

Bike sharing

Analiza danych z wypożyczalni rowerów w Kalifornii

Dane surowe zawierają informacje o ostatnim tygodniu każdego z pięciu pierwszych miesięcy w 2019 roku.

Źródła:

www.lyft.com

www.kaggle.com

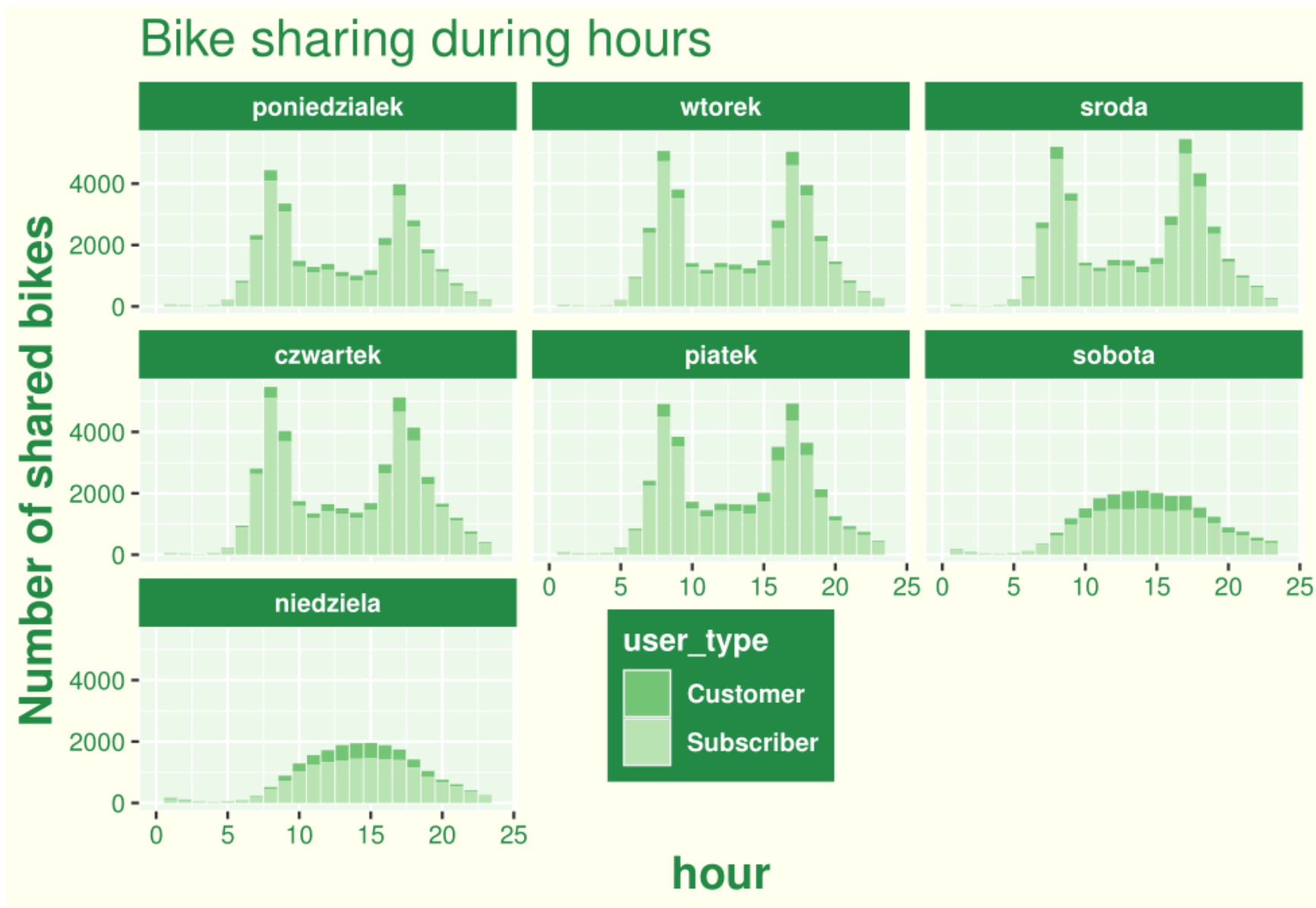
```
## 'data.frame': 238134 obs. of 16 variables:
## $ duration_sec : int 323 272 433 1188 451 252 547 559 376 340 ...
## $ start_time : POSIXct, format: "2019-01-31 23:59:00" "2019-01-31 23:59:00" ...
## $ end_time : POSIXct, format: "2019-02-01 00:05:00" "2019-02-01 00:04:00" ...
## $ start_station_id : int 223 349 266 34 305 121 77 41 25 95 ...
## $ start_station_name : chr "16th St Mission BART Station 2" "Howard St at Mary St" "Parker St at Fulton St" "Father Al
fred E Boeddeker Park" ...
## $ start_station_latitude : num 37.8 37.8 37.9 37.8 37.3 ...
## $ start_station_longitude: num -122 -122 -122 -122 -122 ...
## $ end_station_id : int 129 60 256 146 317 133 349 52 349 58 ...
## $ end_station_name : chr "Harrison St at 20th St" "8th St at Ringold St" "Hearst Ave at Euclid Ave" "30th St at San
Jose Ave" ...
## $ end_station_latitude : num 37.8 37.8 37.9 37.7 37.3 ...
## $ end_station_longitude : num -122 -122 -122 -122 -122 ...
## $ bike_id : int 1976 263 4642 5114 3842 5257 3035 5019 3649 5382 ...
## $ user_type : chr "Subscriber" "Subscriber" "Subscriber" "Subscriber" ...
## $ member_birth_year : int 1991 1993 1996 1984 1992 1995 1988 1991 1990 1989 ...
## $ member_gender : chr "Male" "Male" "Male" "Male" ...
## $ bike_share_for_all_trip: chr "No" "Yes" "No" "No" ... (program dla osób z niskim dochodem)
## - attr(*, "na.action")= 'omit' Named int [1:12536] 33 44 45 63 121 126 132 151 187 188 ...
## ..- attr(*, "names")= chr [1:12536] "33" "44" "45" "63" ...
```

