



**AKADEMIA GÓRNICZO-HUTNICZA IM. STANISŁAWA STASZICA W KRAKOWIE**  
**Wydział Zarządzania**

***Estymacja modeli ANOVA na różnych przykładach***

Autor:	<i>Kacper Prorok</i>
Kierunek studiów:	Informatyka i Ekonometria
Przedmiot:	<i>Statystyczna analiza danych</i>

Kraków, 2024

## Contents

Wprowadzenie .....	3
ANOVA wieloczynnikowa .....	3
Założenia .....	4
1.    Zmienna zależna ma wartości na skali przedziałowej .....	4
2.    Założenia o randomizacji (2,3,4) .....	5
3.    Założenie o normalności danych w każdej próbie .....	5
4.    Założenie o równości wariancji w grupach .....	5
Przeprowadzenie ANOVY wieloczynnikowej .....	6
Analiza post hoc .....	7
Podział na dni(Day) .....	7
Podział na Section .....	7
Dla interakcji Day i Section .....	8
ANOVA z powtarzanymi pomiarami .....	9
Założenia .....	10
1.    Zmienna zależna ma wartości na skali przedziałowej .....	10
2.    Założenia o randomizacji (2,3,4) .....	10
3.    Założenie o normalności danych w każdej próbie .....	10
4.    Założenie o równości wariancji w grupach .....	10
5.    Założenie o sferyczności .....	11
Przeprowadzenie testu ANOVA z powtarzanymi pomiarami .....	11
ANCOVA .....	12
Założenia .....	13
1.    Zmienna zależna ma wartości na skali przedziałowej .....	13
2.    Założenia o randomizacji (2,3,4) .....	13
3.    Założenie o normalności danych w każdej próbie .....	13
4.    Założenie o równości wariancji w grupach .....	13
1.    Liniowość .....	14
2.    Rozkład normalny reszt .....	14
3.    Brak autokorelacji reszt .....	14
4.    Homoskedastyczność .....	15
1.    Zmienne niezależne nieskorelowane .....	15
2.    Jednorodność lub równoległość regresji w obrębie grup .....	15
Przeprowadzenie testu ANCOVA .....	16
Test post hoc .....	16

# Wprowadzenie

ANOVA (Analysis of Variance) to metoda statystyczna, której celem jest analiza różnic pomiędzy średnimi w kilku grupach. Umożliwia ona sprawdzenie, czy co najmniej jedna grupa różni się istotnie od pozostałych, analizując zmienność wyników w obrębie i pomiędzy grupami. Jeśli zmienność pomiędzy grupami jest większa niż wewnątrz grup, sugeruje to, że różnice między średnimi grup są statystycznie istotne.

Metoda ta jest powszechnie stosowana w badaniach eksperymentalnych, gdzie badamy wpływ jednego lub więcej czynników na zmienną zależną. Przykładem może być analiza wpływu różnych rodzajów nawozów na wzrost roślin, gdzie celem jest porównanie średnich wzrostu w grupach nawozowych.

## ANOVA wieloczynnikowa

Test ANOVA z wieloma czynnikami opiera się na tej samej zasadzie co ANOVA jednoczynnikowa. Różnicą jest to, że całkowitą sumę kwadratów  $SS_{\text{total}}$  dzieli się na sumy odchyleń wynikające z wpływu działania: czynnika A ( $SS_A$ ) oraz B ( $SS_B$ ), interakcji A i B ( $SS_{AB}$ ) oraz losowości zjawiska  $SS_E$

$$SS_{\text{total}} = SS_A + SS_B + SS_E$$

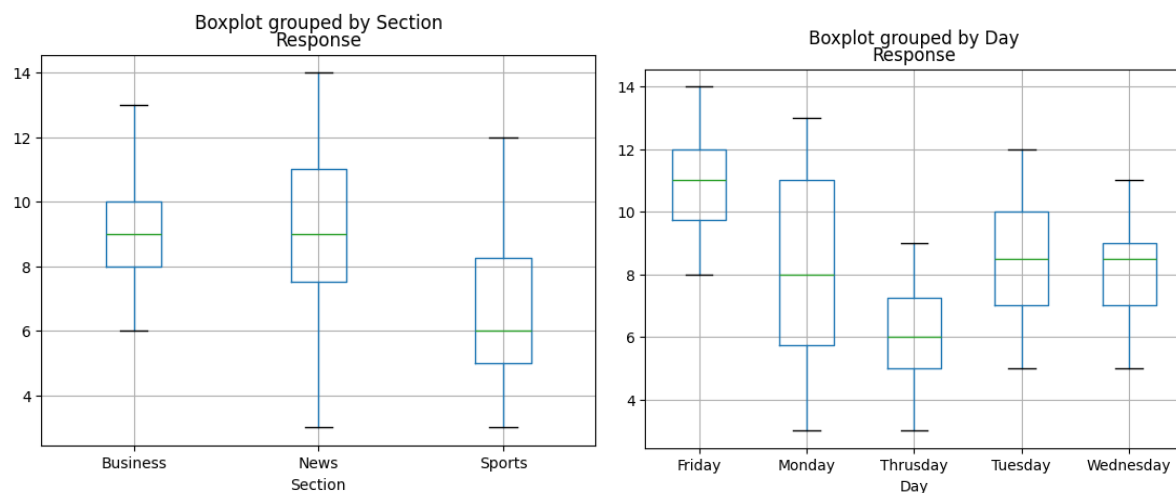
Poniżej przedstawiam kilka pierwszych wierszy zestawu danych, który zapożyczyłem z naszego kursu na UPELu:

	Day	Section	Response
0	Monday	News	11
1	Monday	News	8
2	Monday	News	6
3	Monday	News	8
4	Tuesday	News	9

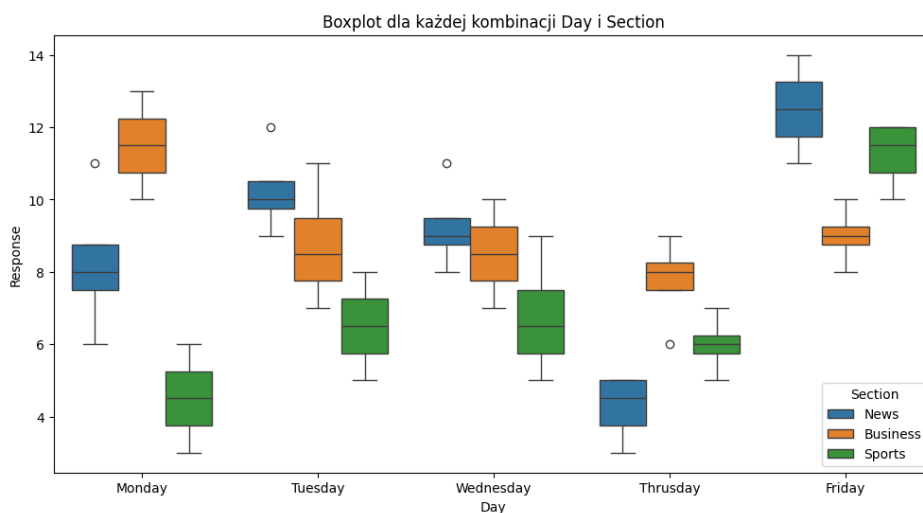
*Tabela 1: Dane ogłoszenia.csv*

Zmienne **Day** (poniedziałek-piątek) oraz **Section** (News, Business, Sports) stanowią zmienne kategoryczne, natomiast zmienna **Response** reprezentuje wynik przypisany do danej kombinacji kategorii.

W ramach analizy **ANOVA wieloczynnikowej** badamy jednoczesny wpływ kilku czynników na zmienną zależną. W tym przypadku zmienna **Response** jest zmienną zależną, natomiast zmienne **Day** i **Section** pełnią rolę czynników, czyli zmiennych niezależnych.



Rysunek 1 Podział na Section oraz Day



Rysunek 2 Podział na Section i Day

Już na pierwszy rzut oka możemy zobaczyć że istnieją spore różnice między niektórymi podgrupami, ale aby stwierdzić że są one statystycznie istotne, musimy przeprowadzić test ANOVA.

## Założenia

1. Zmienna zależna ma wartości na skali przedziałowej
2. Próbkę została wybrana z populacji w sposób losowy
3. Elementy próby zostały przypisane do danych podgrup losowo (II zasada randomizacji)
4. Wszystkie pomiary są niezależne
5. Dane w każdej próbie mają rozkład normalny,
6. Wariancje w podgrupach są równe.

### 1. Zmienna zależna ma wartości na skali przedziałowej

Zmienna Response ma wartości na skali przedziałowej, gdyż jesteśmy w stanie określić różnice między dwoma wartościami i możemy ją interpretować (np. różnica wynosząca 10 będzie spora)

## 2. Założenia o randomizacji (2,3,4)

Zakładam, że autor danych poprawnie je przygotował i dane spełniają założenia o randomizacji oraz że pomiary są od siebie niezależne.

