

# Отчет об использовании прокатных велосипедов в Jersey за 2019 год.

## В целом:

Данные о поездках: <https://www.citibikenyc.com/system-data>

Географии районов и велодорожки: <https://data.jerseycitynj.gov/pages/home/>

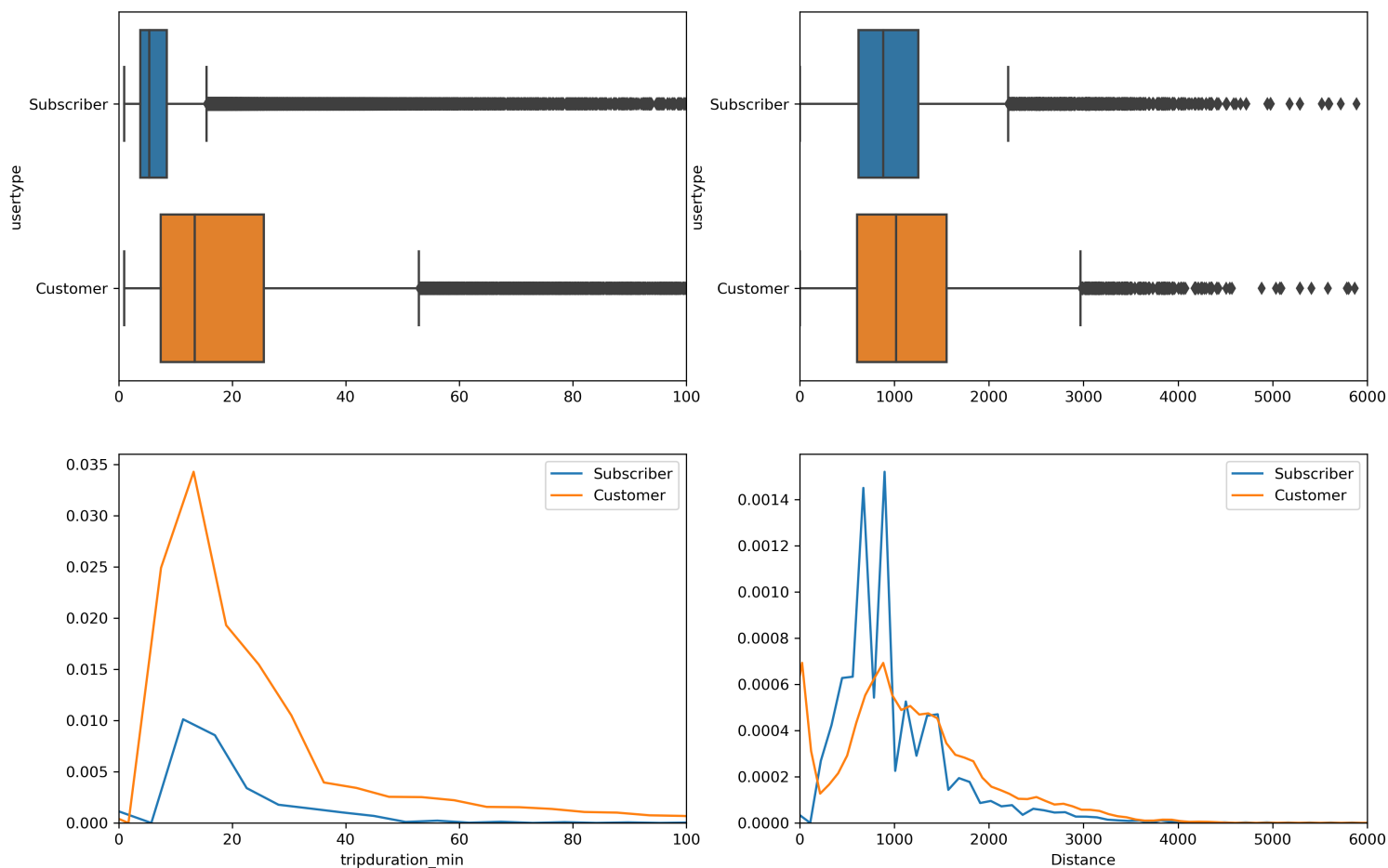
Особые обозначения:

Customer = 24-hour pass or 3-day pass user; Subscriber = Annual Member

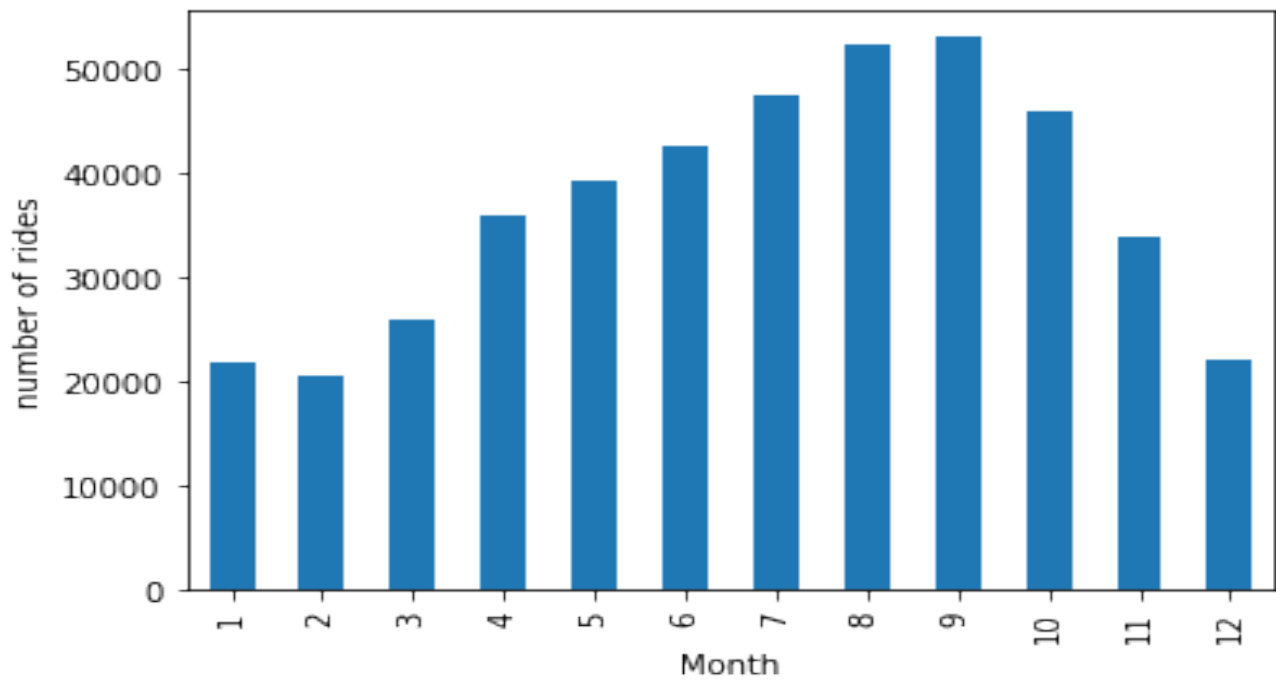
Gender (Zero=unknown; 1=male; 2=female)

Я произвел осмотр датасета по разным показателям на адекватность, представленных данных. Столбец “tripduration” вызвал наибольшие вопросы, потому что там были поездки длительностью более одного месяца. Для очистки данных хотел отсеять все данные больше, чем  $(75 \text{ квантиль} + 1.5 \cdot \text{интерквартильный размах})$ , но так как  $Q3 = 9.7$  минут, то получилось, что мы отседем много правдоподобных данных, поэтому удалим все поездки больше 12 часов = 720 минут.

В итоге мы имеем dataset состоящий из 441 000 поездок. Среднее время поездки равно 9,7 минут, а среднее расстояние 1 км.

