

Przedmiot: Eksploracja danych

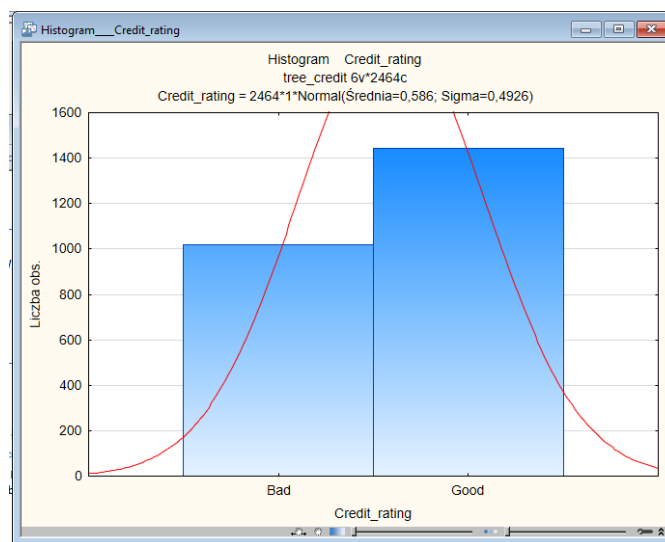
Kierunek: Informatyka – Data Science

Ćwiczenie: Analiza ryzyka kredytowego

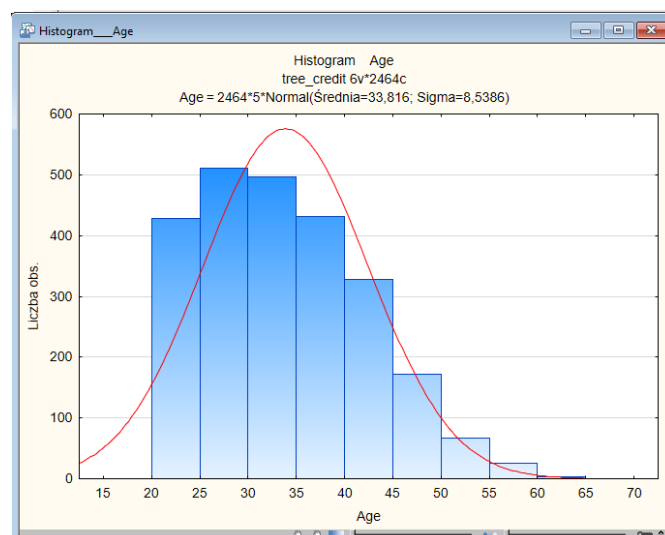
Autor: Bartłomiej Jamiołkowski

Ad 3)

Wnioski:

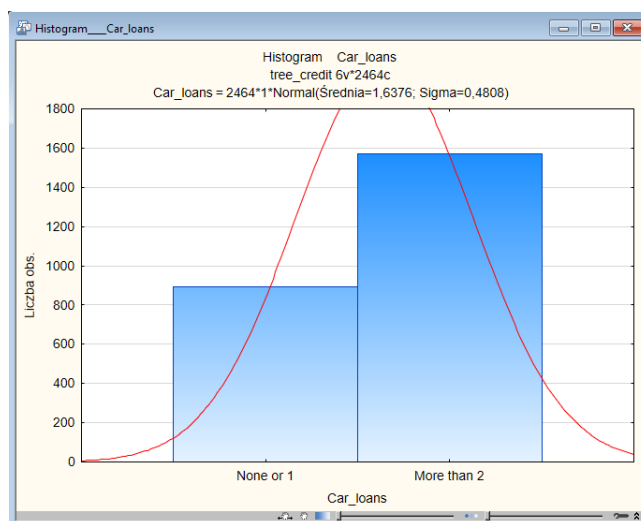


W zmiennej Credit\_rating liczniejszą grupę stanowią osoby posiadające zdolność kredytową (ponad 1400 rekordów). Mniej liczna grupę stanowią osoby nie posiadające zdolności kredytowej (ponad 1000 rekordów).

