

Badanie statystyczne wpływu różnych czynników na poziom samopoczucia

Marcin Dziadosz, Alicja Hołowiecka, Matylda Jankowska

Spis treści

| | | |
|----------|--|-----------|
| 1 | Wstęp | 2 |
| 2 | Opis danych | 2 |
| 2.1 | Zbieranie danych, opis badanej grupy | 2 |
| 2.2 | Import i oczyszczanie danych | 2 |
| 2.3 | Zbiorcza tabelka | 3 |
| 3 | Hipotezy badawcze oraz ich weryfikacja | 7 |
| 3.1 | Hipoteza I | 7 |
| 3.2 | Hipoteza II | 10 |
| 3.3 | Hipoteza III | 12 |
| 3.4 | Hipoteza IV | 15 |
| 3.5 | Hipoteza V | 18 |
| 3.6 | Hipoteza VI | 21 |
| 3.7 | Hipoteza VII | 23 |
| 3.8 | Hipoteza VIII | 25 |
| 3.9 | Hipoteza IX | 28 |
| 3.10 | Hipoteza X | 29 |
| 4 | Dane zsumowane | 30 |
| 4.1 | Sen a samopoczucie | 31 |
| 4.2 | Dieta a samopoczucie | 31 |
| 4.3 | Aktywność fizyczna a samopoczucie | 32 |
| 4.4 | Urządzenia elektroniczne a samopoczucie | 32 |
| 4.5 | Kontakty towarzyskie a samopoczucie | 32 |
| 4.6 | Model | 33 |
| 4.7 | Korelacje (wg testu Spearmana) | 35 |
| 4.8 | Bubble chart | 38 |
| 5 | Wnioski | 38 |
| 5.1 | Wnioski z bardziej szczegółowych hipotez | 38 |
| 5.2 | Wnioski dla pięciu głównych czynników | 39 |
| 6 | Bibliografia | 41 |

1 Wstęp

Celem naszego projektu jest analiza statystyczna wpływu różnych czynników na poziom samopoczucia. Zbadamy wpływ na odczuwanie poszczególnych emocji czynników z pięciu grup: dotyczących snu, diety, aktywności fizycznej, kontaktów towarzyskich oraz używania urządzeń elektronicznych. Sprawdzimy, czy średnie samopoczucie różni się pomiędzy płciami oraz pomiędzy różnymi grupami wiekowymi.

2 Opis danych

2.1 Zbieranie danych, opis badanej grupy

Dane zebraliśmy na podstawie ankiety. Pytania dotyczące emocji pochodzą z artykułu “Skale do Pomiaru Nastroju i Sześciu Emocji” autorstwa Bogdana Wojciszke i Wiesława Baryły. Pozostałe pytania są naszego autorstwa. Ankietę stworzyliśmy za pomocą Formularzy Google i rozesyłaliśmy respondentom drogą internetową.

Przebadana przez nas grupa składała się z 240 osób. Spośród nich 48.3% było kobietami, a 51.7% było mężczyznami. Większość osób (trochę ponad połowa) było w wieku co najwyżej 23 lat (w tym aż 30% w wieku 21 lat).

Pierwsze 10 pytań były to zdania dotyczące aktualnego samopoczucia. Należało ocenić w skali od 1 do 5, jak bardzo zgadzasz się z danym twierdzeniem. Po nich następowały 24 uczucia takie jak radość, złość, depresja itp. Należało ocenić w skali od 1 do 7 (od “nigdy” do “zawsze”) jak często odczuwasz daną emocję.

Potem pojawiały się pytania o czynniki, których wpływ na samopoczucie chcemy zbadać. Każdego z pięciu poniższych czynników dotyczyły trzy pytania:

- sen: liczba godzin przespanych w ciągu doby (do wyboru z 5 odpowiedzi), posiadanie wyznaczonych godzin snu (ocena w skali od 1 do 5), jakość snu (ocena w skali od 1 do 5)
- aktywność fizyczna: ilość godzin poświęcana tygodniowo na sport (do wyboru z 5 odpowiedzi), najczęściej wybierany środek transportu (do wyboru z 5 odpowiedzi, w tym możliwość podania własnej odpowiedzi przy opcji “inne”), kondycja fizyczna (ocena w skali od 1 do 5)
- używanie urządzeń elektronicznych: ilość godzin spędzana w ciągu dnia na korzystaniu z urządzeń elektronicznych w celu pracy/nauki (do wyboru z 5 odpowiedzi), ilość godzin spędzana w ciągu dnia na korzystaniu z urządzeń elektronicznych w celu rozrywki (do wyboru z 5 odpowiedzi), poczucie uzależnienia od urządzeń elektronicznych (ocena w skali od 1 do 5)
- dieta: przeważający składnik w diecie (do wyboru z 5 odpowiedzi), znaczenie jakości posiłków (do wyboru z 5 odpowiedzi), regularność posiłków (do wyboru z 5 odpowiedzi)
- kontakty towarzyskie: posiadanie bliskich osób, na których można polegać (do wyboru z 5 odpowiedzi), liczba bliższych znajomych (do wyboru z 5 odpowiedzi), znaczenie kontaktów z innymi ludźmi (ocena w skali od 1 do 5)

Jak widać, wszystkie te czynniki były oceniane w skali od 1 do 5, a czynniki dotyczące samopoczucia bądź w skali od 1 do 7, bądź w skali od 1 do 5. Zbadamy zarówno zależności pomiędzy pojedynczymi czynnikami, jak również pomiędzy zsumowanymi, na przykład Ogólna ocena samopoczucia.

2.2 Import i oczyszczanie danych

Do R zaimportowaliśmy dane z pliku `.xlsx`. Przy tworzeniu ankiety zastrzeżyliśmy, że na każde pytanie trzeba odpowiedzieć, dlatego nie mieliśmy problemu z brakami danych. W kolumnie `wiek` mieliśmy kilka nieprawidłowo wprowadzonych danych, większość osób podawała wiek jako liczbę, ale niektórzy pisali np. “21 lat”. Poprawiliśmy to i zmieniliśmy typ kolumny na numeryczny. Musieliśmy też poprawić kilka odpowiedzi na pytanie o jakość posiłków, ponieważ zmieniliśmy jedną z możliwych odpowiedzi w ankiecie już po rozpoczęciu zbierania danych.

Wszystkim kolumnom zmieniliśmy nazwy, tak aby pozbyć się polskich znaków i odstępów, oraz żeby ułatwić sobie pracę z danymi.

Typ danych w kolumnach, w których możliwych było 5 odpowiedzi na pytanie, zmieniliśmy na **factor**, i uszeregowaliśmy od odpowiedzi, która według nas w danym pytaniu była najgorsza, do najlepszej.

W pytaniach, gdzie należało oceniać czynnik w skali od 1 do 5 lub od 1 do 7, także zmieniliśmy typ na **factor**.

Dodaliśmy także kolumnę z Ogólną oceną samopoczucia, w której zsumowaliśmy odpowiedzi z pytań o emocje - pozytywne uczucia sumowaliśmy z plusem, a negatywne z minusem.

Te działania zostały wykonane w skrypcie `oczyszczanie_danych.R`, który załączymy do niniejszego raportu. Na koniec zapisaliśmy dane do pliku `dane.RData`.

2.3 Zbiorcza tabelka

Pytania dotyczące aktualnego nastroju były oceniane 5-cio stopniową skalą. Wartości 1-5 oznaczały odpowiednio: “1 - nie zgadzam się”, “2 - raczej się nie zgadzam”, “3 - trochę tak, trochę nie”, “4 - raczej się zgadzam”, “5 - zgadzam się”.

| Cecha | Minimum | Mediana | Średnia | Odchylenie standardowe | Maksimum | Czy rozkład różni się istotnie od rozkładu normalnego? |
|----------------------------------|---------|---------|---------|------------------------|----------|--|
| Jestem w złym humorze | 1 | 2 | 1.95 | 1.13 | 5 | tak |
| Czuję się świetnie | 1 | 3 | 3.3 | 1.14 | 5 | tak |
| Jestem w złym nastroju | 1 | 2 | 2 | 1.15 | 5 | tak |
| Czuję się rozluźniony i spokojny | 1 | 3 | 3.32 | 1.15 | 5 | tak |
| Czuję się szaro i beznadziejnie | 1 | 2 | 1.95 | 1.12 | 5 | tak |
| Jestem w dobrym humorze | 1 | 4 | 3.61 | 1.23 | 5 | tak |
| Jestem pogodny | 1 | 4 | 3.52 | 1.21 | 5 | tak |
| Czuję się przygnębiony | 1 | 2 | 2.12 | 1.19 | 5 | tak |
| Moje samopoczucie jest złe | 1 | 2 | 2.04 | 1.16 | 5 | tak |
| Moje samopoczucie jest dobre | 1 | 4 | 3.63 | 1.23 | 5 | tak |

Przy ocenie częstości odczuwania poszczególnych emocji, do wyboru były odpowiedzi oceniane w skali 7-mio stopniowej. Oznaczały odpowiednio: “1 - nigdy”, “2 - bardzo rzadko”, “3 - rzadko”, “4 - czasami”, “5 -

często”, “6 - bardzo często”, “7 - zawsze”.

| Cecha | Minimum | Mediana | Średnia | Odchylenie standardowe | Maksimum | Czy rozkład różni się istotnie od rozkładu normalnego? |
|---------------|---------|---------|---------|------------------------|----------|--|
| Radość | 1 | 5 | 4.8 | 1.04 | 7 | tak |
| Niepokój | 1 | 4 | 3.61 | 1.42 | 7 | tak |
| Strach | 1 | 3 | 2.8 | 1.43 | 7 | tak |
| Miłość | 1 | 5 | 4.5 | 1.82 | 7 | tak |
| Gniew | 1 | 4 | 3.51 | 1.47 | 7 | tak |
| Wesołość | 2 | 5 | 4.98 | 1.11 | 7 | tak |
| Upokorzenie | 1 | 2 | 2.15 | 1.32 | 7 | tak |
| Smutek | 1 | 3 | 3.45 | 1.43 | 7 | tak |
| Wzburzenie | 1 | 3 | 3.34 | 1.41 | 7 | tak |
| Lęk | 1 | 3 | 2.85 | 1.53 | 7 | tak |
| Żal | 1 | 3 | 2.88 | 1.53 | 7 | tak |
| Obawa | 1 | 3 | 3.54 | 1.53 | 7 | tak |
| Złość | 1 | 3 | 3.38 | 1.43 | 7 | tak |
| Załamanie | 1 | 2 | 2.77 | 1.67 | 7 | tak |
| Szczęście | 1 | 5 | 4.8 | 1.35 | 7 | tak |
| Wściekłość | 1 | 3 | 2.73 | 1.53 | 7 | tak |
| Oddanie | 1 | 4 | 4.13 | 1.77 | 7 | tak |
| Zadowolenie | 1 | 5 | 4.77 | 1.30 | 7 | tak |
| Czułość | 1 | 5 | 4.38 | 1.64 | 7 | tak |
| Wstyd | 1 | 2 | 2.45 | 1.33 | 7 | tak |
| Poczucie winy | 1 | 2 | 2.73 | 1.52 | 7 | tak |
| Przywiązanie | 1 | 5 | 4.69 | 1.73 | 7 | tak |
| Depresja | 1 | 2 | 2.35 | 1.66 | 7 | tak |
| Nieszczęście | 1 | 2 | 2.4 | 1.57 | 7 | tak |

Pytania dotyczące snu oceniane były w skali 5-cio stopniowej. Wartości w pytaniu dotyczącym:

- średniej liczby przesypianych godzin oznaczały: “1 - mniej niż 4”, “2 - 4-6h”, “3 - 6-8h”, “4 - 8-10h”, “5 - więcej niż 10h”;
- posiadania wyznaczonych pór snu określały: “1 - posiadanie regularnych pór snu” do “5 - posiadanie nieregularnych pór snu”;
- oceny jakości snu określały: “1 - bardzo zła jakość” do “5 - bardzo dobra jakość”.

| Cecha | Minimum | Mediana | Średnia | Odchylenie standardowe | Maksimum | Czy rozkład różni się istotnie od rozkładu normalnego? |
|--|---------|---------|---------|------------------------|----------|--|
| Liczba godzin przespanych w ciągu doby | 1 | 3 | 2.96 | 0.69 | 5 | tak |
| Posiadanie wyznaczonych pór snu | 1 | 3 | 3.23 | 1.19 | 5 | tak |
| Jakość snu | 1 | 3 | 3.44 | 0.96 | 5 | tak |

Pytania dotyczące sportu oceniane były w skali 5-cio stopniowej. Wartości w pytaniu dotyczącym:

- liczby godzin poświęconych tygodniowo na sport oznaczały: “1 - nie uprawiam żadnego sportu”, “2 - mniej niż 1h”, “3 - 1-2,5h”, “4 - 2,5-4h”, “5 - więcej niż 4h”;
- najczęściej wybieranego środka transportu oznaczały: “1 - samochód”, “2 - komunikacja miejska”, “3 - rower”, “4 - chodzenie pieszo”, “5 - inny”;
- oceny własnej kondycji fizycznej określały: “1 - bardzo zła” do “5 - bardzo dobra”.

| Cecha | Minimum | Mediana | Średnia | Odchylenie standardowe | Maksimum | Czy rozkład różni się istotnie od rozkładu normalnego? |
|---|---------|---------|---------|------------------------|----------|--|
| Godziny poświęcane tygodniowo na sport | 1 | 3 | 2.98 | 1.41 | 5 | tak |
| Najczęściej wybierany środek transportu | 1 | 3 | 3.27 | 1.41 | 5 | tak |
| Ocena kondycji fizycznej | 1 | 3 | 3.22 | 0.84 | 5 | tak |

Pytania dotyczące korzystania z urządzeń elektronicznych oceniane były w skali 5-cio stopniowej. Wartości w pytaniu dotyczącym:

- średniej liczby godzin poświęcanej w ciągu dnia na korzystanie z u.e. w celu pracy bądź nauki oznaczały: “1 - nie używam/rzadko używam komputera do pracy/nauki”, “2 - mniej niż 3”, “3 - 3-5h”, “4 - 5-8h”, “5 - powyżej 5h”;
- średniej liczby godzin poświęconej w ciągu dnia na korzystanie z u.e w celach rozrywkowych oznaczały: “1 - mniej niż 1h”, “2 - 1-2h”, “3 - 2-3h”, “4 - 3-4h”, “5 - więcej niż 4h”;
- częstości występowania poczucia uzależnienia od u.e. oznaczały: “1 - nigdy”, “2 - rzadko”, “3 - czasami”, “4 - często”, “5 - bardzo często”.

| Cecha | Minimum | Mediana | Średnia | Odchylenie standardowe | Maksimum | Czy rozkład różni się istotnie od rozkładu normalnego? |
|---|---------|---------|---------|------------------------|----------|--|
| Godziny spędzone dziennie na pracy na urządzeniach elektronicznych | 1 | 3 | 3.14 | 1.18 | 5 | tak |
| Godziny spędzone dziennie na rozrywce na urządzeniach elektronicznych | 1 | 3 | 2.81 | 1.25 | 5 | tak |
| Poczucie uzależnienia od urządzeń elektronicznych | 1 | 3 | 3.27 | 1.07 | 5 | tak |

Pytania dotyczące diety oceniane były w skali 5-cio stopniowej. Wartości w pytaniu dotyczącym:

- przeważającego składnika diety oznaczały: “1 - fast food”, “2 - żywność mocno przetworzona (mrożonki, gotowe dania)”, “3 - mięso i ciężkostrawne potrawy”, “4 - warzywa i owoce”, “5 - mam zbilansowaną dietę”;
- wagi przywiązywanej do jakości spożywanych posiłków oznaczały: “1 - wszystko mi jedno”, “2 - nie jest to moim priorytetem”, “3 - staram się jeść zdrowo, ale mi nie wychodzi”, “4 - jest to dla mnie dość ważne”, “5 - jestem tym, co jem!”;
- regularności spożywanych posiłków oznaczały: “1 - jem nieregularnie”, “2 - różnie z tym bywa”, “3 - staram się jeść regularnie”, “4 - zazwyczaj jem regularnie”, “5 - mam stałe pory posiłków”.

| Cecha | Minimum | Mediana | Średnia | Odchylenie standardowe | Maksimum | Czy rozkład różni się istotnie od rozkładu normalnego? |
|--------------------------------|---------|---------|---------|------------------------|----------|--|
| Przeważający składnik w diecie | 1 | 5 | 3.92 | 1.31 | 5 | tak |
| Dbanie o jakość posiłków | 1 | 3 | 3.21 | 0.94 | 5 | tak |
| Regularność posiłków | 1 | 2 | 2.41 | 1.06 | 5 | tak |

Pytania dotyczące kontaktów towarzyskich oceniane były w skali 5-cio stopniowej. Wartości w pytaniu dotyczącym:

- posiadania osoby, na której można polegać w trudnej sytuacji oznaczały: “1 - nie mam nikogo takiego”, “2 - z rodziną/znajomymi nie poruszam poważnych tematów”, “3 - nie czuję się komfortowo rozmawiając na trudne tematy z rodziną/znajomymi”, “4 - na ogół mam do kogo zwrócić się z problemem”, “5 - mam osoby, z którymi mogę porozmawiać o wszystkim”;
- liczby osób, z którymi utrzymuje się stały kontakt oznaczały: “1 - brak takiej osoby”, “2 - jedną osobę”, “3 - 2-3 osoby”, “4 - 4-5 osób”, “5 - 5 i więcej osób”;
- znaczenia, jakie mają dla ankietowanego kontakty z innymi ludźmi oznaczały “1 - brak znaczenia” do “5 - bardzo duże znaczenie”.

| Cecha | Minimum | Mediana | Średnia | Odchylenie standardowe | Maksimum | Czy rozkład różni się istotnie od rozkładu normalnego? |
|---|---------|---------|---------|------------------------|----------|--|
| Posiadanie zaufanej osoby | 1 | 5 | 4.35 | 0.90 | 5 | tak |
| Liczba osób, z którymi utrzymuje się bliski kontakt | 1 | 4 | 3.9 | 1.02 | 5 | tak |
| Jakie znaczenie mają dla Ciebie kontakty z innymi ludźmi? | 1 | 4 | 3.91 | 0.96 | 5 | tak |
| Ogólna ocena samopoczucia | -79 | -0.5 | -2.61 | 28.63 | 55 | tak |

Ogólna ocena samopoczucia została uzyskana poprzez zsumowanie wartości uczuć pozytywnych ze znakiem dodatnim oraz negatywnych z ujemnym.

