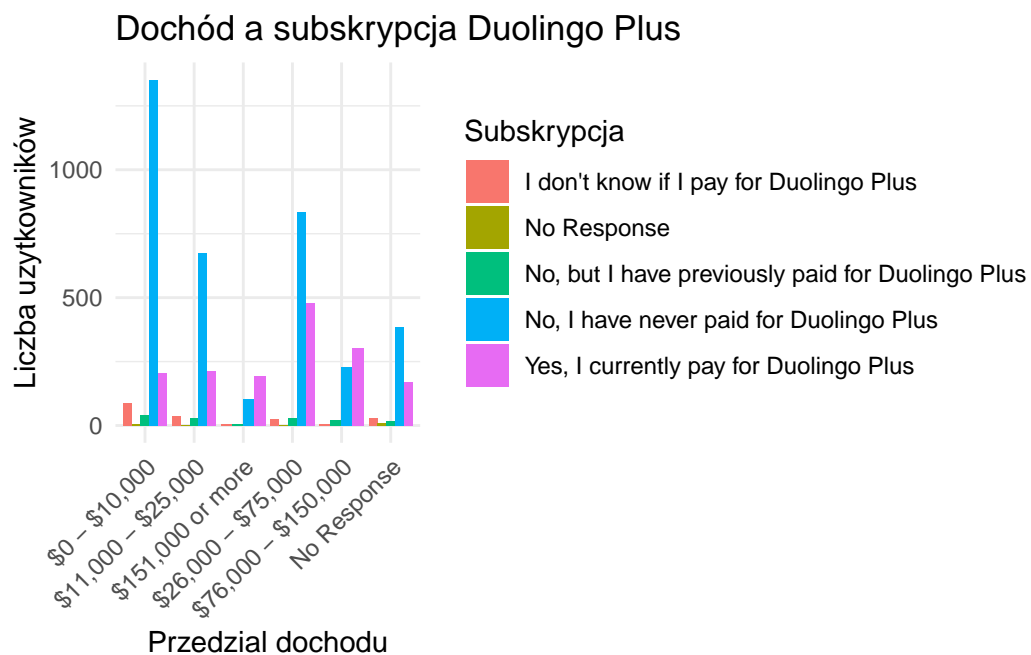
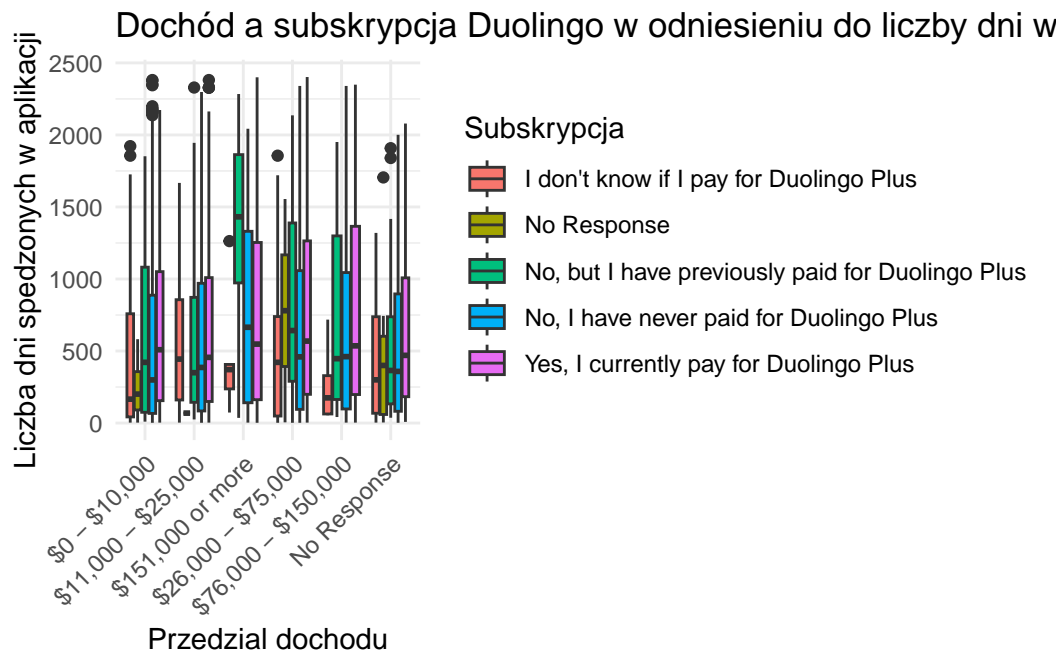


