

# Projekt SQL

## SF Bay Area Bike Share

Tak się jeździ w San Francisco, w pogodę i niepogodę,  
czyli najpopularniejsze trasy w SF i jak (czy) to się zmienia  
w zależności od pogody



**Grupa nonamezzz**

Barbara Gradzik - Hupało

Iwona Onuszko

Michał Mielniczek

<https://github.com/infoshareacademy/nonamezzz>

# Plan prezentacji

- Cel projektu
- Podstawowe analizy pogody oraz przejazdów
- Najpopularniejsze trasy a pogoda
- Analiza ilości przejazdów w zależności od pogody
- Narzędzie biznesowe
- Podsumowanie

# Cel i krótki opis projektu

## Główny cel projektu:

Sprawdzenie czy i w jaki sposób warunki pogodowe wpływają na najpopularniejsze trasy rowerowe w rejonie San Francisco.

## Wykorzystana baza danych:

<https://www.kaggle.com/benhamner/sf-bay-area-bike-share#trip.csv> (Baza obejmuje rekordy od VIII 2013 do VIII 2015)

## Wykorzystane narzędzia:

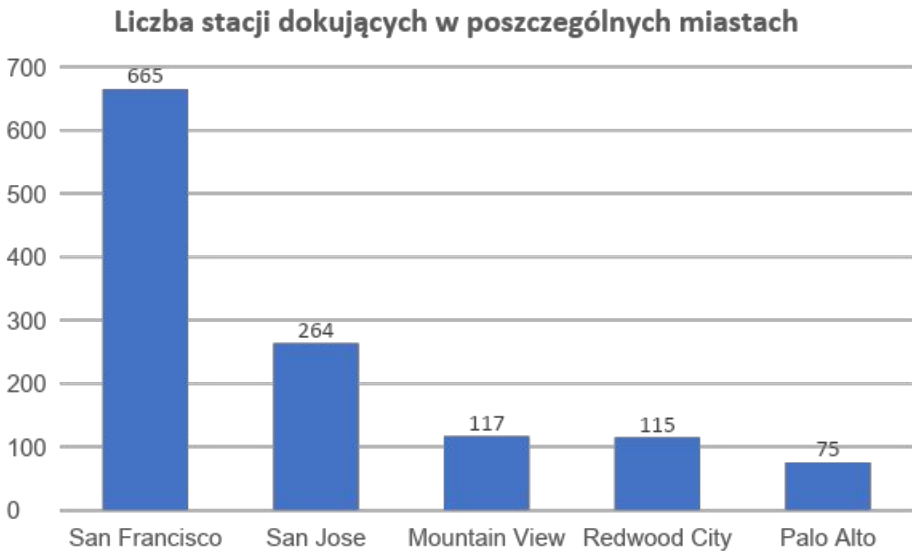
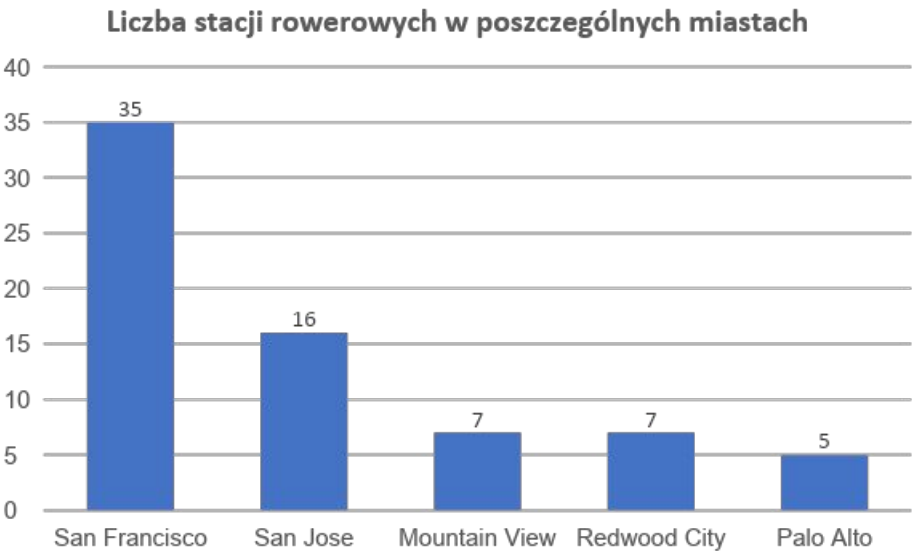
- W projekcie korzystano z języka zapytań SQL (PostgreSQL w narzędziu/aplikacji Dbeaver)
- Wizualizacja danych w Microsoft Excel.

W analizie warunków pogodowych panujących w rejonie San Francisco uwzględniono wszystkie rekordy, natomiast z analizy warunków panujących na najpopularniejszych trasach wykluczono wypożyczenia o czasie dłuższym niż 24h.

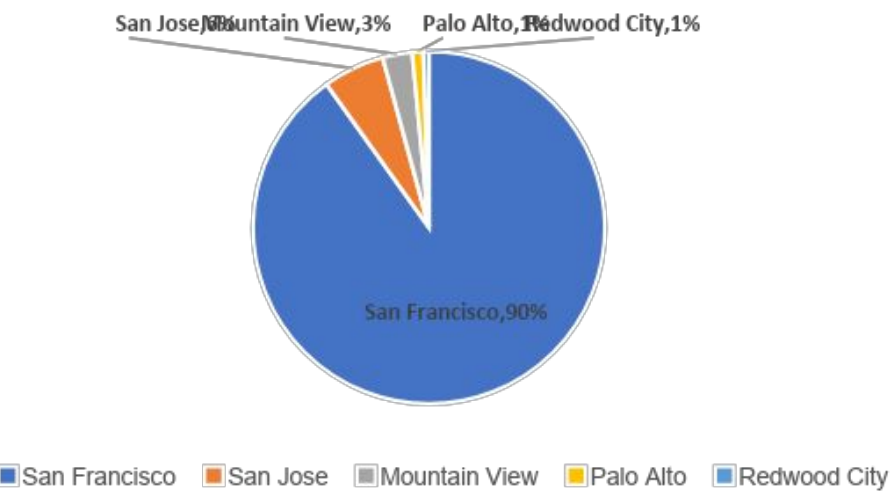
## Wśród zarejestrowanych w poszczególnych stacjach czynników pogodowych do analizy ich wpływu wybrano:

- temperaturę,
- zachmurzenie,
- zjawiska pogodowe (mgła, deszcz, burza),
- prędkość wiatru,
- widoczność,
- opady.

# Podstawowe informacje



**Struktura procentowa rowerów, które wyjechały z danego miasta niezależnie od warunków pogodowych oraz czasu**



Stacje rowerowe, z których wyjechało najwięcej rowerów w ciągu jednego dnia każdego badanego miesiąca:

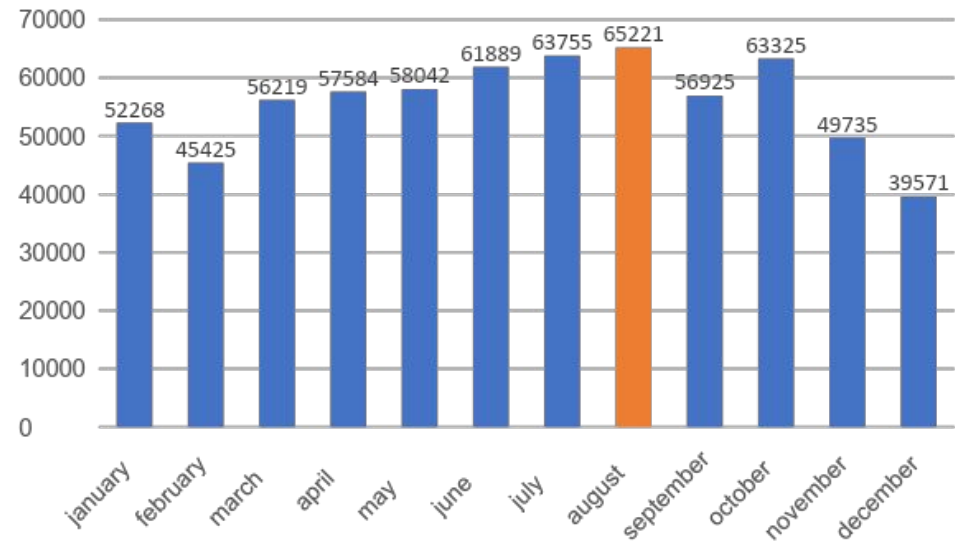
Harry Bridges Plaza (Ferry Building)

