

ЗВІТ

з лабораторної роботи №1

з дисципліни «Аналіз даних»

Склад команди:

- Піковець Артем КМ-22

1 Вступ

Підвищений артеріальний тиск є одним із головних факторів ризику серцево-судинних захворювань. Це дослідження спрямоване на виявлення ключових факторів, що впливають на артеріальний тиск, та їхнього відносного значення.

1. Які ключові демографічні, соціально-економічні та фактори способу життя (вік, стать, дохід, освіта, вага тіла, фізична активність, сон, куріння, дієта) найбільш сильно пов'язані з підвищеним артеріальним тиском?

З віком фізіологічні процеси в організмі змінюються, що може привести до підвищення артеріального тиску. Соціально-економічні фактори, такі як дохід та освіта, можуть впливати на доступ до здорового способу життя та медичної допомоги. Здоровий спосіб життя, що включає регулярну фізичну активність та збалансоване харчування, ймовірно, сприяє підтримці нормального артеріального тиску.

2. Чи існує прямий зв'язок між відсотком жиру в організмі та рівнем артеріального тиску? Чи має значення розподіл жиру, зокрема, який тип жиру більше пов'язаний з вищим артеріальним тиском: андроїдний чи гіноїдний?

Хоча загальна маса тіла вважається потенційним фактором, що впливає на артеріальний тиск, можливо, саме розподіл жирової тканини відіграє більш значущу роль.

2 Опис даних

Дані були отримані з Національного дослідження здоров'я та харчування (NHANES). Щороку близько 5000 дорослих та дітей з різних громад Сполучених Штатів беруть участь у NHANES. Процес вибірки забезпечує, що ця група людей є репрезентативною.

NHANES включає:

- Інтерв'ю щодо здоров'я, харчування та особистих, соціальних та економічних характеристик.
- Відвідування мобільного центру обстеження для вимірювання показників здоров'я та тіла.

2.1 Збір даних

Для аналізу були використані дані NHANES за період з 2011 по 2018 роки. Змінні, що відповідають нашим дослідницьким питанням, були вилучені за допомогою R-пакета `nhanesA`. Код для вилучення початкового набору даних знаходиться у файлі `lab0/00-extract-data.R`.

2.2 Очищення даних

Початковий набір даних складає **37399** рядків та **34** змінних.

```
names(data)
```

```
[1] "seqn"                      "diastolic1"
[3] "diastolic2"                 "diastolic3"
[5] "diastolic4"                 "systolic1"
[7] "systolic2"                  "systolic3"
[9] "systolic4"                  "pulse"
[11] "weight_kg"                 "waist_cm"
[13] "subcutaneous_fat_g"        "visceral_fat_g"
[15] "android_g"                 "android_fat_g"
[17] "android_non_fat_g"         "gynoid_g"
[19] "gynoid_fat_g"              "gynoid_non_fat_g"
[21] "survey_year"               "age"
[23] "gender"                     "race"
[25] "education"                 "poverty_index"
[27] "vigorous_phys_activity"    "vigorous_phys_activity_days"
[29] "moderate_phys_activity"     "moderate_phys_activity_days"
[31] "sleep"                      "smoking_now"
[33] "smoked_100_cigs"           "diet_quality"
```

Змінні артеріального тиску

Спочатку необхідно вилучити нашу основну змінну - вимірювання артеріального тиску. Артеріальний тиск записується у вигляді двох чисел:

- Систолічний (systolic) артеріальний тиск - перше число. Воно вимірює тиск, з яким кров тисне на стінки артерій під час скорочення серця.
- Діастолічний (diastolic) артеріальний тиск - друге число. Воно вимірює тиск, з яким кров тисне на стінки артерій під час розслаблення серцевого м'яза між скороченнями.

Для забезпечення точності було проведено кілька вимірювань систолічного та діастолічного артеріального тиску. Зазвичай проводилося 3 вимірювання, і якщо вони не були узгодженими, проводилося 4-те вимірювання.

Не всім учасникам вимірювали артеріальний тиск, але можна припустити, що підмножина людей, яким його вимірювали, все ще є репрезентативною. Кількість значень NA для вимірювань артеріального тиску показана в наступній таблиці:

```
# A tibble: 8 x 3
  column      na_percent na_count
  <chr>          <dbl>     <int>
1 diastolic1    0.268    10024
```

2	diastolic2	0.246	9206
3	diastolic3	0.248	9260
4	diastolic4	0.951	35557
5	systolic1	0.268	10024
6	systolic2	0.246	9206
7	systolic3	0.248	9260
8	systolic4	0.951	35557

Кількість рядків, що містять принаймні 1 вимірювання діастолічного та 1 вимірювання систолічного тиску, становить **28666**.

