

Praca domowa 2

June 5, 2025

1 Praca domowa nr 2 „Planowanie i analiza eksperymentu”

wykonał Paweł Jan Tłusty s95596 Zestaw zadań dla grupy 1.1

Wersja online: Notes dostępny również w serwisie GitHub w formacie do odtworzenia .ipynb

Link:

<https://github.com/Kotmin/R-Planowanie-Eksperymentu>

Dla wszystkich zadań, o ile nie zaznaczono inaczej, przyjmujemy domyślny poziom istotności $\alpha = 0.05$. W przypadkach, gdzie zostaje on zmieniony (np. $\alpha = 0.01$), informujemy o tym lokalnie przy interpretacji wyników. Z uwagi na problemy z kodowaniem w przypadku exportu pliku do formatu .pdf zamieniono część symboli matematycznych na programistyczne (np. \neq -różne/ operator różny od; $=$ - równe).

$$\alpha = 0.05$$

1.1 Zadanie 1: Porównanie średnich wag mandarynek z dwóch plantacji

Zważono losowo wybrane mandarynki z dwóch różnych plantacji i otrzymano następujące wyniki (w gramach)

Dane:

- Waga1 = c(75, 67, 73, 78, 70, 78, 84, 75, 70, 72, 78)
- Waga2 = c(80, 75, 82, 76, 78, 82, 80, 85, 76, 72)

Cel: Sprawdzić, czy średnia waga mandarynek z dwóch plantacji jest taka sama.

Hipotezy:

- **H0 (hipoteza zerowa):** Średnia waga mandarynek z plantacji 1 == średnia waga z plantacji 2
- **H1 (hipoteza alternatywna):** Średnie są różne

```
[79]: waga1 <- c(75, 67, 73, 78, 70, 78, 84, 75, 70, 72, 78)
      waga2 <- c(80, 75, 82, 76, 78, 82, 80, 85, 76, 72)
```

```
[80]: length(waga1); length(waga2) # Rozmiary prób
```

11

10

```
[81]: mean(waga1); var(waga1)  
      mean(waga2); var(waga2)
```

74.5454545454545

23.2727272727273

78.6

15.3777777777778

```
[82]: waga1
```

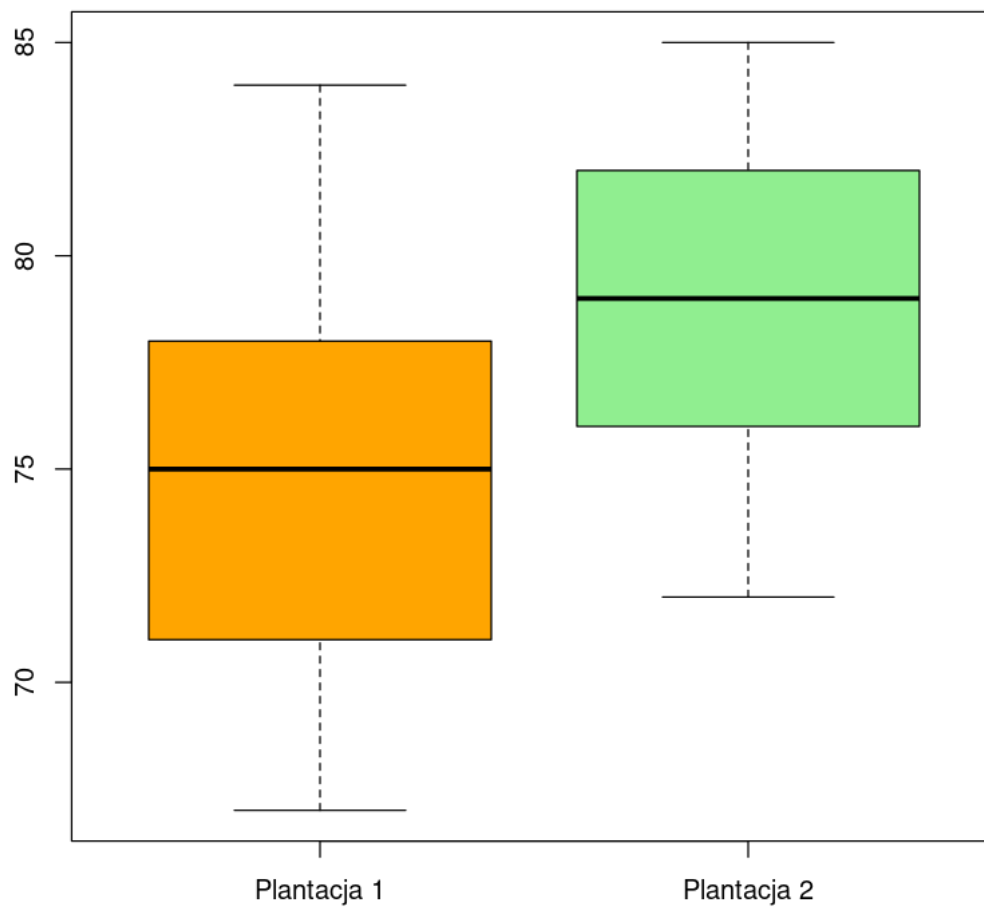
1. 75 2. 67 3. 73 4. 78 5. 70 6. 78 7. 84 8. 75 9. 70 10. 72 11. 78

```
[83]: waga2
```

1. 80 2. 75 3. 82 4. 76 5. 78 6. 82 7. 80 8. 85 9. 76 10. 72

```
[84]: boxplot(waga1, waga2, names = c("Plantacja 1", "Plantacja 2"), col = c("orange", "lightgreen"), main = "Porownanie wag mandarynek")
```

