Dviračių dalijimosi paslaugos skiriasi nuo kito viešojo transporto, kadangi išmanieji įrenginiai nuolatos renka duomenis apie kelionės laiką, starto ir pabaigos pozicijas ir pan. Šie įrankiai paverčia dviračius miesto sensoriais, kurie gali matuoti eismo intensyvumą. Taigi tikėtina, kad patys svarbiausi įvykiai mieste atsispindės ir duomenyse.

Šios dienos tikslas yra suprasti kaip aplinkos faktoriai, tokie kaip temperatūra, drėgmė ir pan. veikia dviračių nuomą.

1. Atlikite pradinę duomenų analizę. Naudokite day.csv duomenis.
2. Identifikuokite koreliacijas ir santykius tarp išnuomotų dviračių skaičiaus ir aplinkos faktorių.

3. Kaip temperatūra veikia dviračių nuomą? Sukurkite tiesinės regresijos modelį ir jį interpretuokite. Extra: Kaip drėgmė ir vėjo greitis keičia išnuomotų dviračių skaičių?

4. Įvertinkite savo modelį ir jo efektyvumą. Naudokite statistinius skaičiavimus.(

5. Interpretuokite savo gautą rezultatą.

Extra:

7. Kaip dienos laikas keičia dviračių nuomos rezultatus? Naudokite hour.csv duomenis.

Stulpelių paaiškinimai:

Both hour.csv and day.csv have the following fields, except hr which is not available in day.csv

- instant: record index

- dteday : date

- season : season (1:springer, 2:summer, 3:fall, 4:winter)

- yr : year (0: 2011, 1:2012)

- mnth : month ( 1 to 12)

- hr : hour (0 to 23)

- holiday : weather day is holiday or not

- weekday : day of the week

- workingday : if day is neither weekend nor holiday is 1, otherwise is 0.

+ weathersit :

- 1: Clear, Few clouds, Partly cloudy, Partly cloudy

- 2: Mist + Cloudy, Mist + Broken clouds, Mist + Few clouds, Mist

- 3: Light Snow, Light Rain + Thunderstorm + Scattered clouds, Light Rain + Scattered clouds

- 4: Heavy Rain + Ice Pallets + Thunderstorm + Mist, Snow + Fog

- temp : Normalized temperature in Celsius. The values are divided to 41 (max)

- atemp: Normalized feeling temperature in Celsius. The values are divided to 50 (max)

- hum: Normalized humidity. The values are divided to 100 (max)

- windspeed: Normalized wind speed. The values are divided to 67 (max)

- casual: count of casual users

- registered: count of registered users

- cnt: count of total rental bikes including both casual and registered