Розподіл цих вимірювань показано на Figure 1.

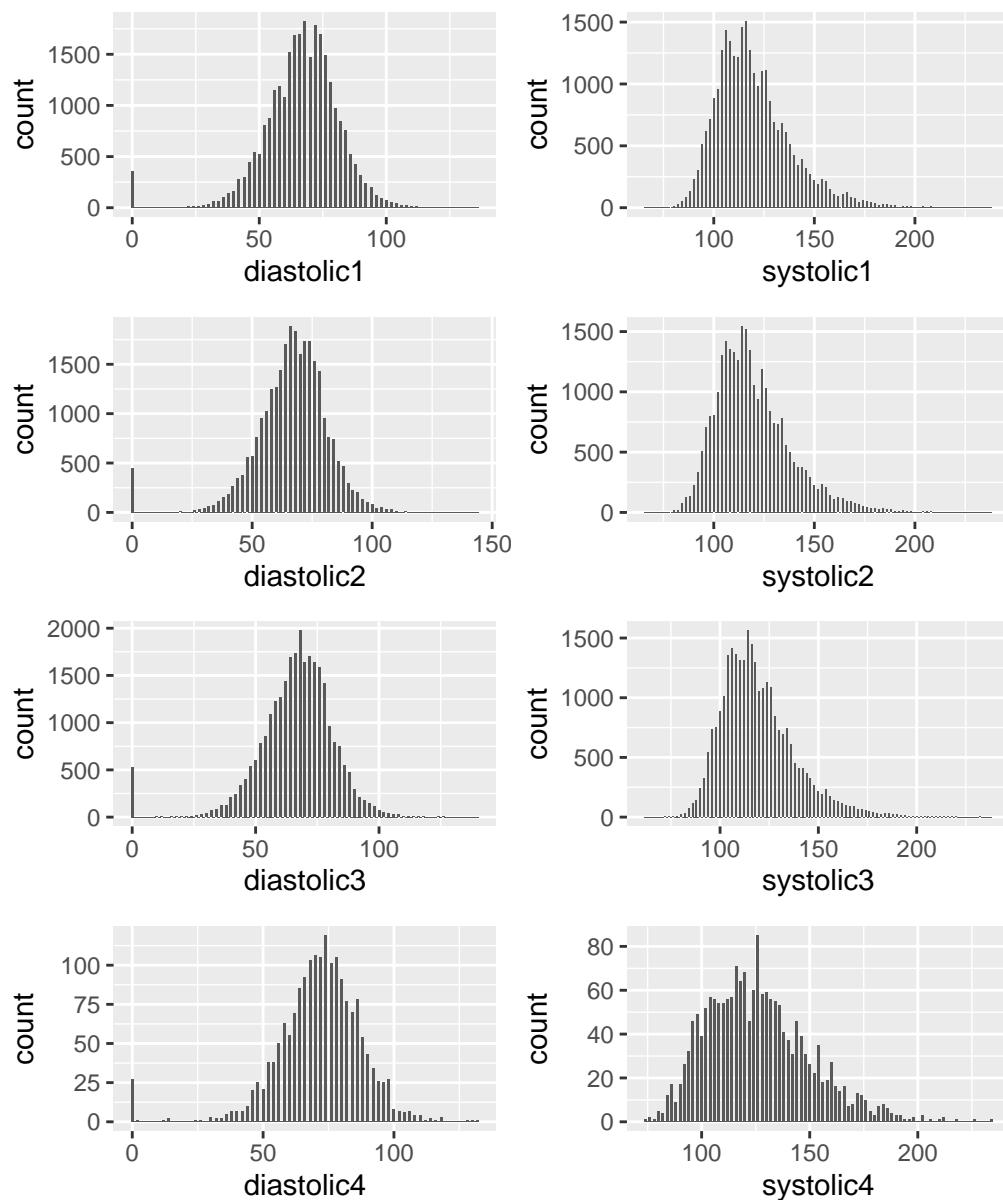
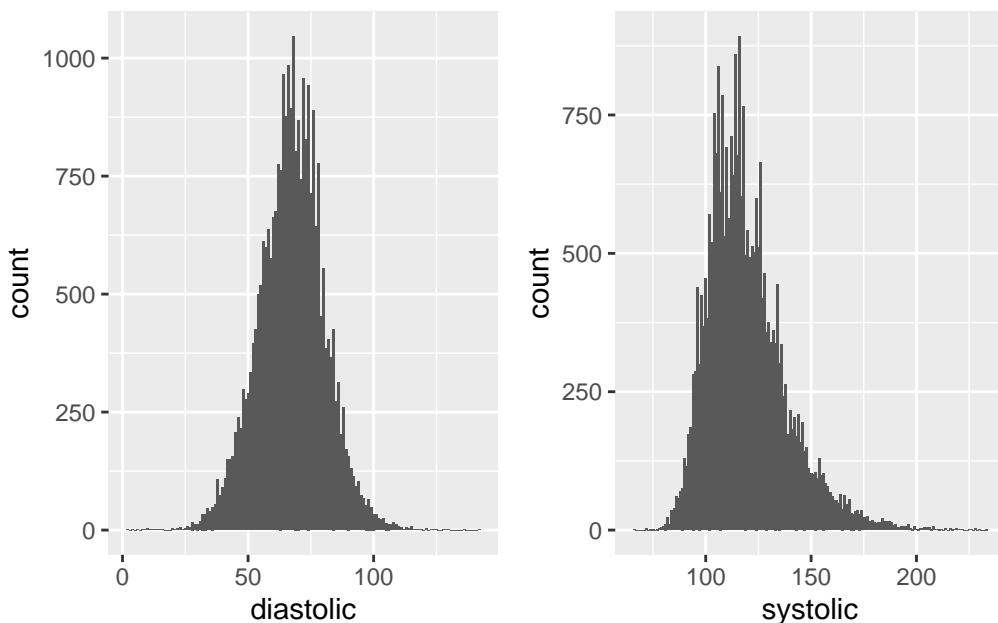


