

# Wpływ Stylu Życia na Jakość Snu

Marcin Górski, Mykhailo Bondar

2025-11-23

## Spis treści

Opis danych	2
Test nr 1: Płeć a długość snu	4
Test nr 2: Czy BMI wpływa na poziom stresu?	5
Część 3: Analiza korelacji i regresji	7
Model regresji liniowej	8

## Spis tablic

1	Statystyki Opisowe . . . . .	2
2	Istotność różnic (p adj) dla par grup . . . . .	5

## Spis rysunków

1	Wizualizacje analizowanych zmiennych . . . . .	3
2	Długość snu w zależności od płci . . . . .	4
3	Poziom stresu w zależności od BMI . . . . .	5

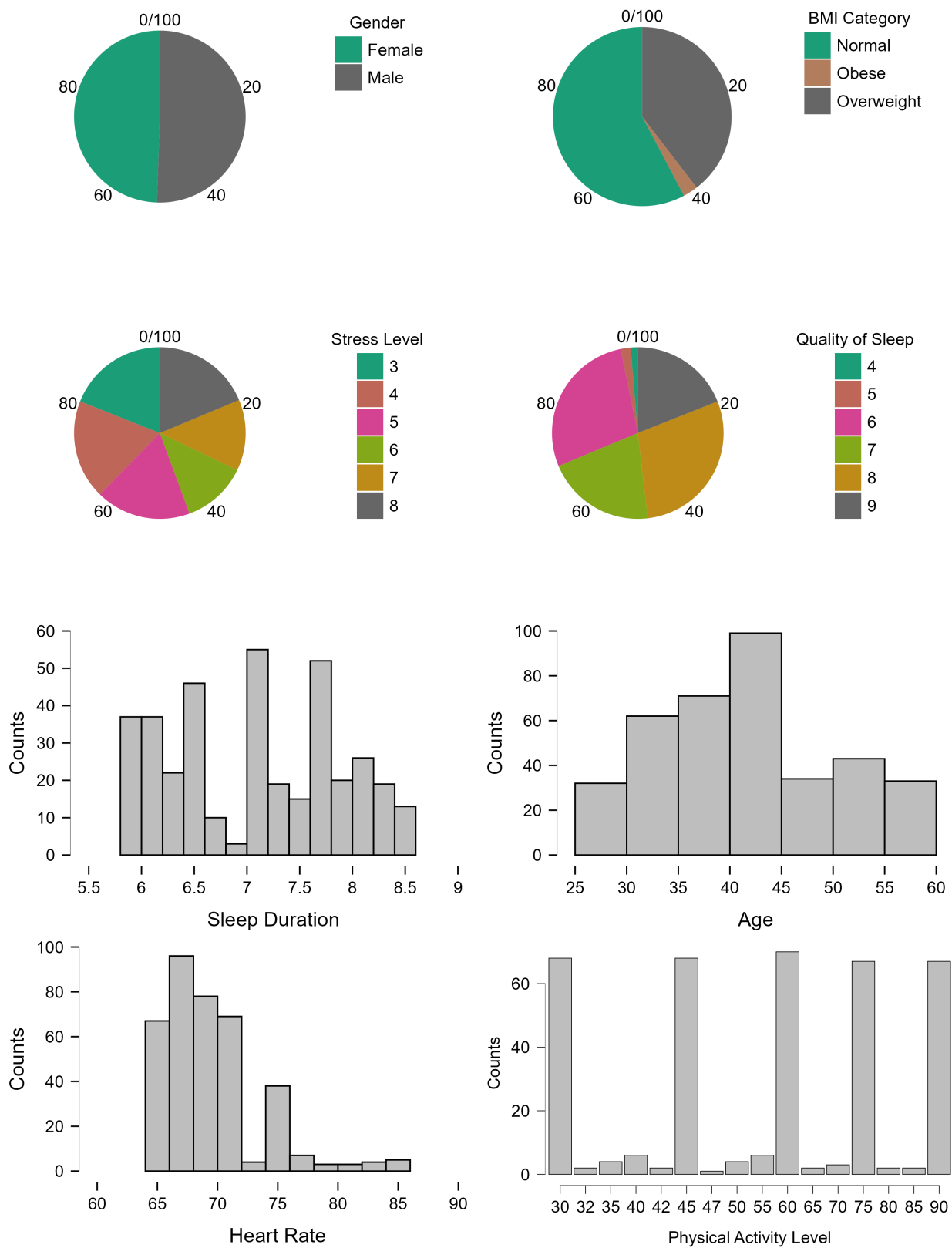
## Opis danych

Projekt przygotowaliśmy w oparciu o dane “Sleep Health and Lifestyle Dataset”, zawierające informacje o blisko 400 osobach, m. in. ich zdrowiu fizycznym i jakości snu. Dane przed przejściem do obliczeń odpowiednio wyczyściliśmy i przygotowaliśmy. W projekcie badaliśmy związki między zmiennymi: Płeć (zmienna nominalna), Kategoria BMI, Jakość snu i Poziom stresu (zmienne porządkowe, przy czym na potrzeby analizy i przeprowadzenia testów dwie ostatnie traktujemy jako ilościowe) oraz Długość snu, Poziom aktywności fizycznej (czas aktywności fizycznej dziennie w minutach), Wiek i Tętno (zmienne ilościowe).

Poniższa tabela przedstawia statystyki opisowe dla badanych zmiennych ilościowych:

Tablica 1: Statystyki Opisowe						
	Sleep Duration	Stress Level	Quality of Sleep	Physical Activity Level	Age	Heart Rate
Mean	7.132	5.385	7.313	59.17	42.18	70.17
Std. Deviation	0.796	1.775	1.197	20.83	8.673	4.136
Minimum	5.800	3.000	4.000	30.00	27.00	65.00
Maximum	8.500	8.000	9.000	90.00	59.00	86.00

Z kolei na następnej stronie znaleźć można wizualne przedstawienie ilości danych nominalnych i porządkowych oraz rozkładów danych ilościowych:



Rysunek 1: Wizualizacje analizowanych zmiennych

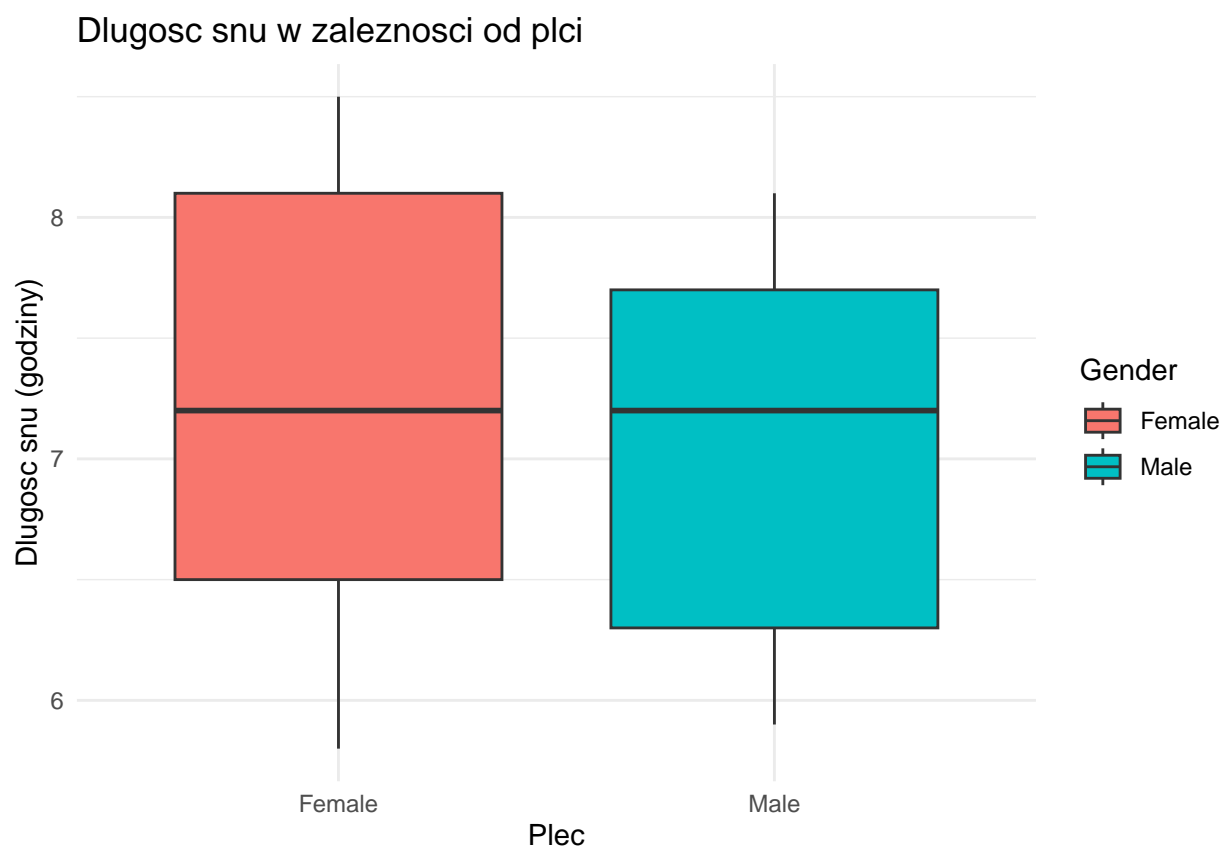
## Test nr 1: Płeć a długość snu

Przeprowadzamy test t-Studenta dla prób niezależnych, sprawdzający równość wartości oczekiwanych. R korzysta w tym przypadku z testu t-Welcha, który nie zakłada równości wariancji w obu populacjach.

Hipoteza zerowa: średnia długość snu jest taka sama u kobiet i mężczyzn.

Hipoteza alternatywna: średnia długość snu kobiet i mężczyzn różni się istotnie.

```
test1 <- t.test(Sleep.Duration ~ Gender, data=dane)
```



Rysunek 2: Długość snu w zależności od płci