San Francisco Caltrain (Townsend at 4th)

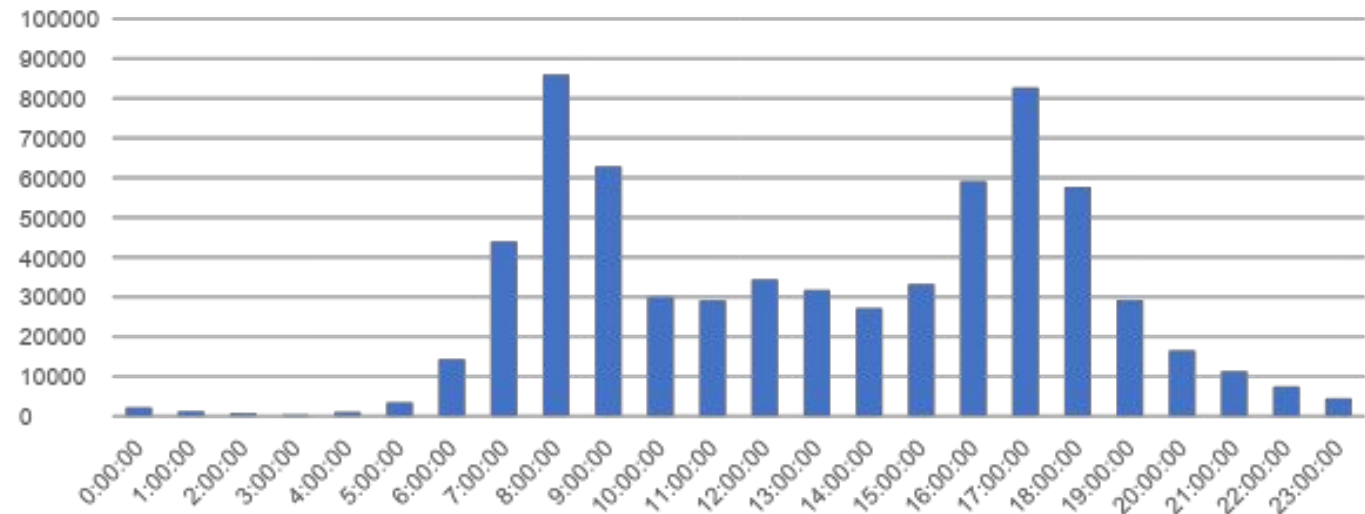
San Francisco Caltrain 2 (330 Townsend)

# Podstawowe informacje

Liczba rowerów wypożyczonych w danym miesiącu



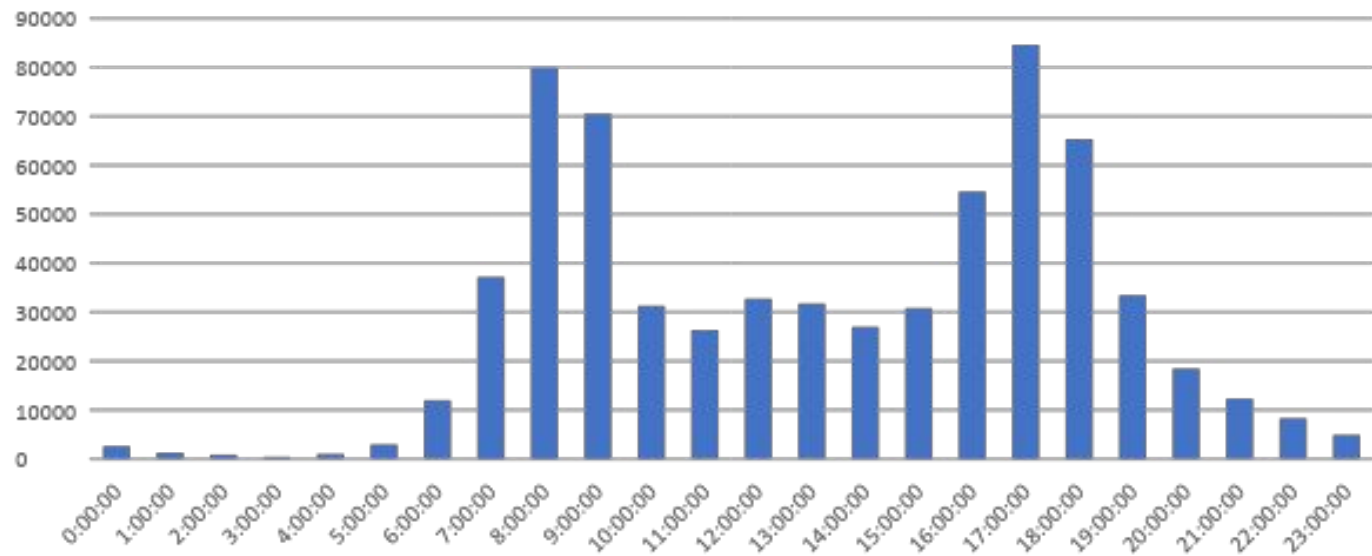
Liczba wypożyczonych rowerów w poszczególnych godzinach