Figure 1

Ми помітили, що вимірювання містять лише парні цілі числа, і виявили, що діастолічні вимірювання, що дорівнюють 0, є помилковими. Після позначення їх як NA, у нас залишилося **28479** рядків з принаймні 1 діастолічним та 1 систолічним вимірюванням.

Ми агрегуємо 4 вимірювання наступним чином. Якщо два найближчі діастолічні вимірювання відрізняються незначно (не більше ніж на 8), то ми беремо їх середнє значення як вимірювання діастолічного артеріального тиску, інакше позначаємо його як NA. Те ж саме робимо для систолічного артеріального тиску. Таким чином, ми отримуємо змінні, що представляють систолічний та діастолічний артеріальний тиск:



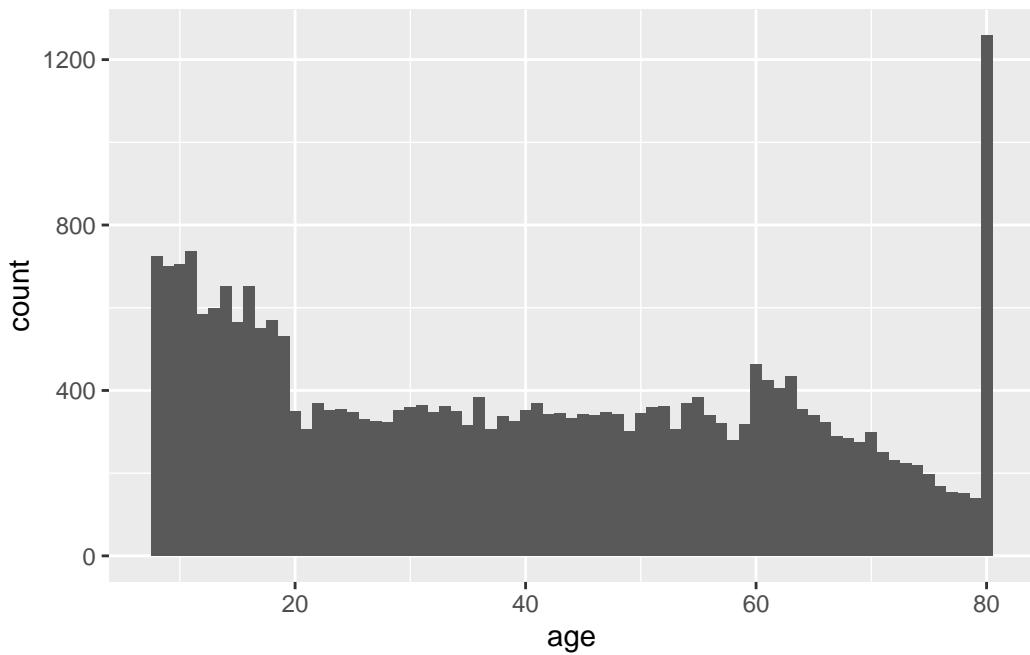
Кількість рядків з наявними агрегованими систолічними та діастолічними вимірюваннями становить **28097**.