Na podstawie przeprowadzonego testu t-Studenta uzyskano wartość **wartość  $p = 0.019$** . Oznacza to, że różnica jest istotna statystycznie (odrzucaamy hipotezę zerową i przyjmujemy alternatywną). Mimo, że na wykresach pudełkowych różnica zdaje się być niewielka, przeprowadzony test potwierdza jej istotność.

## Test nr 2: Czy BMI wpływa na poziom stresu?

Przeprowadzamy jednoczynnikową analizę wariancji ANOVA w celu sprawdzenia, czy masa ciała ma wpływ na odczuwany stres.

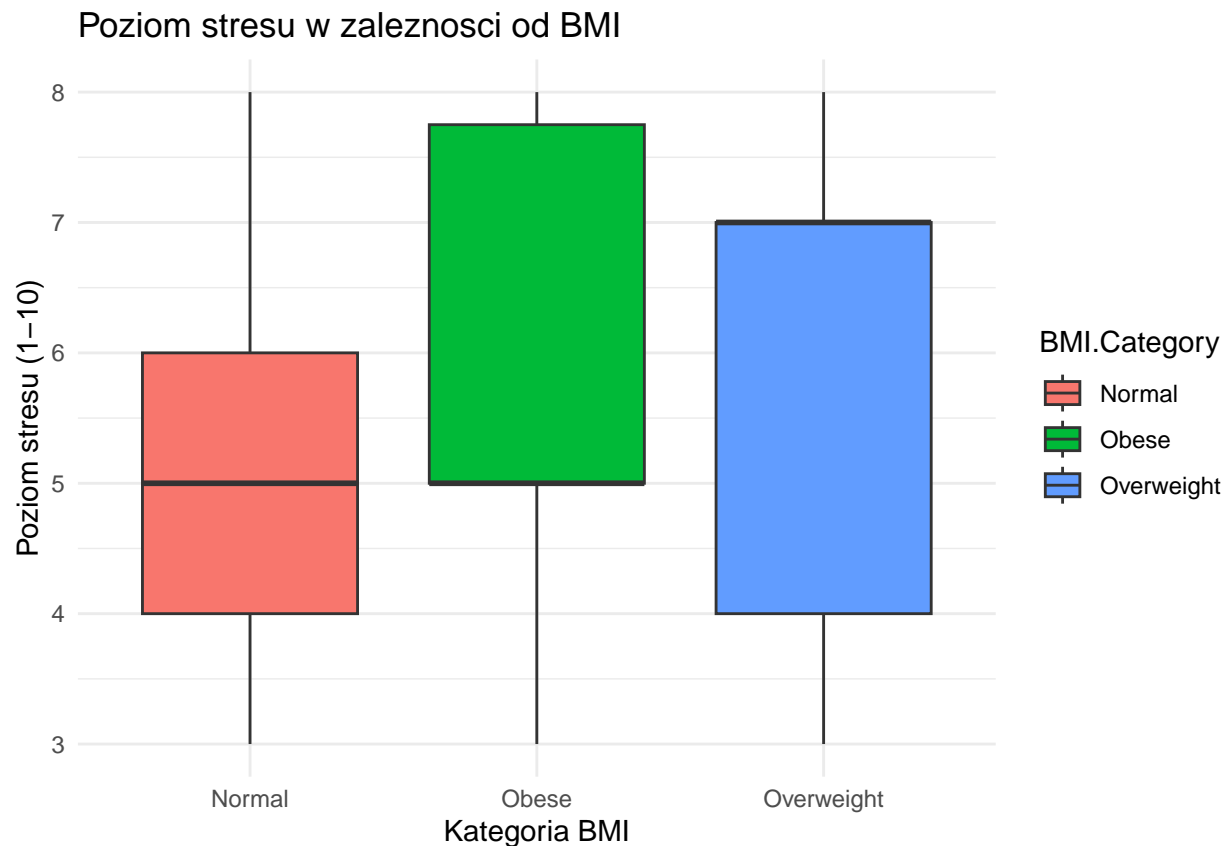
```
test2 <- aov(Stress.Level ~ BMI.Category, data = dane)
```