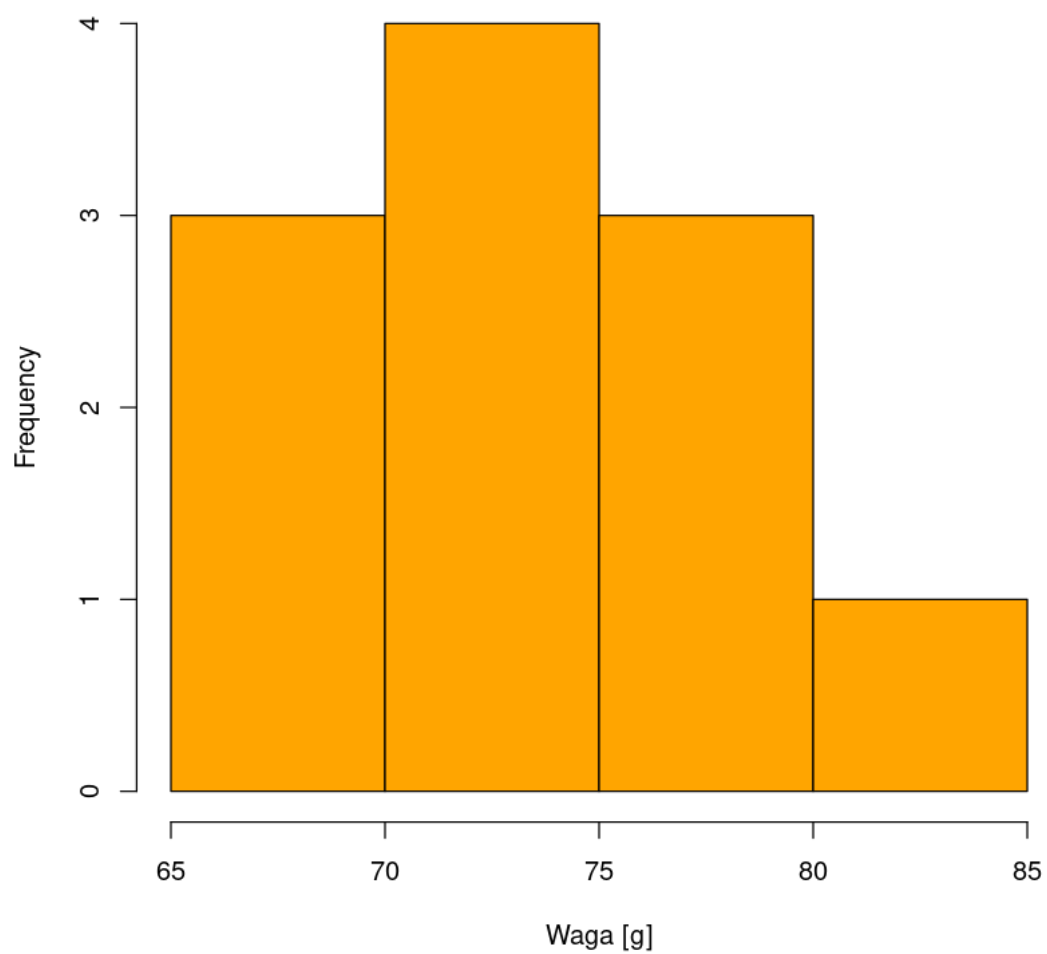
Porównanie wag mandarynek



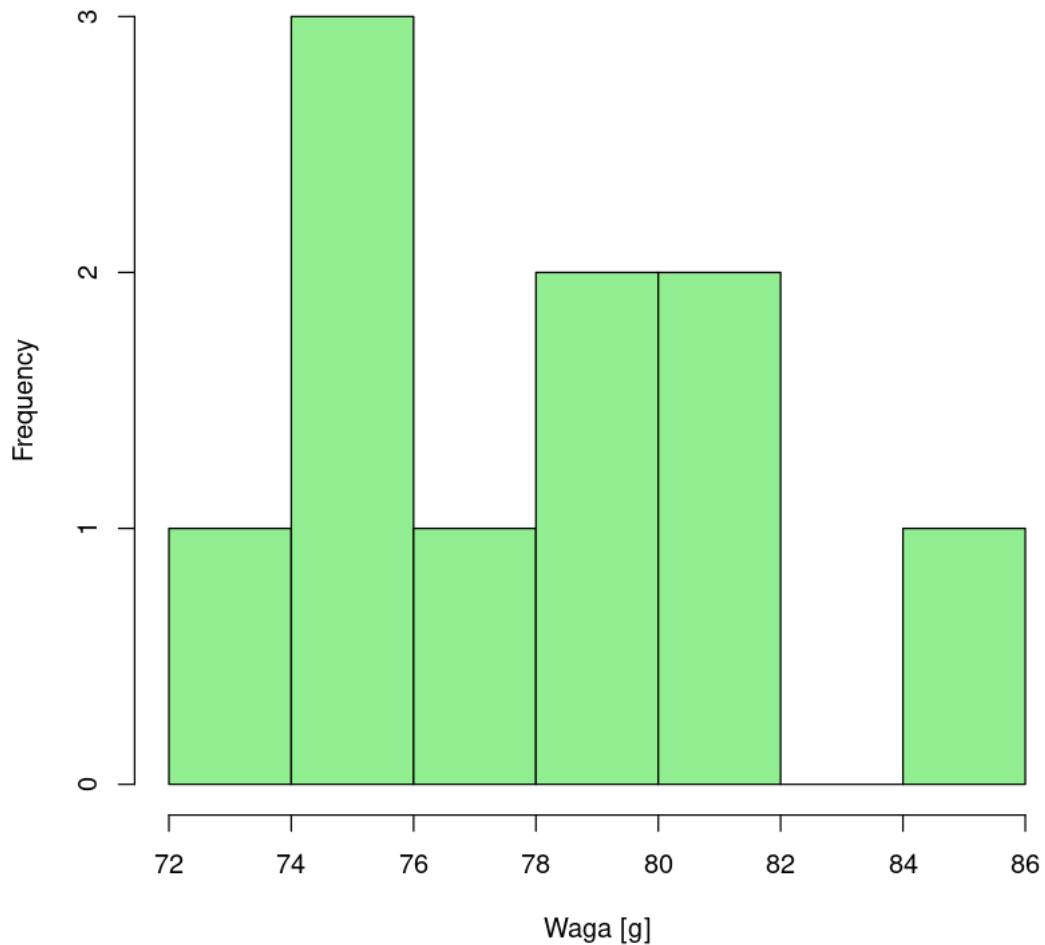
```
[85]: # par(mfrow = c(2, 1))
```

```
[86]: hist(waga1, main = "Histogram: Plantacja 1", col = "orange", xlab = "Waga [g]")  
hist(waga2, main = "Histogram: Plantacja 2", col = "lightgreen", xlab = "Waga [g]")
```