Ми також додаємо нову змінну, що представляє категорію артеріального тиску, наступним чином:

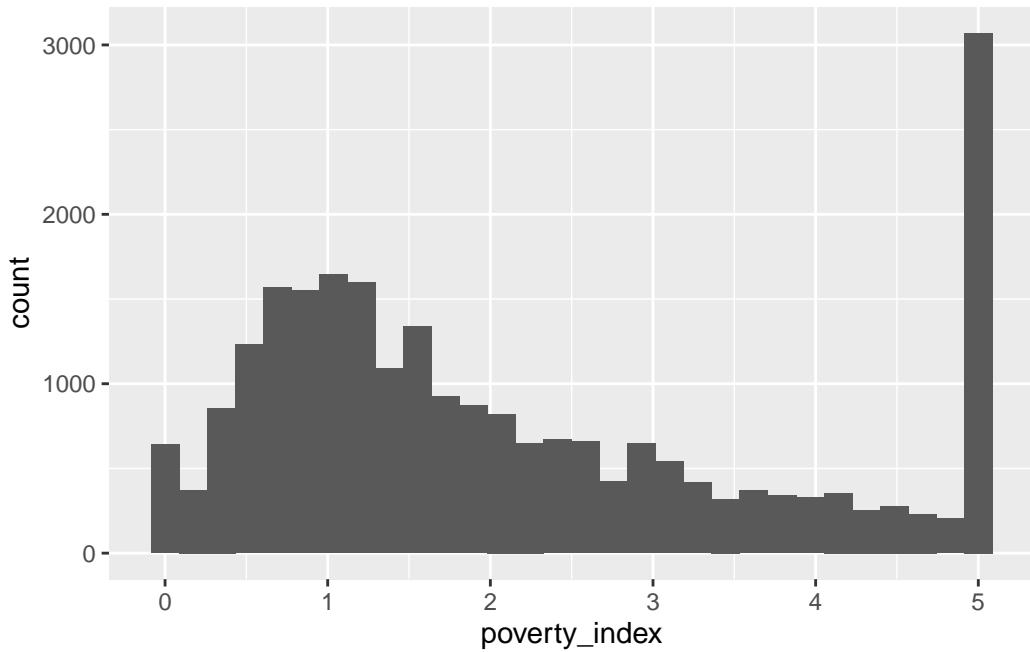
- **Normal**, якщо систолічний < 120 та діастолічний < 80
- **Elevated**, якщо $120 \leq$ систолічний < 130 та діастолічний < 80
- **Hypertension Stage 1**, якщо систолічний ≥ 130 або діастолічний ≥ 80
- **Hypertension Stage 2**, якщо систолічний ≥ 140 або діастолічний ≥ 90

Інші змінні

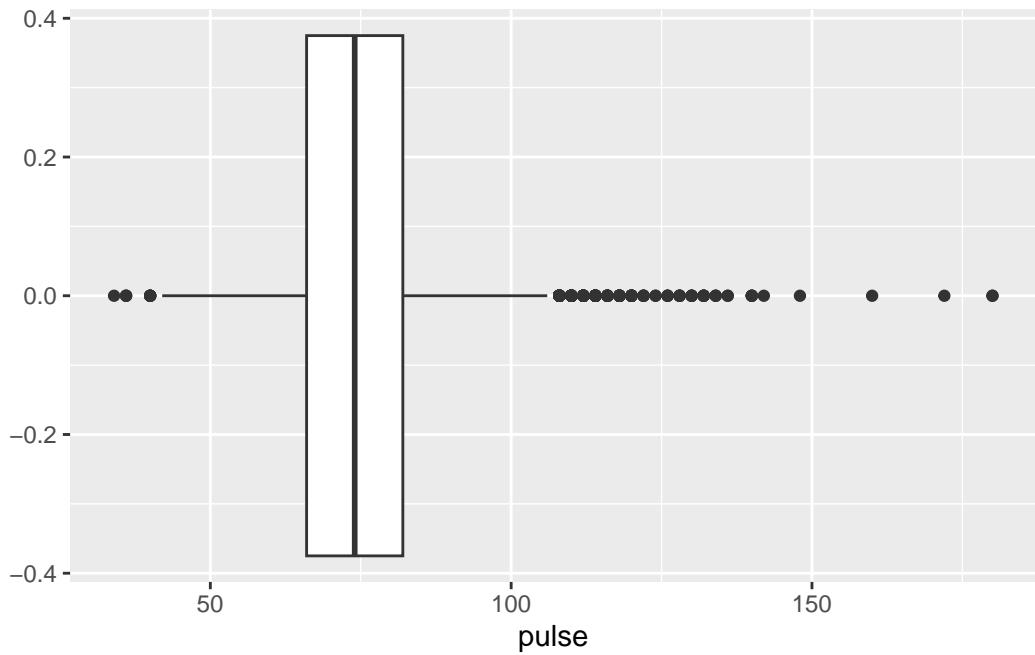
Вимірювання артеріального тиску проводилося лише для осіб віком 8 років і старше. Вік осіб 80 років і старше був закодований як 80.