3 Hipotezy badawcze oraz ich weryfikacja

H_0 : nie istnieje zależność pomiędzy zmiennymi.

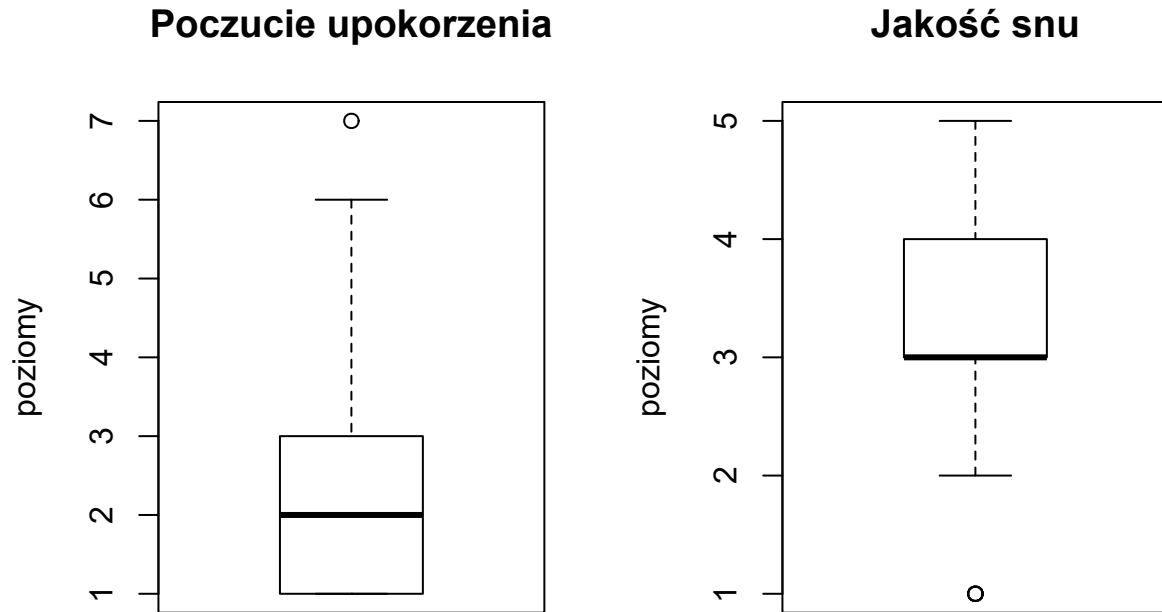
H_1 : istnieje zależność pomiędzy zmiennymi.

Hipotezy badane są na poziomie istotności $\alpha = 0.05$.

3.1 Hipoteza I

Badanie, czy istnieje zależność pomiędzy poczuciem upokorzenia a jakością snu.

3.1.1 Wykresy boxplot



| Cecha | Minimum | Mediana | Średnia | Odchylenie standardowe | Maksimum | Czy rozkład różni się istotnie od rozkładu normalnego? |
|-------------|---------|---------|---------|------------------------|----------|--|
| Upokorzenie | 1 | 2 | 2.15 | 1.32 | 7 | tak |
| Jakość snu | 1 | 3 | 3.44 | 0.96 | 5 | tak |

Upokorzenie było oceniane w skali od 1 do 7 (odczuwanie od “nigdy” do “zawsze”). Jak widać z wykresu boxplot, większość respondentów nie odczuwa upokorzenia nigdy, lub odczuwa je bardzo rzadko lub rzadko, gdzie “bardzo rzadko” jest medianą. Pojawił się co najmniej jeden element odstający - osoby odczuwające upokorzenie zawsze.

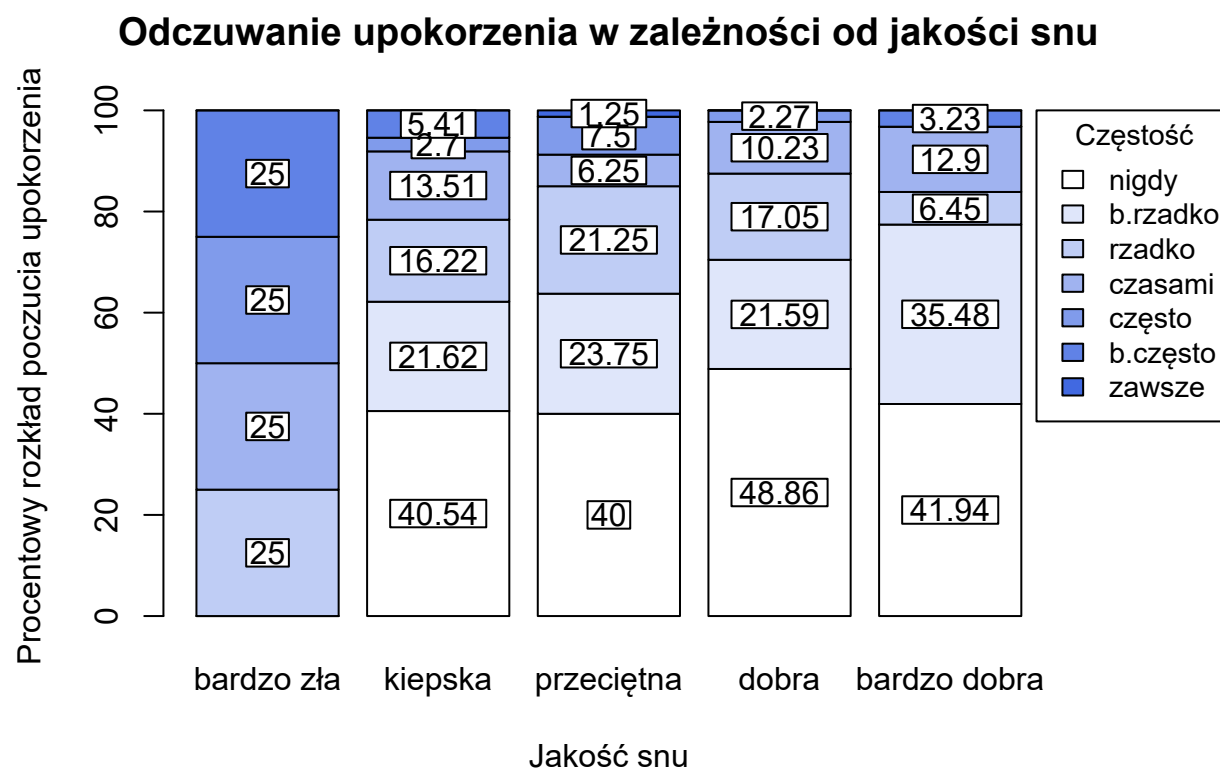
Jakość snu była oceniona w skali od 1 do 5 (od najgorszej do najlepszej). Jakość snu była oceniana głównie na 3 lub 4, gdzie 3 jest medianą. Pojawiały się elementy odstające - osoby, które oceniły swoją jakość snu na 1.

3.1.2 Tabela kontyngencji

| Upokorzenie | Jakość snu | | | | |
|---------------|------------|------|------|------|------|
| | 1 | 2 | 3 | 4 | 5 |
| nigdy | 0.0 | 40.5 | 40.0 | 48.9 | 41.9 |
| bardzo rzadko | 0.0 | 21.6 | 23.8 | 21.6 | 35.5 |
| rzadko | 25.0 | 16.2 | 21.2 | 17.0 | 6.5 |
| czasami | 25.0 | 13.5 | 6.2 | 10.2 | 12.9 |
| często | 25.0 | 2.7 | 7.5 | 2.3 | 0.0 |
| bardzo często | 25.0 | 5.4 | 0.0 | 0.0 | 3.2 |
| zawsze | 0.0 | 0.0 | 1.2 | 0.0 | 0.0 |

3.1.3 Wykresy słupkowe

Poniżej przedstawiony wykres zawiera wartości procentowe, nie liczebności dla poszczególnych odpowiedzi; podobnie jak w przypadku wykresów tego typu przedstawionych dla kolejnych hipotez.



3.1.4 Wyniki testów

| test rho-Spearmana | | test tau-Kendalla | |
|------------------------|---------|------------------------|---------|
| współczynnik korelacji | p-value | współczynnik korelacji | p-value |
| -0.125 | 0.054 | -0.105 | 0.053 |

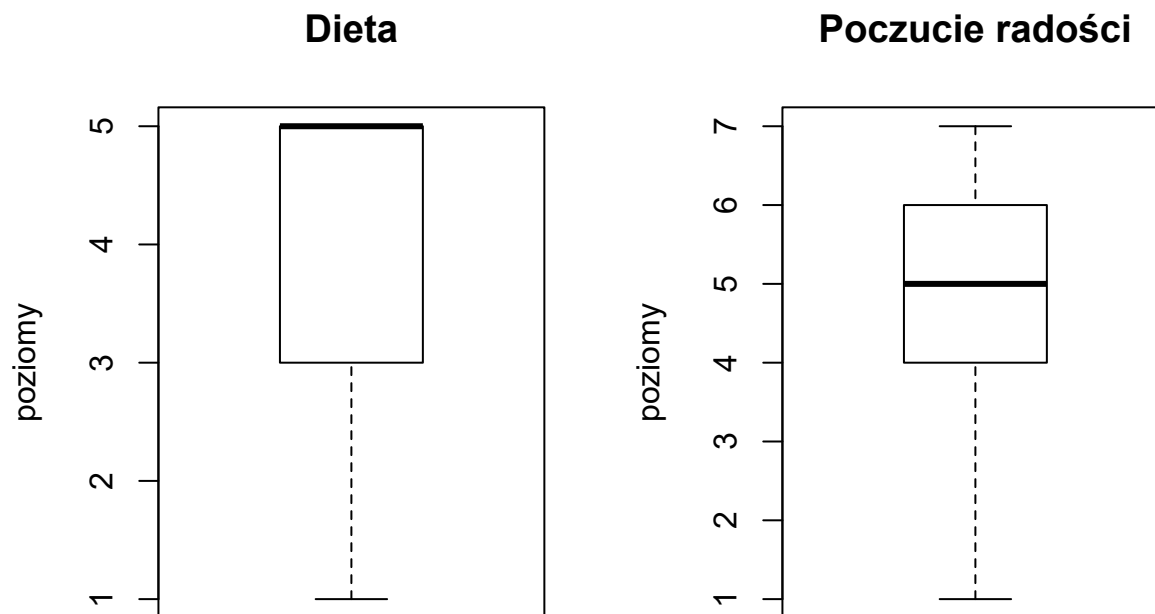
P-value dla testu rho-Spearmana wynosi $p = 0.054$, natomiast dla testu tau-Kendalla wynosi $p = 0.053$. Nie ma podstaw do odrzucenia hipotezy H_0 , a więc nie istnieje istotna zależność pomiędzy poczuciem upokorzenia

a jakością snu. Współczynniki korelacji dla testów **rho-Spearmana** oraz **tau-Kendalla** wynoszą odpowiednio -0.12478 oraz -0.10524, ale na poziomie istotności $\alpha = 0.05$ nie są istotne statystycznie.

3.2 Hipoteza II

Badanie, czy istnieje zależność pomiędzy dietą a poczuciem radości.

3.2.1 Wykresy boxplot



| Cecha | Minimum | Mediana | Średnia | Odchylenie standardowe | Maksimum | Czy rozkład różni się istotnie od rozkładu normalnego? |
|--------------------------------|---------|---------|---------|------------------------|----------|--|
| Przeważający składnik w diecie | 1 | 5 | 3.92 | 1.31 | 5 | tak |
| Radość | 1 | 5 | 4.8 | 1.04 | 7 | tak |

Dieta (a właściwie jej przeważający składnik) była oceniana w skali od 1 do 5, z następująco uszeregowanymi wartościami: - 1 - fast food

- 2 - żywność mrożona, produkty mocno przetworzone
- 3 - mięso i potrawy ciężkostrawne
- 4 - dieta wegetariańska

- 5 - dieta zbilansowana

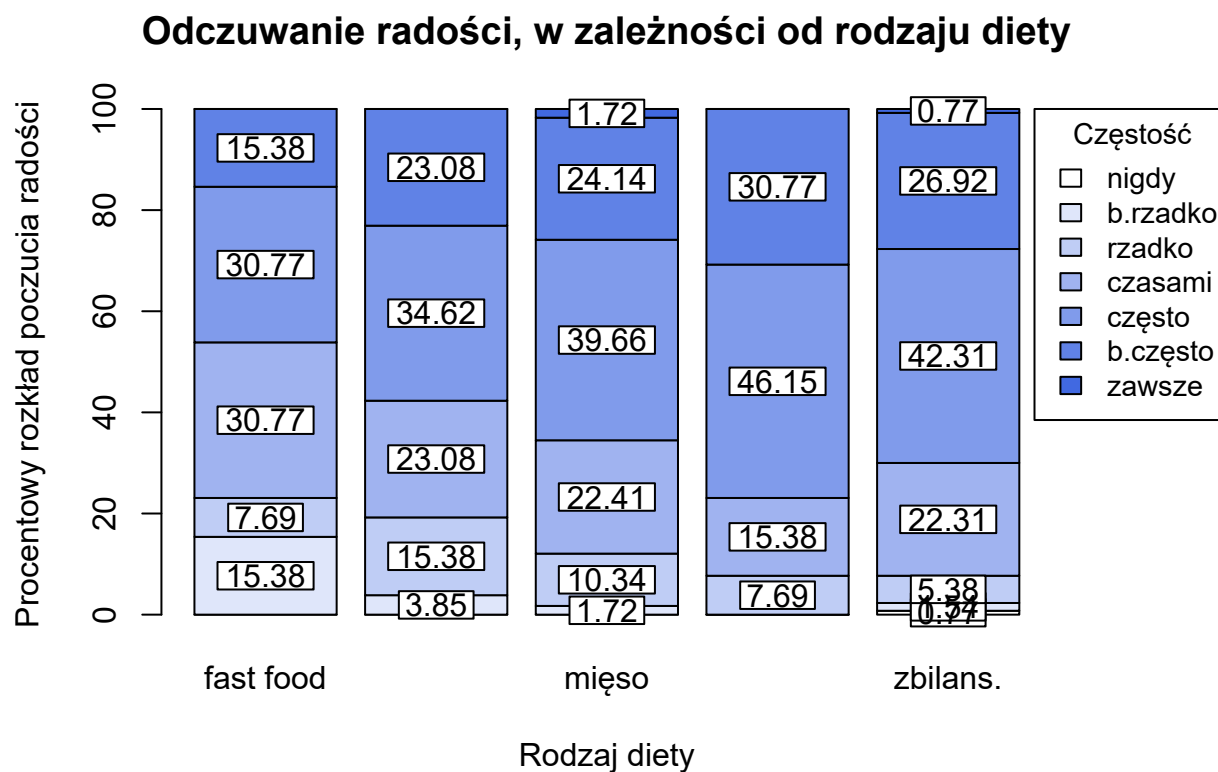
Większość ocen wahała się od 3 do 5, gdzie 5 było medianą.