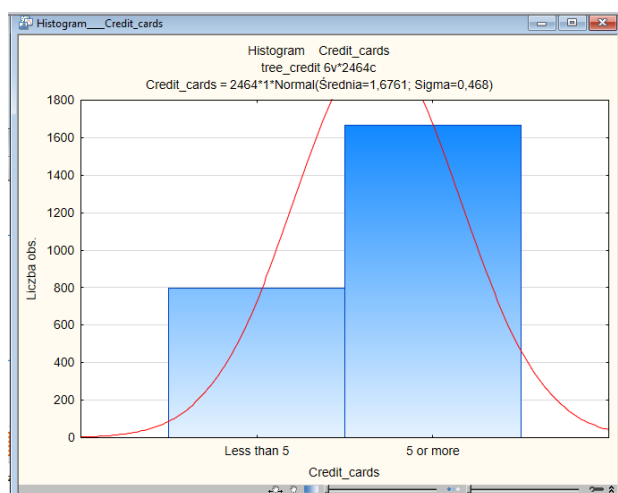


Histogram dla zmiennej Age ukazuje rozkład pozytywnie skośny. Liczniejsze grupy stanowią osoby ubiegające się o kredyt w młodszym wieku. Średni wiek osób wynosi 33,816 lat. Wraz

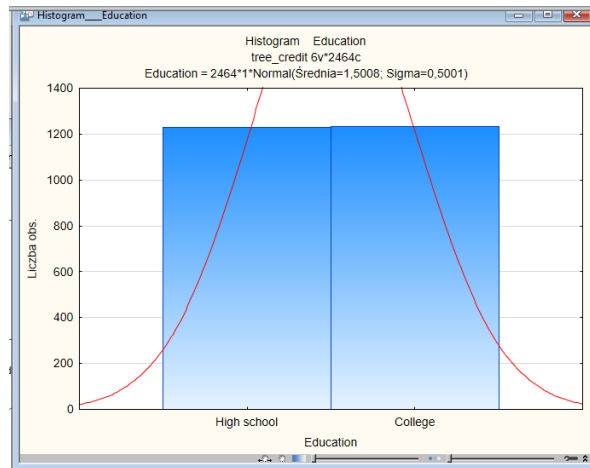
z wiekiem liczba osób ubiegających się o kredyt maleje. Osoby w młodszym wieku są bardziej skłonne do ubiegania się o kredyt.



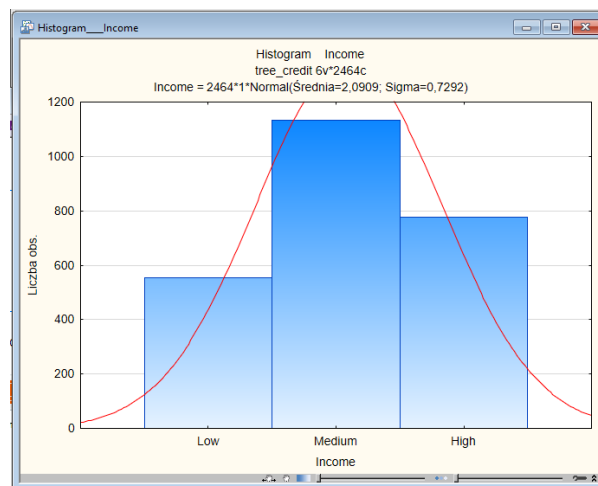
W zmiennej Car\_loans najliczniejszą grupę stanowią osoby, które zaciągnęły więcej niż 2 kredyty samochodowe (około 1600 rekordów). Mniej liczną grupę stanowią osoby, które nie zaciągnęły kredytu samochodowego lub zrobiły to tylko raz (około 900 rekordów).



W zmiennej Credit\_cards najliczniejszą grupę stanowią osoby, które posiadają 5 lub więcej kart kredytowych (ponad 1600 rekordów). Z tą grupą może być związane większe ryzyko nadmiernego zadłużenia. Mniej liczną grupę stanowią osoby, które posiadają mniej niż 5 kart kredytowych.