`poverty_index` - відношення місячного доходу сім'ї до рівня бідності, що залежать від розміру сім'ї. Значення, що перевищують 5, закодовані як 5.



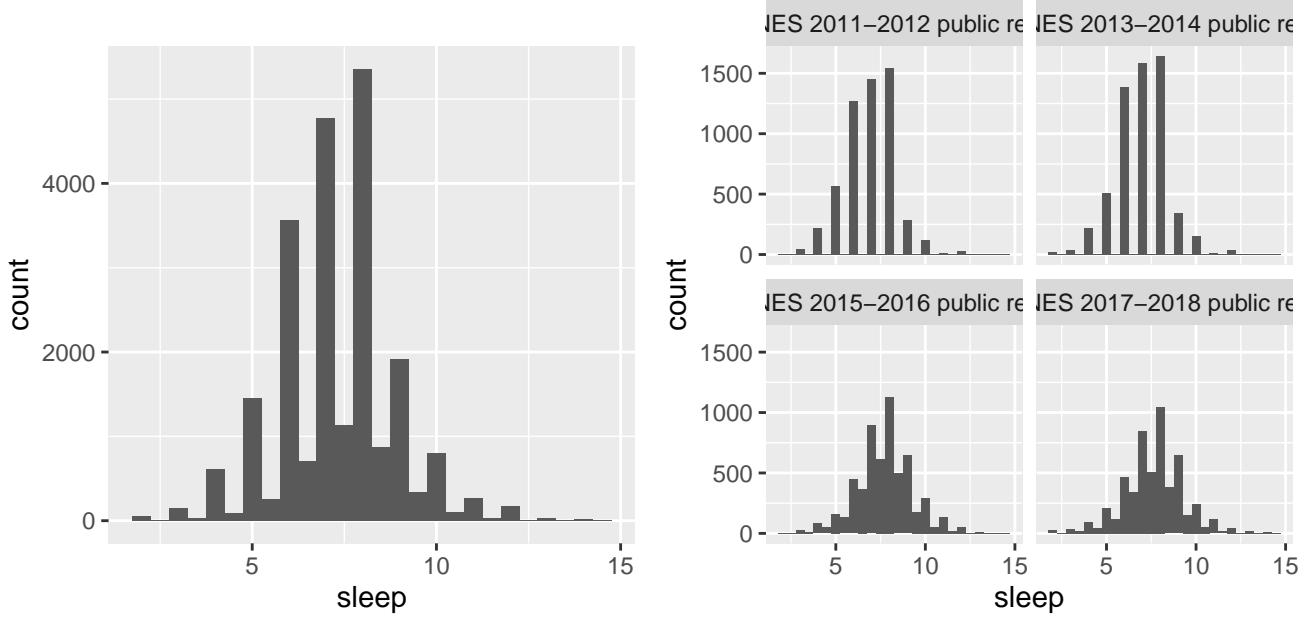
Було виявлено одне спостереження з `pulse = 0`, що є помилкою, тому ми позначили його як NA. Також є викиди з високими значеннями `pulse`, але вони вважаються коректними.



Додано три нові змінні що відображають розподіл жиру в організмі людини:

```
data <- data |>
  mutate(
    android_fat_percent = android_fat_g / android_g,
    gynoid_fat_percent = gynoid_fat_g / gynoid_g,
    android_gynoid_ratio = android_fat_g / gynoid_fat_g
  )
```

Тривалість сну мала дивний розподіл, де числа з дробовою частиною .5 зустрічалися рідше, ніж цілі числа. Це було пояснено різними правилами округлення вимірювань у різні роки опитування.