Odczuwanie radości było oceniane w skali od 1 do 7 (“nigdy” do “zawsze”). Oceny wahały się między 4 a 6, gdzie 5 jest medianą.

3.2.2 Tabela kontyngencji

| Poczucie radości | Jakość diety | | | | |
|------------------|--------------|---|-------------------------------|-----------------|------------------------|
| | fast food | żywność mocno przetworzona (mrożonki, gotowe dania) | mięso i ciężkostrawne potrawy | warzywa i owoce | mam zbilansowaną dietę |
| nigdy | 0.00 | 0.00 | 0.00 | 0.00 | 0.77 |
| bardzo rzadko | 15.38 | 3.85 | 1.72 | 0.00 | 1.54 |
| rzadko | 7.69 | 15.38 | 10.34 | 7.69 | 5.38 |
| czasami | 30.77 | 23.08 | 22.41 | 15.38 | 22.31 |
| często | 30.77 | 34.62 | 39.66 | 46.15 | 42.31 |
| bardzo często | 15.38 | 23.08 | 24.14 | 30.77 | 26.92 |
| zawsze | 0.00 | 0.00 | 1.72 | 0.00 | 0.77 |

3.2.3 Wykresy słupkowe



3.2.4 Wyniki testów

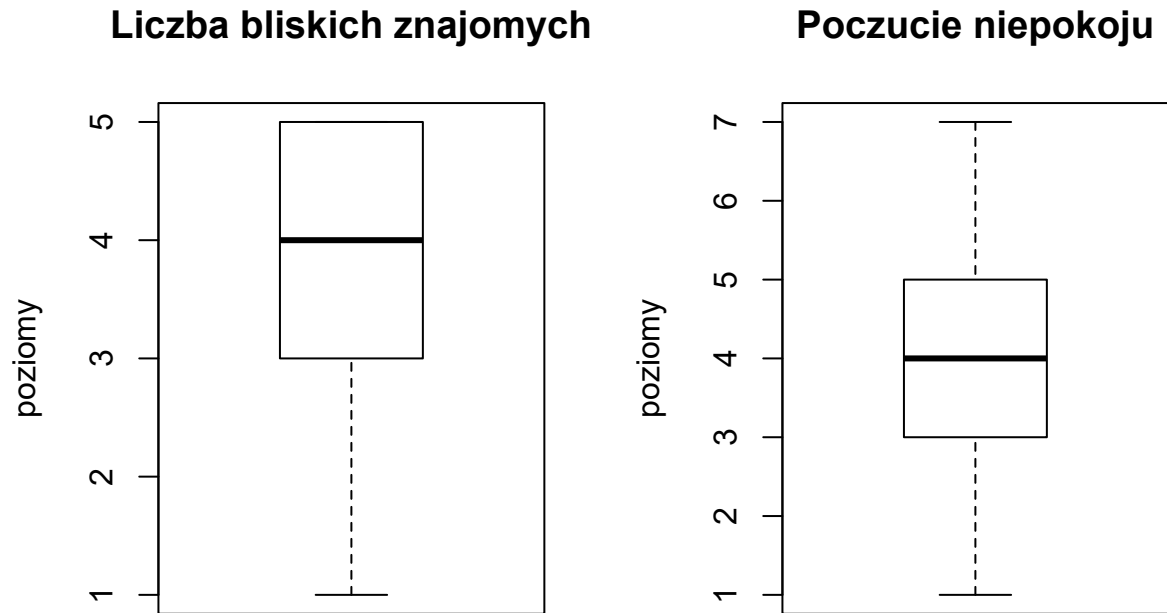
| test rho-Spearmana | | test tau-Kendalla | |
|------------------------|---------|------------------------|---------|
| współczynnik korelacji | p-value | współczynnik korelacji | p-value |
| 0.112 | 0.084 | 0.096 | 0.084 |

P-value dla testu rho-Spearmana wynosi $p = 0.084$, natomiast dla testu tau-Kendalla wynosi $p = 0.084$. Nie ma podstaw do odrzucenia H_0 . Nie istnieje istotna zależność pomiędzy dietą a poczuciem radości.

3.3 Hipoteza III

Badanie, czy istnieje zależność pomiędzy liczbą bliskich znajomych a odczuwaniem niepokoju.

3.3.1 Wykresy boxplot



| Cecha | Minimum | Mediana | Średnia | Odchylenie standardowe | Maksimum | Czy rozkład różni się istotnie od rozkładu normalnego? |
|---|---------|---------|---------|------------------------|----------|--|
| Liczba osób, z którymi utrzymuje się bliski kontakt | 1 | 4 | 3.9 | 1.02 | 5 | tak |
| Niepokój | 1 | 4 | 3.61 | 1.42 | 7 | tak |

Liczba bliskich znajomych była oceniana od 1 do 5 z kolejnymi poziomami:

- 1 - brak bliskich znajomych
- 2 - 1 do 2 bliskich osób
- 3 - 2-3 takich osób
- 4 - 4-5 osób
- 5 - więcej niż 5 bliskich znajomych

Najwięcej ocen było pomiędzy 3 a 5, gdzie 4 jest medianą.

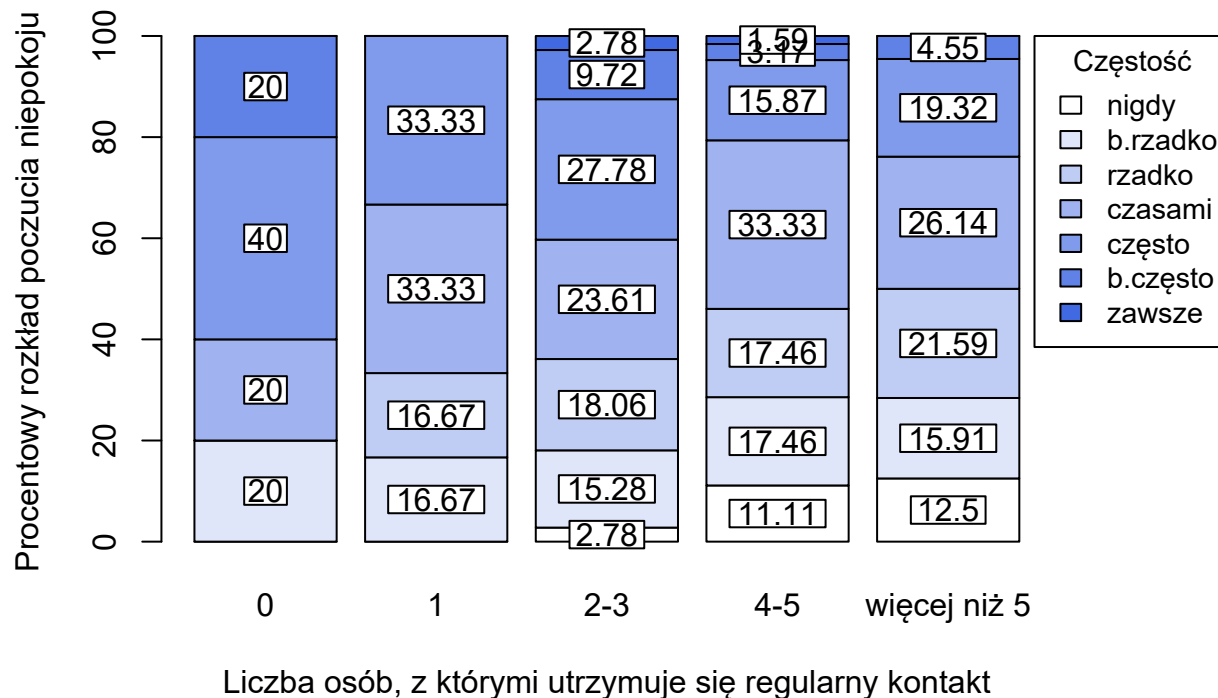
Poczucie niepokoju było oceniane w skali 1 do 7 ("nigdy" do "zawsze"). Najwięcej ocen było pomiędzy 3 a 5, gdzie 4 jest medianą.

3.3.2 Tabela kontyngencji

| Niepokój | Liczba bliskich znajomych | | | | |
|---------------|---------------------------|------|-------|-------|------------|
| | 0 | 1 | 2 - 3 | 4 - 5 | 5 i więcej |
| nigdy | 0.0 | 0.0 | 2.8 | 11.1 | 12.5 |
| bardzo rzadko | 20.0 | 16.7 | 15.3 | 17.5 | 15.9 |
| rzadko | 0.0 | 16.7 | 18.1 | 17.5 | 21.6 |
| czasami | 20.0 | 33.3 | 23.6 | 33.3 | 26.1 |
| często | 40.0 | 33.3 | 27.8 | 15.9 | 19.3 |
| bardzo często | 20.0 | 0.0 | 9.7 | 3.2 | 4.5 |
| zawsze | 0.0 | 0.0 | 2.8 | 1.6 | 0.0 |

3.3.3 Wykresy słupkowe

Odczuwanie niepokoju, w zależności od liczby bliskich znajomych



3.3.4 Wyniki testów

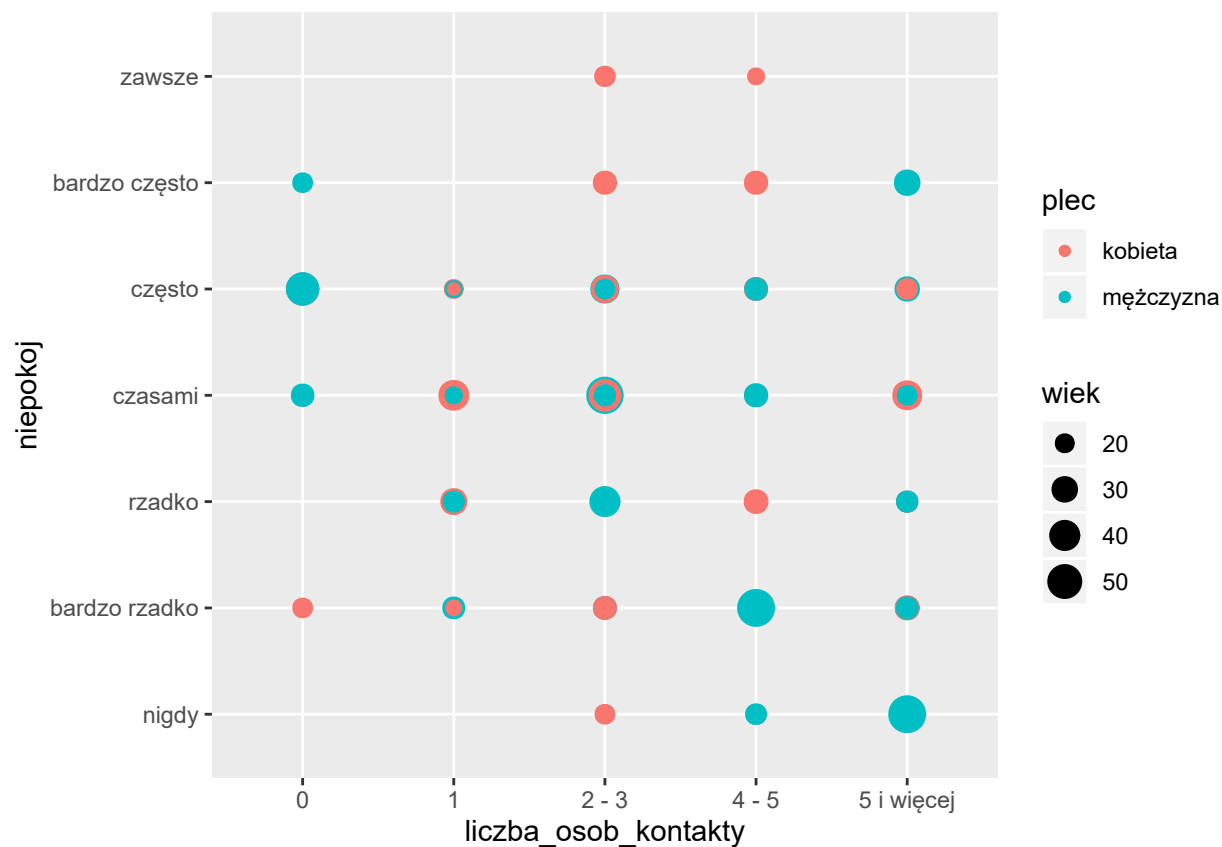
| test rho-Spearmana | | test tau-Kendalla | |
|------------------------|---------|------------------------|---------|
| współczynnik korelacji | p-value | współczynnik korelacji | p-value |
| -0.179 | 0.006 | -0.147 | 0.006 |

P-value dla obu testów wynosi około $p = 0.006$. Należy odrzucić hipotezę H_0 , na rzecz hipotezy alternatywnej H_1 . Istnieje zależność pomiędzy liczbą osób, z którymi utrzymuje się regularny kontakt a odczuwaniem niepokoju.

Współczynniki korelacji dla testów rho-Spearmana i tau-Kendalla wynoszą odpowiednio -0.179 oraz -0.147. Interpretując te współczynniki oraz dane z wykresów, można stwierdzić, że wraz ze wzrostem liczby osób, z którymi utrzymuje się regularny kontakt, spada częstość odczuwania niepokoju.

3.3.5 Bubble!

Ponieważ z testów wynika, że istnieje zależność pomiędzy tymi dwoma czynnikami, można również zobrazować ją wykresem bubble chart, uwzględniając podział na wiek i płeć.



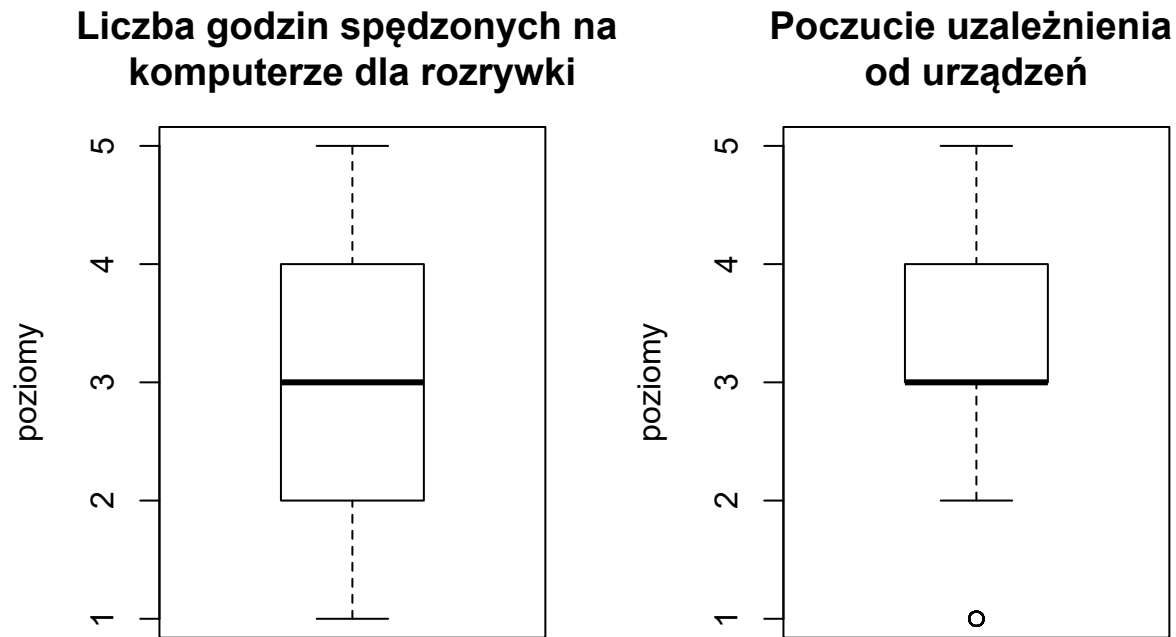
Z powyższego wykresu można wysnuć wniosek, że u mężczyzn brak posiadania stałych kontaktów z innymi ludźmi znacząco wpływa na częstość odczuwania niepokoju; gdy liczba kontaktów wzrasta, ta częstotliwość spada. Również wiek ma tutaj spore znaczenie.