## 3. Założenie o normalności danych w każdej próbie

Dla każdej podgrupy przeprowadziłem test normalności Shapiro-Wilka w celu oceny rozkładu danych. Poniżej przedstawiam wyniki testów:

Section	p-value
Business	0,54129291
News	0,60100137
Sports	0,06013187

Rysunek 4 Wyniki przy podziale na Section

Day	Section	p-value
Friday	Business	0,682962
	News	0,971877
	Sports	0,272453
Monday	Business	0,971877
	News	0,571911
	Sports	0,971877
Thrusday	Business	0,406387
	News	0,272453
	Sports	0,682962
Tuesday	Business	0,849971
	News	0,406387
	Sports	0,971877
Wednesda	Business	0,971877
	News	0,406387
	Sports	0,849971

Rysunek 5 Wyniki przy podziale na Section i Day

Day	p-value
Friday	0,915503
Monday	0,607099
Thrusday	0,854963
Tuesday	0,993397
Vednesda	0,836246

Rysunek 3 Wyniki

Wszystkie rozważane podgrupy spełniają założenie o zgodności z rozkładem normalnym.

## 4. Założenie o równości wariancji w grupach

Założenie o równości wariancji w grupach zostanie zweryfikowane za pomocą testu Bartletta, który przeprowadzę zarówno dla interakcji czynników oraz dla każdego osobno. W teście Barletta hipotezy przyjmują następującą postać:

$$H_0: \sigma^2_1 = \sigma^2_2 = \dots = \sigma^2_k$$
$$H_1: \text{nie wszystkie } \sigma^2_i \text{ są jednakowe}$$

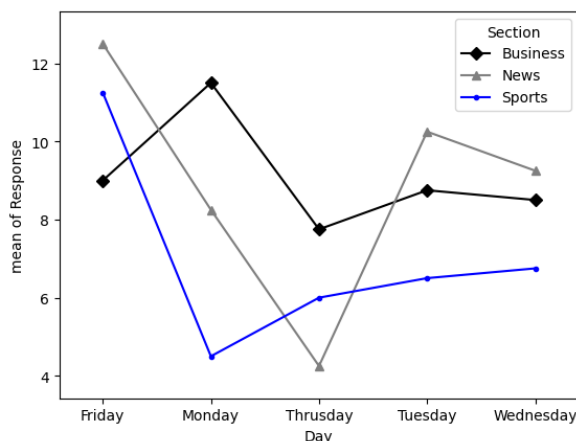
	p-value
Day	0,098205222
Section	0,060687273
Day:Section	0,984181972

Tabela 2 Wyniki testu Bartletta

Dla każdego z czynników wartość p wynosi powyżej 0,05, co oznacza, że nie ma podstaw do odrzucenia hipotezy zerowej. W związku z tym przyjmujemy, że wariancje w każdej kombinacji są równe.

# Przeprowadzenie ANOVY wieloczynnikowej

Przed przystąpieniem do przeprowadzenia testu ANOVA należy zweryfikować, czy występuje efekt interakcji. Poniżej przedstawiam wykres średnich, który ilustruje ten efekt:



Rysunek 6 Wykres interakcji

Linie na wykresie przecinają się, co wskazuje na obecność efektu interakcji. Oznacza to, że między czynnikami zachodzą interakcje. W związku z tym, przeprowadzam test ANOVA uwzględniający interakcje. Hipotezy przyjmują następującą formę:

**$H_0$ : Średnie we wszystkich grupach są równe, tj.  $\mu_1 = \mu_2 = \dots = \mu_k$ .**

**$H_1$ : Przynajmniej jedna średnia różni się od pozostałych.**

Przejdźmy teraz do wyniku testu ANOVA:

	sum_sq	df	F	PR(>F)
C(Day)	146,8333	4	20,90981	8,51754E-10
C(Section)	53,73333	2	15,3038	8,50275E-06
C(Day):C(Section)	135,7667	8	9,66693	1,12469E-07
Residual	79	45		

Tabela 3 Wyniki testu ANOVA wieloczynnikowej

Wyniki są następujące:

- Po podzieleniu danych na grupy według dni (Day) zaobserwowano istotne różnice w średnich
- Po podzieleniu danych na grupy według sekcji (Section) zaobserwowano istotne różnice w średnich
- Gdy uwzględnimy interakcje pomiędzy Section i Day to również obserwujemy istotne różnice w średnich

## Analiza post hoc

Po przeprowadzeniu testu ANOVA, wiemy że różnice między średnimi występują przy podziale danych na czynnik Day, Section oraz interakcji Section:Day. Aby dowiedzieć się, które podgrupy dokładnie spowodowały te różnice musimy przeprowadzić analizę post hoc. Posłużę się do tego testem Tukeya.