Liczba osób dołączających do Duolingo w różnych przedziałach



Ilość osób korzystających z Duolingo na podstawie wieku i płci

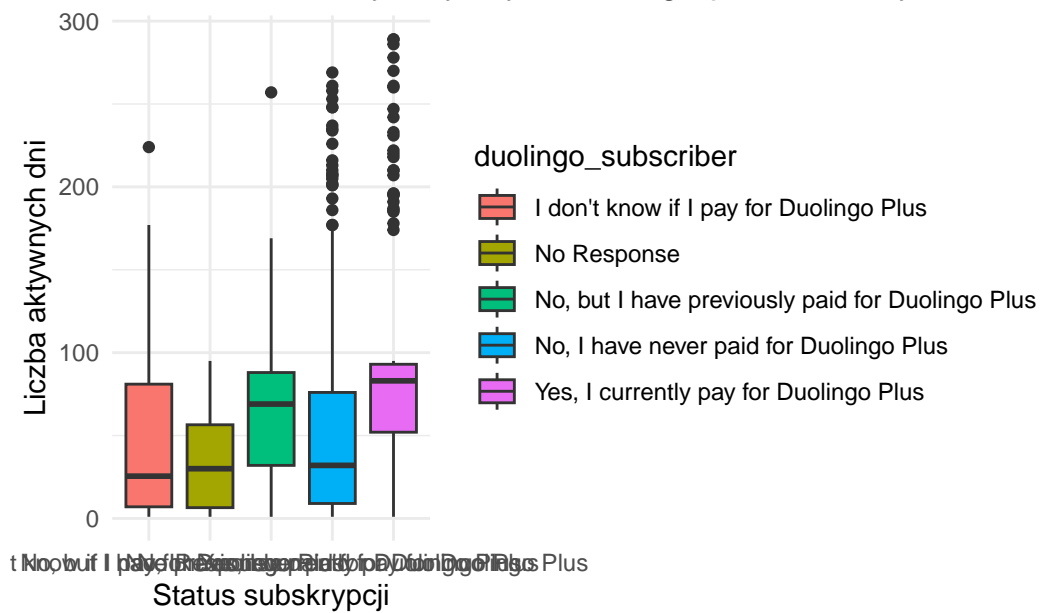




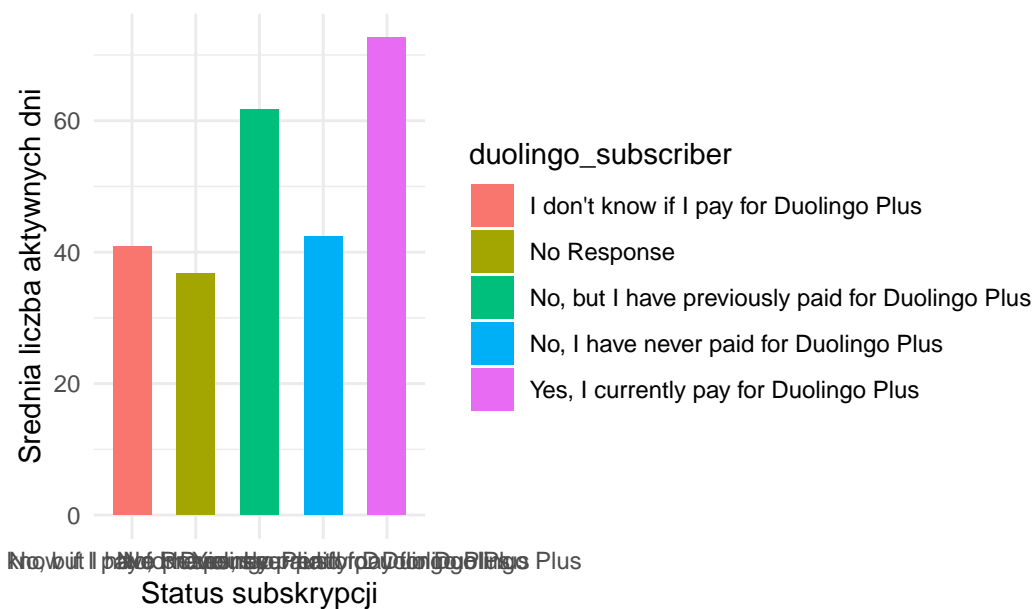
Liczba innych zasobów a liczba aktywnych dni w Duolingo