U kobiet - szczególnie młodszych, niezależnie od liczby bliskich znajomych, odczuwanie niepokoju jest częstszym zjawiskiem (dlatego SWOJĄ kobietę trzeba dużo tulić).

3.4 Hipoteza IV

Badanie, czy istnieje zależność pomiędzy liczbą godzin przeznaczonych na korzystanie z urządzeń elektronicznych a poczuciem uzależnienia od urządzeń.

3.4.1 Wykresy boxplot



| Cecha | Minimum | Mediana | Średnia | Odchylenie standardowe | Maksimum | Czy rozkład różni się istotnie od rozkładu normalnego? |
|-------------|---------|---------|---------|------------------------|----------|--|
| Upokorzenie | 1 | 2 | 2.15 | 1.32 | 7 | tak |
| Jakość snu | 1 | 3 | 3.44 | 0.96 | 5 | tak |

Liczba godzin spędzonych na urządzeniach w celu rozrywki była oceniana w skali od 1 do 5 gdzie:

- 1 - mniej niż 1h
- 2 - od 1 do 2h
- 3 - od 2 do 3h
- 4 - od 3 do 4h
- 5 - powyżej 4h

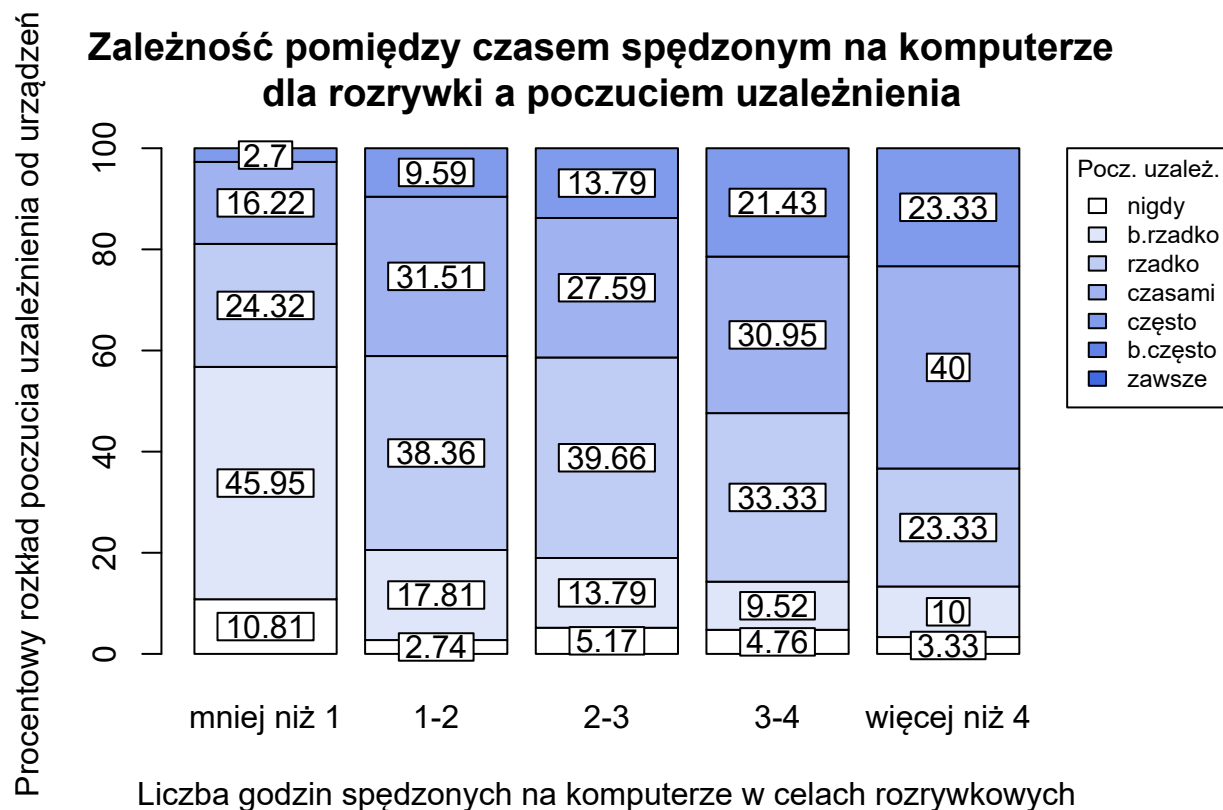
Najwięcej ocen pojawiało się między 2 a 4, gdzie 3 to mediana.

Poczucie uzależnienia było oceniane od 1 do 5 (“nigdy” do “zawsze”). Najwięcej ocen pojawiało się od 3 do 4 gdzie 3 to mediana. Pojawiały się elementy odstające - osoby, które nigdy nie czują się uzależnione.

3.4.2 Tabela kontyngencji

| Poczucie uzależnienia | Godziny spędzone na używaniu urządzeń elektronicznych | | | | |
|-----------------------|---|-------|-------|-------|--------------|
| | mniej niż 1 | 1 - 2 | 2 - 3 | 3 - 4 | więcej niż 4 |
| nigdy | 10.8 | 2.7 | 5.2 | 4.8 | 3.3 |
| rzadko | 45.9 | 17.8 | 13.8 | 9.5 | 10.0 |
| czasami | 24.3 | 38.4 | 39.7 | 33.3 | 23.3 |
| często | 16.2 | 31.5 | 27.6 | 31.0 | 40.0 |
| bardzo często | 2.7 | 9.6 | 13.8 | 21.4 | 23.3 |

3.4.3 Wykresy słupkowe



3.4.4 Wyniki testów

| test rho-Spearmana | | test tau-Kendalla | |
|------------------------|---------|------------------------|---------|
| współczynnik korelacji | p-value | współczynnik korelacji | p-value |
| 0.291 | <0.001 | 0.243 | <0.001 |

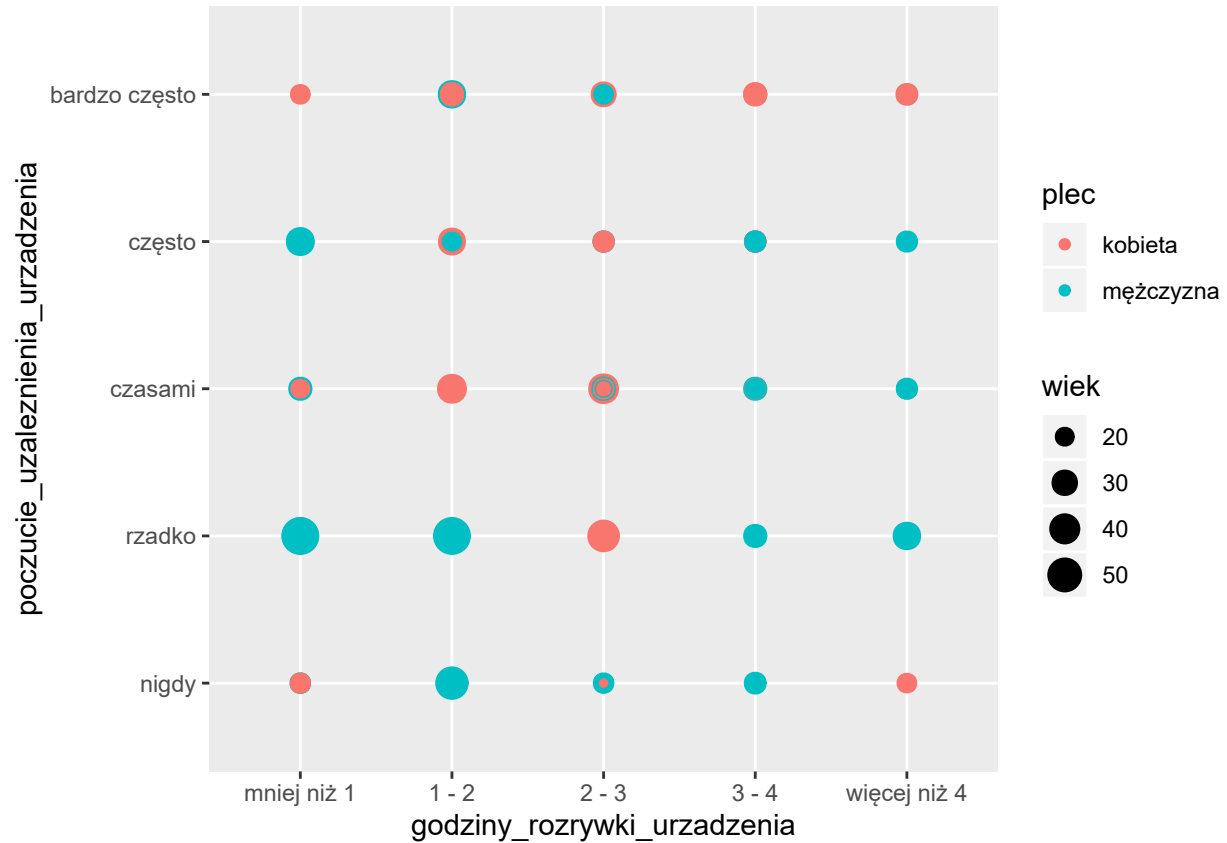
P-value dla obu testów są nieistotnie różne od zera. Należy odrzucić hipotezę H_0 , na rzecz hipotezy alternatywnej H_1 . Istnieje zależność pomiędzy liczbą godzin spędzonych na używaniu urządzeń elektronicznych w celach rozrywkowych a poczuciem uzależnienia od nich.

Współczynniki korelacji dla testów rho-Spearmana i tau-Kendalla wynoszą odpowiednio 0.291 oraz 0.243. Wniosek: wraz ze wzrostem liczby godzin, spędzonych na korzystaniu dla rozrywki z urządzeń elektronicznych,

wzrasta znacznie poczucie uzależnienia od tych urządzeń.

3.4.5 Bubble chart

Poniżej przedstawiony jest wykres **bubble chart** badanej zależności, uwzględniający podział na wiek i płeć.



Poczucie uzależnienia od urządzeń u mężczyzn wzrasta wraz ze wzrostem liczby godzin, które poświęca się na korzystanie z nich dla rozrywki.

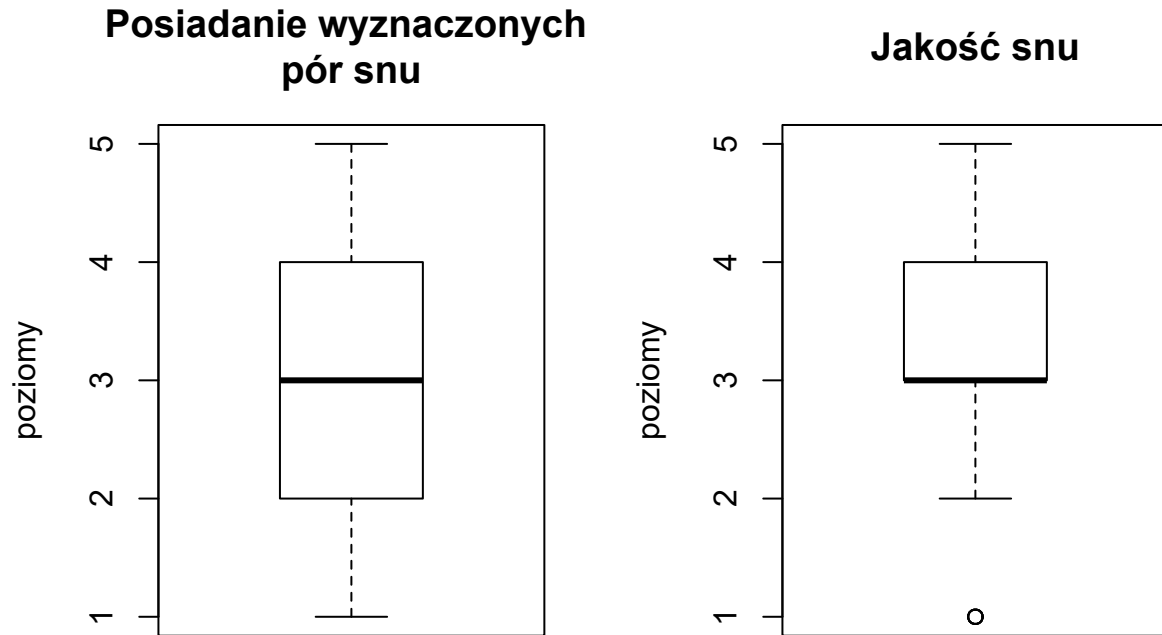
Kobiety, niezależnie od liczby godzin, zdają się mieć to poczucie częściej.

Istnieją również przypadki osób, które nie czują się uzależnione, pomimo bardzo dużej liczby godzin, które spędzają na urządzeniach elektronicznych.

3.5 Hipoteza V

Badanie, czy istnieje zależność pomiędzy posiadaniem wyznaczonych pór snu a jego jakością.

3.5.1 Wykresy boxplot



| Cecha | Minimum | Mediana | Średnia | Odchylenie standardowe | Maksimum | Czy rozkład różni się istotnie od rozkładu normalnego? |
|---------------------------------|---------|---------|---------|------------------------|----------|--|
| Posiadanie wyznaczonych pór snu | 1 | 3 | 3.23 | 1.19 | 5 | tak |
| Jakość snu | 1 | 3 | 3.44 | 0.96 | 5 | tak |

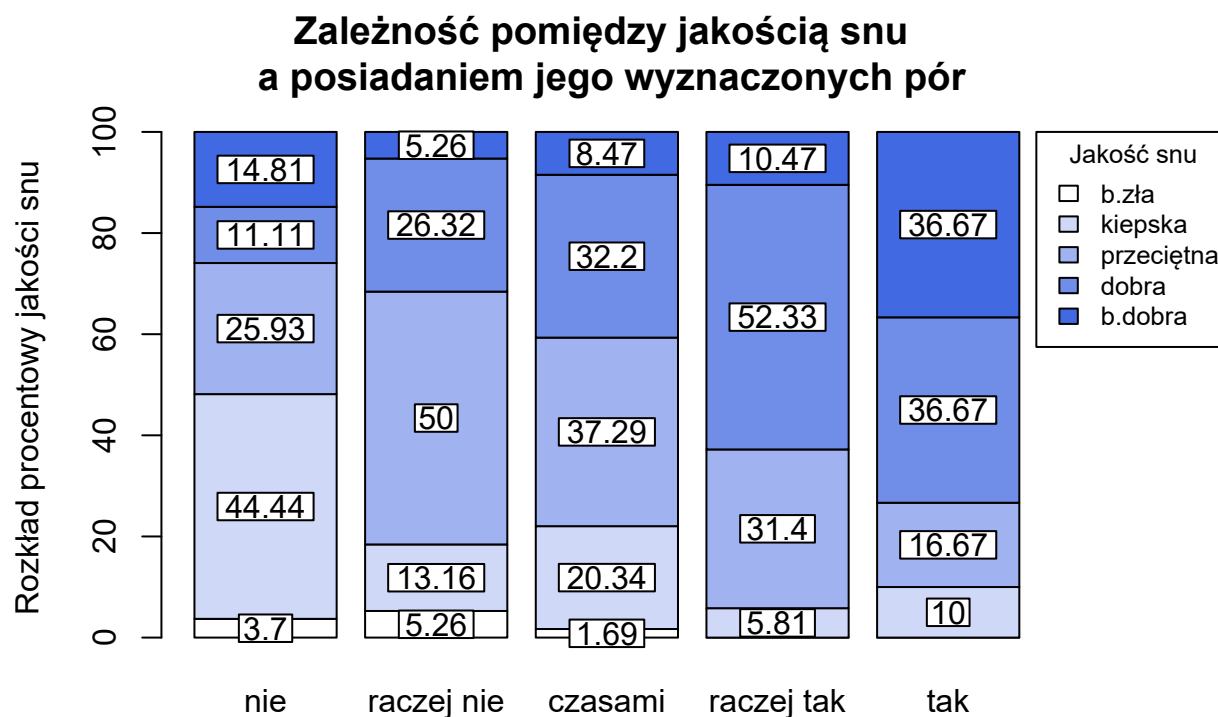
Posiadanie wyznaczonych pór snu było oceniane od 1 do 5 (“nie” do “tak”). Najwięcej ocen było od 2 do 5, gdzie 3 to mediana.