```

## duration_sec start_time end_time
## Min. : 61.0 Min. :2019-01-25 00:03:00 Min. :2019-01-25 00:08:00
## 1st Qu.: 341.0 1st Qu.:2019-02-22 17:11:00 1st Qu.:2019-02-22 17:23:00
## Median : 543.0 Median :2019-03-27 08:48:00 Median :2019-03-27 08:58:00
## Mean : 758.1 Mean :2019-03-24 02:52:14 Mean :2019-03-24 03:04:37
## 3rd Qu.: 846.0 3rd Qu.:2019-04-26 18:09:00 3rd Qu.:2019-04-26 18:23:00
## Max. :86114.0 Max. :2019-05-31 23:59:00 Max. :2019-06-01 10:09:00
##
## start_station_id start_station_name start_station_latitude
## Min. : 3.0 Length:238134 Min. :37.31
## 1st Qu.: 47.0 Class :character 1st Qu.:37.77
## Median :104.0 Mode :character Median :37.78
## Mean :140.1 Mean :37.77
## 3rd Qu.:239.0 3rd Qu.:37.80
## Max. :427.0 Max. :37.88
##
## start_station_longitude end_station_id end_station_name
## Min. : -122.5 Min. : 3.0 Length:238134
## 1st Qu.: -122.4 1st Qu.: 44.0 Class :character
## Median : -122.4 Median :100.0 Mode :character
## Mean : -122.4 Mean :138.1
## 3rd Qu.: -122.3 3rd Qu.:235.0
## Max. : -121.9 Max. :427.0
##
## end_station_latitude end_station_longitude bike_id user_type
## Min. : 0.00 Min. : -122.5 Min. : 11 Length:238134
## 1st Qu.:37.77 1st Qu.: -122.4 1st Qu.:1972 Class :character
## Median :37.78 Median : -122.4 Median :4317 Mode :character
## Mean :37.77 Mean : -122.3 Mean :3758
## 3rd Qu.:37.80 3rd Qu.: -122.3 3rd Qu.:5383
## Max. :37.88 Max. : 0.0 Max. :7196
##
## member_birth_year member_gender bike_share_for_all_trip
## Min. :1920 Length:238134 Length:238134
## 1st Qu.:1980 Class :character Class :character
## Median :1987 Mode :character Mode :character
## Mean :1985
## 3rd Qu.:1992
## Max. :2001
##

```

Podsumowanie danych



Analiza:

- subskrybentów jest znacznie więcej niż zwykłych użytkowników
- maksimum wypożyczeń
 - w godzinach 7-8 i 17-18 w dni robocze, kiedy ludzie przemieszczają się do/z pracy
 - w porze południowej w ciągu weekendu, gdy w dniu wolnym od pracy organizowane są wycieczki rowerowe
- w weekend ogólne użytkowanie rowerów spada, gdyż nie ma porannych i popołudniowych wyjazdów, a w południe zachodzi tylko lekki wzrost dla zwykłych użytkowników
- w dni robocze najmniej wypożyczeń jest w poniedziałek, co prawdopodobnie spowodowane jest mniejszą motywacją do aktywności fizycznej po weekendzie
- w późniejszych porach jeżdżą głównie subskrybenci
- w nocy więcej osób korzysta z rowerów w weekend niż w tygodniu

Długość podróży w zależności od użytkownika i dnia tygodnia

```
## # A tibble: 14 x 3
## # Groups:   week_day [7]
##   week_day    user_type mean_duration_min
##   <ord>      <chr>          <dbl>
## 1 niedziela  Customer          30.2
## 2 sobota    Customer          28.7
## 3 poniedziałek Customer          21.5
## 4 piątek    Customer          20.7
## 5 czwartek  Customer          20.3
## 6 środa     Customer          19.9
## 7 wtorek    Customer          18.3
## 8 sobota    Subscriber         12.9
## 9 niedziela Subscriber         12.8
## 10 piątek   Subscriber         11.2
## 11 środa    Subscriber         10.9
## 12 czwartek Subscriber         10.7
## 13 wtorek   Subscriber         10.7
## 14 poniedziałek Subscriber         10.7
```

- osoby nie będące subskrybentami korzystają z roweru dłużej niż zwykli użytkownicy
- dla obu typów użytkowników w weekendy wypożyczenia są najdłuższe, ze względu na aktywny odpoczynek w wolnym dniu od pracy
- średni czas użytkowania dla subskrybentów jest prawie niezmienny od poniedziałku do czwartku, co oznacza, że są to osoby mające regularne trasy

Podział użytkowników według płci

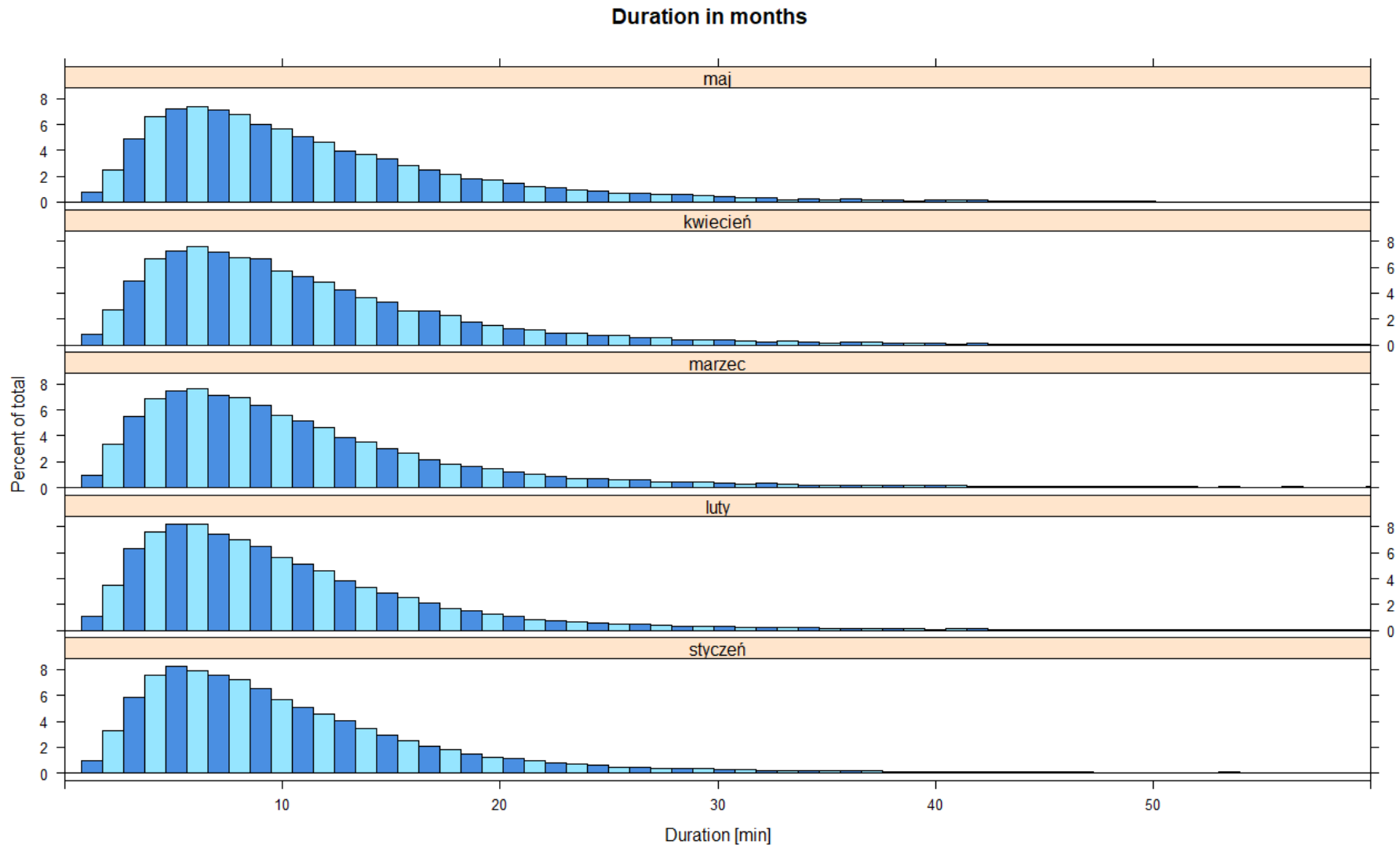
```
## # A tibble: 3 x 2
## # Groups:   member_gender [3]
##   member_gender      n
##   <chr>          <int>
## 1 Male          176089
## 2 Female         57445
## 3 Other          4560
```