Porównanie liczby aktywnych dni w grupach subskrybentów Duolingo Plus

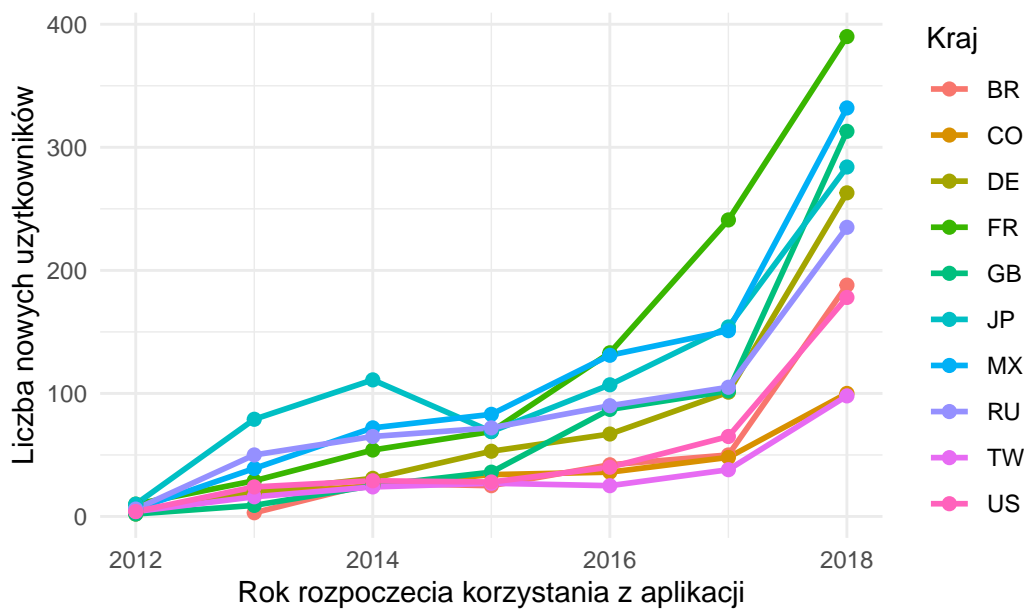


Srednia liczba aktywnych dni w grupach subskrypcyjnych

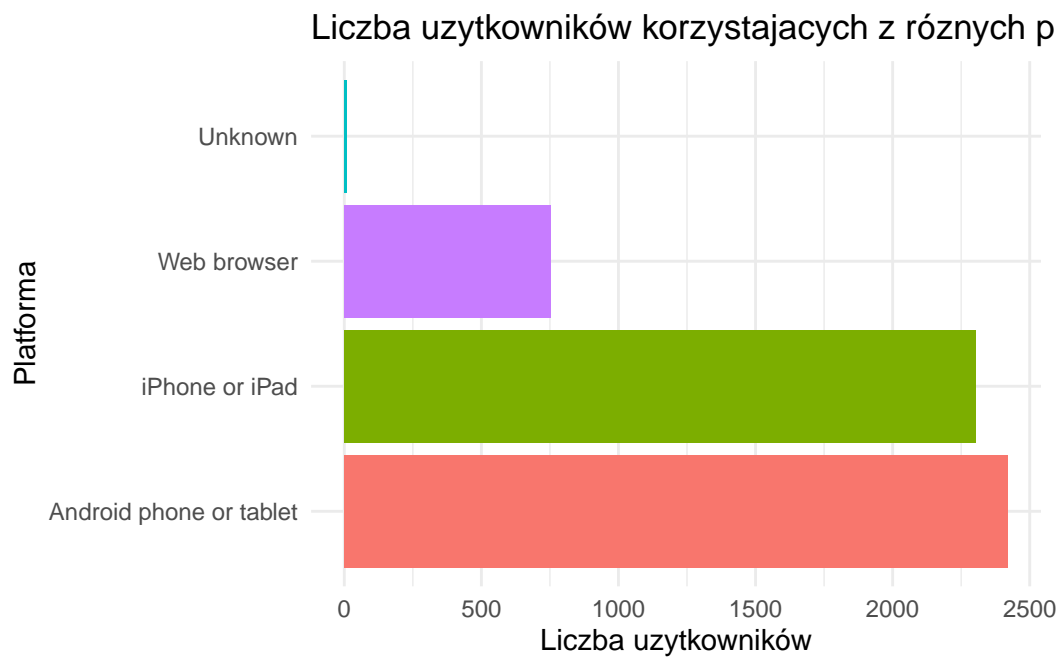


Autor nie zamieścił informacji o licencji na jakiej udostępniono dane jednak zostaną one użyte jedynie do wykonania analizy funkcjonalnej w poniższej pracy. Znalezione dane zostały zebrane za pomocą ankiety przeprowadzonej przez Duolingo, które miały miejsce od 1 sierpnia do 5 listopada 2018 roku.

