1. **Budowa modelu regresji liniowej**
   1. **Zadanie**

Zadanie polegało na budowie kilku modeli regresji liniowej dla zadanych zestawów danych (wycinek danych przedstawiono poniżej). Ponadto polecono spróbowanie budowy modelu nieliniowego i porównanie go z modelem liniowym.

Obraz zawierający tekst, biały

Opis wygenerowany automatycznie

Rys. 1. Fragment danych na podstawie których stworzony został model regresji.

* 1. **Sposób rozwiązania**

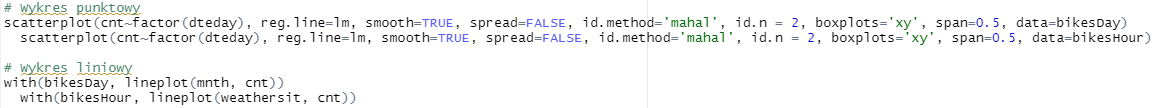
Dla każdego z zadanych zestawów zbudowano po dwa modele liniowe oraz jeden nieliniowy. Model liniowy nr 2 różnił się od modelu nr 1 tym, że budowany był z wykorzystaniem tylko istotnych zmiennych.

* 1. **Implementacja w środowisku R**

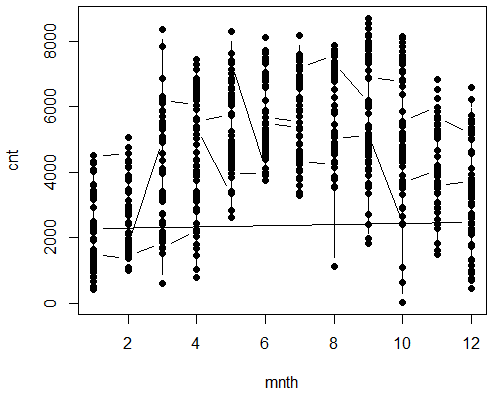
Zadanie rozpoczęto od wczytania danych poddanych analizie.

****

Po załadowaniu postawiono kilka prostych hipotez, które sprawdzono rysując wykresy. Dzięki temu można było lepiej poznać zbiór danych i formułować inne hipotezy.

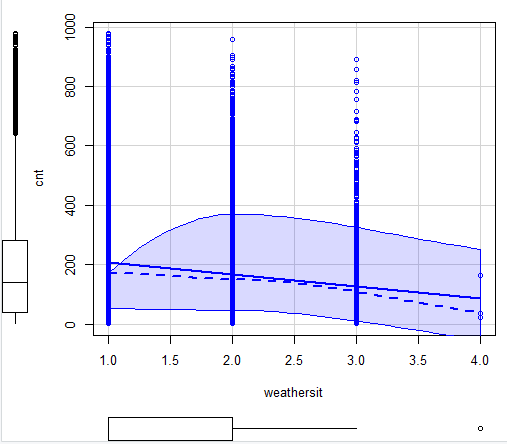


1. Prosta hipoteza nr 1 – w miesiącach zimowych spada liczba wypożyczanych rowerów.



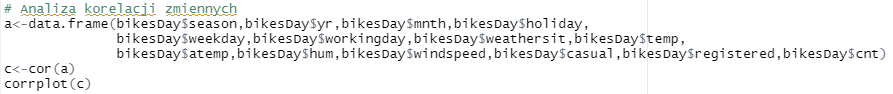
Rys. 2. Zależność miesięczna liczby wypożyczeni rowerów w pierwszym zbiorze danych.