Вибірки даних

Кількість рядків після обробки - **28097**. Кількість відсутніх значень у кожному стовпці наведена нижче:

column	na_percent	na_count
<chr>	<dbl>	<int>
1 seqn	0	0
2 diastolic	0	0
3 systolic	0	0
4 blood_pressure_category	0	0
5 pulse	0.000107	3
6 weight_kg	0.00883	248
7 waist_cm	0.0447	1256
8 subcutaneous_fat_g	0.343	9628
9 visceral_fat_g	0.343	9630
10 android_g	0.343	9640
11 android_fat_g	0.343	9640
12 android_non_fat_g	0.343	9640
13 gynoid_g	0.343	9625
14 gynoid_fat_g	0.343	9625
15 gynoid_non_fat_g	0.343	9625
16 survey_year	0	0
17 age	0	0
18 gender	0	0
19 race	0	0
20 education	0.270	7582

21 poverty_index	0.136	3827
22 vigorous_phys_activity	0.133	3740
23 vigorous_phys_activity_days	0.133	3743
24 moderate_phys_activity	0.133	3744
25 moderate_phys_activity_days	0.134	3752
26 sleep	0.191	5355
27 smoking_now	0.685	19233
28 smoked_100_cigs	0.241	6758
29 diet_quality	0.188	5271
30 android_fat_percent	0.343	9640
31 gynoid_fat_percent	0.343	9625
32 android_gynoid_ratio	0.344	9658

Вибірка людей, яким вимірювали жирову тканину, відрізняється від вибірки людей, яких опитували про їхній спосіб життя. Для кожного з наших дослідницьких питань ми вирішили вибрати підмножину даних, яка не має відсутніх значень у змінних, що стосуються цього дослідницького питання.

1. Виділимо підмножину даних, що не містить пропущених значень в змінних ваги та пульсу, і назовемо її **whole_data**. Вона містить наступні змінні, кожна з яких не має пропущених значень:

```
# A tibble: 10 x 3
  column          na_percent na_count
  <chr>            <dbl>      <int>
1 seqn              0          0
2 diastolic         0          0
3 systolic          0          0
4 blood_pressure_category 0          0
5 pulse              0          0
6 weight_kg          0          0
7 survey_year        0          0
8 age                0          0
9 gender             0          0
10 race              0          0
```

2. Виділимо підмножину даних, що не містить пропущених значень в змінних які стосуються способу життя, і назовемо її **lifestyle_data**. Вона містить наступні змінні, кожна з яких не має відсутніх значень:

```
# A tibble: 18 x 3
  column          na_percent na_count
  <chr>            <dbl>      <int>
1 seqn              0          0
2 diastolic         0          0
3 systolic          0          0
4 blood_pressure_category 0          0
5 weight_kg          0          0
6 survey_year        0          0
```

7 age	0	0
8 gender	0	0
9 race	0	0
10 education	0	0
11 poverty_index	0	0
12 vigorous_phys_activity	0	0
13 vigorous_phys_activity_days	0	0
14 moderate_phys_activity	0	0
15 moderate_phys_activity_days	0	0
16 sleep	0	0
17 smoked_100_cigs	0	0
18 diet_quality	0	0

Важливо зазначити, що про спосіб життя опитували лише людей віком 20 років і старше.

3. Виділимо підмножину даних, що не містить пропущених значень в змінних ваги, об'єму талії та вимірювань жирової тканини, і назовемо її **fat_data**. Вона містить наступні змінні, кожна з яких не має відсутніх значень:

```
# A tibble: 21 x 3
  column      na_percent na_count
  <chr>        <dbl>     <int>
1 seqn          0         0
2 diastolic     0         0
3 systolic       0         0
4 blood_pressure_category 0         0
5 weight_kg      0         0
6 waist_cm       0         0
7 subcutaneous_fat_g 0         0
8 visceral_fat_g 0         0
9 android_g      0         0
10 android_fat_g 0         0
# i 11 more rows
```

Важливо зазначити, що вимірювання жирової тканини проводилося лише для людей віком від 8 до 59 років.

Загальна інформація про вибірки даних надана в наступній таблиці.

Subset name	Number of observations	Number of variables	Age range
whole_data	27846	10	8+
lifestyle_data	17393	18	20+
fat_data	18292	21	8-59

3 Результати EDA

3.1 Як вік, вага, раса та стать впливають на артеріальний тиск?

Для відповіді на це питання використаємо вибірку whole_data.