Wzrost popularności Duolingo w różnych krajach

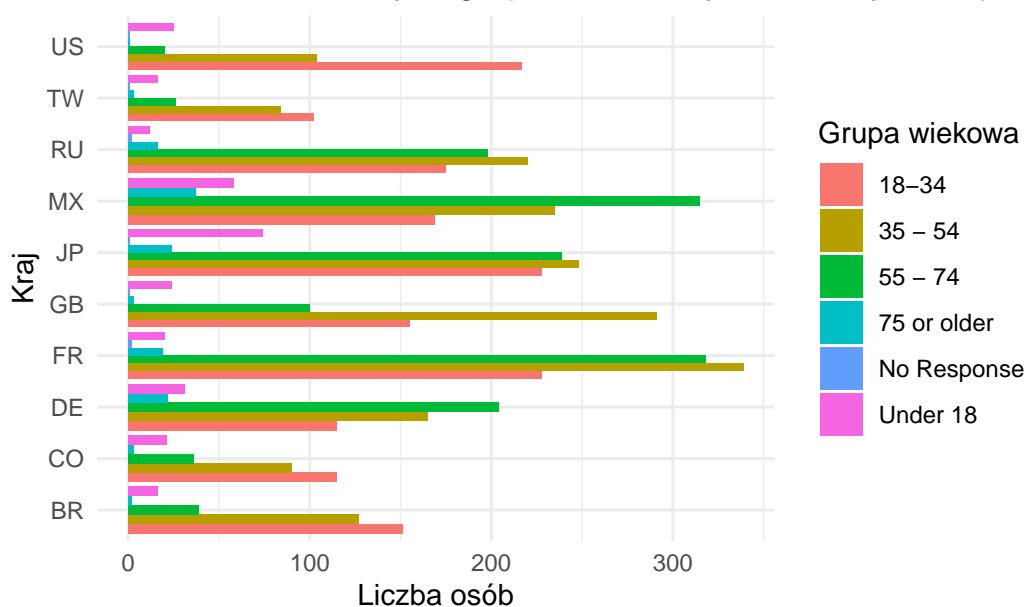


```
<ScaleContinuousPosition>  
Range:  
Limits: 0 -- 1
```

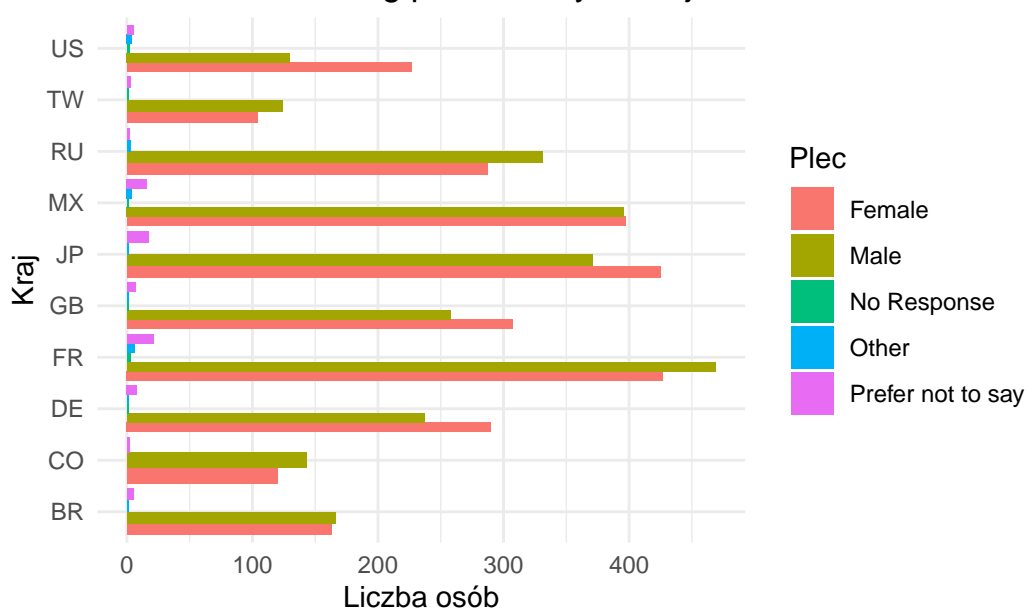


```
# A tibble: 10 x 1  
  country  
  <chr>  
1 Japan  
2 Russian Federation  
3 Colombia  
4 Taiwan  
5 United Kingdom  
6 Brazil  
7 Mexico  
8 Germany  
9 France  
10 United States
```