### Podział na dni(Day)

group1	group2	meandiff	p-adj	lower	upper	reject
Friday	Monday	-2,8333	0,0218	-5,3773	-0,2893	TRUE
Friday	Thursday	-4,9167	0	-7,4607	-2,3727	TRUE
Friday	Tuesday	-2,4167	0,0702	-4,9607	0,1273	FALSE
Friday	Wednesday	-2,75	0,0279	-5,294	-0,206	TRUE
Monday	Thursday	-2,0833	0,1574	-4,6273	0,4607	FALSE
Monday	Tuesday	0,4167	0,9904	-2,1273	2,9607	FALSE
Monday	Wednesday	0,0833	1	-2,4607	2,6273	FALSE
Thursday	Tuesday	2,5	0,0563	-0,044	5,044	FALSE
Thursday	Wednesday	2,1667	0,1301	-0,3773	4,7107	FALSE
Tuesday	Wednesday	-0,3333	0,9959	-2,8773	2,2107	FALSE

Tabela 4 Wyniki testu Tukeya (Dni)

W przypadku podziału danych według dni, różnice w średnich występują pomiędzy grupami: **Friday-Monday**, **Friday-Thursday** oraz **Friday-Wednesday**. Piątek szczególnie wyróżnia się na tle pozostałych grup.

### Podział na Section

group1	group2	meandiff	p-adj	lower	upper	reject
Business	News	-0,2	0,9659	-2,1167	1,7167	FALSE
Business	Sports	-2,1	0,0286	-4,0167	-0,1833	TRUE
News	Sports	-1,9	0,0525	-3,8167	0,0167	FALSE

Tabela 5 Wyniki testu Tukeya (Section)

W przypadku podziału danych według sekcji, różnice w średnich występują między grupami **Business i Sports**, a w podgrupie **News-Sports** obserwuje się wynik bliski granicy odrzucenia hipotezy zerowej.

## Dla interakcji Day i Section

group1	group2	meandiff	p-adj
Friday - Business	Friday - News	3,5	0,0338
Friday - Business	Monday - Sports	-4,5	0,0015
Friday - Business	Thursday - News	-4,75	0,0006
Friday - News	Monday - News	-4,25	0,0034
Friday - News	Monday - Sports	-8	0
Friday - News	Thursday - Business	-4,75	0,0006
Friday - News	Thursday - News	-8,25	0
Friday - News	Thursday - Sports	-6,5	0
Friday - News	Tuesday - Business	-3,75	0,0163
Friday - News	Tuesday - Sports	-6	0
Friday - News	Wednesday - Business	-4	0,0075
Friday - News	Wednesday - Sports	-5,75	0
Friday - Sports	Monday - Sports	-6,75	0
Friday - Sports	Thursday - Business	-3,5	0,0338
Friday - Sports	Thursday - News	-7	0
Friday - Sports	Thursday - Sports	-5,25	0,0001
Friday - Sports	Tuesday - Sports	-4,75	0,0006
Friday - Sports	Wednesday - Sports	-4,5	0,0015
Monday - Business	Monday - Sports	-7	0
Monday - Business	Thursday - Business	-3,75	0,0163
Monday - Business	Thursday - News	-7,25	0
Monday - Business	Thursday - Sports	-5,5	0
Monday - Business	Tuesday - Sports	-5	0,0003
Monday - Business	Wednesday - Sports	-4,75	0,0006
Monday - News	Monday - Sports	-3,75	0,0163
Monday - News	Thursday - News	-4	0,0075
Monday - Sports	Tuesday - Business	4,25	0,0034
Monday - Sports	Tuesday - News	5,75	0
Monday - Sports	Wednesday - Business	4	0,0075
Monday - Sports	Wednesday - News	4,75	0,0006
Thursday - Business	Thursday - News	-3,5	0,0338
Thursday - News	Tuesday - Business	4,5	0,0015
Thursday - News	Tuesday - News	6	0
Thursday - News	Wednesday - Business	4,25	0,0034
Thursday - Sports	Tuesday - Business	2,75	0,2197
Thursday - Sports	Tuesday - News	4,25	0,0034
Tuesday - News	Tuesday - Sports	-3,75	0,0163
Tuesday - News	Wednesday - Sports	-3,5	0,0338

Tabela 6 Wyniki testu Tukeya (Section+Dni)

Z uwagi na fakt że powstało bardzo dużo kombinacji grup powyżej ukazuje tylko te, w którym test post hoc wykazał że występują istotne różnice.