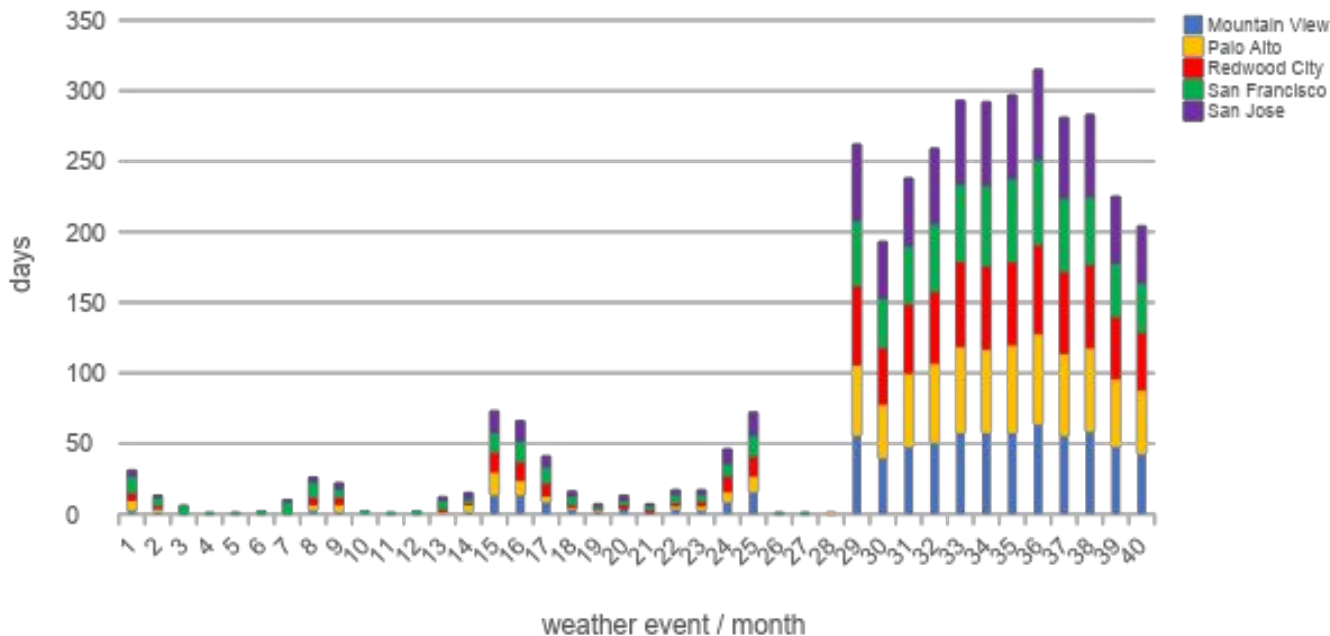
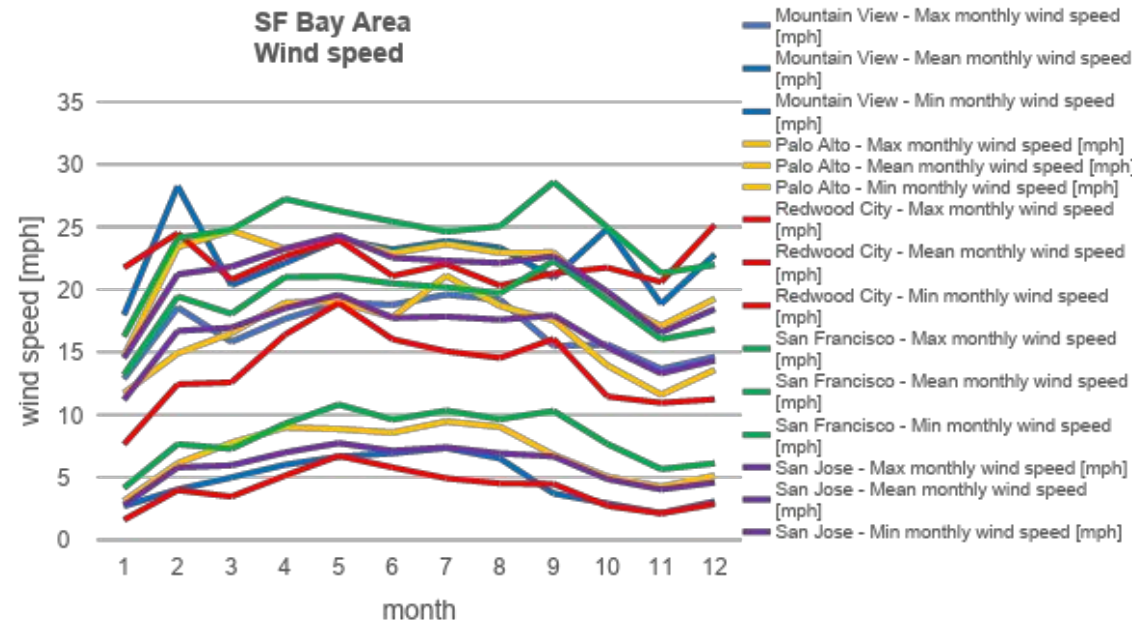
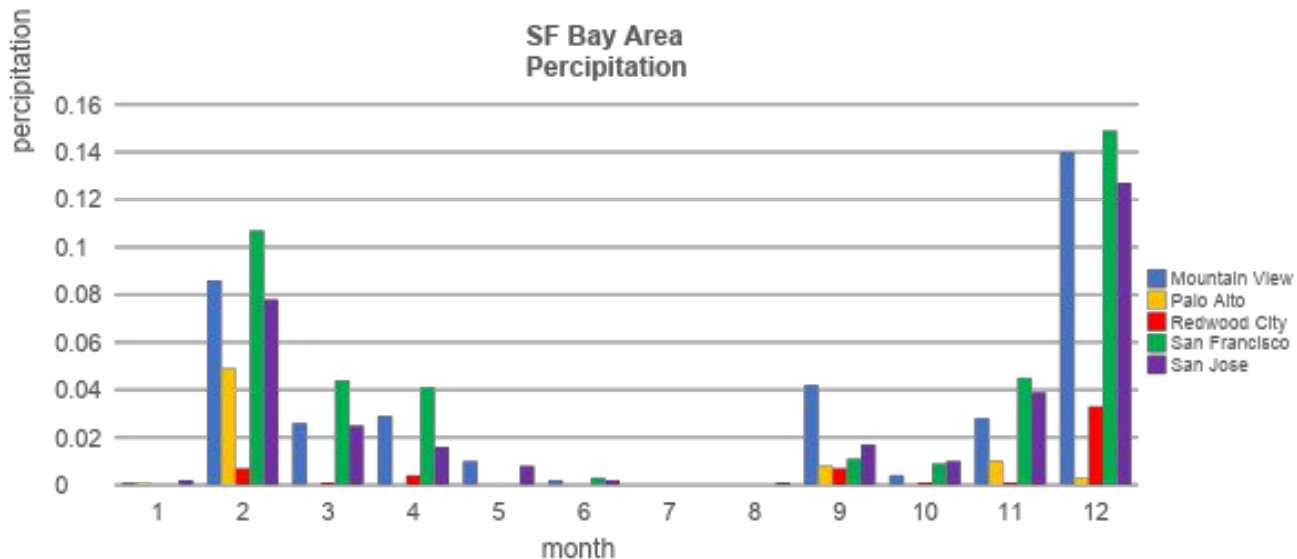
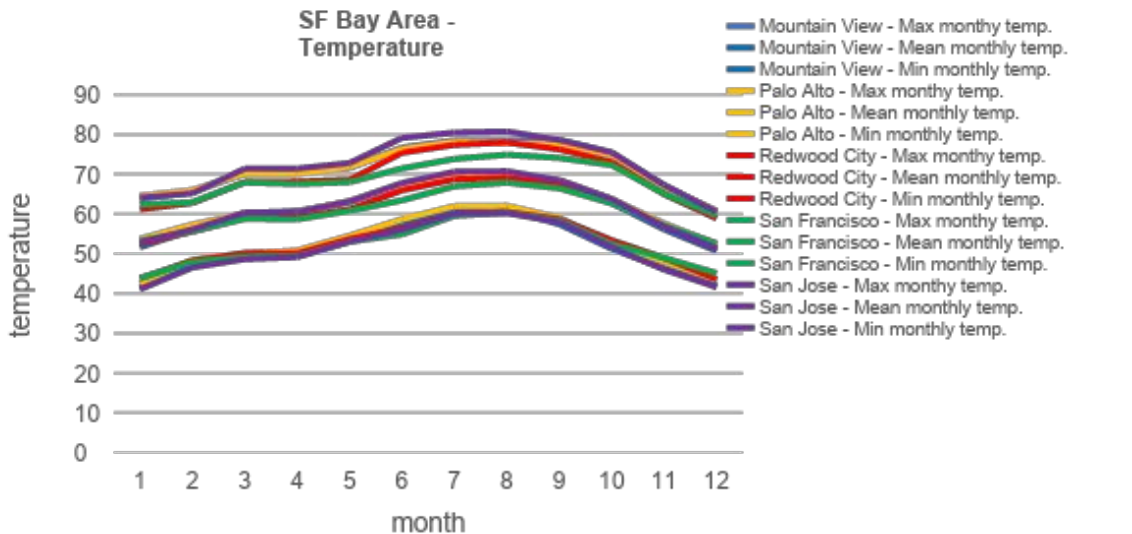
Średnia liczba rowerów, która została wypożyczona  
w badanym okresie w poszczególne dni tygodnia

day	avg_bikes
tuesday	19.02
wednesday	18.62
thursday	18.49
monday	18.02
friday	17.00
saturday	8.01
sunday	6.99

Liczba oddanych rowerów w poszczególnych godzinach

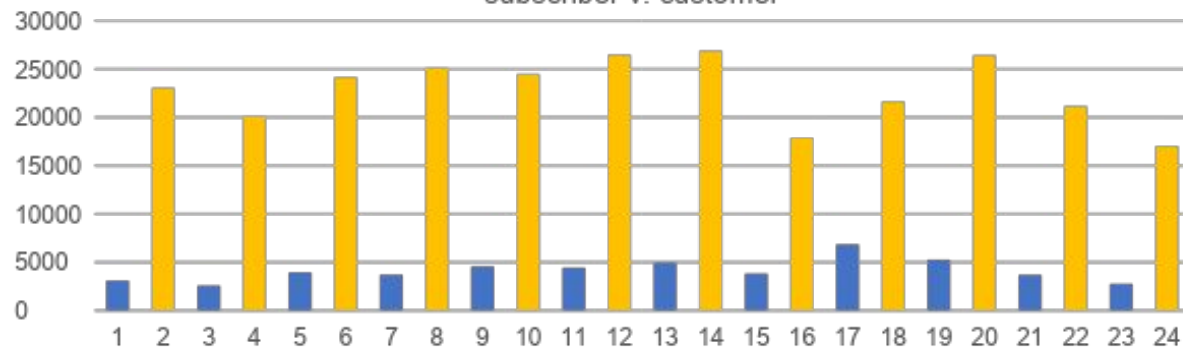


# Analiza warunków pogodowych panujących w rejonie San Francisco w zależności lokalizacji stacji pogodowej



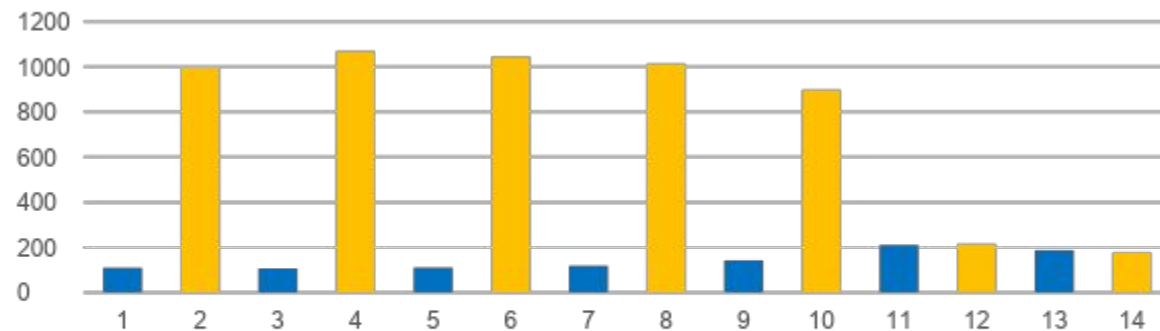
# Analiza użytkowników SF Bay Area Bike Share