- mężczyźni to aż 74 % osób wypożyczających rowery, kobiety- 24%

Liczba wypożyczeń w godzinach 22-4 według płci

```
##   member_gender      n
##   <chr>          <int>
## 1 Male           1803
## 2 Female          368
## 3 Other           52
```

- kobiety rzadziej jeżdżą nocą, stanowią tylko 17% użytkowników, mężczyźni-81%



Analiza:

- przeważają wypożyczenia o czasie trwania ok. 5 minut, gdyż celem miejskich rowerów jest jazda w obrębie centrum miasta
- liczba wypożyczeń o czasie trwania ponad 30 minut jest znikoma

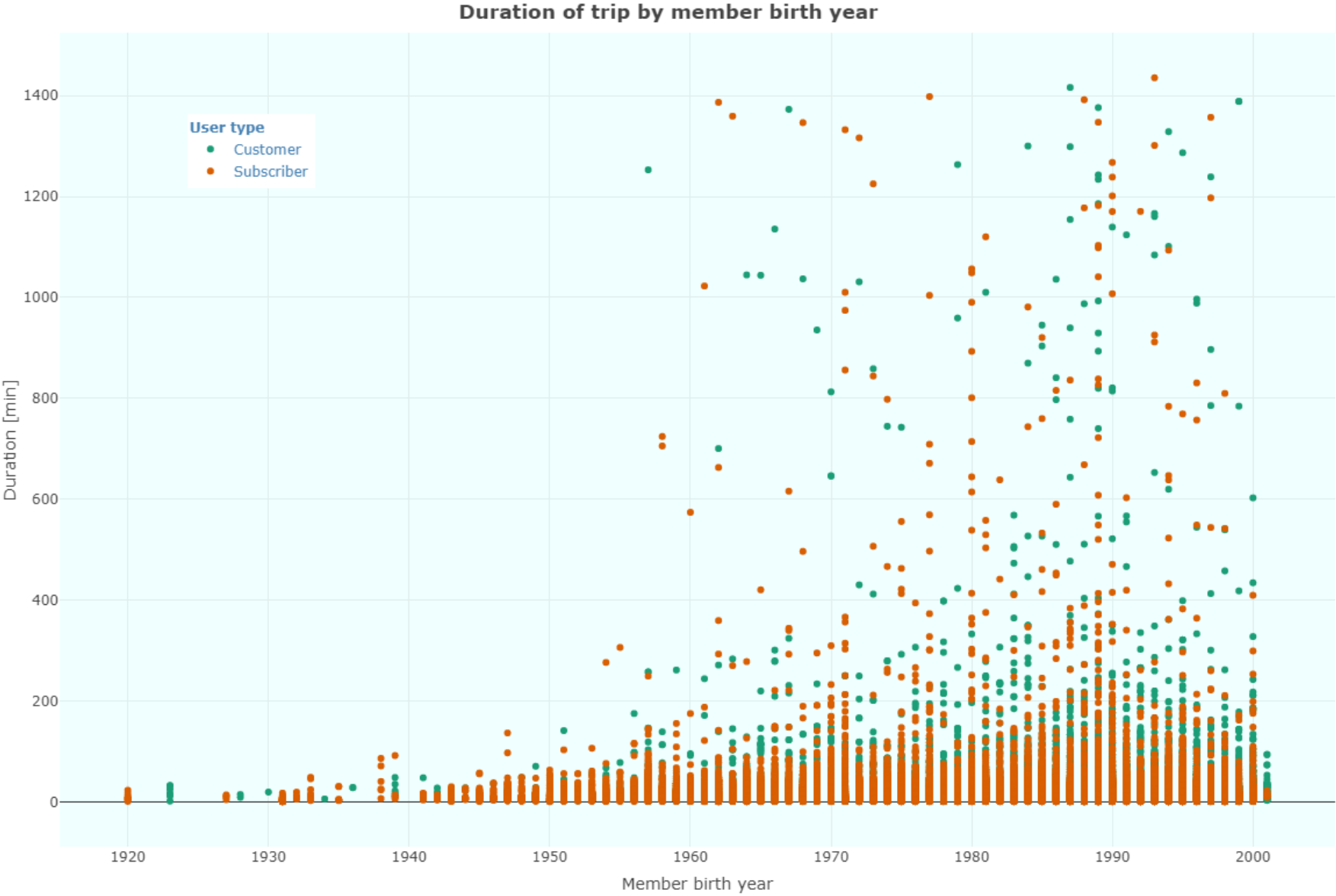
Mediana czasu trwania wypożyczeń rowerów według miesięcy

```
## # A tibble: 5 x 2
##   month      median_duration_min
##   <ord>          <dbl>
## 1 maj              9.62
## 2 kwiecień         9.4
## 3 marzec           9.15
## 4 styczeń          8.72
## 5 luty             8.58
```

- w maju podróże są najdłuższe, ze względu na najwyższą temperaturę powietrza
- najkrótsze wypożyczenia rejestrowane są w lutym, gdyż jest to najzimniejszy miesiąc

```
## # A tibble: 5 x 2
## # Groups:   month [5]
##   month      n
##   <ord>   <int>
## 1 marzec  57967
## 2 styczeń 53517
## 3 luty    46948
## 4 kwiecień 43145
## 5 maj     36517
```

- najwięcej wypożyczeń występuje w marcu, przypuszczalnie przez zmianę pory roku i rozpoczęcie sezonu rowerowego
- najrzadziej korzysta się z rowerów miejskich w maju



Analiza:

- mediana wieku użytkowników to 32 lata
- dłuższe podróże rowerem odbywają zwykli użytkownicy
- najdłuższe wypożyczenia trwają prawie całą dobę

Mediana czasu trwania wypożyczeń w zależności od roku urodzenia i płci

```
## # A tibble: 10 x 3
## # Groups:   member_gender [3]
##   member_gender member_birth_year median_duration_min
##   <chr>          <int>          <dbl>
## 1 Female         1939           92.3
## 2 Male           1938           33.5
## 3 Other          1954           33.4
## 4 Female         1948           30.3
## 5 Female         1936           28.8
## 6 Other          2000           28.5
## 7 Other          1961           28.0
## 8 Female         1944           25.2
## 9 Other          1949           24.8
## 10 Male          1923           21.2
```