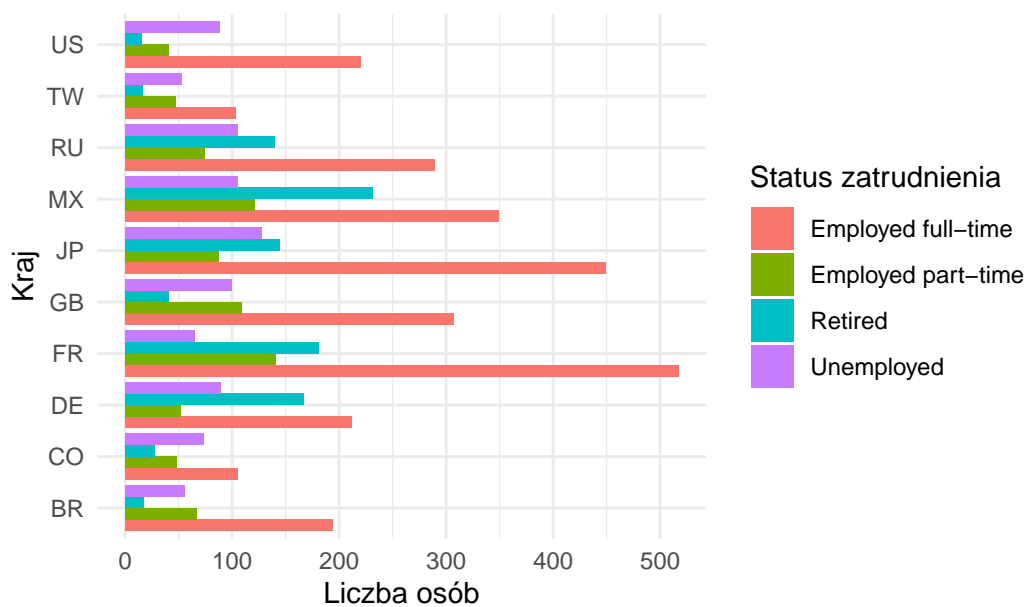
Liczba osób w różnych grupach wiekowych w różnych krajach



Liczba osób według płci w różnych krajach



Liczba osób według statusu zatrudnienia w różnych krajach



8 Śmietnik

8.1 Wprowadzenie

We wprowadzeniu napisać z jakich dat są zebrane dane i zaznaczyć daty widoczne w zestawieniu kiedy uczestnicy podlegający spisowi rozpoczęli korzystanie z aplikacji odnieść się do daty wprowadzenia aplikacji na rynek i mniej więcej kiedy aplikacja stała się popularna można dołączyć dane dotyczące dynamiki wejścia na rynki w różnych krajach w szczególności tych uwzględnionych w zestawieniu

[1] 4

8.2 Kto sięga po aplikację (ogólna charakterystyka użytkowników)?

Analiza jakie kraje są brane pod uwagę w każdym z nich możemy zestawić ile jest osób zatrudnionych w jakim wymiarze i ilu w jakim wieku jakiej płci (ogólnie info)

Profil użytkownika

Analiza profilu użytkowników dokładna pod kątem używania aplikacji tu zestawić: -> zaangażowanie -> motywację -> czy miał już styczność z językiem -> poziom deklarowany vs wyznaczony w aplikacji

Dynamika rozwoju przy użyciu aplikacji -> Zestawienie liczby rozpoczętych lekcji w zestawieniu z zakończonymi -> liczba dni na platformie vs liczba aktywnych dni -> najdłuższa seria