Histogram: Plantacja 1



Histogram: Plantacja 2



```
[87]: # sprawdzimy czy dane mają rozkład normalny
# używamy testu Shapiro-Wilka (próbki są małe)
# test SW jest skonstruowany z hipotezami:
# H0: dane pochodzą z rozkładu normalnego
# H1: dane nie pochodzą z rozkładu normalnego / dane są istotnie różne od
#     ↳ rozkładu normalnego

shapiro.test(waga1)
shapiro.test(waga2)
```

Shapiro-Wilk normality test

```
data: waga1
W = 0.96419, p-value = 0.8227
```

Shapiro-Wilk normality test

```
data: waga2
W = 0.97536, p-value = 0.9356
```

```
[88]: # p-value większe od 0.05 --dla obu przypadków.
# nie znaleźliśmy podstaw do odrzucenia hipotezy zerowej. Dane pochodzą z
↳ rozkładu normalnego
# gdyby dane nie pochodziły z r.norm. użylibyśmy zapewne testu wilcozona dla
↳ średnich, dla zgodności z rozkładem normalnym użyjemy
# dobranej wersji t-test
# w celu doboru sprawdzimy wariancję
var.test(waga1, waga2)
# h0 wariancje są równe
# h1 wariancje się różnią
```

F test to compare two variances

```
data: waga1 and waga2
F = 1.5134, num df = 10, denom df = 9, p-value = 0.545
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.381799 5.719082
sample estimates:
ratio of variances
      1.5134
```

```
[89]: # wariancje są znane, nie znaleźliśmy podstaw do odrzucenia hipotezy zerowej.
↳ Wariancje nie różnią się istotnie statystycznie
t.test(waga1, waga2, var.equal = TRUE)
```

Two Sample t-test

```
data: waga1 and waga2
t = -2.0996, df = 19, p-value = 0.04935
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -8.09632069 -0.01277022
sample estimates:
```

mean of x mean of y
74.54545 78.60000

[90]: # p-value mniejsze od 0.05 - odrzucamy hipotezę zerową

1.1.1 Podsumowanie Z1: Porównanie średniej wagi mandarynek z dwóch plantacji

Czy można twierdzić, że średnia waga mandarynek jest taka sama dla każdej z plantacji?

– Nie.

Hipotezy:

- Hipoteza zerowa (H_0): średnia waga mandarynek jest taka sama dla obu plantacji ($\bar{x} = \bar{y}$)
- Hipoteza alternatywna (H_1): średnia waga mandarynek różni się ($\bar{x} \neq \bar{y}$)

Założenia:

- Próbkki są niezależne – pochodzą z dwóch różnych plantacji
- Dane mają rozkład normalny (sprawdzone testem Shapiro-Wilka)
- Wariancje są równe (sprawdzone testem Fishera)

Wyniki testu t:

- Statystyka t: -2.0996
- p-value: 0.04935
- Przedział ufności 95%: [-8.10, -0.01]
- Średnia (plantacja 1): 74.55 g
- Średnia (plantacja 2): 78.60 g

Wniosek: Na poziomie istotności $\alpha = 0.05$, odrzucamy hipotezę H_0 – istnieje statystycznie istotna różnica średnich wag mandarynek między plantacjami.

p-value bardzo bliskie 0.05 \rightarrow wynik interpretować ostrożnie.

1.2 Zadanie 2: Nowy lek

Polecenie:

Pewna firma farmaceutyczna przetestowała działanie nowego leku na losowo wybranych pacjentach. Celem leku jest obniżenie poziomu pewnego składnika X we krwi. Wyniki badania przed podaniem i po podaniu leku są następujące (w odpowiednich jednostkach dla ustalonej objętości krwi)

Sprawdzić, czy średnio poziom składnika X spada po podaniu leku.

Czy można wnioskować, że lek obniża średnio poziom X ?

Dane:

```
- przed = c(160, 205, 230, 245, 180, 280, 230, 200, 170, 210)
- po     = c(150, 210, 240, 230, 170, 260, 240, 180, 190, 200)
```

Hipotezy:

- H0: Średnia różnica (przed – po) = 0 → brak efektu lub wzrost
- H1: Średnia różnica (przed – po) > 0 → poziom X **spada** po leczeniu

Z konstrukcji zadania wynika, że wartości na kolejnych pozycjach odpowiadają sobie wzajemnie przed/po.

```
[91]: przed <- c(160, 205, 230, 245, 180, 280, 230, 200, 170, 210)
      po    <- c(150, 210, 240, 230, 170, 260, 240, 180, 190, 200)
```

```
[92]: przed
```