Trips per month  
subscriber v. customer



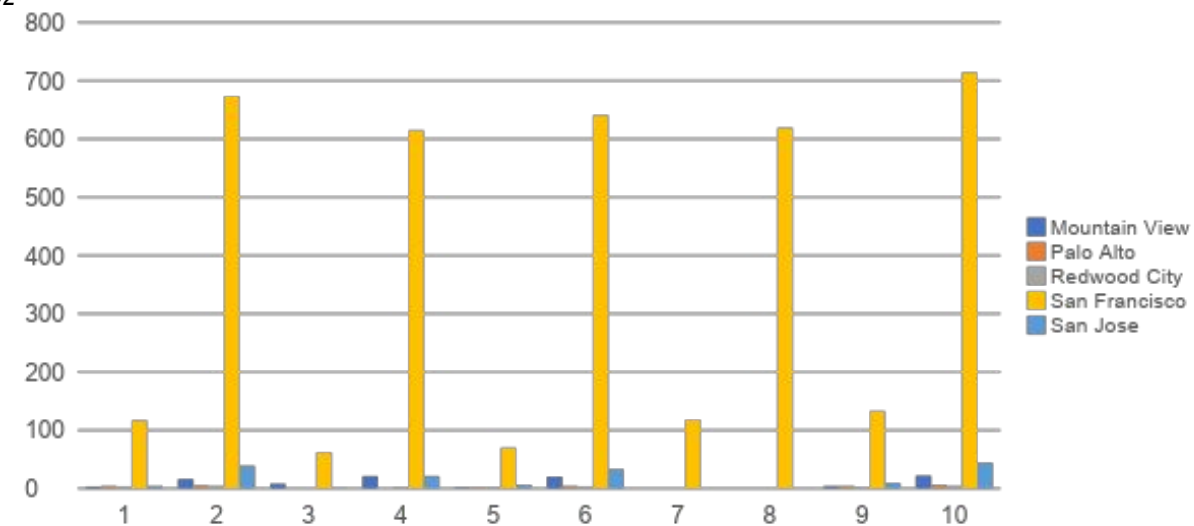
month

Trips per weekday  
subscriber v. customer



Współczynnik korelacji  
zdarzenie pogodowe i  
ilość tras w ciągu dnia:  
0,002

Trips per day with weather events



Trips duration time [min] per weekday  
subscriber v. customer





Do analizy wpływu warunków pogodowych wybrano 10 najpopularniejszych tras:

Start station name	End station name	No of trips	Avg duration [min]
San Francisco Caltrain 2 (330 Townsend)	Townsend at 7th	6 215	5
Harry Bridges Plaza (Ferry Building)	Embarcadero at Sansome	6 164	21
Townsend at 7th	San Francisco Caltrain (Townsend at 4th)	5 041	5
2nd at Townsend	Harry Bridges Plaza (Ferry Building)	4 839	10
Harry Bridges Plaza (Ferry Building)	2nd at Townsend	4 357	11
Embarcadero at Sansome	Steuart at Market	4 269	9
Embarcadero at Folsom	San Francisco Caltrain (Townsend at 4th)	3 966	12
Steuart at Market	2nd at Townsend	3 903	10
2nd at South Park	Market at Sansome	3 627	9
San Francisco Caltrain (Townsend at 4th)	Harry Bridges Plaza (Ferry Building)	3 622	14

SUBSCRIBER

Start station name	End station name	No of trips	Avg duration [min]
San Francisco Caltrain 2 (330 Townsend)	Townsend at 7th	6 049	5
Townsend at 7th	San Francisco Caltrain (Townsend at 4th)	4 864	4
2nd at Townsend	Harry Bridges Plaza (Ferry Building)	4 319	9
Harry Bridges Plaza (Ferry Building)	2nd at Townsend	3 926	10
Embarcadero at Folsom	San Francisco Caltrain (Townsend at 4th)	3 815	11
Embarcadero at Sansome	Steuart at Market	3 801	7
Steuart at Market	2nd at Townsend	3 718	9
2nd at South Park	Market at Sansome	3 516	9
Harry Bridges Plaza (Ferry Building)	Embarcadero at Sansome	3 455	8
Temporary Transbay Terminal (Howard at Beale)	San Francisco Caltrain (Townsend at 4th)	3 445	11

CUSTOMER

Start station name	End station name	No of trips	Avg duration [min]
Harry Bridges Plaza (Ferry Building)	Embarcadero at Sansome	2 709	38
Embarcadero at Sansome	Embarcadero at Sansome	1 843	80
Harry Bridges Plaza (Ferry Building)	Harry Bridges Plaza (Ferry Building)	1 489	114
Embarcadero at Sansome	Harry Bridges Plaza (Ferry Building)	1 237	27
Embarcadero at Vallejo	Embarcadero at Sansome	1 006	44
University and Emerson	University and Emerson	831	110
Harry Bridges Plaza (Ferry Building)	Embarcadero at Vallejo	650	54
Steuart at Market	Embarcadero at Sansome	630	36
Market at 4th	Market at 4th	604	145
Powell at Post (Union Square)	Powell at Post (Union Square)	576	181



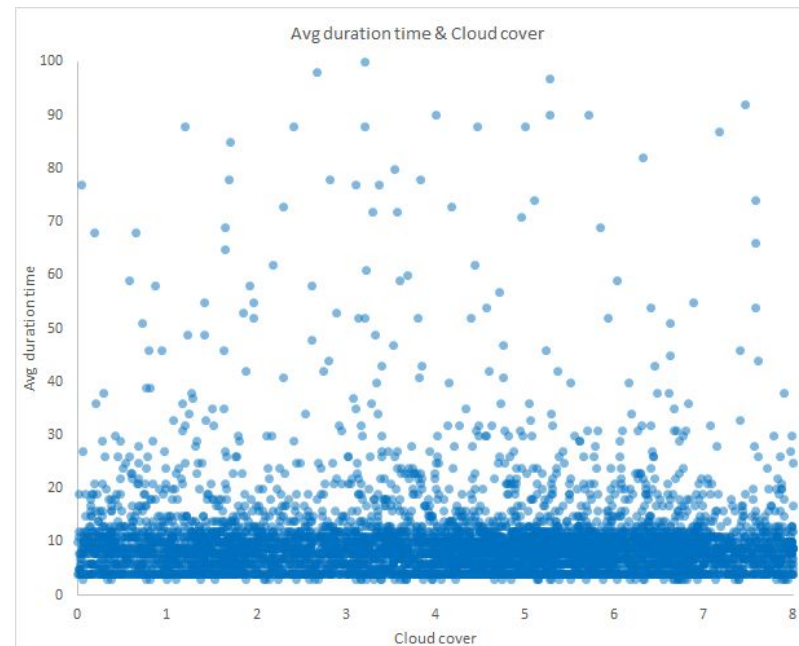
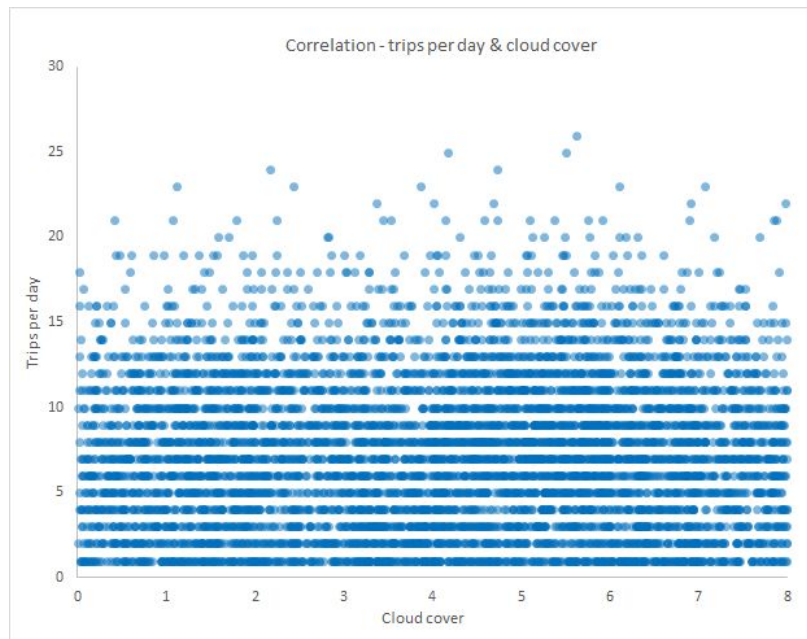
# Analiza korelacji danych w zależności od czynnika pogodowego