Jakość snu była oceniana w skali od 1 do 5 (od najgorszej do najlepszej). Najwięcej ocen było od 3 do 4, gdzie 3 to mediana. Pojawiały się elementy odstające - osoby, które swoją jakość snu oceniły na 1.

3.5.2 Tabela kontyngencji

| Jakość snu | Posiadanie wyznaczonych pór snu | | | | |
|------------|---------------------------------|------|------|------|------|
| | 1 | 2 | 3 | 4 | 5 |
| 1 | 3.7 | 5.3 | 1.7 | 0.0 | 0.0 |
| 2 | 44.4 | 13.2 | 20.3 | 5.8 | 10.0 |
| 3 | 25.9 | 50.0 | 37.3 | 31.4 | 16.7 |
| 4 | 11.1 | 26.3 | 32.2 | 52.3 | 36.7 |
| 5 | 14.8 | 5.3 | 8.5 | 10.5 | 36.7 |

3.5.3 Wykresy słupkowe



3.5.4 Wyniki testów

| test rho-Spearmana | | test tau-Kendalla | |
|------------------------|---------|------------------------|---------|
| współczynnik korelacji | p-value | współczynnik korelacji | p-value |
| 0.354 | <0.001 | 0.306 | <0.001 |

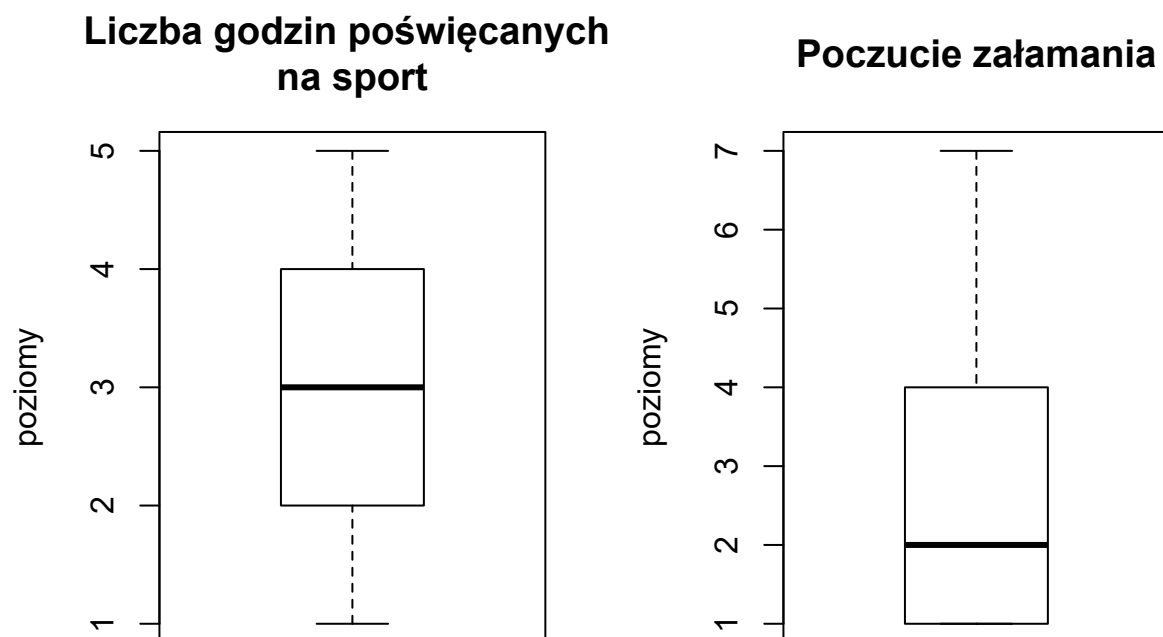
P-value dla obu testów są nieistotnie różne od zera. Należy odrzucić hipotezę H_0 , na rzecz hipotezy alternatywnej H_1 . Istnieje zależność pomiędzy posiadaniem wyznaczonych pór snu a jego jakością.

Współczynniki korelacji dla testów rho-Spearmana i tau-Kendalla są dość duże i wynoszą odpowiednio 0.354 oraz 0.306. Im bardziej regularne są pory snu, tym lepsza jest jego jakość.

3.6 Hipoteza VI

Badanie, czy istnieje zależność pomiędzy liczbą godzin przeznaczonych na uprawianie sportu a odczuwaniem załamania.

3.6.1 Wykresy boxplot



| Cecha | Minimum | Mediana | Średnia | Odchylenie standardowe | Maksimum | Czy rozkład różni się istotnie od rozkładu normalnego? |
|--|---------|---------|---------|------------------------|----------|--|
| Godziny poświęcane tygodniowo na sport | 1 | 3 | 2.98 | 1.41 | 5 | tak |
| Załamanie | 1 | 2 | 2.77 | 1.67 | 7 | tak |

Liczba godzin poświęcanych tygodniowo na sport była oceniona w skali od 1 do 5. Liczby godzin poświęconych tygodniowo na sport oznaczały:

- 1 - nie uprawiam żadnego sportu
- 2 - mniej niż 1h
- 3 - 1-2,5h
- 4 - 2,5-4h

- 5 - więcej niż 4h

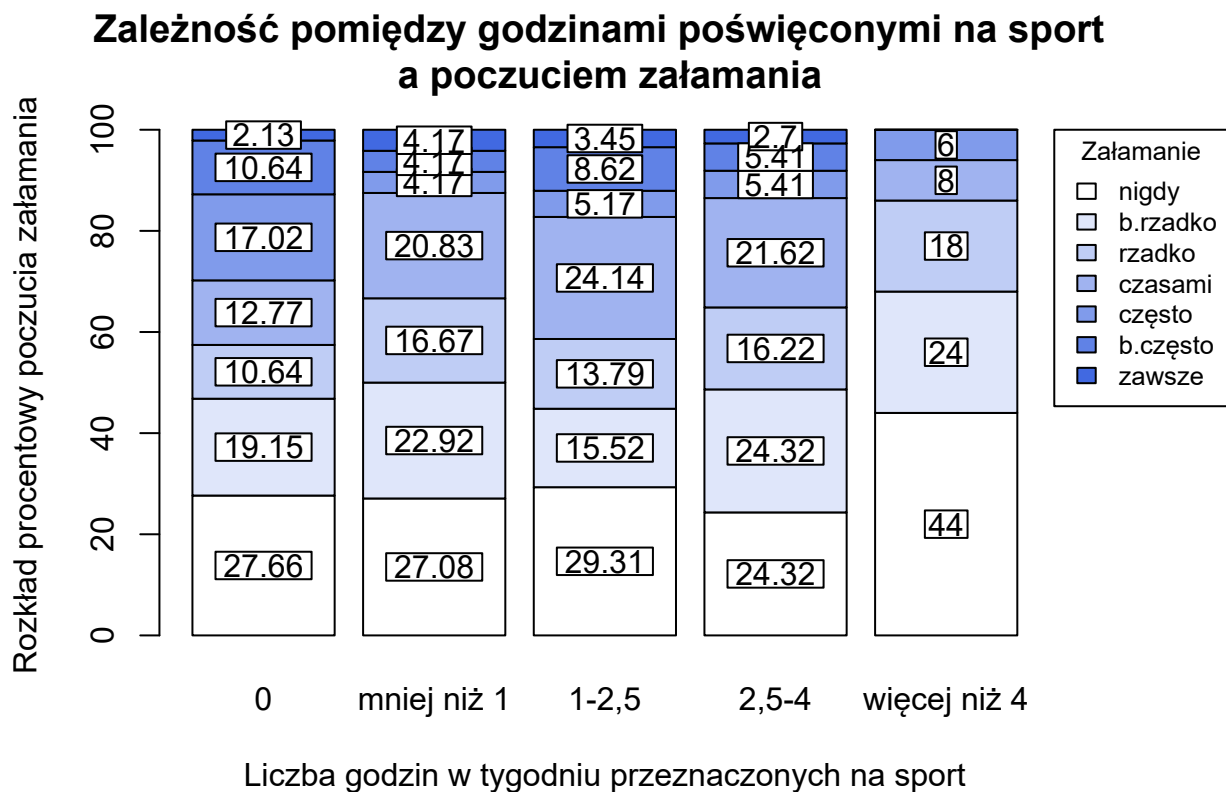
Najwięcej ocen pojawiało się od 2 do 4, gdzie 3 to mediana.

Poczucie załamania było oceniane w skali od 1 do 7 (nigdy do zawsze). Najwięcej ocen pojawiało się od 1 do 4, gdzie 2 to mediana.

3.6.2 Tabela kontyngencji

| Załamanie | Liczba godzin przeznaczonych na sport | | | | |
|---------------|---------------------------------------|-------------|---------|---------|--------------|
| | nie uprawiam żadnego sportu | mniej niż 1 | 1 - 2,5 | 2,5 - 4 | więcej niż 4 |
| nigdy | 27.7 | 27.1 | 29.3 | 24.3 | 44.0 |
| bardzo rzadko | 19.1 | 22.9 | 15.5 | 24.3 | 24.0 |
| rzadko | 10.6 | 16.7 | 13.8 | 16.2 | 18.0 |
| czasami | 12.8 | 20.8 | 24.1 | 21.6 | 8.0 |
| często | 17.0 | 4.2 | 5.2 | 5.4 | 6.0 |
| bardzo często | 10.6 | 4.2 | 8.6 | 5.4 | 0.0 |
| zawsze | 2.1 | 4.2 | 3.4 | 2.7 | 0.0 |

3.6.3 Wykresy słupkowe



3.6.4 Wyniki testów

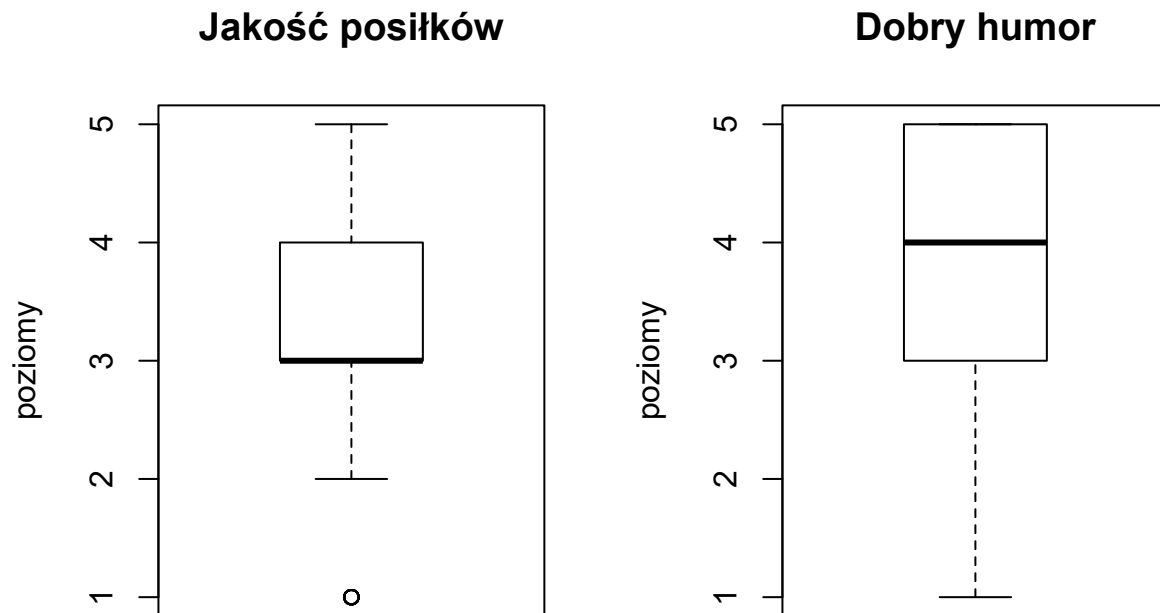
| test rho-Spearmana | | test tau-Kendalla | |
|------------------------|---------|------------------------|---------|
| współczynnik korelacji | p-value | współczynnik korelacji | p-value |
| -0.163 | 0.011 | -0.132 | 0.011 |

P-value dla testu **rho-Spearmana** wynosi 0.011, natomiast dla testu **tau-Kendalla** wynosi 0.011. Należy odrzucić hipotezę H_0 , na rzecz hipotezy alternatywnej H_1 . Istnieje zależność pomiędzy liczbą godzin przeznaczonych na uprawianie sportu a odczuwaniem załamania.

Współczynniki korelacji dla testów **rho-Spearmana** i **tau-Kendalla** wynoszą odpowiednio -0.163 oraz -0.132. Im więcej godzin poświęca się tygodniowo na sport, tym rzadziej odczuwa się załamanie.

3.7 Hipoteza VII

Badanie, czy istnieje zależność pomiędzy jakością posiłków a dobrym humorem.



| Cecha | Minimum | Mediana | Średnia | Odchylenie standardowe | Maksimum | Czy rozkład różni się istotnie od rozkładu normalnego? |
|--------------------------|---------|---------|---------|------------------------|----------|--|
| Jestem w dobrym humorze | 1 | 4 | 3.61 | 1.23 | 5 | tak |
| Dbanie o jakość posiłków | 1 | 3 | 3.21 | 0.94 | 5 | tak |

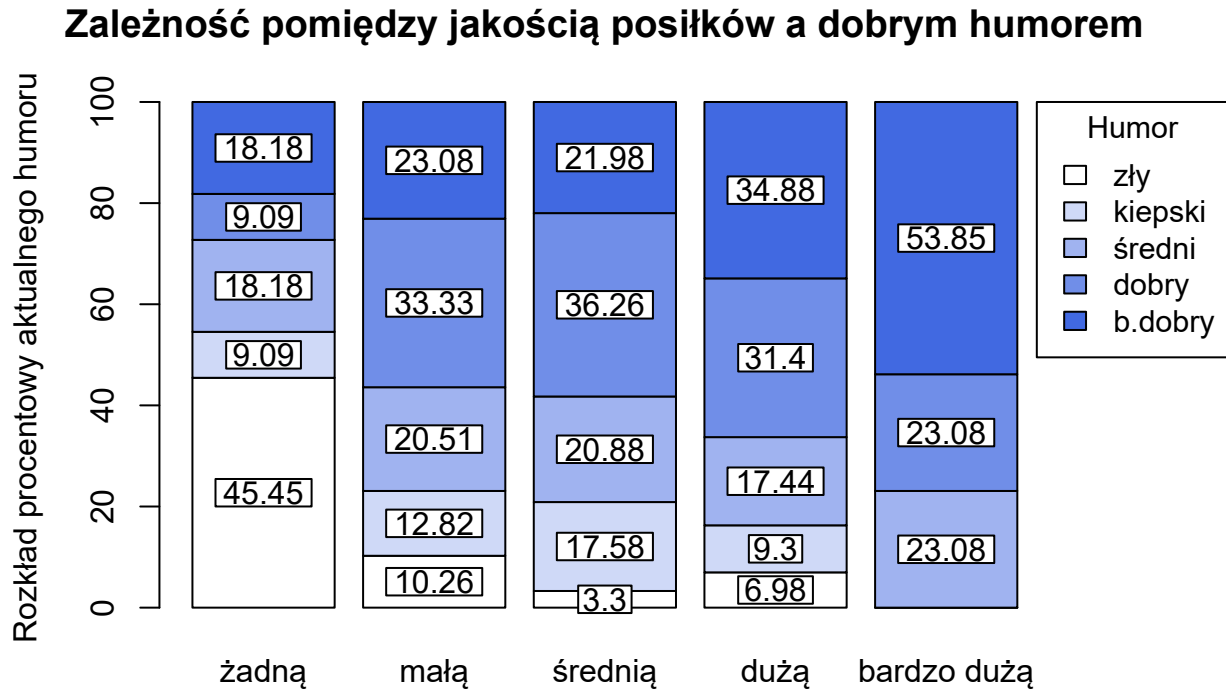
Jakość posiłków była oceniana w skali od 1 do 5 (od najgorszej do najlepszej). Najwięcej ocen pojawiało się między 3 a 4, gdzie 3 to mediana. Osoby, które oceniły jakość swoich posiłków na 1, są na wykresie zaznaczone jako element odstający.