Дескриптивні статистики

Числові змінні

variable name	min	1st quartile	mean	median	3rd quartile	max
diastolic	2.0	59.0	67.36	68	76.0	130.0
systolic	66.0	106.0	119.44	116	129.0	233.0
age	8.0	18.0	39.70	38	59.0	80.0
weight_kg	16.9	58.8	74.95	73	88.9	222.6

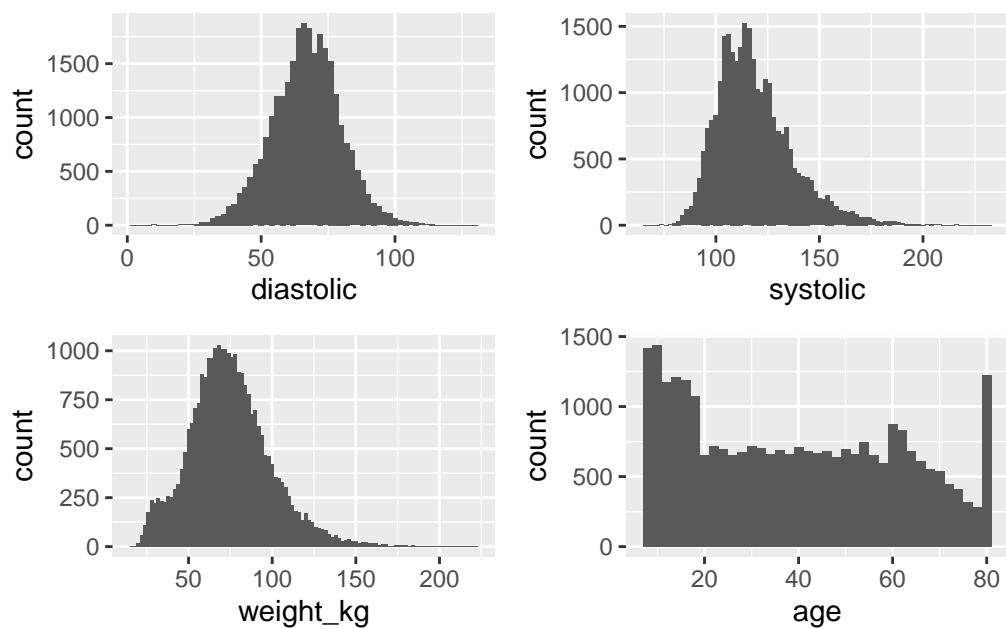
Категорійні змінні

```
# A tibble: 4 x 3
  blood_pressure_category percentage     n
  <ord>                  <dbl> <int>
1 Normal                 0.555 15452
2 Elevated                0.139 3858
3 Hypertension Stage 1   0.160 4462
4 Hypertension Stage 2   0.146 4074

# A tibble: 2 x 3
  gender percentage     n
  <chr>      <dbl> <int>
1 Female     0.509 14178
2 Male       0.491 13668

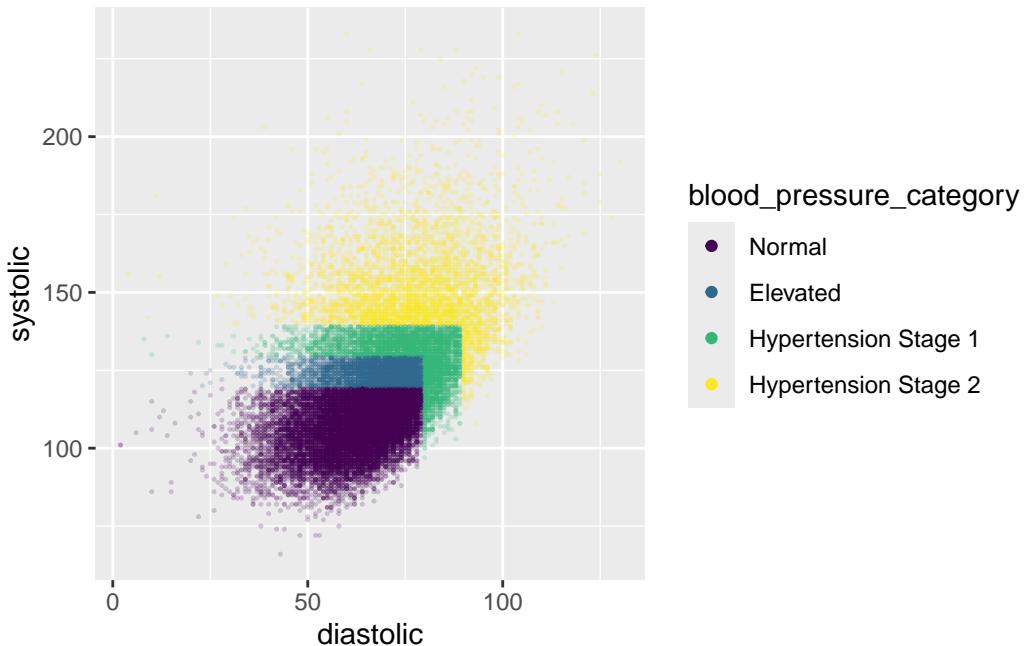
# A tibble: 6 x 3
  race          percentage     n
  <chr>         <dbl> <int>
1 Mexican American  0.153  4273
2 Non-Hispanic Asian 0.122  3391
3 Non-Hispanic Black 0.234  6512
4 Non-Hispanic White 0.342  9519
5 Other Hispanic    0.105  2913
6 Other Race        0.0445 1238
```