	trips per day	avg duration of trips	trips per day	avg duration of trips	trips per day	avg duration of trips	trips per day	avg duration of trips	trips per day	avg duration of trips	trips per day	avg duration of trips	trips per day	avg duration of trips
	10 the most popular routes		10 the most popular routes subscriber		10 the most popular routes customer		10 the most popular routes weekend		10 the most popular routes weekdays		10 the most popular routes weekend & customer		10 the most popular routes weekdays & customer	
cloud cover	0,007	-0,022	0,007	-0,010	-0,003	-0,022	0,009	-0,022	0,007	-0,010	0,000	-0,019	-0,003	-0,022
min temperature	0,031	0,012	0,030	0,015	0,040	0,017	0,041	0,022	0,027	0,013	0,041	0,020	0,040	0,017
mean temperature	0,032	0,017	0,038	0,017	0,051	0,023	0,048	0,026	0,036	0,016	0,051	0,026	0,051	0,023
max temperature	0,027	0,019	0,039	0,015	0,049	0,025	0,045	0,025	0,038	0,014	0,050	0,026	0,049	0,025
mean wind speed	0,006	-0,004	-0,008	0,013	0,000	0,004	0,004	0,005	-0,013	0,011	-0,001	0,006	0,000	0,004
max wind speed	-0,005	-0,002	-0,022	0,009	-0,013	-0,002	-0,013	-0,003	-0,027	0,008	-0,015	0,000	-0,013	-0,002
mean visibility	0,003	0,022	0,008	0,016	0,016	0,031	0,008	0,040	0,006	0,014	0,014	0,035	0,016	0,031
precipitation	-0,025	-0,003	-0,032	-0,004	-0,038	-0,011	-0,034	-0,007	-0,033	-0,004	-0,039	-0,010	-0,038	-0,011
rain	-0,014	-0,012	-0,021	-0,012	-0,028	-0,023	-0,023	-0,019	-0,020	-0,011	-0,029	-0,021	-0,028	-0,023
rain-thunderstorm	0,001	-0,001	0,003	0,001	0,001	-0,002	0,005	-0,003	0,003	0,000	0,002	-0,003	0,001	-0,002
fog	0,001	-0,003	-0,002	-0,005	-0,006	-0,001	-0,006	-0,004	-0,001	-0,001	-0,007	-0,004	-0,006	-0,001
fog-rain	-0,004	-0,007	-0,007	-0,005	-0,009	-0,007	-0,009	-0,013	-0,005	-0,005	-0,008	-0,007	-0,009	-0,007

# Analiza korelacji danych w zależności od czynnika pogodowego

	trips per day	avg duration of trips	trips per day	avg duration of trips	trips per day	avg duration of trips	trips per day	avg duration of trips	trips per day	avg duration of trips	trips per day	avg duration of trips	trips per day	avg duration of trips	trips per day	avg duration of trips	trips per day	avg duration of trips	trips per day	avg duration of trips	trips per day	avg duration of trips
	ALL TRIPS		San Francisco Caltrain 2 (330 Townsend) --> Townsend at 7th		Harry Bridges Plaza (Ferry Building)-->Embarcadero at Sansome		Townsend at 7th-->San Francisco Caltrain (Townsend at 4th)		2nd at Townsend-->Harry Bridges Plaza (Ferry Building)		Harry Bridges Plaza (Ferry Building)-->2nd at Townsend		Embarcadero at Sansome-->Steuart at Market		Embarcadero at Folsom-->San Francisco Caltrain (Townsend at 4th)		Steuart at Market-->2nd at Townsend		2nd at South Park-->Market at Sansome		San Francisco Caltrain (Townsend at 4th)-->Harry Bridges Plaza (Ferry Building)	
cloud cover	0,007	-0,022	0,060	0,056	-0,005	-0,051	0,024	-0,032	0,038	-0,001	0,031	-0,085	0,066	-0,026	-0,008	-0,103	-0,039	-0,017	-0,021	0,108	-0,034	-0,023
min temperature	0,031	0,012	-0,018	0,032	0,292	-0,012	-0,088	0,063	0,104	-0,015	0,059	-0,037	0,162	-0,008	0,073	0,049	0,074	0,053	0,019	-0,038	-0,057	0,079
mean temperature	0,032	0,017	-0,002	0,013	0,329	-0,004	-0,082	0,066	0,156	-0,017	0,114	0,009	0,142	0,002	0,119	0,106	0,157	0,059	0,070	-0,092	-0,009	0,092
max temperature	0,027	0,019	0,016	-0,008	0,298	0,000	-0,060	0,055	0,173	-0,008	0,143	0,046	0,099	0,005	0,138	0,135	0,202	0,055	0,106	-0,126	0,038	0,086
mean wind speed	0,006	-0,004	-0,069	0,072	-0,069	0,072	-0,188	0,027	-0,063	0,034	-0,075	-0,071	0,086	-0,008	-0,091	-0,052	-0,050	0,002	-0,046	-0,003	-0,112	0,015
max wind speed	-0,005	-0,002	-0,091	0,081	-0,043	-0,013	-0,151	-0,012	-0,122	0,057	-0,106	-0,054	-0,001	0,016	-0,154	-0,044	-0,067	-0,005	-0,038	-0,053	-0,116	0,003
mean visibility	0,003	0,022	-0,024	0,008	0,141	0,024	-0,107	0,036	-0,038	0,030	-0,029	0,016	0,033	0,055	0,060	0,032	0,050	0,022	0,052	-0,034	-0,005	0,069
precipitation	-0,025	-0,003	-0,086	-0,027	-0,086	-0,027	-0,019	-0,017	-0,091	-0,007	-0,106	-0,029	-0,115	-0,033	-0,179	-0,026	-0,192	-0,013	-0,136	0,321	-0,081	-0,043
rain	0,210	-0,011	-0,004	0,005	-0,142	-0,039	0,028	-0,024	-0,051	-0,004	-0,059	-0,047	-0,059	-0,058	-0,112	-0,048	-0,116	-0,022	-0,048	0,166	-0,046	-0,071
rain-thunderstorm	0,035	-0,001	-0,011	-0,009	-0,026	-0,015	-0,053	-0,003	0,002	0,148	0,087	-0,008	0,064	-0,007	0,034	-0,003	-0,034	-0,003	-0,025	-0,003	0,019	-0,008
fog	0,178	-0,002	-0,004	0,024	-0,045	-0,005	0,053	-0,027	0,044	-0,023	0,059	-0,022	-0,069	0,003	-0,022	0,008	-0,017	0,146	0,019	-0,031	0,015	-0,024
fog-rain	0,066	-0,006	-0,008	-0,016	-0,025	-0,023	0,037	-0,009	-0,039	-0,014	-0,015	-0,019	0,020	-0,018	-0,040	-0,016	0,006	-0,007	-0,064	-0,013	0,017	-0,016