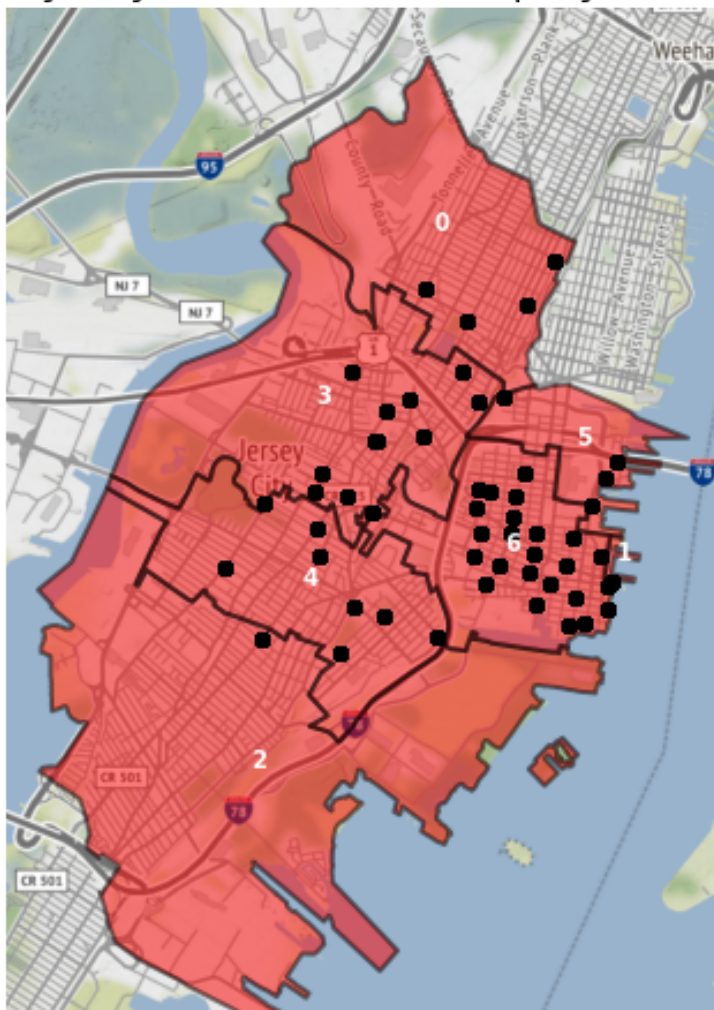


У нас прослеживается ярко выраженная сезонность. Декабрь – март спрос падает вдвое:



### Анализ по районам:

#### Jersey