

Projekt 3. - prezentacja

Mateusz Kwiatkowski i Antoni Szustakowski

09.06.2020

Wstęp

Projekt 3. –
prezentacja

Mateusz
Kwiatkowski i
Antoni
Szustakowski

Do naszych badań użyliśmy ramek danych z Jersey City z lipców 2017-2019 oraz z wakacji 2019 (czerwiec-sierpień).

Wśród danych znalazły się, między innymi, informacje o:

- czasie trwania podróży
- wieku i płci użytkownika
- stacji początkowej i końcowej
- typie użytkownika (gość, subskrybent)

Najpierw postanowiliśmy zbadać ogólne zależności dla podstawowych informacji, a następnie zajęliśmy się szczegółową analizą.

Tworzenie profilu statystycznego klienta

Projekt 3. –
prezentacja

Mateusz
Kwiatkowski i
Antoni
Szustakowski

Na początku przeanalizowaliśmy podstawowe informacje o użytkownikach względem ich płci. Dane pochodzą z wakacji 2019.

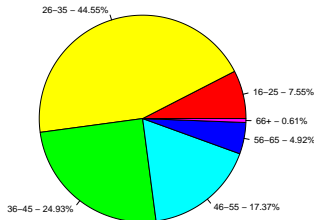
płeć	przejazdy	średni_czas	stacja_końcowa
nieznana	10955	45	Grove St PATH
mężczyzna	90692	11	Grove St PATH
kobieta	30240	14	Grove St PATH

Jak widać, najpopularniejsza stacja końcowa jest wspólna dla wszystkich płci. Wynika to z położenia *Grove ST Path* - jest to stacja blisko centrum Jersey. Zdecydowana większość przejazdów to przejazdy mężczyzn, a najdłużej jeżdżą ci, którzy nie podali swojej płci.

Udział poszczególnych grup wiekowych w liczbie przejazdów

Sprawdziliśmy również, które grupy wiekowe jeżdżą najczęściej.

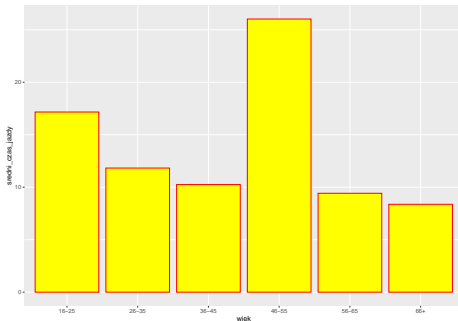
Udział procentowy w przejazdach w zależności od wieku



Zdecydowana większość użytkowników to osoby w wieku od 26 do 45 lat, wynika to z faktu, że są one bardziej aktywne fizycznie.

Średni czas jazdy w zależności od wieku

Zbadaliśmy, które grupy wiekowe jeżdżą najdłużej.

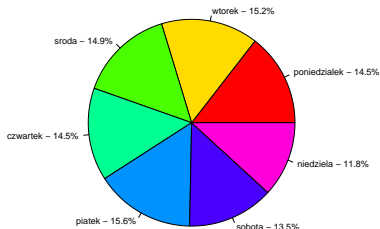


Z wykresu można odczytać, że na najdłuższe przejażdżki decydują się osoby w wieku 46-55 oraz 16-25. Można mniemać, iż te grupy korzystają z rowerów w celach bardziej rekreacyjnych niż grupy 26-35 i 36-45.

Najpopularniejsze dni tygodnia

Zaciekawiło nas, czy istnieją różnice pomiędzy liczbą przejazdów w poszczególne dni tygodnia.

Udział procentowy dni tygodnia



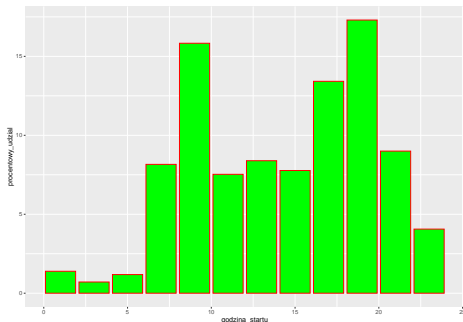
Okazuje się, że liczba przejazdów w dni robocze jest minimalnie większa niż w weekendy.

W jakim celu ludzie korzystają z rowerów?

Projekt 3. -
prezentacja

Mateusz
Kwiatkowski i
Antoni
Szustakowski

Można podejrzewać, że spora część użytkowników korzysta z rowerów, aby dostać się do pracy. W celu potwierdzenia tej tezy, zbadaliśmy, które godziny są najpopularniejsze.



Badanie najpopularniejszych godzin

Projekt 3. -
prezentacja

Mateusz
Kwiatkowski i
Antoni
Szustakowski

wiek	godzina	procentowy_udział
16-25	8-10	6.87
26-35	8-10	45.95
36-45	8-10	28.24
46-55	8-10	15.19
56-65	8-10	4.19
66+	8-10	0.45
16-25	18-20	7.28
26-35	18-20	46.74
36-45	18-20	25.34
46-55	18-20	13.91
56-65	18-20	4.74
66+	18-20	0.43

Wnioski

Projekt 3. –
prezentacja

Mateusz
Kwiatkowski i
Antoni
Szustakowski

Na podstawie powyższych wykresów oraz tabeli możemy stwierdzić, że duża część przejazdów odbywa się w dni robocze w godzinach od 8 do 10 oraz od 16 do 20. Łącząc to z popularnością tych godzin wśród grup wiekowych 26-35 oraz 36-45, możemy wywnioskować, że rowery są popularnym środkiem transportu wśród pracowników. Dodatkowym bodźcem wpływającym korzystnie na liczbę przejazdów może być pogoda, gdyż dane pochodzą z lata 2019.