1. 160 2. 205 3. 230 4. 245 5. 180 6. 280 7. 230 8. 200 9. 170 10. 210

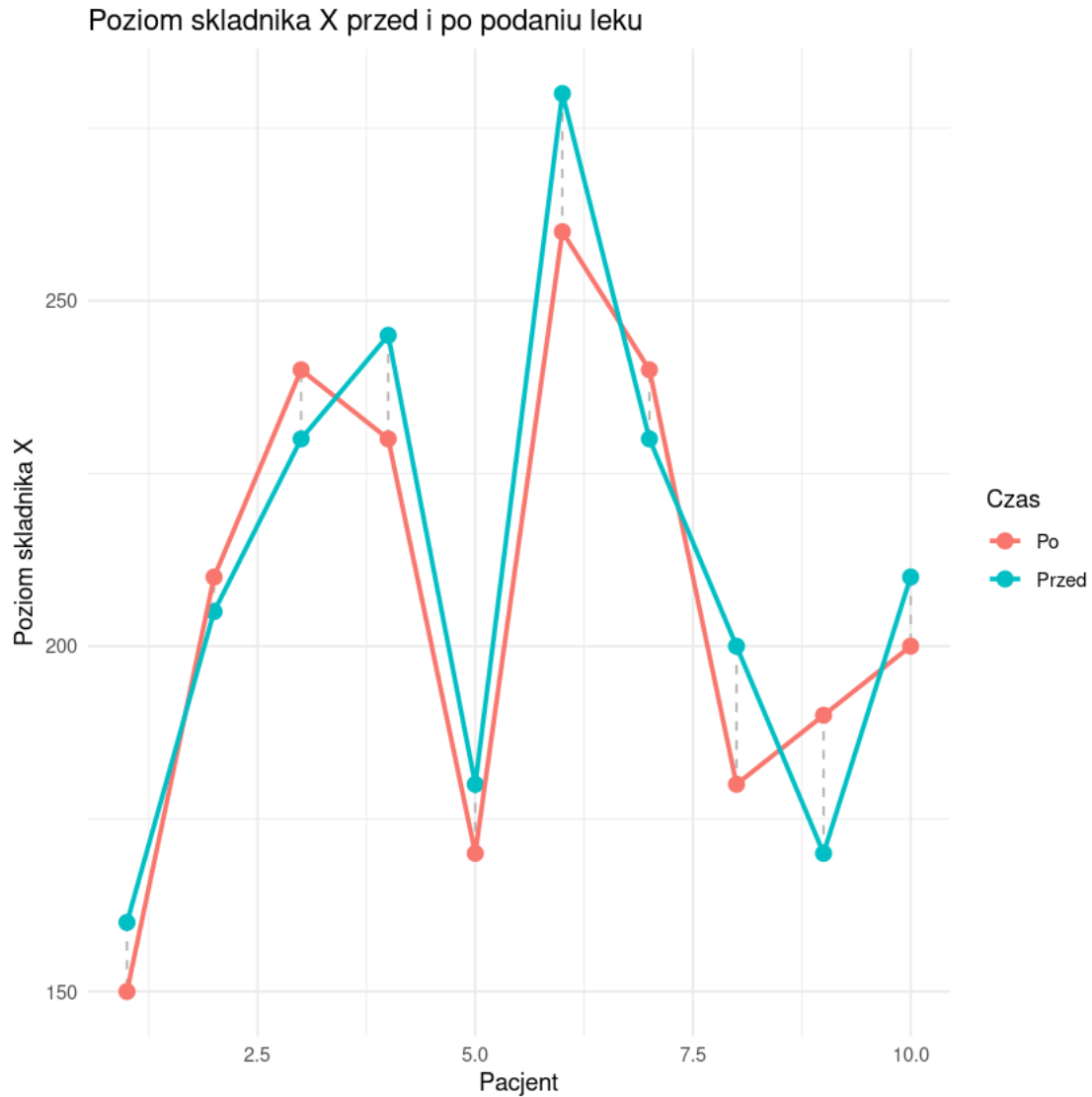
```
[93]: po
```

1. 150 2. 210 3. 240 4. 230 5. 170 6. 260 7. 240 8. 180 9. 190 10. 200

```
[124]: # chcemy zwizualizować dane
pacjenci <- 1:length(przed)
df <- data.frame(
  Pacjent = pacjenci,
  Przed = przed,
  Po = po
)

library(tidyr)
df_long <- pivot_longer(df, cols = c(Przed, Po), names_to = "Czas", values_to = "Wartosc")

# Wykres
library(ggplot2)
ggplot(df_long, aes(x = Pacjent, y = Wartosc, color = Czas, group = Czas)) +
  geom_line(aes(group = Pacjent), color = "gray70", linetype = "dashed") +
  geom_point(size = 3) +
  geom_line(size = 1) +
  labs(title = "Poziom skladnika X przed i po podaniu leku",
       x = "Pacjent",
       y = "Poziom skladnika X",
       color = "Czas") +
  theme_minimal()
```

[95]: *# wykres liniowy zdaje się być tutaj niepoprawny. Chcieliśmy zobaczyć czy ↪ kształt zmian będzie miał podobny przebieg.*
bardziej odpowiednia wizualizacja poniżej