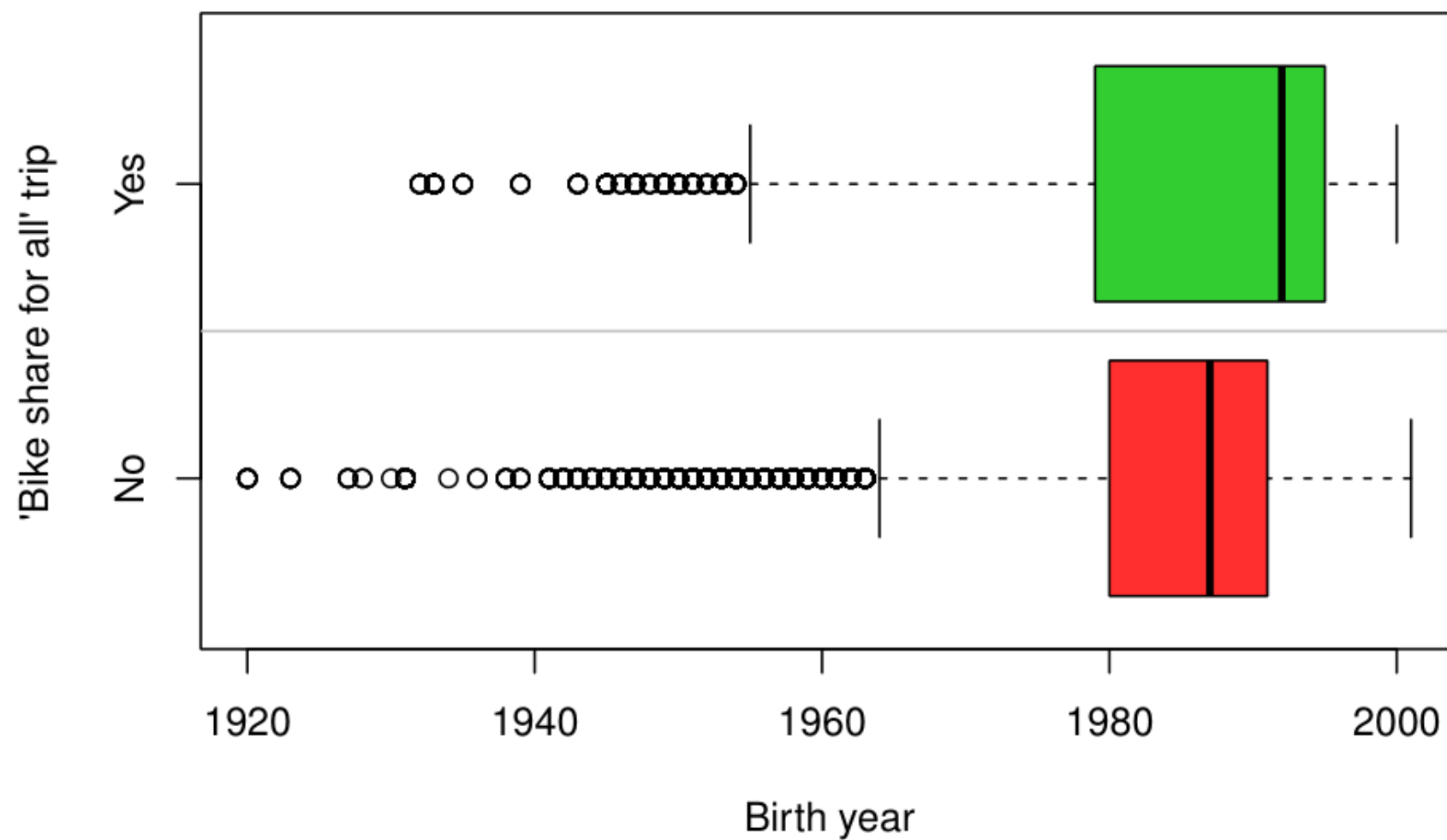
- najdłuższe podróże odbywają starsze osoby, zapewne przez to, że mają mniej siły i jadą wolniejszym tempem
- najdłuższe wycieczki mają 80-letnie kobiety

Długość podróży w zależności od płci

```
## # A tibble: 3 x 2
##   member_gender median_duration_min
##   <chr>          <dbl>
## 1 Female         10.2
## 2 Other           9.88
## 3 Male           8.7
```

- kobiety wypożyczają rowery na dłużej

Box plot



Analiza:

Mediana roku urodzenia osób korzystających z programu 'Bike share for all'

```
## # A tibble: 2 x 2
##   bike_share_for_all_trip median_birth_year
##   <chr>                  <int>
## 1 Yes                    1992
## 2 No                     1987
```