Dobry humor był oceniany w skali od 1 do 5 (najrzadziej do najczęściej odczuwany). Najwięcej ocen jest pomiędzy 3 a 5, gdzie 4 to mediana.

3.7.1 Tabela kontyngencji

| Dobry humor | Jakość posiłków | | | | |
|------------------------|-------------------|------------------------------|---|-----------------------------|---------------------|
| | wszystko mi jedno | nie jest to moim priorytetem | staram się jeść zdrowo, ale mi nie wychodzi | jest to dla mnie dość ważne | jestem tym, co jem! |
| nie zgadzam się | 45.5 | 10.3 | 3.3 | 7.0 | 0.0 |
| raczej się nie zgadzam | 9.1 | 12.8 | 17.6 | 9.3 | 0.0 |
| trochę tak, trochę nie | 18.2 | 20.5 | 20.9 | 17.4 | 23.1 |
| raczej się zgadzam | 9.1 | 33.3 | 36.3 | 31.4 | 23.1 |
| zgadzam się | 18.2 | 23.1 | 22.0 | 34.9 | 53.8 |

3.7.2 Wykresy słupkowe



3.7.3 Wyniki testów

| test rho-Spearmana | | test tau-Kendalla | |
|------------------------|---------|------------------------|---------|
| współczynnik korelacji | p-value | współczynnik korelacji | p-value |
| 0.204 | 0.002 | 0.174 | 0.001 |

P-value dla testu rho-Spearmana wynosi $p = 0.002$, natomiast dla testu tau-Kendalla wynosi $p = 0.001$. Należy odrzucić hipotezę H_0 , na rzecz hipotezy alternatywnej H_1 . Istnieje zależność pomiędzy jakością posiłków a dobrym humorem.

Zarówno współczynniki korelacji dla testów rho-Spearmana i tau-Kendalla wynoszące odpowiednio 0.204 oraz 0.174, jak i wykresy słupkowe jasno wskazują na istotną zależność pomiędzy badanymi zmiennymi. Im większe znaczenie ma jakość spożywanych posiłków, tym lepszy humor się odczuwa.

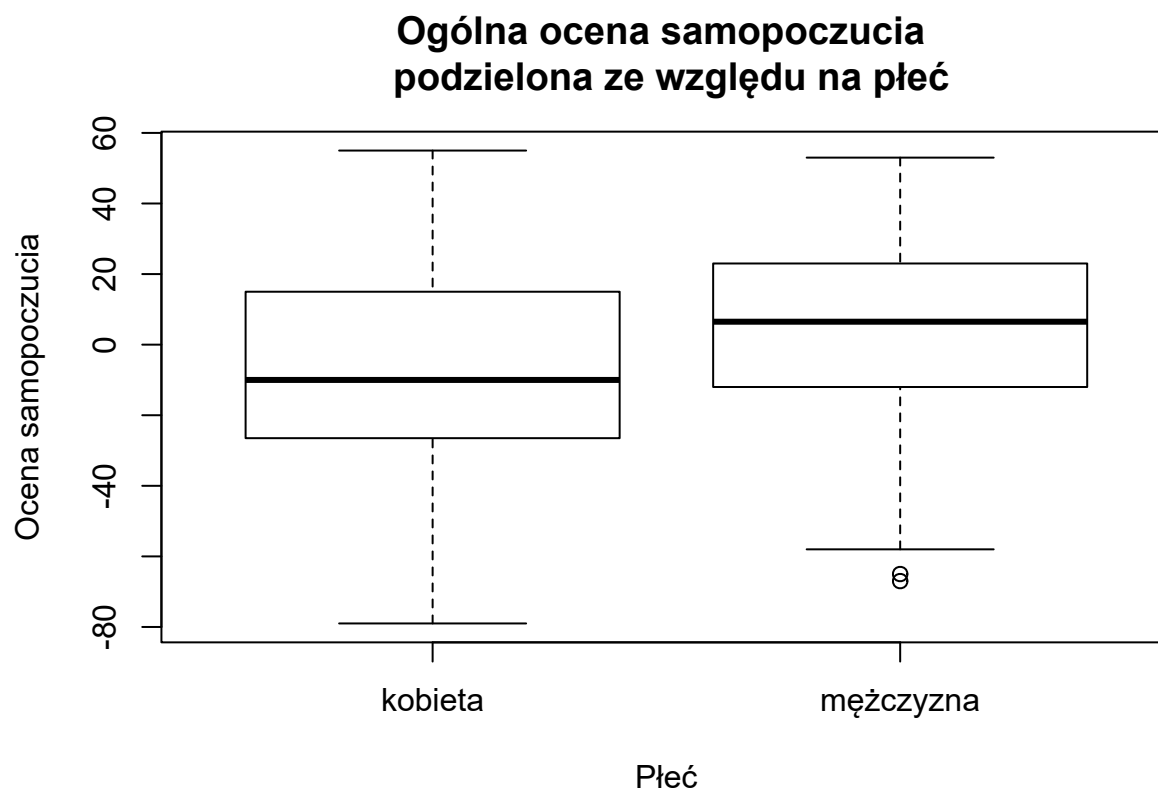
3.8 Hipoteza VIII

Badanie zależności między płcią a ogólną oceną samopoczucia.

3.8.1 Średnia ocena samopoczucia w grupie kobiet i mężczyzn

| | Płeć | |
|---------------------------|---------|-----------|
| | kobieta | mężczyzna |
| Ogólna ocena samopoczucia | -8.0 | 3.2 |

3.8.2 Wykres boxplot



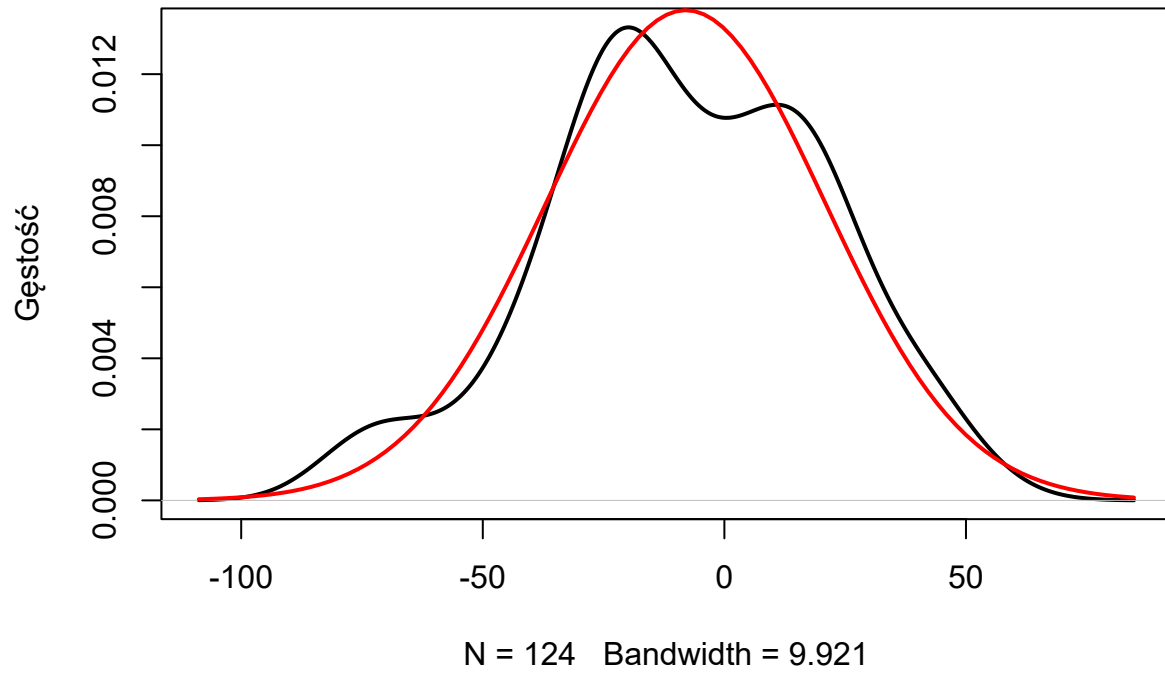
| | Minimum | Mediana | Średnia | Odchylenie standardowe | Maksimum | P-value z testu Shapiro-Wilka |
|-----------|---------|---------|---------|------------------------|----------|-------------------------------|
| Kobiety | -79 | -10 | -8.008 | 28.908 | 55 | 0.178 |
| Mężczyźni | -67 | 6.5 | 3.164 | 27.287 | 53 | 0.014 |

Z wykresu **ramka-wąsy** można wyczytać, że wyniki u kobiet były o wiele bardziej zróżnicowane - “wąsy”, wskazujące na najwyższy i najniższy wynik, niebędący obserwacją odstającą, są znacząco dłuższe, niż w przypadku wykresu dla mężczyzn. Również mediana na pierwszym wykresie jest niżej (poniżej zera). Możliwe jest również, że rozkład gęstości samopoczucia u kobiet charakteryzowałby się asymetrią.

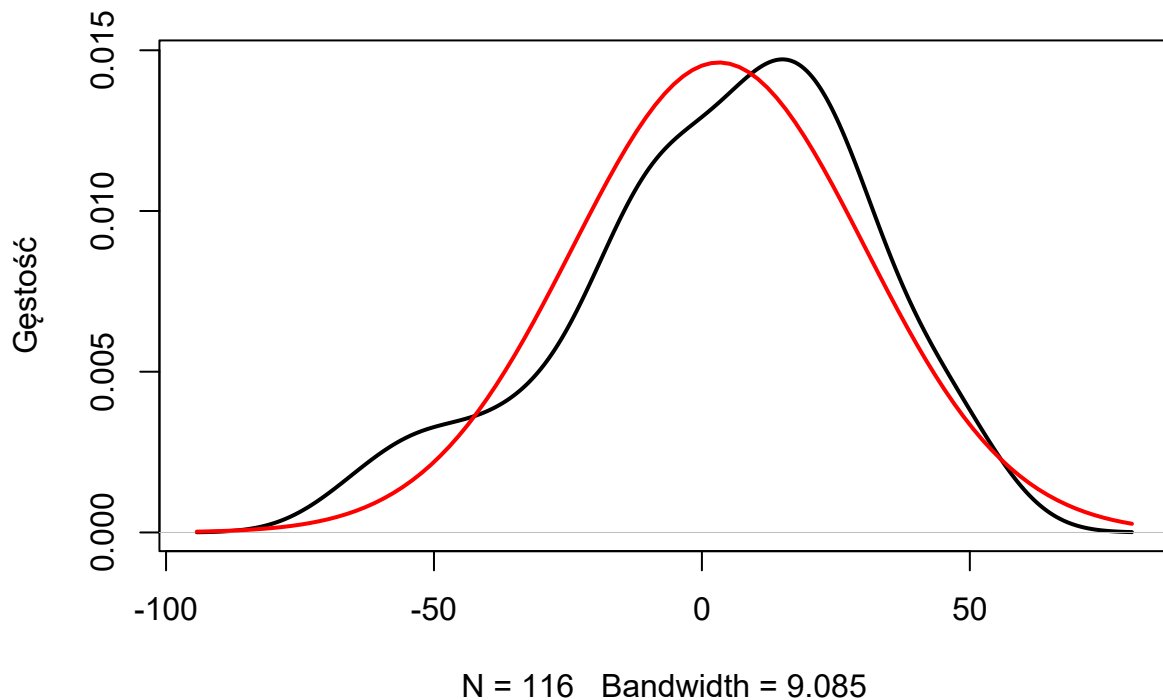
W przypadku mężczyzn, wykres wygląda dość symetrycznie. Można jednak zauważyć, że najniższy wynik niebędący obserwacją odstającą znajduje się dość daleko od pierwszego kwartyła (dolnej części “pudełka”) oraz że tym razem istnieją obserwacje odstające (dolne obserwacje ekstremalne).

Dla lepszego zilustrowania tej zależności posłużymy się wykresami gęstości rozkładu.

Rozkład oceny samopoczucia dla kobiet



Rozkład oceny samopoczucia dla mężczyzn



Rozkład oceny samopoczucia u mężczyzn jest zbliżony do rozkładu normalnego; u kobiet można pokusić się o stwierdzenie dwumodalności rozkładu, a na pewno stwierdzić jego asymetryczność.

3.8.3 Wyniki testów

P-value z testu Shapiro-Wilka dla cechy Ogólna ocena samopoczucia wynosi $p = 0.007$, zatem ta cecha nie ma rozkładu normalnego.

Zaskakujące jest natomiast, że w przypadku podziału tej zmiennej na płeć, okazuje się, że p-value testu Shapiro-Wilka dla kobiet wynosi $p = 0.178$, natomiast dla mężczyzn $p = 0.014$, co - pomimo wcześniejszych hipotez stawianych na podstawie wykresów - nie pozostawia wątpliwości. Rozkład samopoczucia u mężczyzn nie jest normalny, natomiast u kobiet nie ma podstaw do odrzucenia hipotezy o normalności.

Stosujemy test Manna - Whitney'a - Wilcoxona.

P-value z testu Manna - Whitney'a - Wilcoxona wynosi $p = 0.001$. Istnieje zależność pomiędzy płcią a ogólną oceną samopoczucia.

3.9 Hipoteza IX

Badanie zależności między wiekiem a ogólną oceną samopoczucia. Zmienne nie mają rozkładów normalnych, więc wybieramy test Kruskala-Wallisa.

P-value wynosi około 0.063, więc decyzja jest niejednoznaczna, ale na wybranym przez nas poziomie istotności 0.05 nie ma podstaw do odrzucenia hipotezy zerowej.

| test rho-Spearmana | | test tau-Kendalla | |
|------------------------|---------|------------------------|---------|
| współczynnik korelacji | p-value | współczynnik korelacji | p-value |
| 0.172 | 0.008 | 0.119 | 0.01 |

Testy Spearmana oraz Kendalla jednoznacznie wskazują, że należy odrzucić hipotezę zerową (p-value odpowiednio 0.008 oraz 0.01).

Współczynniki korelacji z tych testów wynoszą odpowiednio 0.172 i 0.119, zatem wzrost wieku powoduje wzrost oceny ogólnej samopoczucia.

