Adathalmaznak bicikli bérlésekhez kapcsolódó adatokat használtam. Minden mérés leírja az adott órában bérelt bicikliket, valamint az adott órához kapcsolódó tulajdonágokat, például időjárás és ünnepek. A mérések két év adataiból származnak ami 17414 bejegyzésnek felel meg. A mért adatok között nem szerepelnek hiányzóak. Az adatokat az alábbi forrásról töltöttem le: <https://www.kaggle.com/hmavrodiev/london-bike-sharing-dataset> .

Fontosabb oszlopok az adathalmazban:

timestamp: a mérés időbélyege

cnt: az adott pillanatban kibérelt biciklik száma

t1: valós hőmérséklet az adott időpontban, Celsiusban

t2: hőérzet az adott időpontban, Celsiusban

hum: páratartalom az adott időpontban százalékban megadva

wind\_speed: szél erőssége, km/h-ban

is\_holiday: munkaszüneti naphoz tartozik-e a mért adat, nullával és egyel jelölve

is\_weekend: hétévéhez tartozik-e a mért adat, nullával és egyel jelölve

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | cnt | t1 | t2 | hum | wind\_speed |
| count | 17414 | 17414 | 17414 | 17414 | 17414 |
| mean | 1143.101642 | 12.468091 | 11.520836 | 72.324954 | 15.913063 |
| std | 1085.108068 | 5.571818 | 6.615145 | 14.313186 | 7.894570 |
| min | 0 | -1.5 | -6 | 20.5 | 0 |
| 25% | 257 | 8 | 6 | 63 | 10 |
| 50% | 844 | 12 | 12 | 74.5 | 15 |
| 75% | 1671.75 | 16 | 16 | 83 | 20 |
| max | 7860 | 34 | 34 | 100 | 56 |

season: évszak, nullától háromig számmal jelölve

A főbb változókat leíró alapvető statisztikai értékek:

Chart, histogram

Description automatically generated

A korrelációt ábrázoló diagrammon látszik, hogy erős összefüggés van a valós és érzékel hőmérséklet között. Ezen felül, ebből kiindulva ennek a kettőnek hatása lehet a bicikli kölcsönzésekre is bizonyos fokig. A páratartalom és az évszak között is jelez a diagramm gyenge összefüggést. Emellett a páratartalom negatív összefüggésben van a diagramm alapján a darabszámmal és mindkét hőmérséklettel.

kérdések:

Hogyan befolyásolják a a bérlések számát a környzeti változók?

Hogyan oszlik el hétvégére és hétközbenre a bérlések száma?

A bérlés órája hogyan befolyásolja a bérlések számát összesen és havonta?

A havi eloszlások mennyire igazodnak az időszaknak megfelelő napkeltének és napnyugtának?

Hogyan oszlik el a bérlések száma hétvégére és hétközbenre, ha a érzékelt hőmérséklet magas, közepes és alacsony.

Többen bérelnek e biciklit munkaszüneti napon, jó időben mint rossz időben?

Egyes eseményekhez (például: választáshoz) kötődő napokon az értékeke hogyan viszonyúlnak a hasonló napokhoz tartozó értékekehez?

Rossz időhöz( magasabb páratartalom, szél, alacsony hőmérséklet) kapcsolódód értékek mennyiben térek el az átlagos és jó időhöz kapcsolódó értékektől.