Wartość  $p$  wyniosła  $p = 0.0058$ , co oznacza, że przynajmniej jedna z grup różni się istotnie od pozostałych (odrzucaamy  $H_0$ ). By sprawdzić, która lub które ze średnich spowodowały taki wynik, przeprowadzamy dalszą analizę post-hoc, korzystając z **testu Tukeya** (HSD-honestly significant difference).

Tablica 2: Istotność różnic ( $p$  adj) dla par grup

	Wartość $p$
Obese-Normal	0.5794
Overweight-Normal	0.0045
Overweight-Obese	0.9985

Powyższa tabela przedstawia  $p$ -wartości dla porównania grup parami. Stąd widać, że jedyna statystycznie istotna różnica istnieje między **osobami z nadwagą a osobami o wadze prawidłowej**. Różnice między pozostałymi grupami nie są istotne.



Rysunek 3: Poziom stresu w zależności od BMI

Wykresy pudełkowe wizualizują poziom stresu w zależności od BMI. Widać tu różnice między grupami.

Wniosek: zależność między BMI a stresem istnieje, ale nie jest liniowa. Na poziom stresu mogą wpływać inne czynniki lub liczebność niektórych grup (np. otyłych) jest zbyt mała, by wykazać istotność różnic z pozostałymi kategoriami.

## Część 3: Analiza korelacji i regresji

### Problem badawczy

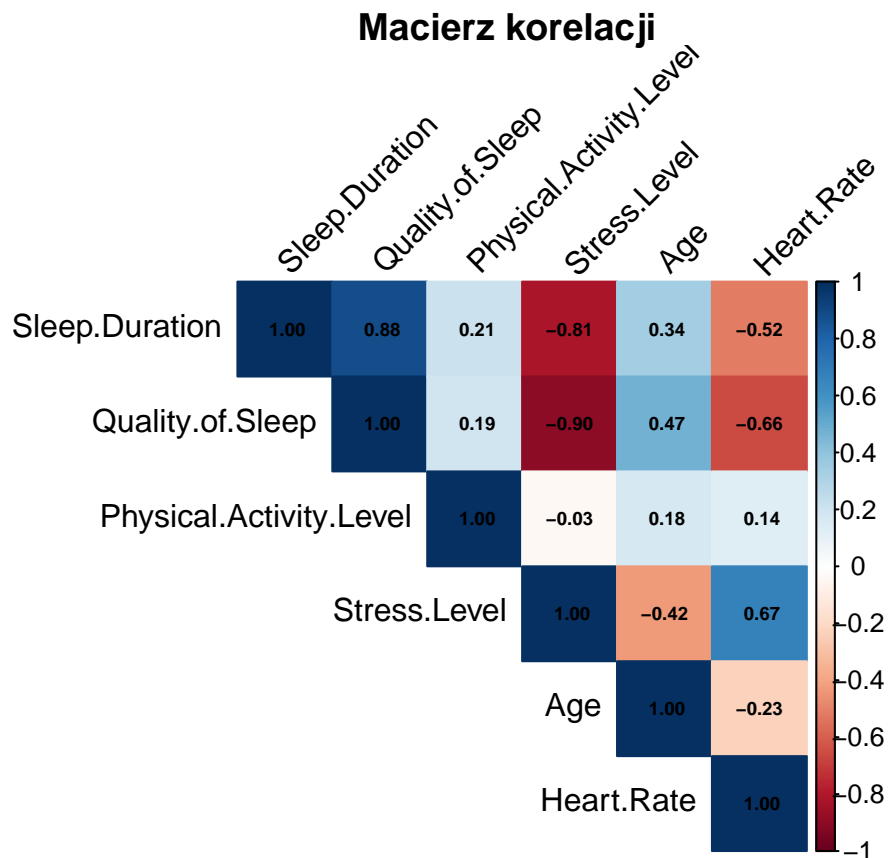
Celem tej części analizy było sprawdzenie, czy styl życia (reprezentowany przez poziom aktywności fizycznej) oraz wiek badanych mają istotny statystycznie wpływ na subiektywną ocenę jakości snu.