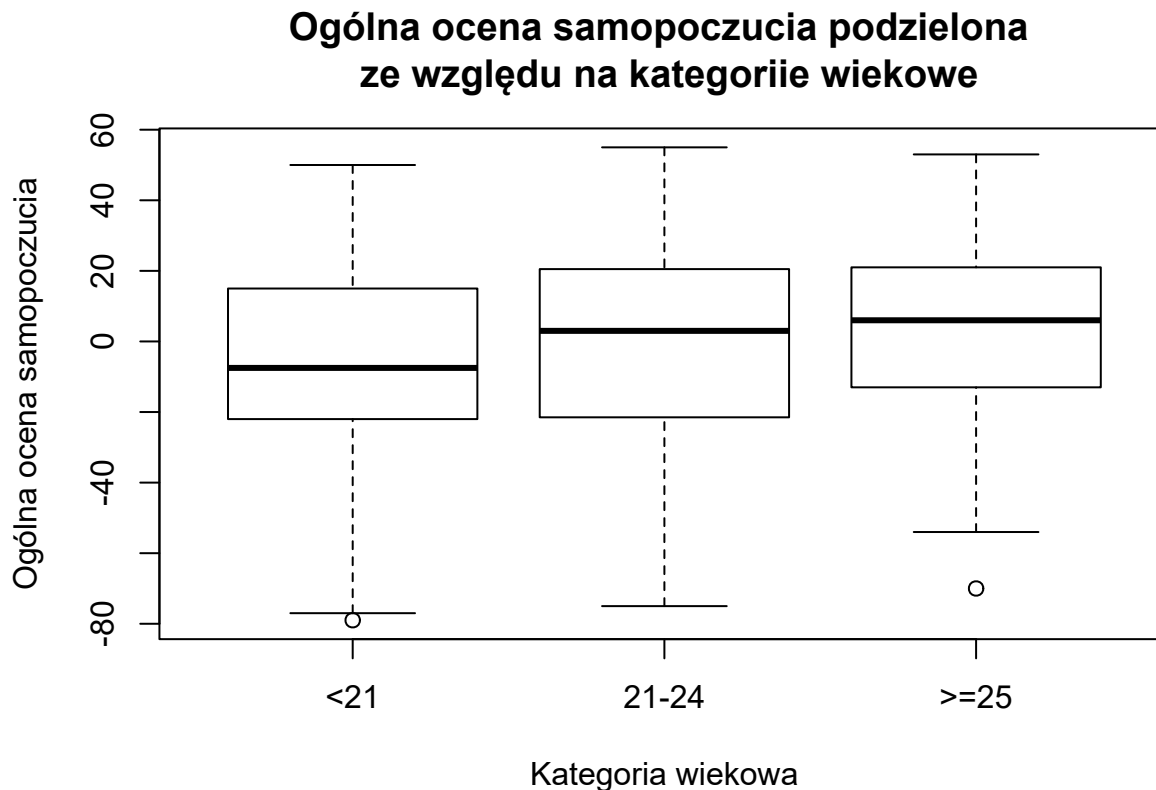
3.10 Hipoteza X

Zbadamy, czy średnie ogólnej oceny samopoczucia różnią się w trzech grupach wiekowych. Ze względu na silną prawostronną asymetrię, wybraliśmy następujące grupy wiekowe:

- poniżej 21 lat
- pomiędzy 21 a 24 lata
- 25 i więcej lat

Poszczególne grupy wiekowe mają następującą liczebność:

| Grupa wiekowa | Liczebność |
|---------------|------------|
| <21 | 70 |
| 21-24 | 131 |
| >=25 | 39 |



Z wykresu boxplot można przypuszczać, że raczej nie ma bardzo istotnych różnic pomiędzy grupami.

Z testu ANOVA otrzymujemy p-value około 0.246, zatem grupa wiekowa nie różnicuje istotnie oceny ogólnej samopoczucia.

4 Dane zsumowane

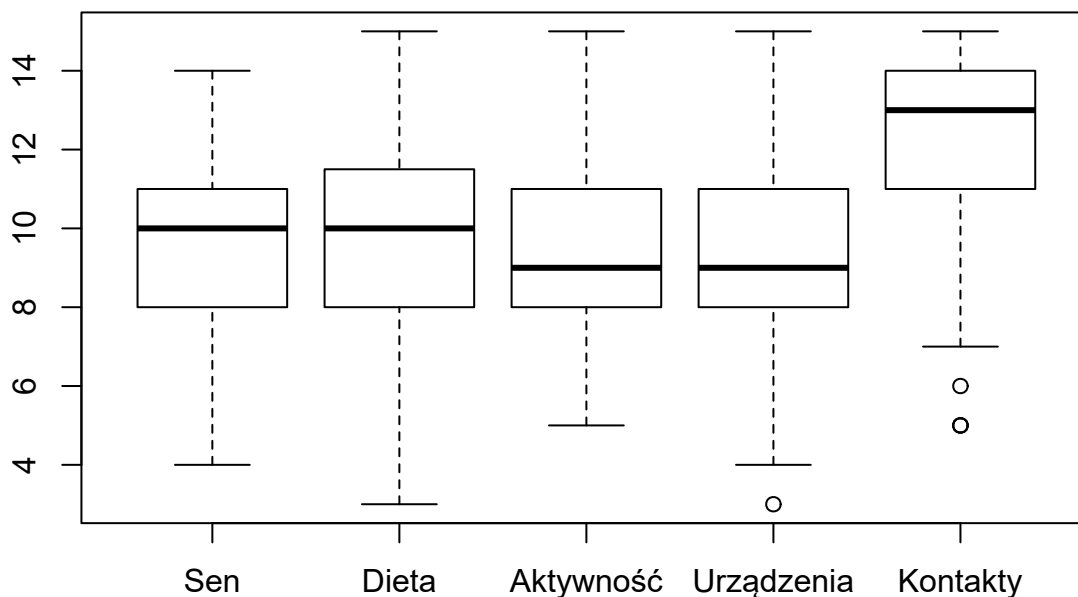
Wcześniej pracowaliśmy osobno na odpowiedziach z ankiety. Teraz będziemy pracować na danych, w których odpowiedzi z danych kategorii zostały zsumowane, tj. zmienne sen, dieta, aktywność, urządzenia i kontakty są sumą trzech odpowiednich zmiennych wymienionych wcześniej:

- na ocenę snu składa się ocena jego jakości, posiadania wyznaczonych pór snu oraz liczby przespanych godzin
- na ocenę diety składa się ocena jakości posiłków, przeważający składnik diety oraz regularność posiłków
- na ocenę aktywności fizycznej składa się ocena kondycji, liczba godzin poświęcanych tygodniowo na sport oraz preferowany środek transportu
- na ocenę czasu spędzanego z urządzeniami elektronicznymi składa się dzienna liczba godzin spędzana na pracy z urządzeniami, dzienna liczba godzin poświęcana na rozrywkę przy użyciu urządzeń oraz poczucie uzależnienia od urządzeń
- na ocenę kontaktów towarzyskich składa się ocena ważności kontaktów międzyludzkich dla danej osoby, posiadanie zaufanej osoby oraz liczba osób, z którymi utrzymuje się regularny kontakt

Wcześniej poszczególne czynniki były ocenione w skali od 1 do 5, więc teraz skala wynosi od 1 do 15.

Oprócz tego do nowego zbioru danych włączamy zmienne płeć, wiek oraz ogólna ocena samopoczucia.

| Cecha | Min | Mediana | Średnia | Odchylenie standardowe | Max | Czy rozkład różni się istotnie od rozkładu normalnego? |
|--------------|-----|---------|---------|------------------------|-----|--|
| Wiek | 14 | 21 | 22.57 | 6.29 | 59 | tak |
| Sen | 4 | 10 | 9.62 | 2.06 | 14 | tak |
| Dieta | 3 | 10 | 9.55 | 2.5 | 15 | tak |
| Aktywność | 5 | 9 | 9.46 | 2.32 | 15 | tak |
| Urządzenia | 3 | 9 | 9.22 | 2.34 | 15 | tak |
| Kontakty | 5 | 13 | 12.16 | 2.06 | 15 | tak |
| Samopoczucie | -79 | -0.5 | -2.61 | 28.63 | 55 | tak |



Z wykresu boxplot widać, że najlepsze oceny pojawiały się w kategorii Kontakty. W przypadku pozostałych czynników, oceny są na podobnym poziomie, z medianami pomiędzy 8 a 10.

4.1 Sen a samopoczucie

| test rho-Spearmana | | test tau-Kendalla | |
|------------------------|---------|------------------------|---------|
| współczynnik korelacji | p-value | współczynnik korelacji | p-value |
| 0.261 | <0.001 | 0.191 | <0.001 |

Z testów Kendalla oraz Spearmana otrzymujemy p-value praktycznie równe zero. Oznacza to, że zależność pomiędzy snem a samopoczuciem jest istotna statystycznie. Współczynniki korelacji wynoszą 0.261 dla Spearmana i 0.191 dla Kendalla. Oznacza to, że im lepiej śpimy, tym lepsze samopoczucie mamy :)

4.2 Dieta a samopoczucie

| test rho-Spearmana | | test tau-Kendalla | |
|------------------------|---------|------------------------|---------|
| współczynnik korelacji | p-value | współczynnik korelacji | p-value |
| 0.322 | <0.001 | 0.232 | <0.001 |

Z testów Spearmana oraz Kendalla otrzymujemy p-value praktycznie równe zero. Istnieje zależność pomiędzy dietą a samopoczuciem. Współczynniki korelacji wynoszą odpowiednio 0.322 dla Spearmana oraz 0.232 dla Kendalla. Im lepsza jakość naszej diety, tym lepsze nasze samopoczucie :)

4.3 Aktywność fizyczna a samopoczucie

| test rho-Spearmana | | test tau-Kendalla | |
|------------------------|---------|------------------------|---------|
| współczynnik korelacji | p-value | współczynnik korelacji | p-value |
| 0.163 | 0.011 | 0.116 | 0.011 |

Z testów Spearmana oraz Kendalla otrzymujemy p-value około 0.011, co pozwala odrzucić hipotezę zerową o braku zależności. Współczynniki korelacji wynoszą odpowiednio 0.163 dla Spearmana i 0.116 dla Kendalla. Im więcej mamy aktywności, tym lepsze nasze samopoczucie :)

4.4 Urządzenia elektroniczne a samopoczucie

| test rho-Spearmana | | test tau-Kendalla | |
|------------------------|---------|------------------------|---------|
| współczynnik korelacji | p-value | współczynnik korelacji | p-value |
| -0.162 | 0.012 | -0.115 | 0.012 |

P-value z testów Spearmana oraz Kendalla wynoszą około 0.012, a więc odrzucamy hipotezę o braku zależności. Współczynniki korelacji wynoszą odpowiednio -0.162 dla Spearmana oraz -0.115 dla Kendalla. Im więcej korzystamy z urządzeń elektronicznych, tym gorsze nasze samopoczucie :(Moje na przykład pogorszyło się podczas robienia tego projektu.

4.5 Kontakty towarzyskie a samopoczucie

| test rho-Spearmana | | test tau-Kendalla | |
|------------------------|---------|------------------------|---------|
| współczynnik korelacji | p-value | współczynnik korelacji | p-value |
| 0.455 | <0.001 | 0.329 | <0.001 |

P-value z obu testów są praktycznie równe zero. Istnieje zależność pomiędzy kontaktami towarzyskimi a samopoczuciem. Współczynniki korelacji wynoszą odpowiednio 0.455 dla Spearmana i 0.329 dla Kendalla. Im lepsze nasze kontakty towarzyskie, tym lepsze nasze samopoczucie :)

4.6 Model

Tablica 1:

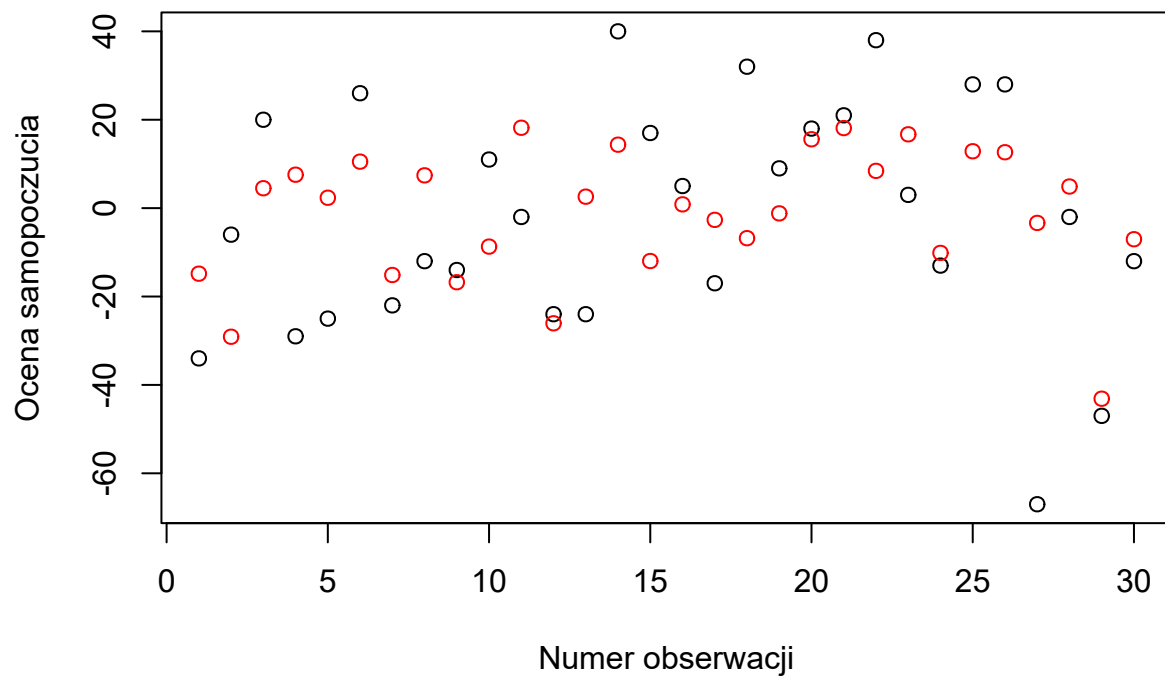
| <i>Dependent variable:</i> | |
|--|-------------------------|
| samopoczucie | |
| sen | 1.522* (0.804) |
| dieta | 2.329*** (0.667) |
| aktywnosc | 1.413** (0.687) |
| urządzenia | -1.651** (0.679) |
| kontakty | 5.577*** (0.775) |
| Constant | -105.469*** (14.664) |
| Observations | 240 |
| R ² | 0.323 |
| Adjusted R ² | 0.309 |
| Residual Std. Error | 23.799 (df = 234) |
| F Statistic | 22.374*** (df = 5; 234) |
| <i>Note:</i> *p<0.1; **p<0.05; ***p<0.01 | |

Stworzyliśmy aplikację okienkową w języku `Python`, w której można wypełnić 15 pytań z naszej ankiety, dotyczących snu, aktywności fizycznej, korzystania z urządzeń elektronicznych, diety oraz kontaktów fizycznych. Na podstawie odpowiedzi użytkownika oraz wyestymowanego przed chwilą modelu, aplikacja wysnuwa przypuszczenie, jakie wg modelu powinno być samopoczucie użytkownika.

Możemy teraz wykonać predykcję Twojego samopoczucia - wystarczy że wypełnisz ankietę na temat 5 czynników :)

Sprawdźmy, jak spisuje się model na naszych danych.