[96]: `library(ggplot2)`

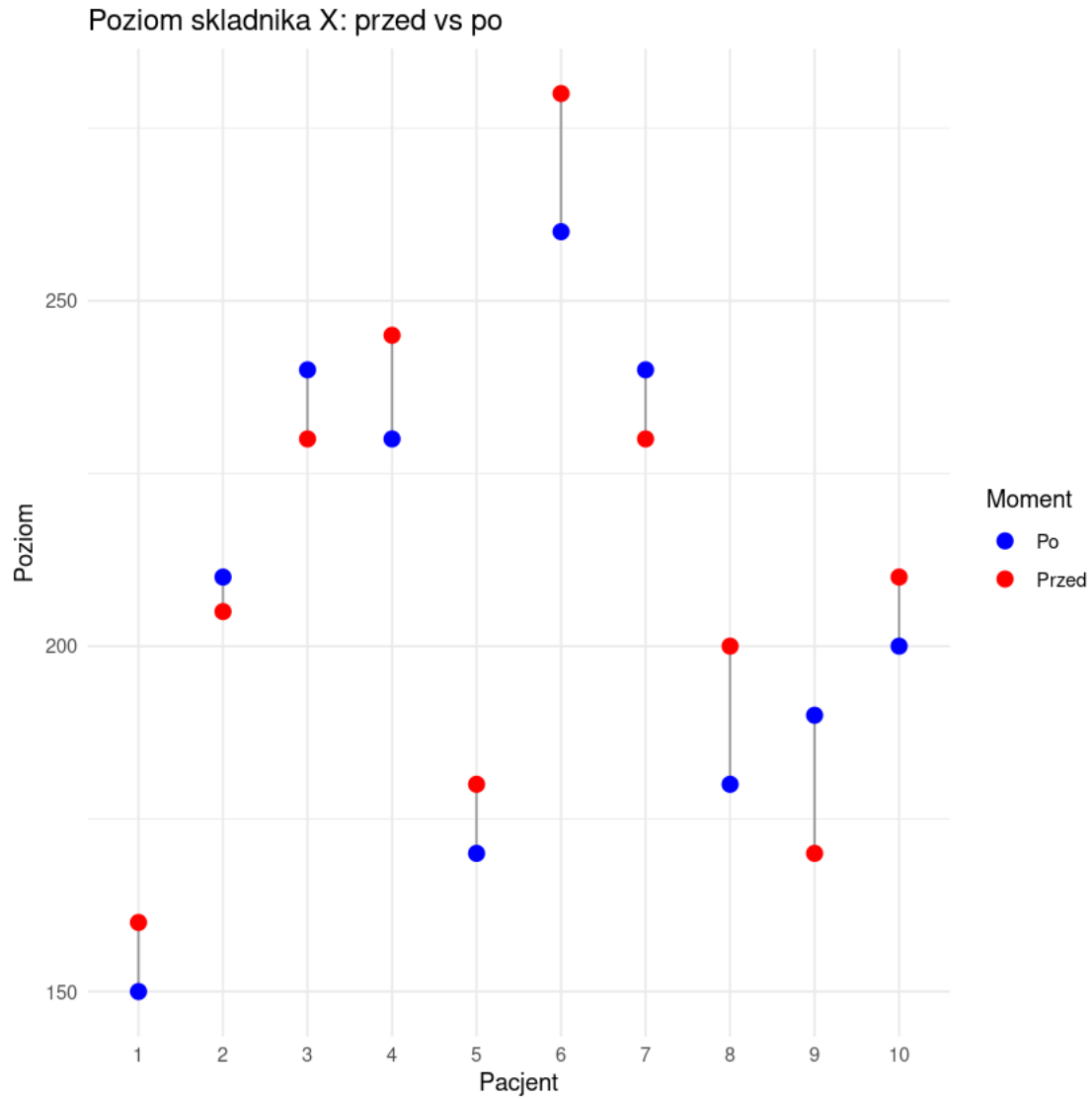
```
# Ramka danych
df <- data.frame(
  Pacjent = factor(1:length(przed)),
  Przed = przed,
  Po = po
)
```

```

# Długa wersja do wykresu
library(tidyr)
df_long <- pivot_longer(df, cols = c(Przed, Po), names_to = "Moment", values_to_
  ↪= "Poziom")

# Wykres z liniami łączącymi przed i po
ggplot(df_long, aes(x = Pacjent, y = Poziom, group = Pacjent)) +
  geom_line(color = "gray60") +
  geom_point(aes(color = Moment), size = 3) +
  scale_color_manual(values = c("Przed" = "red", "Po" = "blue")) +
  labs(
    title = "Poziom składnika X: przed vs po",
    x = "Pacjent",
    y = "Poziom",
    color = "Moment"
  ) +
  theme_minimal()

```



[97]: *# Czy lek obniża X po zażyciu: wykres pokazuje, że zachodzą takie przypadki. ▮*
↪ Należy to jednak zweryfikować

```
roznice <- przed - po
```

[98]: `shapiro.test(roznice)`

Shapiro-Wilk normality test

```
data: roznice
W = 0.89467, p-value = 0.1913
```

```
[99]: # poruszamy się w obszarze testów dla dwóch średnich, w tym przypadku zależnych
# nie znaleźliśmy testem Shapiro-Wilka by różnice miały statystycznie istotną
# różnicę w dystrybucji od rozkładu normalnego
# wobec tego możemy użyć t.test, jednostronny (szukamy odpowiedzi na pytanie
# przed > po)

t.test(przed, po, paired = TRUE, alternative = "greater")
```

Paired t-test

```
data: przed and po
t = 0.89692, df = 9, p-value = 0.1966
alternative hypothesis: true mean difference is greater than 0
95 percent confidence interval:
 -4.175127      Inf
sample estimates:
mean difference
4
```

```
[100]: # p-value jest większe od 0.05 wobec tego nie mamy podstawy do odrzucenia
# hipotezy zerowej.
# Wobec czego w mocy utrzymuje się H0: Lek nie obniża poziomu składnika X
```

1.2.1 Podsumowanie Z2: Czy lek obniża poziom składnika X

Czy można wnioskować, że lek obniża średnio poziom X?

– Nie.

Hipotezy:

- Hipoteza zerowa (H0): Średnia różnica ≤ 0 (lek nie obniża poziomu składnika X)
- Hipoteza alternatywna (H1): Średnia różnica > 0 (lek obniża poziom składnika X)

Założenia:

- Dane pochodzą od tych samych pacjentów (próby zależne)
- Różnice mają rozkład normalny (sprawdzone testem Shapiro-Wilka)

Wyniki testu t (dla prób zależnych, jednostronny):

- Statystyka t: 0.8969
- p-value: 0.1966
- Przedział ufności 95%: $[-4.18, \infty)$

- Średnia różnica (przed - po): 4 jednostki

Wyniki testu t:

- Statystyka t: -2.0996
- p-value: 0.04935
- Przedział ufności 95%: [-8.10, -0.01]
- Średnia (plantacja 1): 74.55 g
- Średnia (plantacja 2): 78.60 g

Wniosek: Wnioskujemy, że nie ma statystycznie istotnych dowodów na to, że lek obniża poziom składnika X we krwi.

1.3 Zadanie 3: Monety

Polecenie:

Pierwszą monetą rzucono 100 razy i otrzymano 46 orłów, drugą monetą rzucono 50 razy i otrzymano 21 orłów, trzecią monetą rzucono 60 razy i otrzymano 26 orłów.