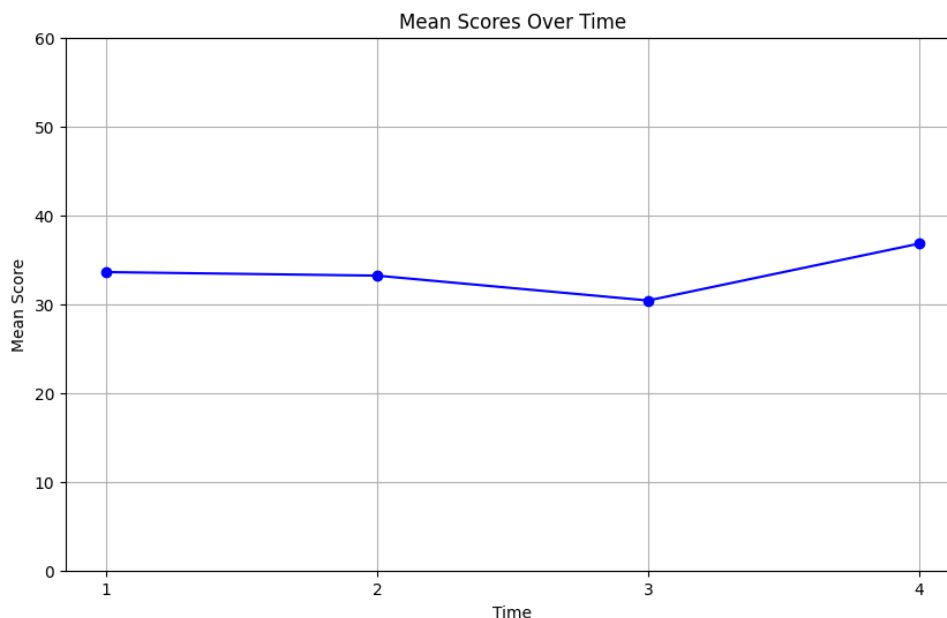
# ANOVA z powtarzanymi pomiarami

Procedura ANOVA dla powtarzanych pomiarów analizuje grupy powiązanych zmiennych zależnych, które reprezentują różne pomiary tego samego atrybutu. Szczególnie test ten przydaje się na przykład przy badaniu terapii w czasie lub przy pomiarach wzrostu roślin. Kluczowym czynnikiem w badaniu jest czas, bo tak naprawdę to istotność tego czynnika chcemy zmierzyć (np. czy terapia faktycznie poprawiła wyniki pacjentów).

Do przeprowadzenia badania posłużę się przykładowym zestawem danych, który przedstawia wyniki testu terapeutycznego, do którego przystąpiło pięciu pacjentów. Wyniki są podzielone w czasie – testy były przeprowadzane co miesiąc, co przedstawiają odpowiednie kolumny.

Czas	1	2	3	4
Pacjent				
1	36	38	30	29
2	34	38	30	29
3	34	28	38	32
4	38	34	20	44
5	26	28	34	50

*Tabela 7 Dane o wynikach testu terapeutycznego*



*Rysunek 8 Wykres średnich*

Po wykresie średnich możemy od razu zauważyć, że różnice nie wydają się być wielkie. Będziemy mogli to stwierdzić dopiero po przeprowadzeniu testu.

## Założenia

Założenia są podobne jak w przypadku testu ANOVA jedno-czynnikowej z jednym dodatkowym - kowariancja pomiarów u tej samej osoby musi być jednorodna (tj. kowariancje są w przybliżeniu równe). Nie bierzemy pod uwagę również założenia o niezależności danych, ponieważ dane ze względu na swoją budowę (np. pobierane od jednego pacjenta) są od siebie zależne,

### 1. Zmienna zależna ma wartości na skali przedziałowej

Wynik testu ma wartości na skali przedziałowej, gdyż jesteśmy w stanie określić różnice między dwoma wartościami i możemy ją interpretować (np. różnica wynosząca 30 będzie spora)

### 2. Założenia o randomizacji (2,3,4)

Zakładam, że autor danych poprawnie je przygotował i dane spełniają założenia o randomizacji.

### 3. Założenie o normalności danych w każdej próbie

Należy zbadać zgodność rozkładu każdej z grup (czyli okresów czasowych) z rozkładem normalnym. Poniżej przedstawiam wyniki testu Shapiro-Wilka:

Variable	Statistic	p-value
Time 1	0,859908	0,22792
Time 2	0,81432	0,105475
Time 3	0,932799	0,615581
Time 4	0,833279	0,147184

Tabela 8 Wyniki testu Shapiro-Wilka

P-value w żadnej grupie nie wyniosło poniżej 0.5, co oznacza że **założenie o zgodności z rozkładem normalnym jest spełnione w każdej grupie.**

### 4. Założenie o równości wariancji w grupach

Levene's test sprawdza, czy wariancje w różnych grupach są jednorodne (homogeniczne). Hipotezy wyglądają następująco:

**$H_0$ : Wariancje w różnych grupach są równe (homogeniczne).**

**$H_1$ : Wariancje w różnych grupach są różne (brak homogeniczności).**

**P-value wyniosło 0.54**, więc stwierdzam że istnieje brak podstaw do odrzucenia hipotezy zerowej. Oznacza to, że **wariancje w grupach są równe, czyli założenie jest spełnione.**

## 5. Założenie o sferyczności

Założenia o jednorodności kowariancji może być zamienione na nieco mniej wymagające, tzn. równość wariancji dla różnic na każdym poziomie, czyli założenie o sferyczności. Do weryfikacji tego założenia przeprowadzę test Mauchley'a. Hipotezy wyglądają następująco:

**$H_0$ : Różnice między parami pomiarów mają równą wariancję, a kowariancje są równe zero**

**$H_1$ : Różnice między parami pomiarów mają różne wariancje, a kowariancje są różne od zera**

W przypadku naszego zestawu danych **p-value wyniosło 0.659**, co oznacza że istnieje brak podstaw do odrzucenia hipotezy zerowej – **macierz kowariancji spełnia założenie o sferyczności**.