# Analiza rozrzutu danych w zależności od czynnika pogodowego

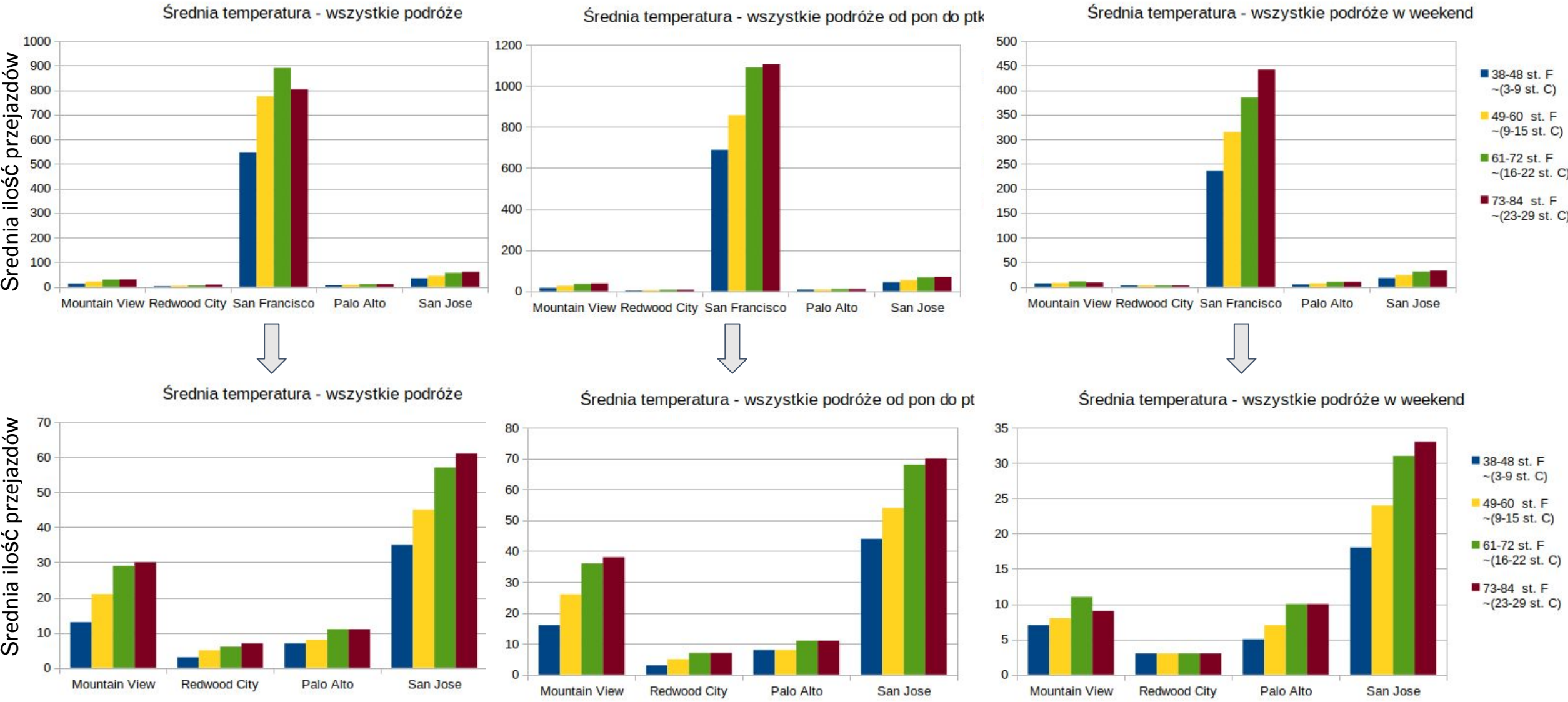


# Analiza korelacji danych w zależności od czynnika pogodowego

Znaleźliśmy trasy ze słabą ujemną korelacją w zależności od czynnika pogodowego.  
Trasy te jednak nie należą do najpopularniejszych w rejonie San Francisco.

Trip name	No of trips	Cloud cover corr
Mountain View Caltrain Station --> Mountain View City Hall	1105	-0.449
Mountain View City Hall --> Mountain View Caltrain Station	979	-0.476

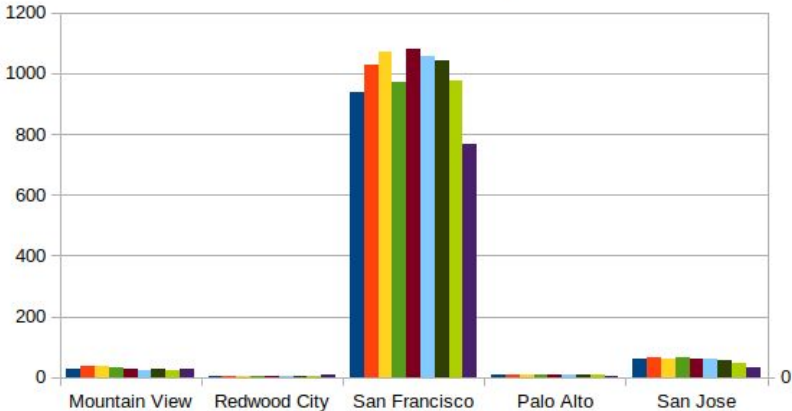
# Analiza Średniej ilości przejazdów - Średnia dzienna temperatura



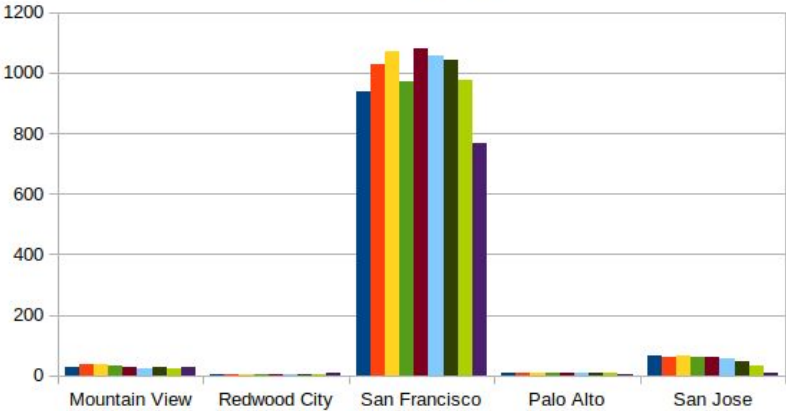
# Analiza Średniej ilości przejazdów - poziom zachmurzenia