Zweryfikować hipotezę, że prawdopodobieństwo wyrzucenia orła jest takie samo dla wszystkich monet.

kolejne pozycje w strukturze reprezentują kolejne monety

Dane:

- Liczba orłów: `orly = c(46, 21, 26)` - Liczba rzutów: `rzuty = c(100, 50, 60)` - Liczba reszek: `reszki = rzuty - orly`

Hipotezy:

- H0: Prawdopodobieństwo wyrzucenia orła jest takie samo dla wszystkich monet ($p_1 = p_2 = p_3$)
- H1: Co najmniej jedna moneta ma inne prawdopodobieństwo ($p_k \neq p_j$)

Z konstrukcji zadania wynika, że wartości na kolejnych pozycjach odpowiadają sobie wzajemnie przed/po.

```
[101]: orly <- c(46, 21, 26)
      rzuty <- c(100, 50, 60)
      reszki <- rzuty - orly
```

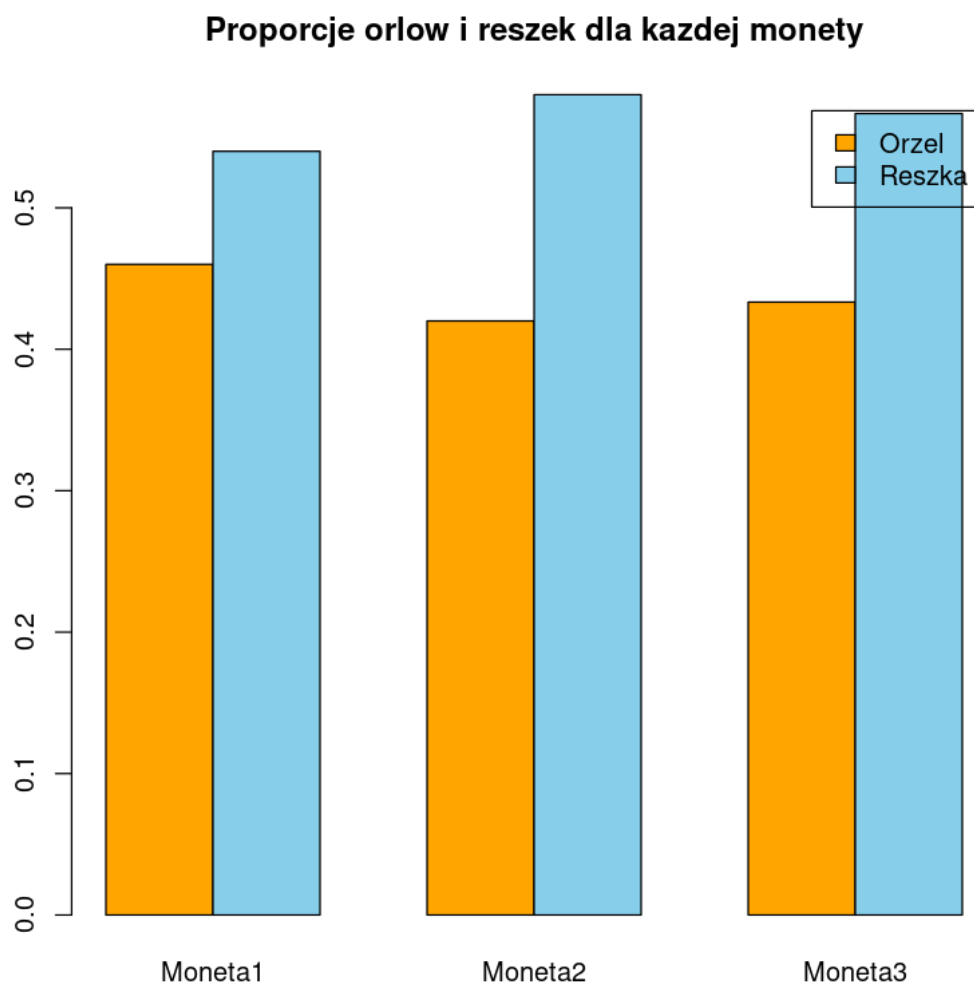
```
[102]: dane <- rbind(orly, reszki)
```

```
[103]: colnames(dane) <- c("Moneta1", "Moneta2", "Moneta3")
      rownames(dane) <- c("Orzel", "Reszka")
```

```
[104]: dane
```

		Moneta1	Moneta2	Moneta3
A matrix: 2 x 3 of type dbl	Orzel	46	21	26
	Reszka	54	29	34

```
[105]: dane_prop <- prop.table(dane, 2) # proporcje w kolumnach
barplot(dane_prop, beside = TRUE,
        col = c("orange", "skyblue"),
        legend = rownames(dane),
        main = "Proporcje orlow i reszek dla kazdej monety")
```



```
[106]: # wizualnie żadna z monet zdaje się nie odbiegać znacznie od reszty
```

```
[107]: test_chi <- chisq.test(dane)
```

```
[108]: test_chi
```

Pearson's Chi-squared test

data: dane

X-squared = 0.24704, df = 2, p-value = 0.8838

```
[109]: # oczekiwane liczności  
test_chi$expected
```

		Moneta1	Moneta2	Moneta3
A matrix: 2 x 3 of type dbl	Orzel	44.28571	22.14286	26.57143
	Reszka	55.71429	27.85714	33.42857

```
[110]: # żadna liczność oczekiwana nie jest mniejsza od 1  
# nie więcej niż 20-25% liczości oczekiwanych jest mniejszych niż 5  
# warunki dla testu chi-kwadrat zostały spełnione  
# nie znaleźliśmy statystycznie istotnego powodu dla odrzucenia H0 -
```

1.3.1 Podsumowanie Z3: Rzut monetą

Czy można twierdzić, że prawdopodobieństwo wyrzucenia orła jest takie samo dla wszystkich monet?

– Tak, brak podstaw do odrzucenia hipotezy zerowej.