## Przeprowadzenie testu ANOVA z powtarzanymi pomiarami

W analizie ANOVA z powtarzanymi pomiarami biorą udział następujące hipotezy zerowe i alternatywne:

**$H_0$ :  $\mu_{czas1} = \mu_{czas2} = \mu_{czas3} = \mu_{czas4}$**

**$H_1$ : przynajmniej jedna średnia populacji różni się od reszty**

Do przeprowadzenia tego testu używam funkcji `AnovaRM(depvar='Score', subject='Patient', within=['Time'])` z pakietu statsmodels w Pythonie. Wyniki są następujące:

F Value	Num DF	Den DF	Pr > F
0,567884	3	12	0,646647

Tabela 9 Wyniki testu ANOVA z powtarzanymi pomiarami

P-value wynosi więcej niż 0.05, więc nie ma podstaw do odrzucenia  $H_0$ . Oznacza to, że **nie występują istotne różnice między okresami**. Możemy powiedzieć, że działanie terapii na przestrzeni 4 miesięcy nie przyniosło istotnego wpływu na wyniki badań u pacjentów.

# ANCOVA

Analiza kowariancji (ANCOVA) to metoda statystyczna łącząca cechy analizy wariancji (ANOVA) i regresji liniowej. ANCOVA pozwala na badanie wpływu jednego lub więcej czynników (zmiennych niezależnych kategoryalnych) na zmienną zależną, jednocześnie kontrolując wpływ zmiennych współzmiennych (covariates), które są ilościowe i mogą wprowadzać dodatkową wariancję.

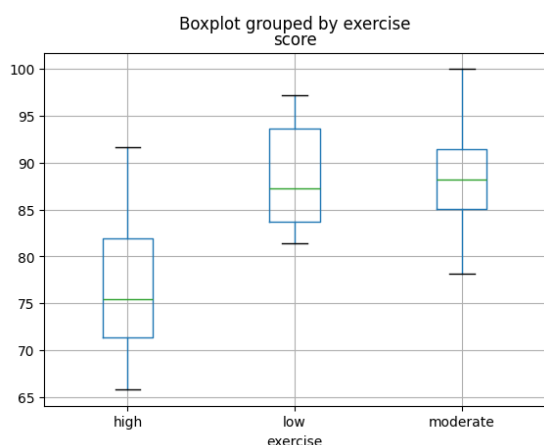
Główne cele ANCOVY to:

1. Usunięcie wpływu współzmiennych na zmienną zależną, aby uzyskać bardziej precyzyjne oszacowanie efektów głównych i interakcji.
2. Sprawdzenie, czy różnice w zmiennej zależnej między grupami są istotne, po uwzględnieniu współzmiennych.

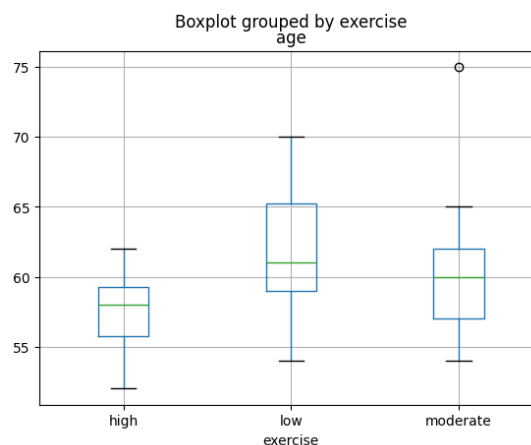
Analizę kowariancji przeprowadzę na danych *stress* z pakietu *datarium*. Przedstawiają one badanie mające na celu ustalić wpływ nowej terapii oraz ćwiczeń na wynik stresu. Dodatkowo mamy informację o wieku każdego pacjenta. **Celem mojego badania jest sprawdzenie wpływu ilości ćwiczeń na poziom stresu (wynik testu).**

	id	score	treatment	exercise	age
0	1	95,6	yes	low	59
1	2	82,2	yes	low	65
2	3	97,2	yes	low	70
3	4	96,4	yes	low	66
4	5	81,4	yes	low	61

Tabela 10 Zestaw danych *stress*



Rysunek 9 Rozkład poziomu stresu dla każdej grupy ćwiczeniowej



Rysunek 10 Rozkład wieku dla każdej grupy ćwiczeniowej

Po analizie wykresów, możemy zauważyć że w grupie która ćwiczyła najwięcej poziom stresu był znacząco niższy. Jednocześnie możemy zauważyć, że tej grupie średnia wieku jest najniższa. Może to sugerować, że wiek również istotnie wpływa na poziom stresu – osoby młodsze mniej się stresują.