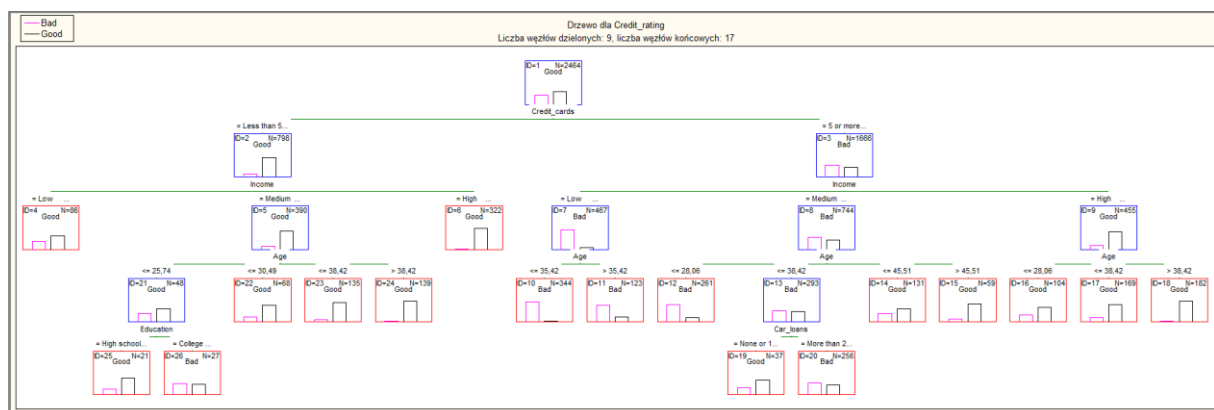
W zmiennej Education grupy osób, które ukończyły High school i College są do siebie zbliżone pod względem liczności (ponad 1200 rekordów w grupie).



W zmiennej Income najliczniejszą grupę stanowią osoby ze średnim dochodem (około 1100 rekordów). Drugą najliczniejszą grupą są osoby z wysokimi dochodami (około 800 rekordów). Najmniej liczną grupę stanowią osoby z niskimi dochodami (około 600 rekordów).

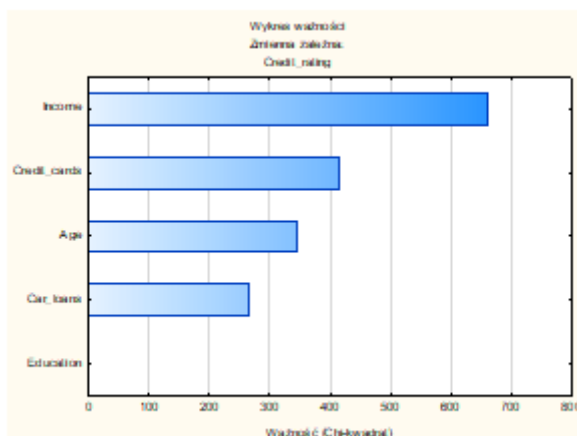
Ad 10)

### Analiza drzewa:

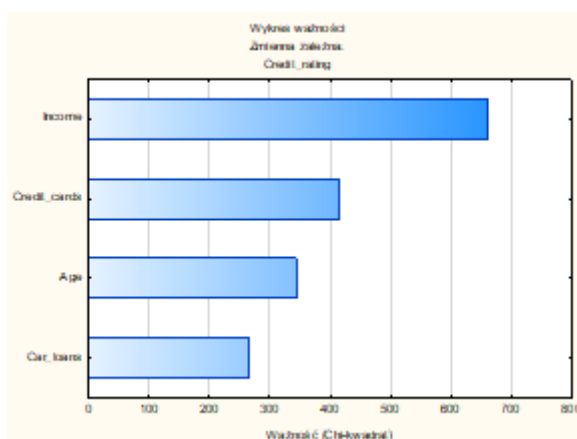


Preferowane przy przyznawaniu kredytu są osoby, które mają mniej niż 5 kart kredytowych. W takim przypadku jedynym dyskwalifikującym czynnikiem byłaby ukończona edukacja w College (w tym punkcie nie nastąpiła jeszcze eliminacja zmiennej Education). Osoba, która posiada więcej niż 5 kart kredytowych, aby otrzymać kredyt, musiałaby mieć wysoki dochód. Osoba ze średnim dochodem również miałaby szansę na otrzymanie kredytu, ale musiałaby być starsza niż 38, 42 lat oraz nie posiadać lub mieć tylko 1 kredyt samochodowy.

Wykres ważności.



Wykres ważności po wyeliminowaniu zmiennej Education.



Najważniejszą zmienną dla zmiennej predykcyjnej jest zmienna Income.