city bike stations map by districts

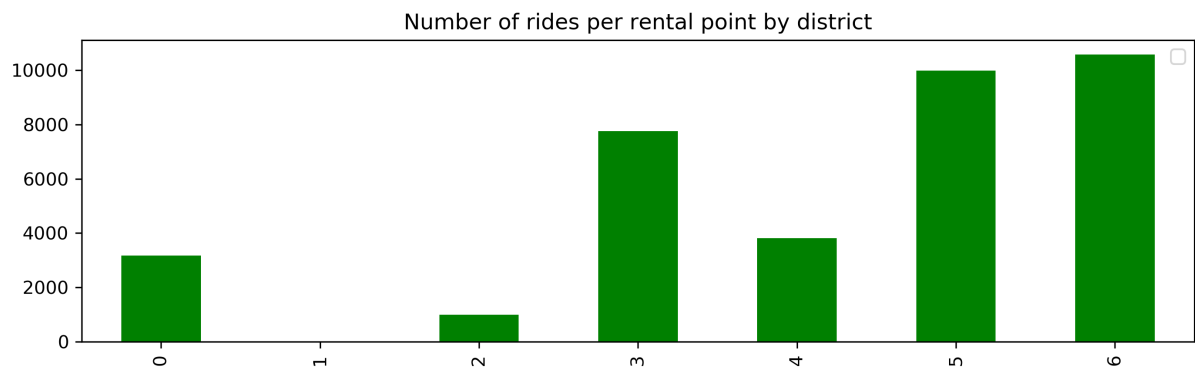
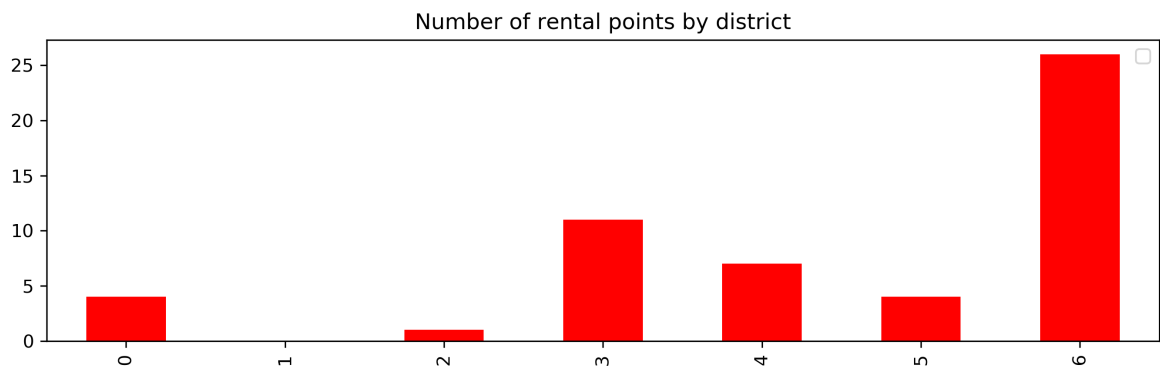
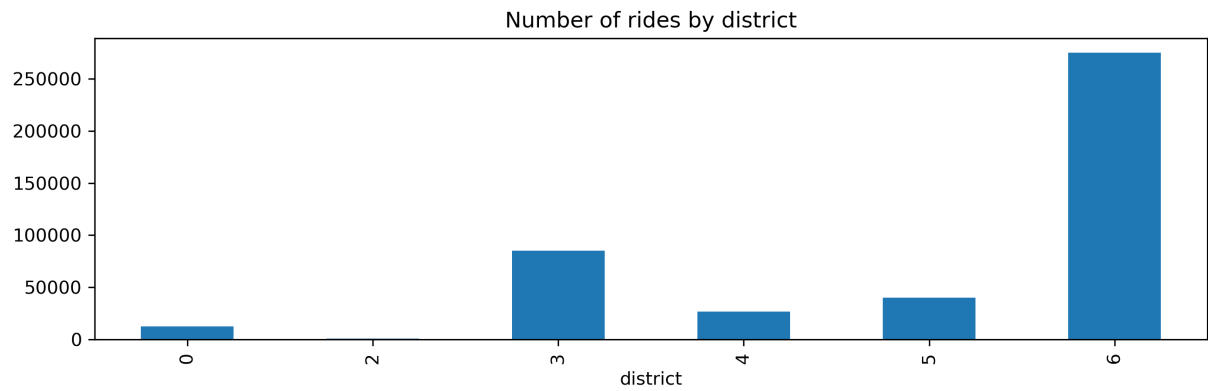