1. Prosta hipoteza nr 2 – trudne warunki atmosferyczne redukują liczbę wypożyczeni w sposób drastyczny.

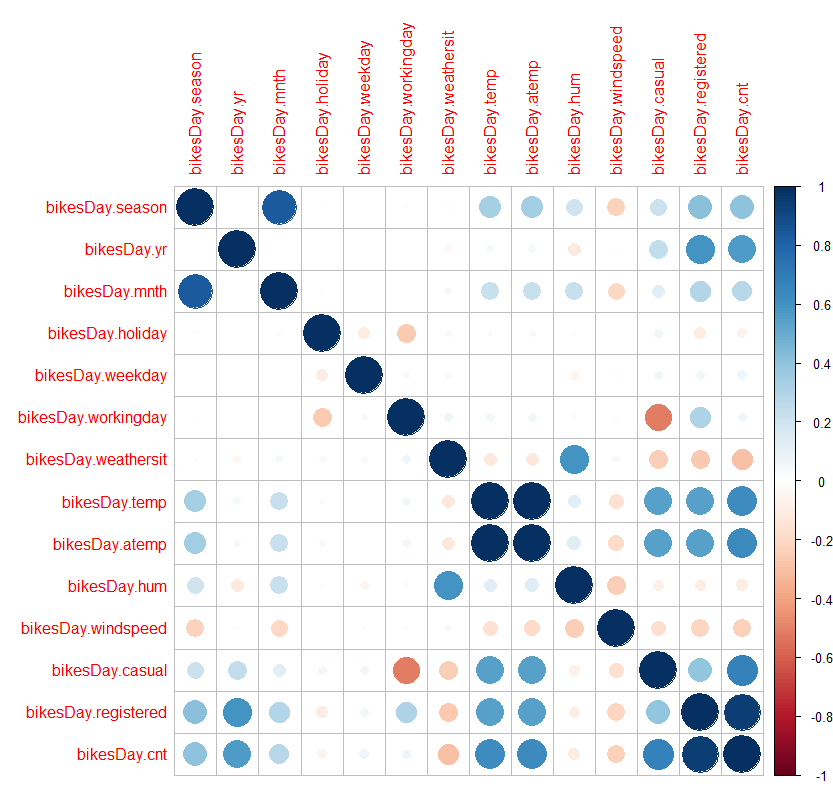


Rys. 3. Zależność wypożyczeni od pogody (1 – czysta pogoda, 2 – pochmurnie, 3 – lekkie opady i pochmurnie, 4 – burze, grad, zachmurzone)

Zbadano również korelacje pomiędzy zmiennymi występującymi w zbiorze danych.



Na podstawie powyższego kodu otrzymano następujący wykres korelacji.



Rys. 4. Korelacja pomiędzy zmiennymi zbioru danych.

1. **Zbudowane modele**

Dla każdego ze zbioru zbudowano modele regresji liniowej.

Obraz zawierający tekst

Opis wygenerowany automatycznie

Fragment kodu przedstawiony powyżej tworzy dla każdego zestawu danych dwa modele liniowe.

1. Model nr 1 dla zbioru danych nr 1

Obraz zawierający tekst, paragon

Opis wygenerowany automatycznie

1. Model nr 2 dla zbioru danych nr 1

Obraz zawierający tekst, paragon

Opis wygenerowany automatycznie

1. Model nr 1 dla zbioru danych nr 2

Obraz zawierający tekst, paragon

Opis wygenerowany automatycznie

1. Model nr 2 dla zbioru danych nr 2

Obraz zawierający tekst, paragon

Opis wygenerowany automatycznie

Po stworzeniu modeli liniowych zbudowano próbne modele nieliniowe. W tym celu należało najpierw przeprowadzić zlogarytmizowanie zmiennych.

Obraz zawierający tekst

Opis wygenerowany automatycznie

1. Model nieliniowy dla zbioru danych nr 1

Obraz zawierający tekst, paragon

Opis wygenerowany automatycznie

1. Model nieliniowy dla zbioru danych nr 2

Obraz zawierający tekst, paragon

Opis wygenerowany automatycznie

1. **Wnioski**

Budowanie modeli regresji w środowisku R jest bardzo przejrzyste i proste   
w obsłudze. Analizując otrzymane w trakcie pracy wyniki możemy stwierdzić kilka rzeczy. Po pierwsze, dla obu zestawu danych redukcja zmiennych nieistotnych w modelu nie wpłynęła negatywnie na jego jakość. Potwierdza to wskaźnik R2. Po drugie, w obu przypadkach modele nieliniowe okazały się znacząco gorsze od modeli liniowych. Niemniej jednak warto było przeprowadzić ich pobieżną analizę.