### Metodologia

Wykorzystano współczynnik korelacji Pearsona do zbadania siły zależności liniowych między zmiennymi ilościowymi oraz zbudowano model regresji liniowej wielorakiej, gdzie zmienną objaśnianą była Quality of Sleep, a zmiennymi objaśniającymi Physical Activity Level oraz Age.

### Analiza korelacji

```
zmienne_numeryczne <- dane %>%  
  select(Sleep.Duration, Quality.of.Sleep, Physical.Activity.Level, Stress.Level, Age, Heart.Rate)  
  
korelacje <- cor(zmienne_numeryczne)  
  
corrplot(korelacje, method = "color", type = "upper",  
          addCoef.col = "black",  
          tl.col = "black", tl.srt = 45,  
          number.cex = 0.6,  
          title = "Macierz korelacji", mar=c(0,0,1,0))
```



## Model regresji liniowej

Poniżej przedstawiono wyniki modelu regresji liniowej:  $\text{Quality.of.Sleep} \sim \text{Physical.Activity.Level} + \text{Age}$ .

```
model_regresji <- lm(Quality.of.Sleep ~ Physical.Activity.Level + Age, data = dane)
summ <- summary(model_regresji)
print(summ)
```

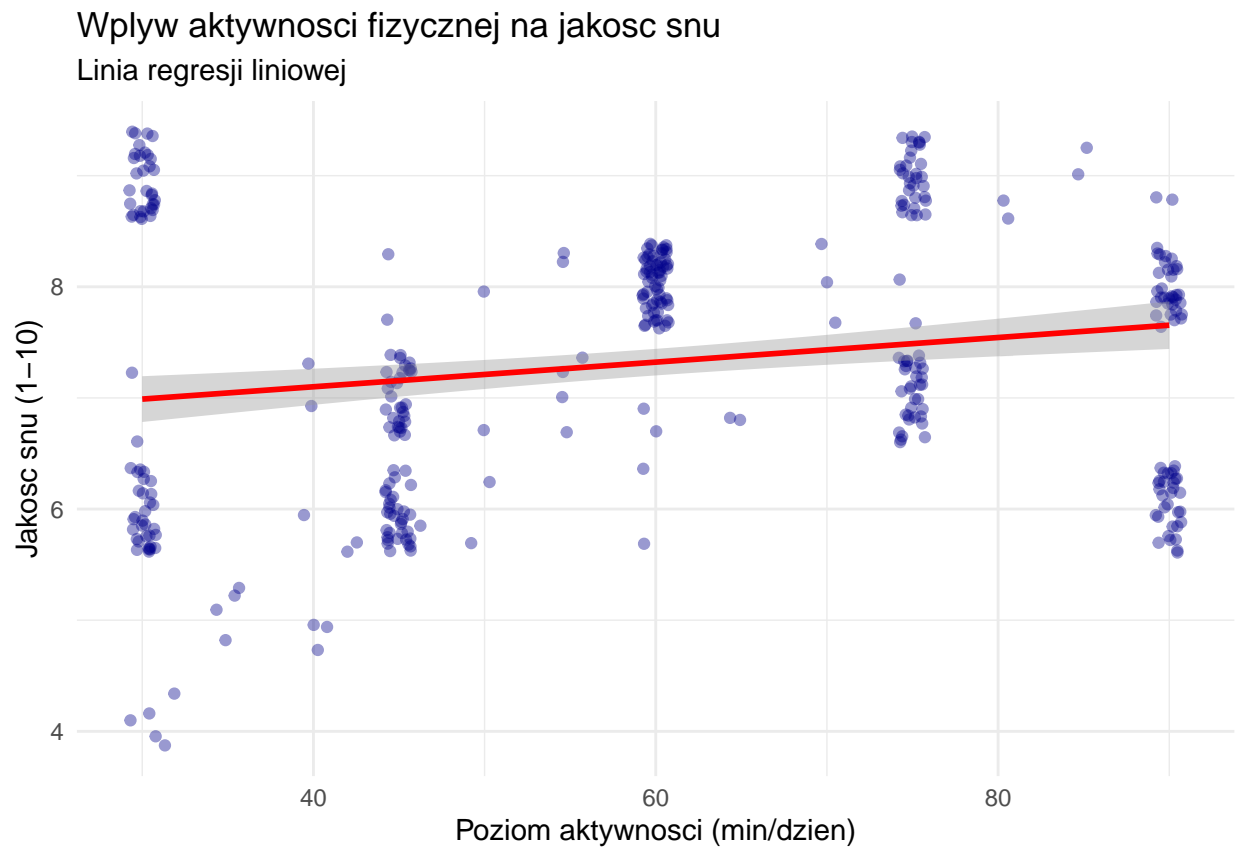
```
##
## Call:
## lm(formula = Quality.of.Sleep ~ Physical.Activity.Level + Age,
##     data = dane)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.6260 -0.5505  0.3486  0.9439  1.7529
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    4.291528   0.290710  14.762  <2e-16 ***
## Physical.Activity.Level 0.006417   0.002649   2.422  0.0159 *
## Age             0.062620   0.006363   9.841  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