В городе у нас всего 7 районов. Точки аренды велосипедов находятся в 6 из 7 районов, но они расположены очень неоднородно. Мы видим основное их кол-во в районах №6, 3 и 4.

Проанализируем:

- 1) кол-во точек аренды
- 2) кол-во поездок
- 3) кол-во поездок на одну точку аренды

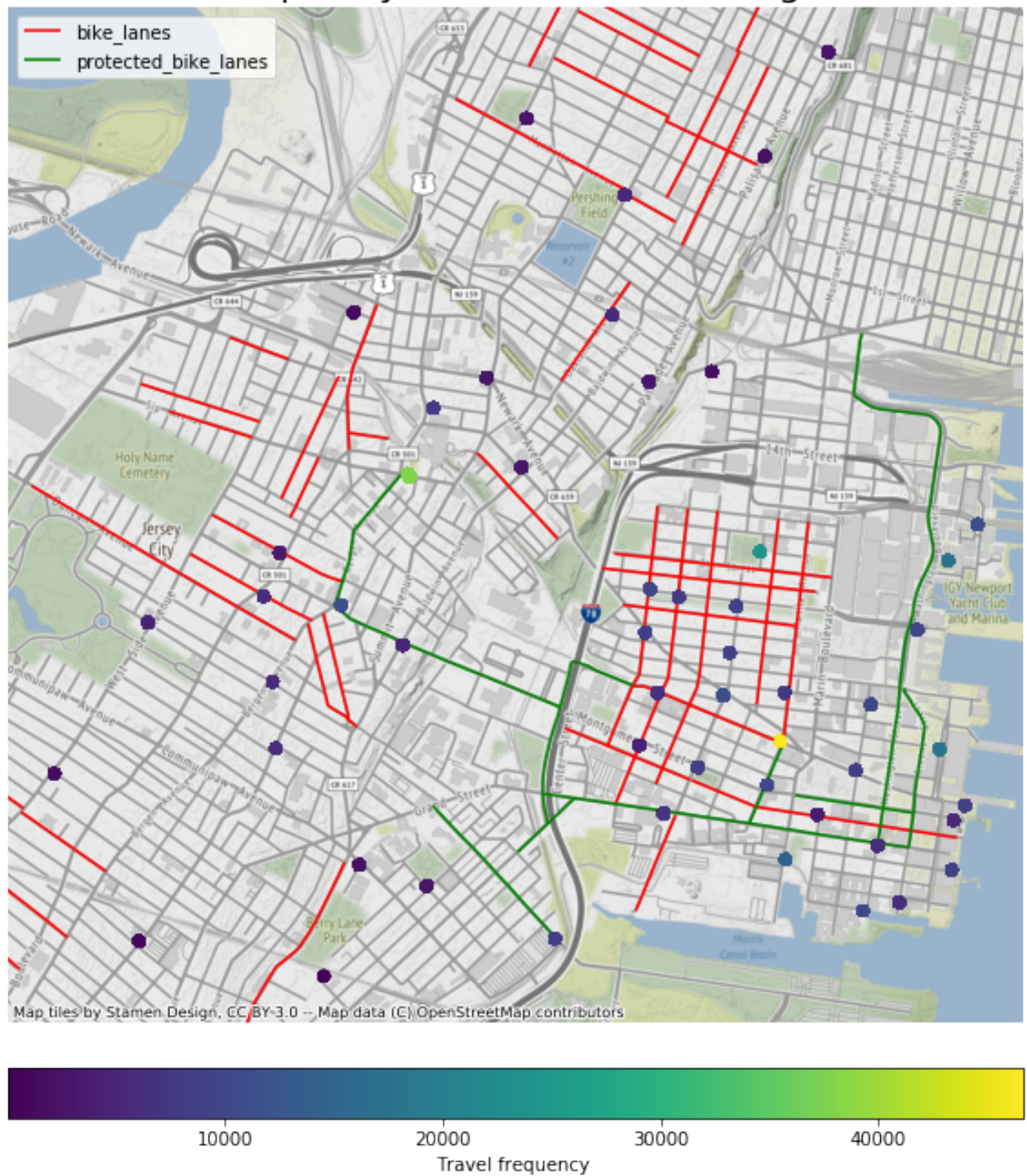
Смотреть график ниже



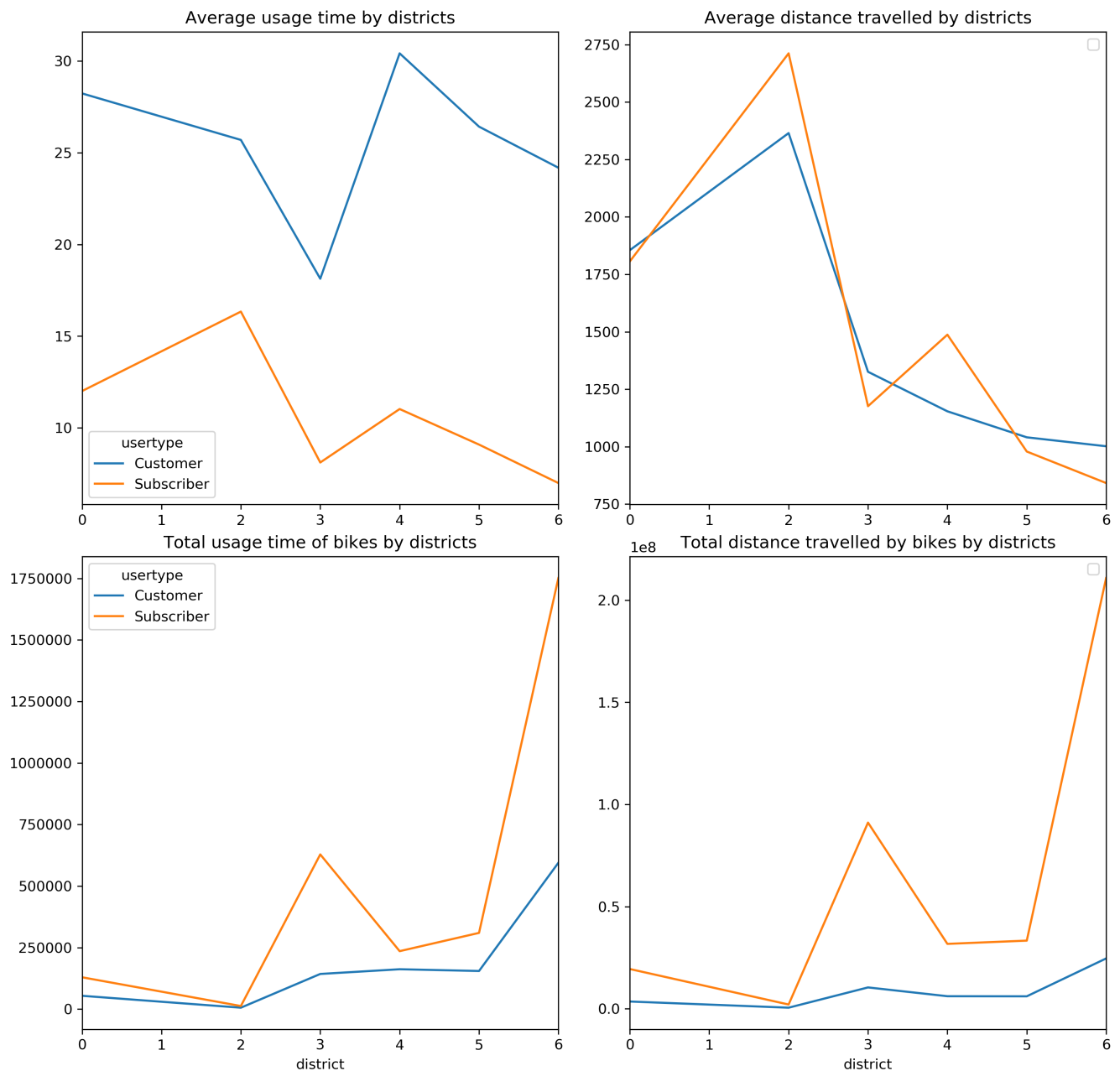
В целом мы видим, что кол-во аренд прямо пропорциональна кол-ву точек проката велосипедов. Единственное открытие здесь – чрезвычайно высокая эффективность точек проката района №5. Возможно имеет смысл увеличить плотность точек проката в этой области. И низкая эффективность в районе №0.

Теперь посмотрим на эффективность работы каждой точки проката на карте и их отношение по отношению к вело-инфраструктуре.

### Frequency of bike station's usage



Становится понятно, что высокая популярность аренды в районе 3, 5, 6, в том числе может быть связана с уровнем развития велосипедной инфраструктуры. Здесь наибольшая концентрация, как велосипедных дорожек, так и пунктов проката.



1. По верхним двум графикам мы видим сильную разницу по времени использования велосипедов между Покупателями и Подписчиками (верхний слева). Но при этом они проезжают одинаковое расстояние (верхний справа). Можно предположить, что:

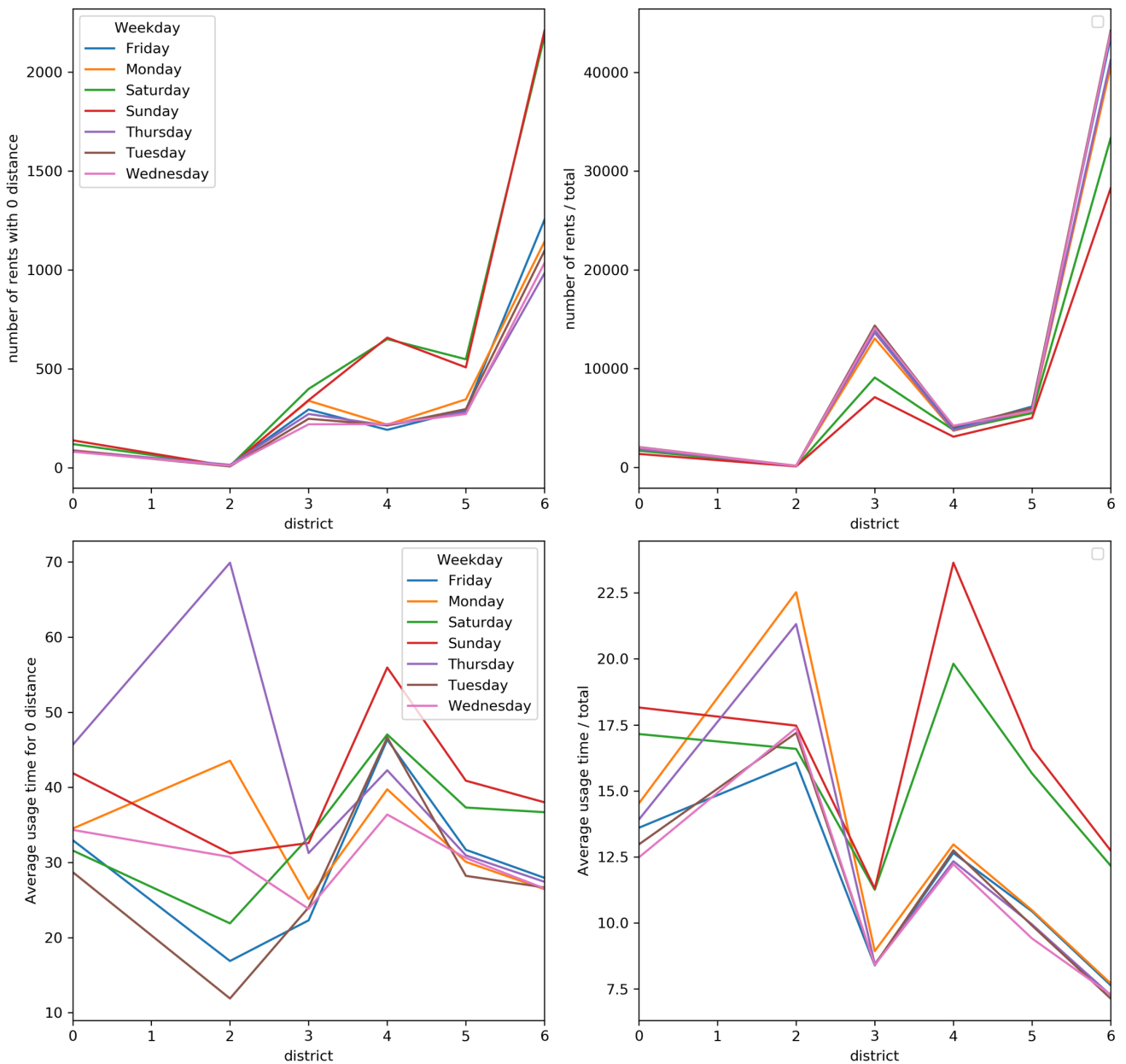
- а) Покупатели просто реже пользуются инфраструктурой и им нужно больше времени, чтобы разобраться и найти точку финиша.
- б) Покупатели используют велосипед для целей отдыха и неторопливой поездки,

2. Среднее время и расстояние поездки падает в районах с высокой плотностью точек аренды, но при этом растет их число (№3,4,5,6)

3. Основной прирост в районах с активным использованием велосипедов происходит за счет подписчиков.



Обращаем внимание на поездки с нулевым пробегом, то есть велосипед взяли и вернули на одну станцию. Можно предположить, что это поездки, совершенные с рекреационными целями.



Интересно, что данные поездки составляют всего 4% от общей массы по кол-ву, но по общему времени поездок они составляют 14%. При этом кол-во поездок и их среднее время сильно растет в субботу и воскресенье, по трём наиболее активным районам (4,5,6), кроме 3 ого. Но надо отметить, что среднее время поездки в выходные больше, чем в будни практически во всех районах (0, 3, 4, 5, 6), при этом в районе (1,4) самое длинное время поездки в понедельник и четверг.

**Портрет покупателя по районам**  
**По максимальному времени поездки и**  
**по максимальному кол-ву поездок**

max_duration				max_rides			
district	gender	birth year		district	gender	birth year	
0	1	1991	9302.4	0	1	1991	890
	2	1989	4602.7		2	1995	316
2	1	1988	3052.1	2	1	1988	194
	2	1991	1474.8		2	1991	82
3	1	1988	36546.5	3	1	1988	3674
	2	1990	13231.4		2	1990	1597
4	1	1988	19314.0	4	1	1988	1620
	2	1986	6964.5		2	1986	566
5	1	1988	17009.3	5	1	1988	1886
	2	1989	8073.9		2	1989	735
6	1	1989	87152.1	6	1	1989	11503
	2	1989	38661.7		2	1988	4284