Hipotezy:

- Hipoteza zerowa (H0): Prawdopodobieństwo wyrzucenia orła jest **takie samo** dla wszystkich trzech monet
- Hipoteza alternatywna (H1): Co najmniej jedna moneta ma **inne** prawdopodobieństwo wyrzucenia orła

Założenia:

- Dane są niezależne (każda próba dotyczy innej monety)
- Liczności oczekiwane spełniają warunki testu chi-kwadrat:
 - żadna < 1
 - nie więcej niż 25% < 5

Dane:

- Moneta 1: 46 orłów / 100 rzutów
- Moneta 2: 21 orłów / 50 rzutów

- Moneta 3: 26 orłów / 60 rzutów

Wniosek: Wnioskujemy, że nie ma statystycznie istotnych dowodów na różnice między monetami. Możemy przyjąć, że **wszystkie trzy monety mają to samo prawdopodobieństwo wyrzucenia orła**.

1.4 Zadanie 4: Czas świecenia żarówki

Polecenie:

Zmierzono długości czasów świecenia czterech typów żarówek. Celem jest sprawdzenie, czy wszystkie typy mają ten sam średni czas świecenia.

Dane (w godzinach):

```
- typ1 = c(910, 960, 940, 1040, 980, 1020, 1060)
- typ2 = c(950, 1010, 1045, 980, 930, 1035)
- typ3 = c(1020, 1070, 1010, 980, 1005, 950)
- typ4 = c(1010, 930, 950, 990, 1040)
```

Czy średnie czasy świecenia różnią się pomiędzy typami żarówek?

Hipotezy: - H0: średni czas świecenia jest taki sam dla wszystkich grup żarówek ($1 = 2 = 3 = 4$) - H1: co najmniej jedna para średnich różni się

```
[111]: typ1 <- c(910, 960, 940, 1040, 980, 1020, 1060)
      typ2 <- c(950, 1010, 1045, 980, 930, 1035)
      typ3 <- c(1020, 1070, 1010, 980, 1005, 950)
      typ4 <- c(1010, 930, 950, 990, 1040)
```

```
[112]: czas <- c(typ1, typ2, typ3, typ4)
      typ <- factor(rep(c("typ1", "typ2", "typ3", "typ4"), times = c(length(typ1),
      ↪length(typ2), length(typ3), length(typ4))))
```

```
[113]: dane <- data.frame(Czas = czas, Typ = typ)
      head(dane)
```

```
A data.frame: 6 x 2
```

	Czas <dbl>	Typ <fct>
1	910	typ1
2	960	typ1
3	940	typ1
4	1040	typ1
5	980	typ1
6	1020	typ1

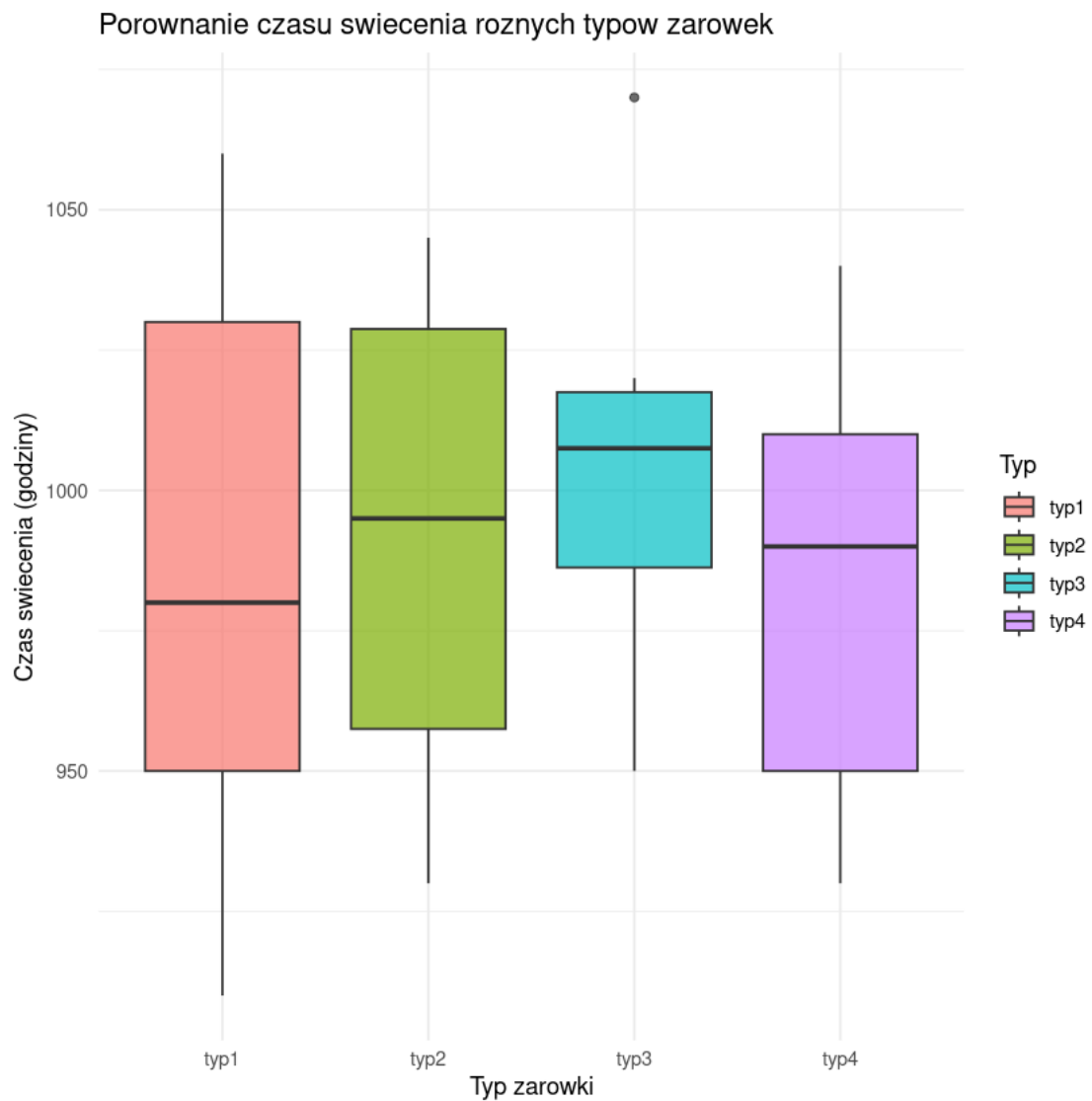
```
[114]: #dane # zakomentowano dla czytelności
```

```
[115]: library(ggplot2)

      ggplot(dane, aes(x = Typ, y = Czas, fill = Typ)) +
```



```
geom_boxplot(alpha = 0.7) +
labs(
  title = "Porównanie czasu świecenia różnych typów żarówek",
  y = "Czas świecenia (godziny)",
  x = "Typ żarówki"
) +
theme_minimal()
```