Średnia ilość przejazdów

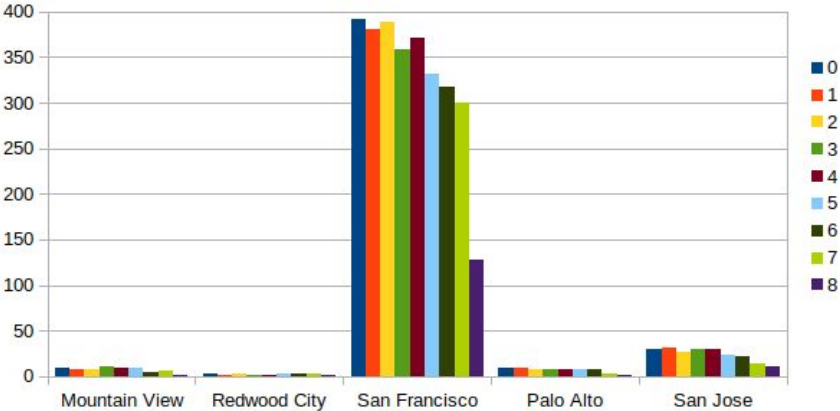
Poziom zachmurzenia - wszystkie podróże



Poziom zachmurzenia - wszystkie podróże od pon do ptk

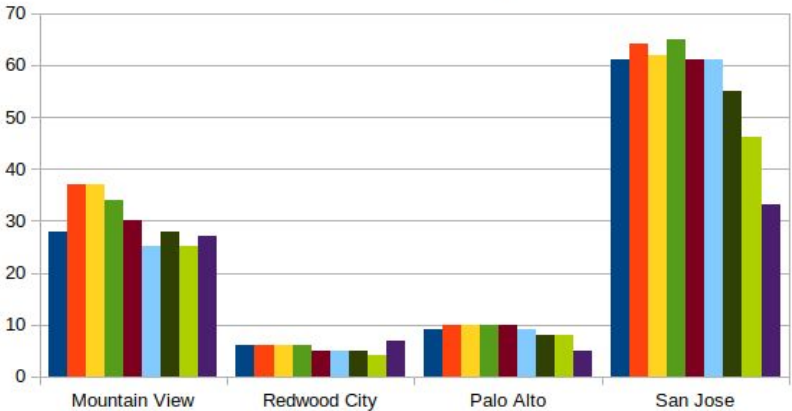


Poziom zachmurzenia - wszystkie podróże w weekend

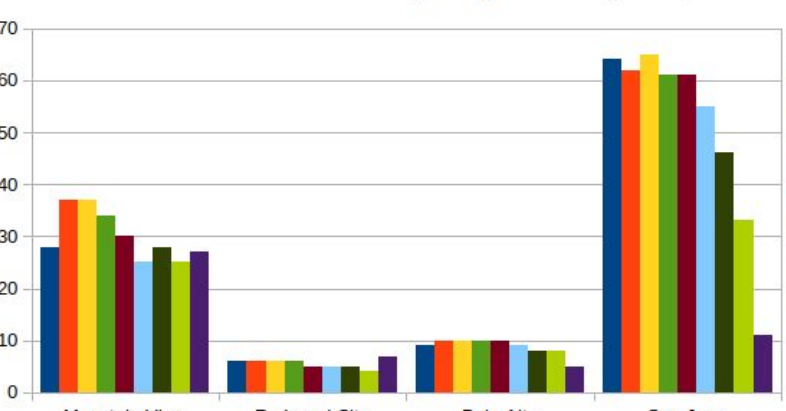


Średnia ilość przejazdów

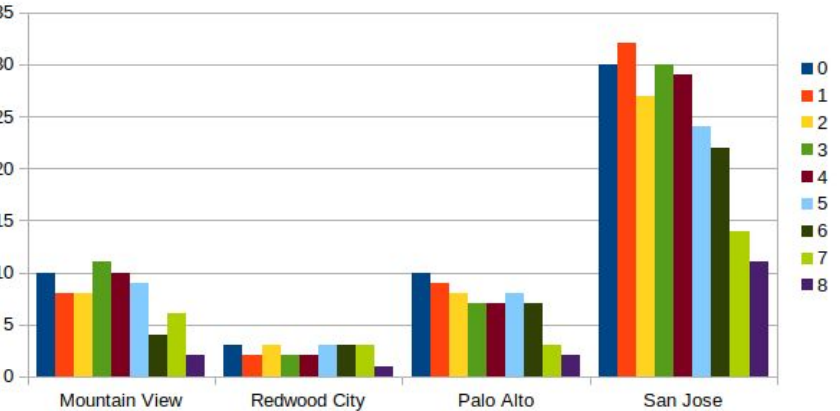
Poziom zachmurzenia - wszystkie podróże



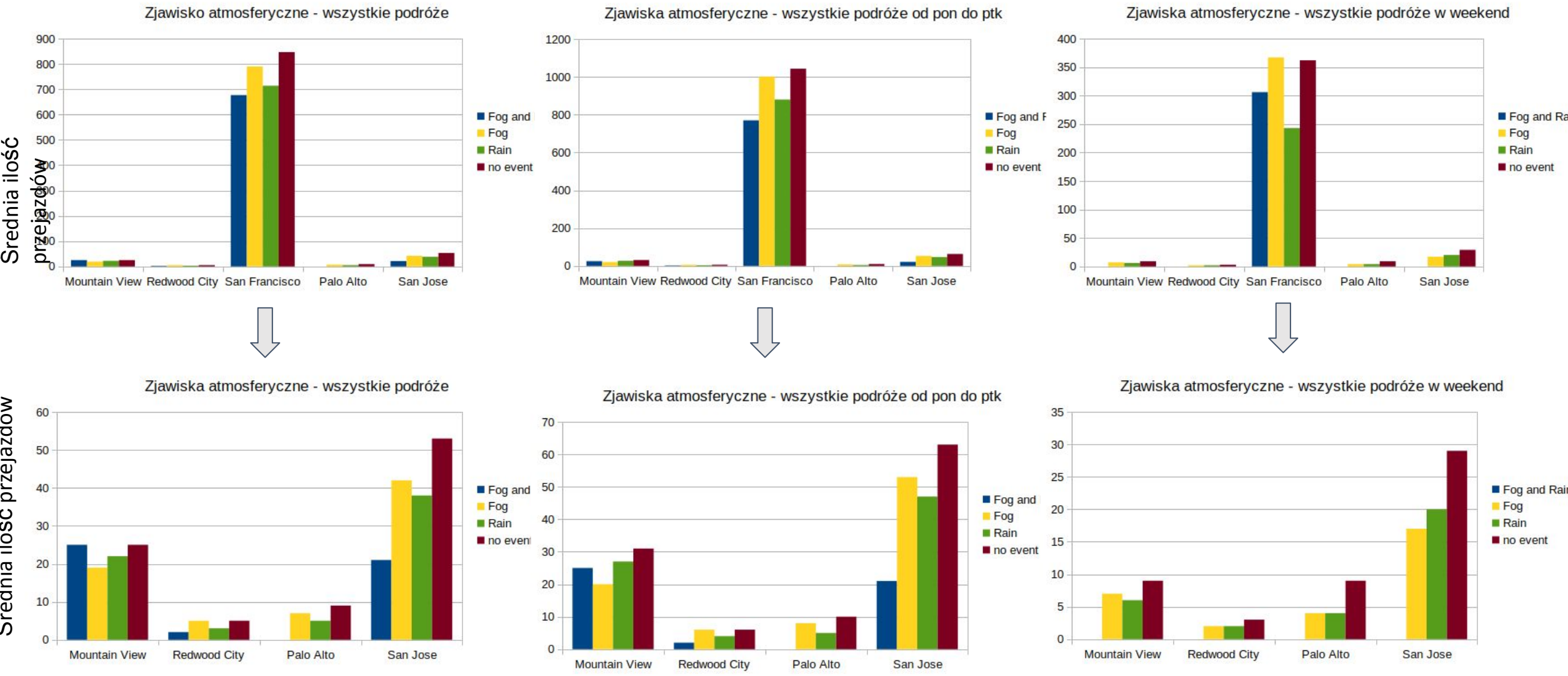
Poziom zachmurzenia - wszystkie podróże od pon do ptk



Poziom zachmurzenia - wszystkie podróże w weekend



# Analiza Średniej ilości przejazdów- zjawiska atmosferyczne





# Rankingi - poniedziałek - piątek vs weekend

Start station name	End station name	Total trips	Duration (min)	Ratio
San Francisco Caltrain 2 (330 Townsend)	Townsend at 7th	5851	5	0.01
Townsend at 7th	San Francisco Caltrain (Townsend at 4th)	4718	4	0.008
Harry Bridges Plaza (Ferry Building)	Embarcadero at Sansome	4614	17	0.0079
2nd at Townsend	Harry Bridges Plaza (Ferry Building)	4424	9	0.0075
Embarcadero at Sansome	Steuart at Market	3958	8	0.0067
Harry Bridges Plaza (Ferry Building)	2nd at Townsend	3956	10	0.0067
Embarcadero at Folsom	San Francisco Caltrain (Townsend at 4th)	3849	12	0.0066
Steuart at Market	2nd at Townsend	3739	9	0.0064
2nd at South Park	Market at Sansome	3531	9	0.006
Temporary Transbay Terminal (Howard at Beale)	San Francisco Caltrain (Townsend at 4th)	3503	11	0.006



Start station name	End station name	Total trips	Duration (min)	Ratio
Harry Bridges Plaza (Ferry Building)	Embarcadero at Sansome	1550	34	0.0186
Embarcadero at Sansome	Harry Bridges Plaza (Ferry Building)	907	19	0.0109
Embarcadero at Sansome	Embarcadero at Sansome	873	74	0.0105
Harry Bridges Plaza (Ferry Building)	Harry Bridges Plaza (Ferry Building)	841	115	0.0101
Embarcadero at Bryant	Embarcadero at Sansome	483	37	0.0058
Embarcadero at Bryant	Harry Bridges Plaza (Ferry Building)	459	15	0.0055
Embarcadero at Vallejo	Embarcadero at Sansome	451	47	0.0054
University and Emerson	University and Emerson	448	166	0.0054
2nd at Townsend	Harry Bridges Plaza (Ferry Building)	415	18	0.005
Powell Street BART	Market at 10th	412	20	0.005