```
##
## Residual standard error: 1.049 on 371 degrees of freedom
## Multiple R-squared:  0.2365, Adjusted R-squared:  0.2324
## F-statistic: 57.46 on 2 and 371 DF,  p-value: < 2.2e-16
```

```
ggplot(dane, aes(x = Physical.Activity.Level, y = Quality.of.Sleep)) +
  geom_point(alpha = 0.4, color = "darkblue", position = "jitter") +
  geom_smooth(method = "lm", color = "red", se = TRUE) +
  labs(title = "Wpływ aktywności fizycznej na jakość snu",
       subtitle = "Linia regresji liniowej",
       x = "Poziom aktywności (min/dzień)",
       y = "Jakość snu (1-10)") +
  theme_minimal()
```

```
## `geom_smooth()` using formula = 'y ~ x'
```



## Interpretacja wyników i wnioski

```
#Przygotowanie zmiennych do tekstu
r2 <- round(summ$r.squared, 3)
r2_proc <- round(summ$r.squared * 100, 1)

coef_act <- coef(model_regresji)["Physical.Activity.Level"]
```

```
pval_act <- summ$coefficients["Physical.Activity.Level", "Pr(>|t|)"]
dir_act <- if(coef_act > 0) "wzrostem" else "spadkiem"
istotnosc_act <- if(pval_act < 0.05) "ma istotny statystycznie" else "nie ma istotnego statystycznie"
```

**Analiza korelacji:** Wykres macierzy korelacji wskazuje na bardzo silną pozytywną zależność między czasem trwania snu a jego jakością. Zauważono również, że poziom stresu jest silnie ujemnie skorelowany z jakością snu (im wyższy stres, tym gorszy sen).

**Model regresji:** Model regresji liniowej okazał się istotny statystycznie (p-value < 0.05).

**Dopasowanie modelu:** Współczynnik determinacji ( $R^2$ ) wyniósł 0.236, co oznacza, że model wyjaśnia 23.6% zmienności jakości snu.

**Wpływ aktywności:** Zmienna Physical Activity Level ma istotny statystycznie wpływ na jakość snu. Interpretacja współczynnika sugeruje, że wzrost aktywności fizycznej o jednostkę wiąże się ze wzrostem oceny jakości snu o 0.0064 punktu (przy stałym wieku).

**Wnioski końcowe:** Analiza sugeruje, że wyższa aktywność fizyczna sprzyja lepszej ocenie jakości snu. Jest to ważna wskazówka w kontekście higieny snu, sugerująca, że ruch może być czynnikiem wspierającym zdrowy wypoczynek.