

REGRESJA LINIOWA 2 - ZADANIA.

ZAJĘCIA NR 6.

Wszystkie ćwiczenia pochodzą z podręczników Wooldridge oraz Stock & Watson.

1. Poniższa tabela zawiera informacje o połowie sardeli (w milionach ton) oraz średniej cenie ryb (w \$ za tonę) w latach 1965-1978.

	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978
Cena (y)	190	160	134	129	172	197	167	239	542	372	245	376	454	410
Połów (x)	7.23	8.53	9.82	10.26	8.96	12.27	10.28	4.45	1.78	4.0	3.3	4.3	0.8	0.5

Wprowadź dane do R. Korzystając z wcześniej poznanych formuł, znajdź oszacowanie regresji ceny w zależności od połowu. Oblicz i zinterpretuj R^2 .

2. Dane zebrano z losowej próby 220 sprzedanych domów z pewnej miejscowości w 2013 roku. Niech *Price* oznacza cenę sprzedaży (w \$1000), *BDR* oznacza liczbę sypialni, *Bath* oznacza liczbę łazienek, *Hsize* oznacza rozmiar domu (w mkw.), *Lsize* oznacza rozmiar działki (w mkw.), *Age* oznacza wiek domu w latach, *Poor* oznacza zmienną zerojedynekową, która przyjmuje wartość 1 jeżeli stan domu jest słaby. Oszacowano regresję objaśniającą cenę domu:

$$Price = 119.2 + 0.485BDR + 23.4Bath + 0.156Hsize + 0.002Lsize + 0.090Age - 48.8Poor$$

$$\bar{R}^2 = 0.72, SER = 41.5.$$

- (i) Przyjmijmy że właściciel przerobi część pokoju dziennego na jedną łazienkę. Jaki jest oczekiwany wzrost wartości domu?
 - (ii) Przypuśćmy że właściciel doda nową łazienkę do swojego domu, co zwiększy jego rozmiar o 100 mkw. Jaki jest oczekiwany wzrost wartości domu?
 - (iii) Jaki będzie spadek wartości domu, jeżeli właściciel zaniedba go, tak że jego stan będzie słaby?
 - (iv) Oblicz R^2 i zinterpretuj go.
3. Używając danych **CHARITY** (opis danych) odpowiedz na poniższe pytania:

- (i) Oszacuj poniższe równanie:

$$gift = \beta_0 + \beta_1 mailsyear + \beta_2 giftlast + \beta_3 propresp + u$$

metodą Najmniejszych Kwadratów i zapisz wynik w postaci równania, podaj również rozmiar próby oraz R^2 . Porównaj R^2 z oszacowanej regresji oraz z prostej regresji, która pomija zmienne *giftlast* oraz *propresp*?

- (ii) Zinterpretuj współczynnik przy *mailsyear*. Czy jest on większy czy mniejszy niż jego odpowiednik z prostej regresji?
- (iii) Zinterpretuj współczynnik przy *propresp* (sprawdź w jakich jednostkach jest wyrażona ta zmienna)
- (iv) Teraz dodaj zmienną *avggift* do równania. Co stało się z oszacowanym efektem zmiennej *mailsyear*?
- (v) W równaniu z podpunktu (iv), co stało się ze współczynnikiem przy zmiennej *giftlast*? Co może to powodować?

4. (W, C2) Skorzystaj z danych **HPRICE1** (opis danych) do oszacowania modelu:

$$price = \beta_0 + \beta_1 sqft + \beta_2 bdrms + u$$

gdzie *price* to cena domu w tys, dolarów.

- (i) Zapisz wynik w formie równania.
 - (ii) Jaki jest szacowany przyrost ceny domu z dodatkową sypialnią, przy niezmienionej powierzchni?
 - (iii) Jaki jest szacowany przyrost ceny domu z dodatkową sypialnią i o powierzchni 140 metrów kw.? Porównaj odpowiedź z odpowiedzią z punktu (ii).
 - (iv) Jaki procent zmienności w cenie jest objaśniany powierzchnią (w metrach kw.) oraz liczbą sypialni?
 - (v) Pierwszy dom w próbie ma $sqft = 2,438$ oraz $bdrms = 4$. Znajdź przewidywaną cenę dla tego domu, implikowaną przez MNK.
 - (vi) Prawdziwa cena pierwszego domu w próbie wynosi \$300,000 (więc $price = 300$). Znajdź resztę dla tego domu. Czy ktoś zapłacił za dużo, czy za mało?
5. Skorzystaj z danych **ATTEND** (opis danych).

- (i) Podaj minimum, maksimum oraz średnie wartości dla zmiennych *atndrte*, *priGPA*, oraz *ACT*.
- (ii) Oszacuj model:

$$atndrte = \beta_0 + \beta_1 priGPA + \beta_2 ACT + u$$

i zapisz wynik w formie równania. Zinterpretuj oszacowanie wyrazu wolnego. Czy ma sensowne znaczenie?

- (iii) Omów współczynniki. Czy zauważasz niespodzianki?
- (iv) Jaka jest przewidywana wartość *atndrte* jeżeli $priGPA = 3.65$ oraz $ACT = 20$? Co zrobić z tym wynikiem? Czy są studenci w próbie, z takimi wartościami zmiennych objaśniających?
- (v) Jeśli student A ma $priGPA = 3.1$ i $ACT = 21$ oraz student B ma $priGPA = 2.1$ i $ACT = 26$, to jaka jest przewidywana różnica w ich frekwencji?