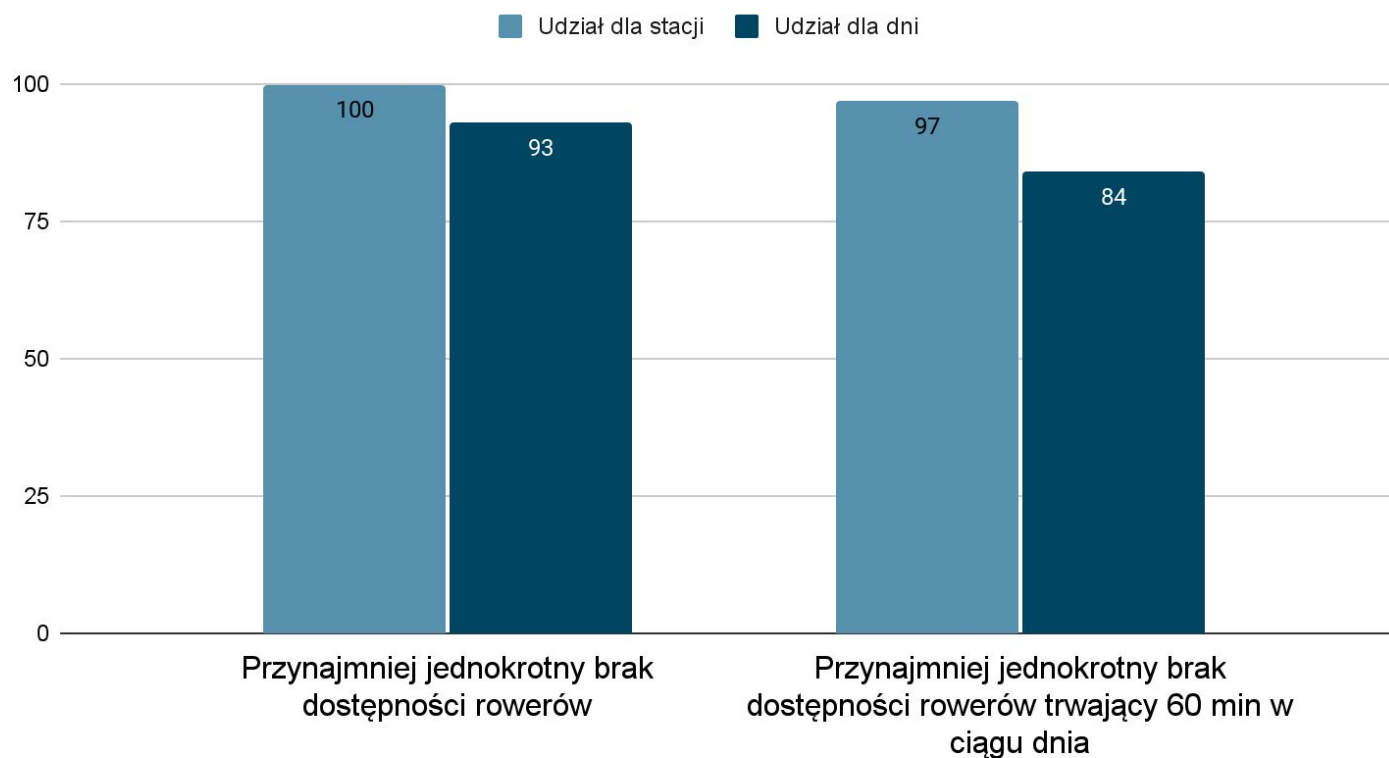
# Brak dostępnych rowerów - statystyki

## TOP 20 braku dostępności rowerów

Station ID	Date	Amount of time where no bikes are available / minutes
10	10-11-2014	1044
34	16-1-2015	964
32	28-5-2015	850
35	29-8-2014	707
76	10-9-2014	680
4	11-3-2015	651
49	15-1-2015	607
63	14-10-2014	598
32	31-3-2015	594
48	19-6-2014	594
64	29-7-2014	578
84	30-5-2015	570
82	7-10-2014	568
6	22-9-2013	566
6	31-8-2013	564
76	11-3-2015	562
64	11-6-2014	559
48	5-8-2015	559
48	4-8-2015	558
6	8-11-2014	555

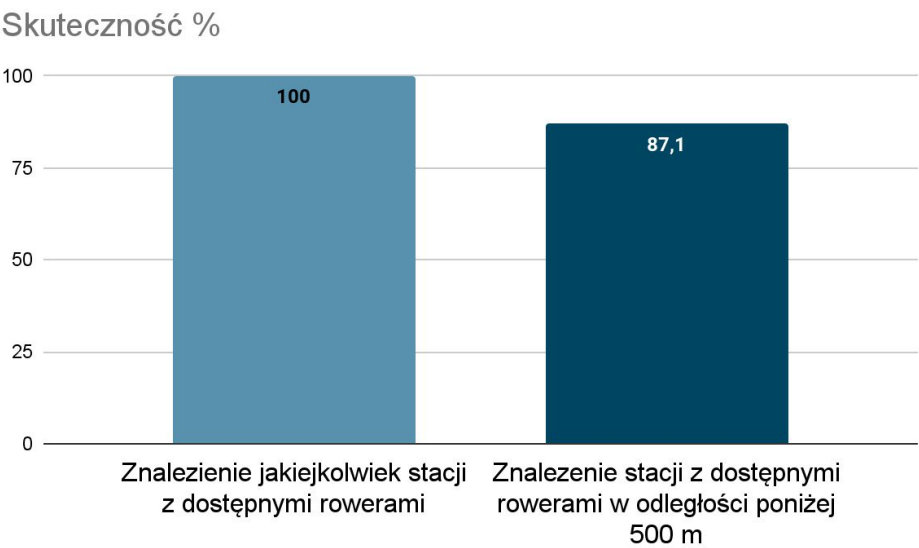
- Średnia długość braku dostępności rowerów - **61 minut**
- Długość braku dostępności rowerów/stację/dzień - **10,3 minuty**

Udział %



# Brak dostępnych rowerów - proponowane rozwiązanie

- Dzięki wartościom współrzędnych geograficznych wyliczone zostały odległości między stacjami
- Każda stacja w San Francisco ma sąsiadującą stację w odległości do 1000 m, a 94% stacji sąsiaduje z inną stacją w odległości do 500 m.



Date and Time	Station id with no bikes	Bikes available	Alternative station id	Bikes available	Distance between stations
2013-08-29 12:08:00	4	0	6	1	312.83
2013-08-29 12:08:00	4	0	5	10	323.79
2013-08-29 12:08:00	4	0	14	10	480.72
2013-08-29 12:08:00	7	0	11	9	258.51
2013-08-29 12:08:00	7	0	12	5	291.65
2013-08-29 12:08:00	7	0	3	8	389.49
2013-08-29 12:08:00	7	0	10	8	400.05
2013-08-29 12:08:00	7	0	8	8	415.96
2013-08-29 12:08:00	13	0	10	8	336.45
2013-08-29 12:08:00	13	0	6	1	465.4
2013-08-29 12:08:00	70	0	69	4	18.2

# Podsumowanie

- Najwięcej stacji rowerowych, a także mieszczących się w nich stacji dokujących, zlokalizowanych jest w San Francisco. Z tego miasta wyjechało także najwięcej rowerów.
- Rowery były chętniej wypożyczane w okresie wakacyjnym, chociaż ich intensywne użytkowanie zauważa się w ciągu całego badanego okresu.
- Zauważalna jest różnica w użytkowaniu rowerów pomiędzy subskrybentami a klientami - dotyczy to zarówno wybieranych przez nich tras, długości przejazdów jak i dni użytkowania. *Osoby posiadające subskrypcję* wykorzystują rowery regularnie, głównie w dni powszechne i są to krótkie przejazdy. W opisie bazy można znaleźć informację, że system oferuje 30 minut darmowego wypożyczenia dla subskrybenta. Użytkownicy sklasyfikowani jako *customer* wypożyczają rowery na dłuższy okres czasu i korzystają z nich częściej w weekend niż w dni robocze.
- Popularność tras zmienia się w zależności od tego czy są to dni powszechne czy weekendy.
- Przeprowadzona analiza korelacji wskazuje na brak związku pomiędzy ilością wypożyczeń rowerów oraz czasem trwania przejazdów z warunkami pogodowymi dla 10 najpopularniejszych tras, dotyczy to również rodzaju użytkowników oraz poszczególnych dni tygodnia.
- Istnieje związek pomiędzy średnią liczbą przejazdów a panującymi warunkami atmosferycznymi - szczególnie jeżeli chodzi o weekendy, gdzie zauważalny jest spadek liczby przejazdów w złych warunkach atmosferycznych.

Dziękujemy za uwagę!