W celu wyeliminowania wpływu wieku na nasze badanie przeprowadzę test ANCOVA, w którym zmienną zależną będzie 'score', czynnikiem będzie zmienna 'exercise', a zmienną współzależną (covariant) będzie 'age'.

## Założenia

Dane muszą spełniać zarówno założenia ANOVY jak i regresji liniowej. Zaczę od sprawdzenia założeń ANOVY:

### 1. Zmienna zależna ma wartości na skali przedziałowej

Zmienna Score ma wartości na skali przedziałowej, gdyż jesteśmy w stanie określić różnice między dwoma wartościami i możemy ją interpretować (np. różnica wynosząca 50 będzie spora)

### 2. Założenia o randomizacji (2,3,4)

Zakładam, że autor danych poprawnie je przygotował i dane spełniają założenia o randomizacji oraz że pomiary są od siebie niezależne.

### 3. Założenie o normalności danych w każdej próbie

Dla każdej podgrupy przeprowadziłem test normalności Shapiro-Wilka w celu oceny rozkładu danych. Poniżej przedstawiam wyniki testów:

Exercise	p-value
Low	0.07
Moderate	0.9
High	0.445

*Tabela 11 Wyniki testu Shapiro-Wilka*

**Wszystkie rozważane podgrupy spełniają założenie o zgodności z rozkładem normalnym.**

### 4. Założenie o równości wariancji w grupach

Założenie o równości wariancji w grupach zostanie zweryfikowane za pomocą testu Barletta. W teście Barletta hipotezy przyjmują następującą postać:

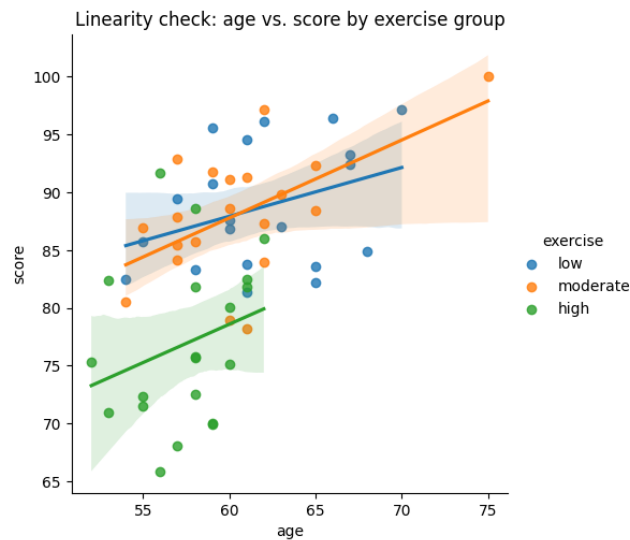
$$H_0: \sigma^2_1 = \sigma^2_2 = \dots = \sigma^2_k$$

$$H_1: \text{nie wszystkie } \sigma^2_i \text{ są jednakowe}$$

**P-value wyniosło 0.397, co oznacza że istnieje brak podstaw do odrzucenia  $H_0$ , czyli dane spełniają założenie o równości wariancji w grupach**

Teraz przejdziemy do badania założeń związanych z regresją liniową:

## 1. Liniowość



Rysunek 11 Wykres zależności liniowych

Widzimy, że istnieje zależność liniowa między zmienną zależną (score) i zmienną współzależną (age) w każdej z podgrup zmiennej niezależnej (exercise).

