

1 Praca domowa nr.10 - Olaf Skrabacz

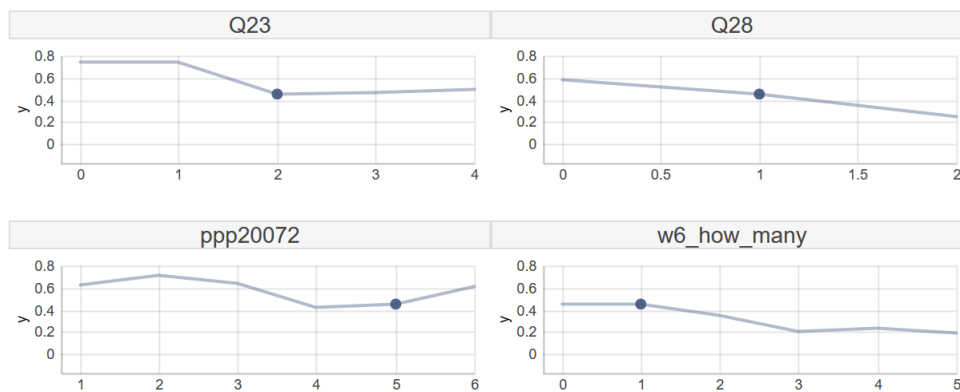
2 Dane

Dane dotyczą osób dorosłych, którzy w ankietach odpowiadali na pytania dotyczące swojego życia oraz swoich relacji interpersonalnych. Dalej zostanie zbudowany model, który przewiduje czy dana osoba jest zameżna. Do przewidywania tej cechy wybrany został model Lasu Losowego.

2.1 Obserwacja

By lepiej zrozumieć działanie modelu, wybierzemy losową próbkę ze zbioru, czyli jakąś losową osobę która wypełniła tę ankietę. By lepiej utożsamić się z tą obserwacją nazwiemy ją Janek. Z punktu widzenia modelu Janek: - zarabia tyle samo co swój partner (Q23), - znał rodziców partnera, przed jego poznaniem (Q28), - raz na rok lub rzadziej bierze udział w obrzędach religijnych (ppp20072), - przed poznaniem swojego obecnego partnera spotykał się tylko z jedną osobą (w6_how_many). Model ocenił szansę Janka na bycie osobą zameżną na około 45%.

Następujący raport pozwoli Jankowi odpowiedzieć na pytanie, co może zmienić by miał większą szansę na bycie zameżnym, a nam pozwoli zgłębić i wyjaśnić odpowiedzi modelu.

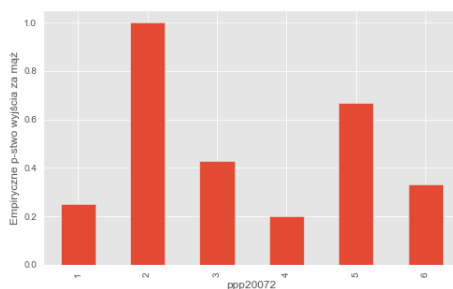


Powyżej zaprezentowane są wykresy ceteris paribus dla każdej zmiennej wchodzącej w skład modelu. Czy możemy z nich wywnioskować co Janek może zrobić by zwiększyć swoją szansę na bycie zameżnym? Największa wariancja co do zmiennej celu występuje w zmiennych ppp20072 oraz Q23. Kropką oznaczamy odpowiedzi Janka, które są wyświetlone wyżej. Jeśli Janek chciałby zmaksymalizować odpowiedź modelu dotyczącą jego osoby musiałby:

- zmienić wartość zmiennej ppp20072 na wartość 2, która odpowiada braniu udziału w obrzędach religijnych około raz w tygodniu,
- zmienić wartość zmiennej Q23 na wartość 0 lub 1 odpowiadającym braku odpowiedzi na pytanie lub zarabianiu więcej niż partner.

2.2 Podobne obserwacje

Oprócz patrzenia tylko na Janka warto zobaczyć obserwację ludzi podobnych do Janka. Sprawdźmy jak wyglądają średnie odpowiedzi zmiennej celu w zależności od wartości zmiennej ppp20072. Wybierzmy tylko te obserwacje, które na pozostałych zmiennych mają te same wartości co Janek i sprawdźmy jak wygląda ich empiryczne p-stwo bycia osobą zameżną.



Jak widać odpowiedzi te różnią się od wykresu Ceteris Paribus dla Janka co oznacza, że nasz model bierze pod uwagę interakcję między zmiennymi.