[116]: *# wizualnie widzimy jakieś różnice - czy statystycznie istotne?*
gdybyśmy szukali odpowiedzi na pytanie „którą żarówkę najpewniej najpóźniej
→musielibyśmy wymienić’’ mielibyśmy silniejszą poszlakę

```
[117]: # właściwym podejściem zdaje się wykorzystanie ANOVA
# musimy sprawdzić jej założenia: normalność & homogeniczność wariancji
# wywołujemy test Shapiro-Wilka dla każdej grupy (sprawdzamy normalność)
by(dane$Czas, dane$Typ, shapiro.test)
```

```
dane$Typ: typ1
```

```
Shapiro-Wilk normality test
```

```
data: dd[x, ]
```

```
W = 0.96381, p-value = 0.8507
```

```
-----
dane$Typ: typ2
```

```
Shapiro-Wilk normality test
```

```
data: dd[x, ]
```

```
W = 0.94008, p-value = 0.6598
```

```
-----
dane$Typ: typ3
```

```
Shapiro-Wilk normality test
```

```
data: dd[x, ]
```

```
W = 0.97159, p-value = 0.9029
```

```
-----
dane$Typ: typ4
```

```
Shapiro-Wilk normality test
```

```
data: dd[x, ]
```

```
W = 0.96873, p-value = 0.867
```

```
[118]: # dla żadnej z grup p-value nie było mniejsze od 0.05, nie mamy podstaw do
# odrzucenia hipotezy zerowej - o normalności
# dla danych o rozkł. norm dla sprawdzenia homogeniczności wariancji możemy
# użyć testu Bartletta (gdyby tak nie było - Levene)
bartlett.test(Czas ~ Typ, data = dane)
```

```
Bartlett test of homogeneity of variances
```

```
data: Czas by Typ
```

Bartlett's K-squared = 0.52124, df = 3, p-value = 0.9142

```
[119]: # H0 dla t.Bartletta - Wariancje we wszystkich grupach są równe
# p-value większe od 0.05, nie znaleźliśmy powodu do odrzucenia H0

# Założenia ANOVA zostały spełnione
```

```
[120]: model <- aov(Czas ~ Typ, data = dane)
summary(model)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Typ	3	1632	544	0.242	0.866
Residuals	20	44917	2246		

```
[121]: # p-value - Pr(>F)
# nie znaleziono podstaw do odrzucenia H0
# Nie mamy podstaw żeby przeprowadzać analizę post-hoc (ANOVA nie pokazał
↳ istotnych różnic)
```

```
[122]: TukeyHSD(model)
```

Tukey multiple comparisons of means
95% family-wise confidence level

Fit: aov(formula = Czas ~ Typ, data = dane)

\$Typ		diff	lwr	upr	p adj
typ2-typ1	4.523810	-69.27180	78.31942	0.9981423	
typ3-typ1	18.690476	-55.10513	92.48608	0.8923859	
typ4-typ1	-3.142857	-80.81049	74.52478	0.9994619	
typ3-typ2	14.166667	-62.41465	90.74799	0.9538484	
typ4-typ2	-7.666667	-87.98583	72.65250	0.9931118	
typ4-typ3	-21.833333	-102.15250	58.48583	0.8709607	

```
[123]: # Średnie różnice w godzinach między różnymi typami zasadniczo wahały się w
↳ przybliżeniu od 4 do 20 godzin.
# przedział lwr, upr zawiera 0 - różnica średnich nie ma znaczenia
# p adj mniejsze od 0.05 mówiłoby o istotnej różnicy między grupami
```

1.4.1 Podsumowanie Z4: Żarówki

Czy można twierdzić, że średnie czasy świecenia są takie same dla wszystkich typów żarówek?

– Tak, brak podstaw do odrzucenia hipotezy zerowej.

Hipotezy:

- Hipoteza zerowa (H_0): Średni czas świecenia jest taki sam dla wszystkich typów żarówek ($\mu_1 = \mu_2 = \mu_3 = \mu_4$)
- Hipoteza alternatywna (H_1): Istnieją różnice między co najmniej dwoma średnimi

Założenia:

- Dane są niezależne i pochodzą z czterech niezależnych grup
- Dane mają rozkład normalny w każdej grupie (sprawdzone testem Shapiro-Wilka)
- Wariancje są równe (sprawdzone testem Levene'a lub Bartlett'a)

Wyniki analizy wariancji (ANOVA):

- Statystyka $F = 0.242$
- $p\text{-value} = 0.866$
- Brak podstaw do odrzucenia hipotezy zerowej

Wyniki testu post-hoc (Tukey HSD):

- Żaden z porównań par typów żarówek nie wykazał istotnej różnicy
- Wszystkie $p\text{-value} > 0.87$
- Wszystkie przedziały ufności zawierały 0

Wniosek: Test ANOVA nie wykazał istotnych różnic (w czasie średniego czasu świecenia żarówek)
– analiza post-hoc nieuzasadniona statystycznie