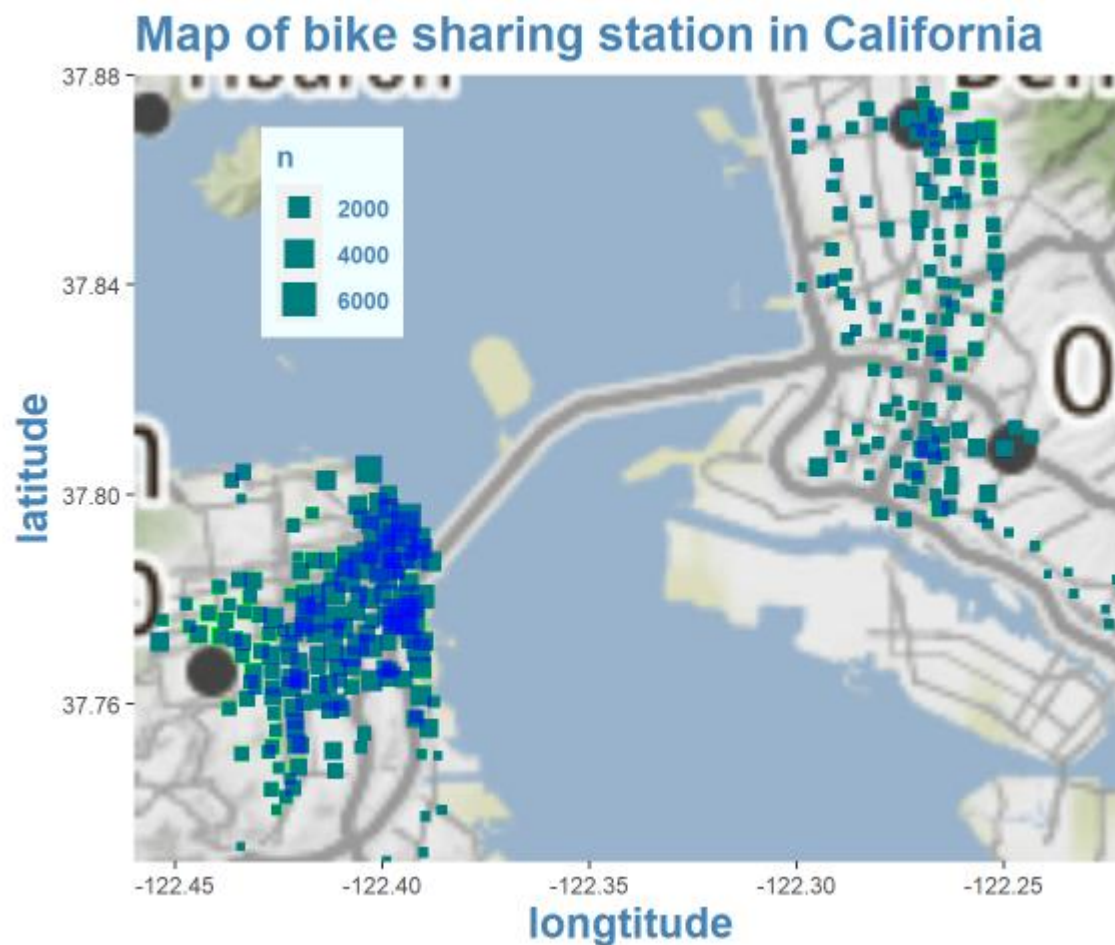
- osoby młodsze częściej korzystają z programu dla osób z małym dochodem- mediana to 27 lat
- osoby po skończeniu 80 lat głównie nie są zapisane do programu

Liczba użytkowników korzystających z programu 'Bike share for all'

```
## # A tibble: 2 x 2
## # Groups:   bike_share_for_all_trip [2]
##   bike_share_for_all_trip      n
##   <chr>                  <int>
## 1 No                    216071
## 2 Yes                    22023
```

- 9 % osób korzysta z programu 'Bike share for all'

Pakiet ggmap



Mapa przedstawia miejsca rozpoczęcia i zakończenia wypożyczeń w *San Francisco* (lewa strona) i *Oakland* (prawa strona) według ilości ich wystąpień w danych

kolor zielony- zarejestrowana stacja początkowa

kolor niebieski- zarejestrowana stacja końcowa

Analiza:

- w San Francisco częściej korzysta się z rowerów miejskich niż w Oakland
- najwięcej wypożyczeń odbywa się przy głównych ulicach oraz przy wybrzeżu, gdzie skonstruowane są liczne ścieżki rowerowe
- jazda rowerem najczęściej rozpoczyna się na obrzeżach miast, a kończy się przy wybrzeżach