Udział stacji

Zbadaliśmy, jak popularne podczas wakacji 2019 były poszczególne stacje końcowe.

stacja_końcowa	procentowy_udział
Grove St PATH	13.784
Hamilton Park	5.536
Harborside	4.678
Newport PATH	4.596
Sip Ave	4.101
Marin Light Rail	3.935

Najpopularniejsze były stacje położone blisko centrum kulturowego lub komunikacyjnego. Co ciekawe, 17 stacji miało marginalny (około 0,001%) udział w liczbie przejazdów.

Dzień Niepodległości

Projekt 3. –
prezentacja

Mateusz
Kwiatkowski i
Antoni
Szustakowski

Uznaliśmy, że warto sprawdzić, jak Amerykanie w Jersey korzystają z rowerów w trakcie święta narodowego (Dzień Niepodległości - 4. lipca). Zbadaliśmy, który cel podróży jest najpopularniejszy.

Nazwa stacji końcowej	2017	2018	2019
Grove St PATH	83	111	104
Hamilton Park	58	64	73
Liberty Light Rail	55	52	50
Marin Light Rail	54	39	66
Morris Canal	46	41	43
Newport PATH	45	48	43
Newport Pkwy	48	51	53
Van Vorst Park	38	44	36
Warren St	57	51	45

Porównanie święta do zwykłego dnia

Zastanowiło nas, czy wyniki dotyczące popularności pewnych stacji w Dzień Niepodległości się różnią od pozostałych dni.

stacja_końcowa	udział_4.07	udział_wakacje
Grove St PATH	8.125	13.784
Hamilton Park	5.703	5.536
Harborside	5.000	4.678
Marin Light Rail	5.156	3.935
Newport PATH	3.359	4.596
Newport Pkwy	4.141	3.128
Paulus Hook	3.203	3.093

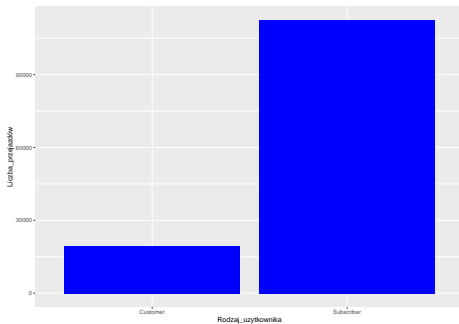
Jak widać, popularność stacji *Grove ST Path* maleje, a popularność mniejszych stacji, delikatnie rośnie. Z tego wynika, że cele podróży w święto narodowe są bardziej zróżnicowane.

Rodzaje użytkowników

Projekt 3. -
prezentacja

Mateusz
Kwiatkowski i
Antoni
Szustakowski

Jak wiemy, dane zostały podzielone na użytkowników jednorazowych (*Customer*) oraz subskrybentów (*Subscriber*). Zbadaliśmy, ile przejazdów wykonali zarówno jedni, jak i drudzy.



Analiza gości

Mimo, że ogromną część użytkowników stanowią subskrybenci, to postanowiliśmy przyjrzeć się bliżej profilowi użytkownika jednorazowego. Im więcej o nich wiemy, tym łatwiej ich będzie można przekonać do zakupu subskrypcji.

wiek	płeć	przejazdy	czas_jazdy
16-25	nie podano	52	25
16-25	mężczyzna	1954	26
16-25	kobieta	1047	43
26-35	nie podano	81	17
26-35	mężczyzna	3698	36
26-35	kobieta	2078	31
36-45	nie podano	50	18
36-45	mężczyzna	1023	28
36-45	kobieta	345	39

Analiza gości cd.

Projekt 3. –
prezentacja

Mateusz
Kwiatkowski i
Antoni
Szustakowski

	wiek	płeć	przejazdy	czas_jazdy
10	46-55	nie podano	8532	54
11	46-55	mężczyzna	246	31
12	46-55	kobieta	125	37
13	56-65	nie podano	0	brak przejazdów
14	56-65	mężczyzna	104	25
15	56-65	kobieta	48	34
16	66+	nie podano	0	brak przejazdów
17	66+	mężczyzna	10	44
18	66+	kobieta	2	58

Zatem, zdecydowanie najpopularniejszą grupą jednorazowych gości są użytkownicy, którzy nie podali płci oraz znajdują się w grupie wiekowej 46-55 oraz mężczyźni z grupy 26-35. Goście jeżdżą średnio znacznie dłużej niż subskrybenci.

Podsumowanie

Projekt 3. –
prezentacja

Mateusz
Kwiatkowski i
Antoni
Szustakowski

Na podstawie naszej analizy można wysnuć następujące wnioski:

- przeciętny użytkownik jest mężczyzną w wieku produkcyjnym
- klienci wykorzystują rowery, przede wszystkim, by poruszać się do pracy
- na dłuższe przejażdżki udaje się młodzież oraz osoby w okolicach 50. roku życia
- zdecydowanie najpopularniejszą stacją jest *Grove ST Path*
- w Dzień Niepodległości użytkownicy częściej udają się na swobodne przejażdżki
- aby zachęcić gości jednorazowych do zakupu subskrypcji warto rozważyć wprowadzenie promocji na dłuższe przejazdy.