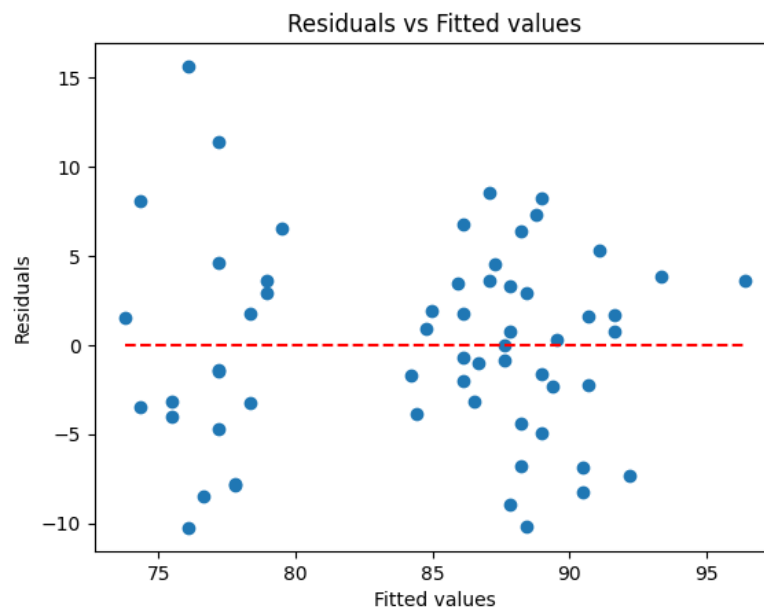
## 2. Rozkład normalny reszt

Stworzyłem model regresji liniowej ('score ~ age + exercise'), aby zbadać czy reszty mają rozkład normalny. P-value w teście Shapiro-Wilka **wyniosło 0.676**, co oznacza że istnieje brak podstaw do odrzucenia  $H_0$ , czyli **reszty modelu są zgodne z rozkładem normalnym**.

## 3. Brak autokorelacji reszt

Wymagamy, by reszty modelu regresji liniowej były nieskorelowane. W tym celu przeprowadzę test Durбина-Watsona, który sprawdza czy występuje autokorelacja reszt w modelu. Przyjmuje się, że wartość statystyki DW w przedziale od 1.5 do 2.5 oznacza, że autokorelacja nie występuje. W naszym przypadku **statystyka DW wyniosła 1.60**, co oznacza że **nie występuje autokorelacja reszt**.

#### 4. Homoskedastyczność



Rysunek 12 Wykres reszt względem dopasowanych wartości

Aby sprawdzić, czy reszty modelu mają stałą wariancję przeprowadzam dodatkowo test Breusch-Pagana o następujących hipotezach:

**$H_0$ : Wariancja reszt jest stała (homoskedastyczność).**

**$H_1$ : Wariancja reszt nie jest stała (heteroskedastyczność).**

Wynik p-value : 0.2558. Oznacza to, że wariancja reszt jest stała.

**Dodatkowo ANCOVA ma dwa własne założenia:**

##### 1. Zmienne niezależne nieskorelowane

Jako że w naszym badaniu występuje tylko jedna zmienna niezależna (exercise), nie istnieje potrzeba sprawdzenia czy zmienne niezależne są skorelowane ze sobą.

##### 2. Jednorodność lub równoległość regresji w obrębie grup

Nachylenie linii regresji(dla zmiennych age i score) powinno być podobne dla każdej z badanych grup(exercise). Innymi słowy badamy, czy występują istotne interakcje między zmienną niezależną(exercise) oraz zmienną towarzyszącą(age). Przeprowadzę ANOVĘ z interakcjami ('score ~ age \* exercise') i otrzymuje następujące wyniki:

	sum_sq	df	F	PR(>F)
exercise	980,9121	2	14,71466	7,93E-06
age	298,4212	1	8,953233	0,004169
age:exercise	13,69964	2	0,205508	0,814867
Residual	1799,88	54		

Tabela 12 Wyniki testu ANOVA

Model ANOVA z interakcjami był nieistotny statystycznie (p-value 0.814), co oznacza że **mamy do czynienia z modelem jednakowych odchyleń**.

## Przeprowadzenie testu ANCOVA

Po sprawdzeniu założeń mogę przystąpić do przeprowadzenia testu ANCOVA. Hipotezy tego testu wyglądają następująco:

**$H_0$ : Grupy nie różnią się istotnie pod względem średniej zmiennej zależnej po uwzględnieniu kowariatu.**

**$H_1$ : Przynajmniej jedna grupa różni się istotnie od pozostałych pod względem średniej zmiennej zależnej po uwzględnieniu kowariatu**

W tym celu użyję funkcji `ancova(dv='score', covar='age', between='exercise')` z pakietu pingouin w Pythonie. Wyniki są następujące:

Source	SS	DF	F	p-unc	np2
exercise	980,9121	2	15,14438	0.000006	0,351016
age	298,4212	1	9,214697	0,003639	0,141298
Residual	1813,58	56			

Tabela 13 Wyniki testu ANCOVA

Wnioski:

- Po uwzględnieniu wpływu zmiennej towarzyszącej 'age' wystąpiły istotne różnice między grupami zmiennej 'exercise'.

## Test post hoc

Aby dowiedzieć się, między którymi grupami wystąpiły istotne różnice przeprowadziłem test post hoc Tukeya:

group1	group2	meandiff	p-adj	lower	upper	reject
high	low	11,835	0	7,2029	16,4671	TRUE
high	moderate	11,225	0	6,5929	15,8571	TRUE
low	moderate	-0,61	0,9462	-5,2421	4,0221	FALSE

Tabela 14 Wyniki testu Tukeya

Istotne różnice wystąpiły między grupami low-high oraz moderate-high. **Możemy więc stwierdzić, że osoby, które najwięcej ćwiczyły miały znacząco mniejszy poziom stresu (znacząco niższy wynik testu), po uwzględnieniu wpływu wieku.**