Wykres rzeczywistej oceny samopoczucia oraz jej oceny na podstawie modelu, dla pierwszych 30 obserwacji (na czarno rzeczywista wartość, na czerwono odpowiedź modelu):



W poniżej tabeli na zielono zostały zaznaczone przypadki, w których wartość bezwzględna błędu wynosiła poniżej 10 (przy skali oceny samopoczucia od -124 do +62, oraz odchyleniu standardowym z próby około 28, uznaliśmy, że pomyłki poniżej 10 dają dostatecznie trafny wynik).

| Numer obserwacji | Samopoczucie | Ocena samopoczucia z modelu | Błąd |
|------------------|--------------|-----------------------------|------------------|
| 1 | -34 | -14.8317320 | -19.168268 |
| 2 | -6 | -29.0992765 | 23.099276 |
| 3 | 20 | 4.5043130 | 15.495687 |
| 4 | -29 | 7.5659449 | -36.565945 |
| 5 | -25 | 2.3548900 | -27.354890 |
| 6 | 26 | 10.5218107 | 15.478189 |
| 7 | -22 | -15.1266310 | -6.873369 |
| 8 | -12 | 7.4202654 | -19.420265 |
| 9 | -14 | -16.7606789 | 2.760679 |
| 10 | 11 | -8.7027278 | 19.702728 |
| 11 | -2 | 18.1887226 | -20.188723 |
| 12 | -24 | -26.0585157 | 2.058516 |
| 13 | -24 | 2.5931121 | -26.593112 |
| 14 | 40 | 14.3556898 | 25.644310 |
| 15 | 17 | -11.9675625 | 28.967562 |
| 16 | 5 | 0.8551247 | 4.144875 |
| 17 | -17 | -2.6524544 | -14.347546 |
| 18 | 32 | -6.7972672 | 38.797267 |
| 19 | 9 | -1.1867838 | 10.186784 |
| 20 | 18 | 15.5822921 | 2.417708 |
| 21 | 21 | 18.1147723 | 2.885228 |
| 22 | 38 | 8.4325614 | 29.567439 |
| 23 | 3 | 16.7026555 | -13.702656 |
| 24 | -13 | -10.1329848 | -2.867015 |
| 25 | 28 | 12.8727459 | 15.127254 |
| 26 | 28 | 12.6472973 | 15.352703 |
| 27 | -67 | -3.3472897 | -63.652710 |
| 28 | -2 | 4.8891393 | -6.889139 |
| 29 | -47 | -43.1224679 | -3.877532 |
| 30 | -12 | -7.0350373 | -4.964963 |

4.7 Korelacje (wg testu Spearmana)

Poniżej przedstawiamy tabelę z korelacjami poszczególnych badanych przez nas cech. Korelacje zostały wyliczone z testu Spearmana, w nawiasach jest podane p-value dla każdego współczynnika korelacji.

| | Sen | Dieta | Aktywność | Urządzenia | Kontakty | Samopoczucie |
|--------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| Sen | 1 | 0.324 (<0.001) | -0.016 (0.809) | -0.112 (0.085) | 0.19 (0.003) | 0.261 (<0.001) |
| Dieta | 0.324 (<0.001) | 1 | 0.151 (0.019) | -0.131 (0.042) | 0.171 (0.008) | 0.322 (<0.001) |
| Aktywność | -0.016 (0.809) | 0.151 (0.019) | 1 | -0.156 (0.015) | 0.033 (0.611) | 0.162 (0.011) |
| Urządzenia | -0.112 (0.085) | -0.131 (0.042) | -0.156 (0.015) | 1 | 0.083 (0.2) | -0.162 (0.012) |
| Kontakty | 0.19 (0.003) | 0.171 (0.008) | 0.033 (0.611) | 0.083 (0.2) | 1 | 0.455 (<0.001) |
| Samopoczucie | 0.261 (<0.001) | 0.322 (<0.001) | 0.163 (0.011) | -0.162 (0.012) | 0.455 (<0.001) | 1 |

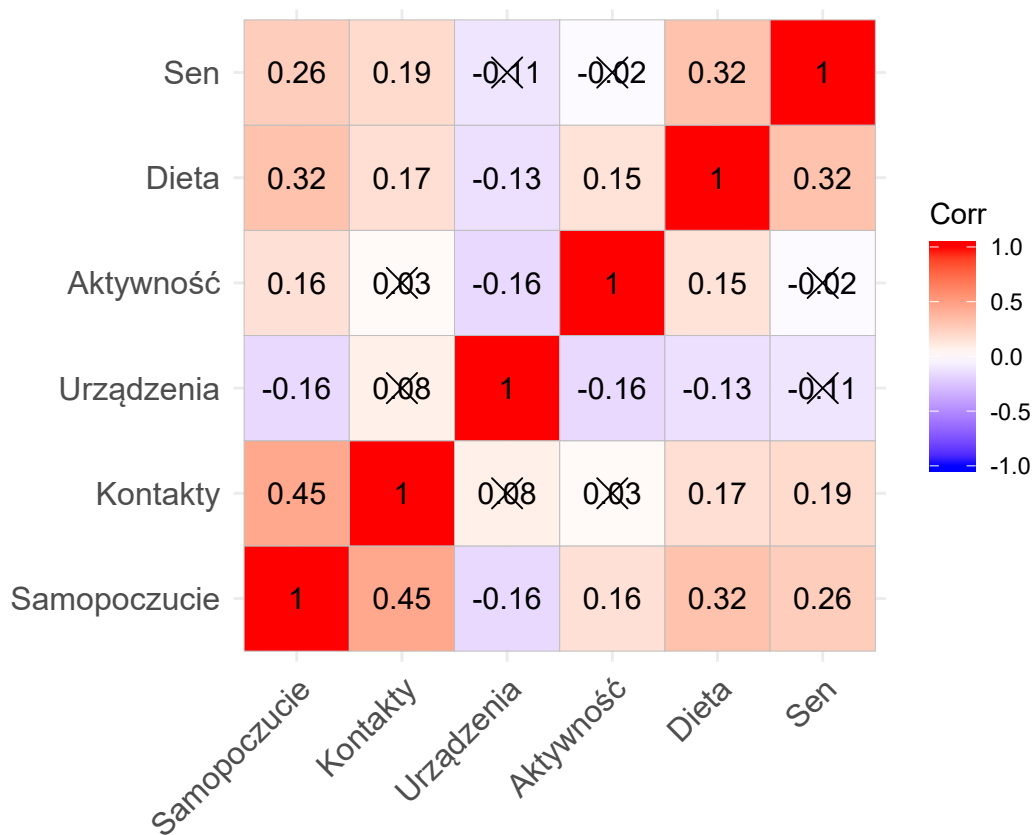
Z istotnych statystycznie korelacji mamy zatem:

- sen i dieta 0.324
- sen i kontakty towarzyskie 0.19
- dieta i kontakty towarzyskie 0.171
- dieta i korzystanie z urządzeń -0.131
- dieta i aktywność fizyczna 0.151
- aktywność fizyczna i korzystanie z urządzeń -0.156

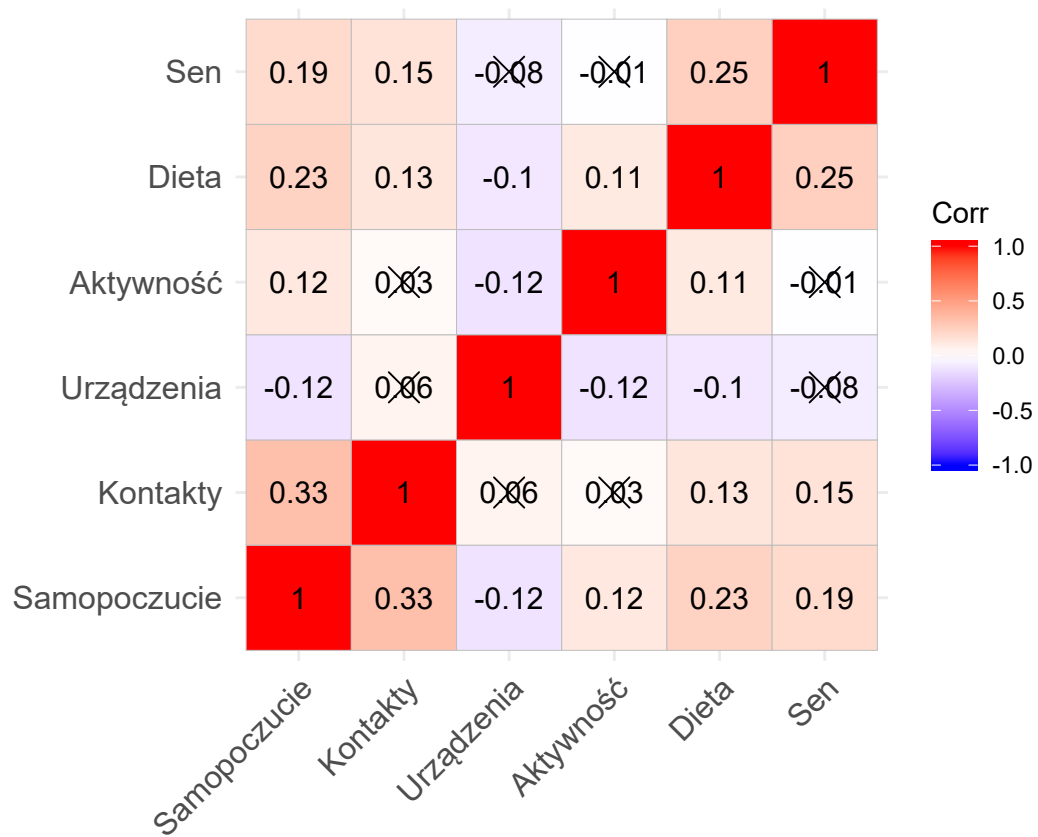
Blisko istotności znalazła się także korelacja między snem a korzystaniem z urządzeń -0.112 z p-value 0.085.

Poza tym istotne są wszystkie korelacje związane z samopoczuciem, co zostało wykazane wcześniej zarówno za pomocą testów Spearmana jak i Kendalla.

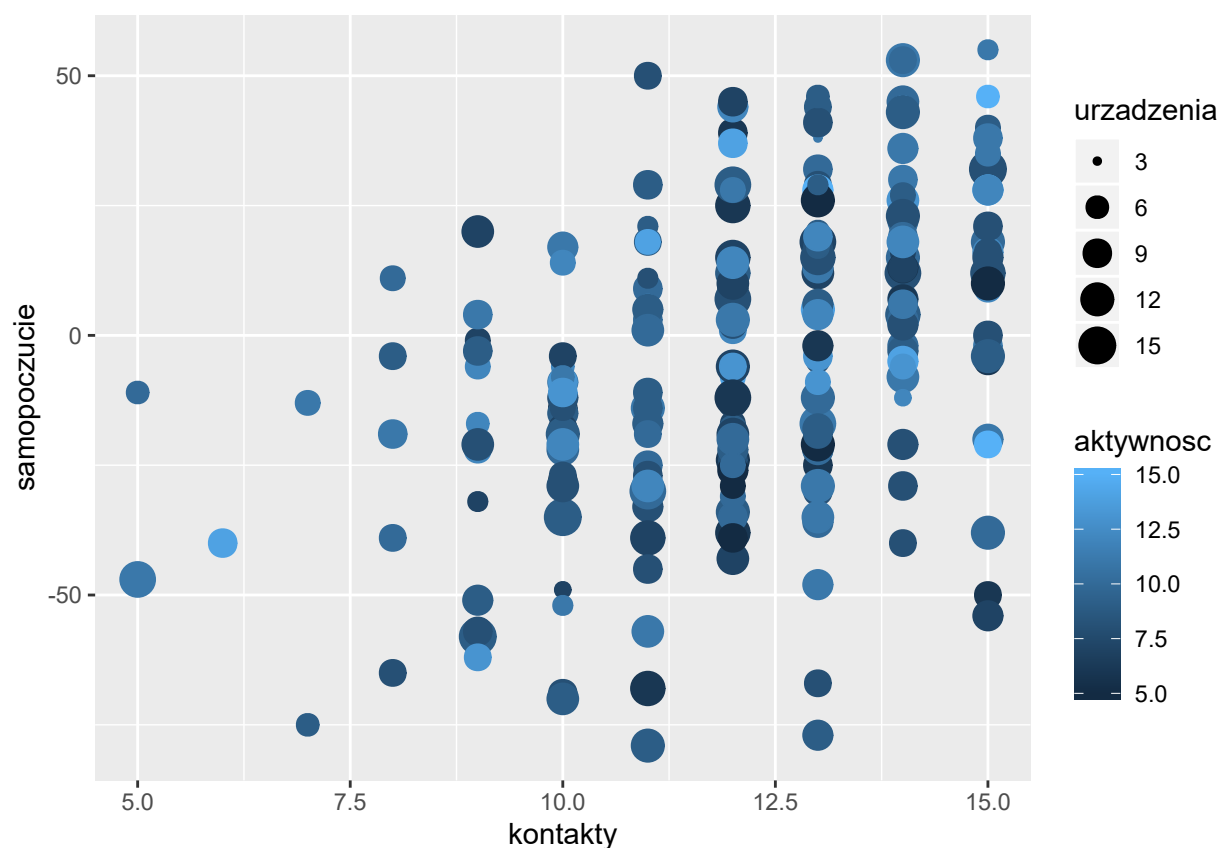
To samo możemy przedstawić także jako kolorową macierz korelacji.



Nieco inne wyniki co do wartości współczynników korelacji uzyskujemy, używając testu Kendalla. Jednakże nie zmienia się istotność korelacji.



4.8 Bubble chart



Wykres ten przedstawia zależność samopoczucia od kontaktów towarzyskich (zaznaczone na osiach x i y). Zgodnie z wcześniejszymi wnioskami, im więcej mamy kontaktu towarzyskiego, tym lepsze nasze samopoczucie. Widać, że dla osób, które dostały ocenę za 'Kontakty' 5, 6 lub 7, samopoczucie było ujemne (poniżej średniej, która wynosiła -2.61). Dopiero przy ocenie za 'Kontakty' 8 pojawiają się oceny dodatnie samopoczucia. Dla osób, które ocenę za kontakty miały w okolicach 13-15, ocena samopoczucia sięgała nawet 50 (maksimum wynosiło 55).

Im jaśniejsza jest kropka, tym lepsza ocena za aktywność wystąpiła, z kolei im większa kropka, tym wyższa ocena za urządzenia.

Nie widać tutaj wyraźnej zależności, ale można zauważyć, że w okolicy średniej oceny za kontakty, oceny za aktywność i urządzenia także są średnie, bo kropki są i średniej wielkości i pośredniego koloru.

5 Wnioski

5.1 Wnioski z bardziej szczegółowych hipotez

Wykazaliśmy następujące związki:

- wraz ze wzrostem liczby bliskich znajomych, spada częstość odczuwania niepokoju
- im więcej godzin spędzamy na korzystaniu z urządzeń elektronicznych dla rozrywki, tym bardziej czujemy się od nich uzależnieni (co wydaje się oczywistym związkiem, i według testów statystycznych jest prawie pewne, ponieważ p-value było bardzo małe dla tej hipotezy, praktycznie zero)
- z podobną pewnością jak powyższą hipotezę, wykazaliśmy, że im bardziej regularne pory snu, tym lepsza jest jego jakość

- im więcej godzin tygodniowo poświęca się na sport, tym rzadziej odczuwa się załamanie
- im większą wagę ktoś przywiązuje do jakości swoich posiłków, tym lepszy ma humor

Wykazaliśmy także, że nie zachodzą związki:

- pomiędzy jakością snu a poczuciem upokorzenia - chociaż tutaj wynik był na krawędzi, ocierał się o wniosek, że im gorsza jakość snu, tym częściej odczuwamy upokorzenie
- pomiędzy dietą a poczuciem radości

Porównując wyniki w różnych podgrupach, odkryliśmy, że:

- średnia ocena ogólna samopoczucia różniła się między grupą kobiet a mężczyzn, w dodatku była znacznie niższa (-8) w grupie kobiet niż mężczyzn (3.2)
- wiek ma wpływ na samopoczucie, ale zaproponowane przez nas grupy wiekowe (poniżej 21 lat, 21-24 i powyżej 24 lat) nie różnicują istotnie samopoczucia

5.2 Wnioski dla pięciu głównych czynników

Na następnej stronie znajduje się infografika przedstawiająca czynniki, które mają wpływ na nasze samopoczucie. Pod nazwą czynnika została podana przybliżona wartość współczynnika korelacji. Na zielonym tle są podane korelacje dodatnie, a na czerwonym ujemne.

Ta infografika jest skrótownym przedstawieniem wniosków, jakie dostaliśmy przy badaniu pięciu wybranych przez nas czynników wpływających na samopoczucie.

CO WPŁYWA NA SAMOPOCZUCIE?

STATYSTYCZNE SPOJRZENIE NA TEMAT

SEN

0.191 do 0.261



Im lepiej się wysypiamy, tym bardziej poprawia się nasze samopoczucie.

DIETA

0.232 do 0.322

Dobra jakość naszej diety owocuje poprawą ogólnego nastroju.



AKTYWNOŚĆ FIZYCZNA

0.116 do 0.163



Aktywność fizyczna sprawia, że lepiej się czujemy.

URZĄDZENIA ELEKTRONICZNE

-0.115 do -0.162

Dłuższe godziny korzystania z urządzeń elektronicznych pogarszają samopoczucie.



KONTAKTY TOWARZYSKIE

0.329 do 0.455



Kontakty towarzyskie pozytywnie wpływają na nasz stan ducha.

6 Bibliografia

(Bogdan Wojciszke 2005) (Wątroba 2009)

Bogdan Wojciszke, Wiesław Baryła. 2005. *Skale do Pomiaru Nastroju i Sześciu Emocji*.

Wątroba, Janusz. 2009. *Wspomaganie Statystycznej Analizy Wyników Badań Empirycznych w STATISTICA 9*. StatSoft Polska Sp. z o. o.