Розподіл числових змінних



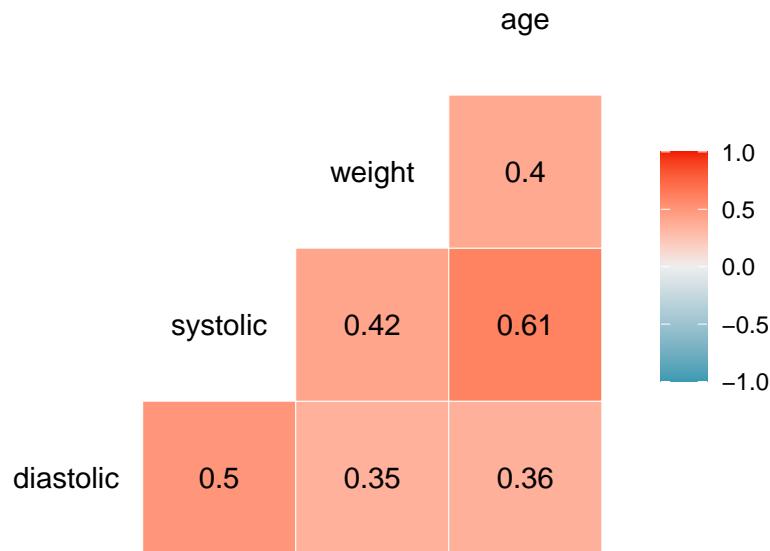
На розподілах віку та ваги помітно перепредставлену групу дітей. Тому варто бути обережним з інтерпретацією результатів дослідження.

Розподіл систолічного та діастолічного тиску

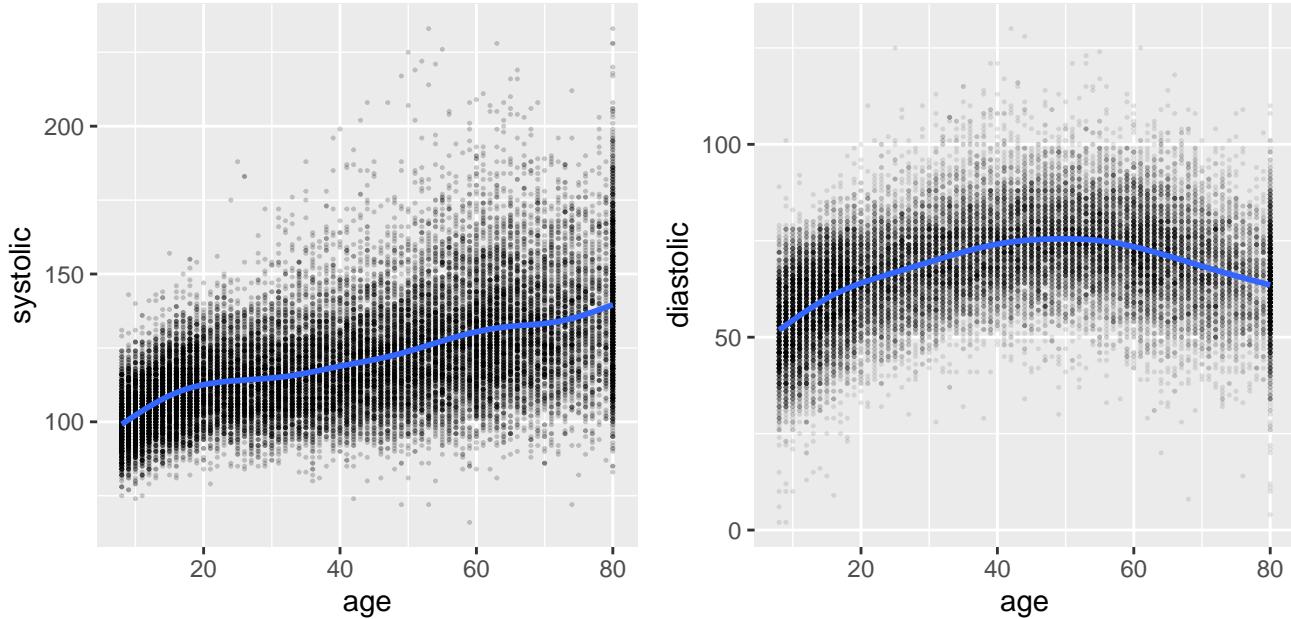


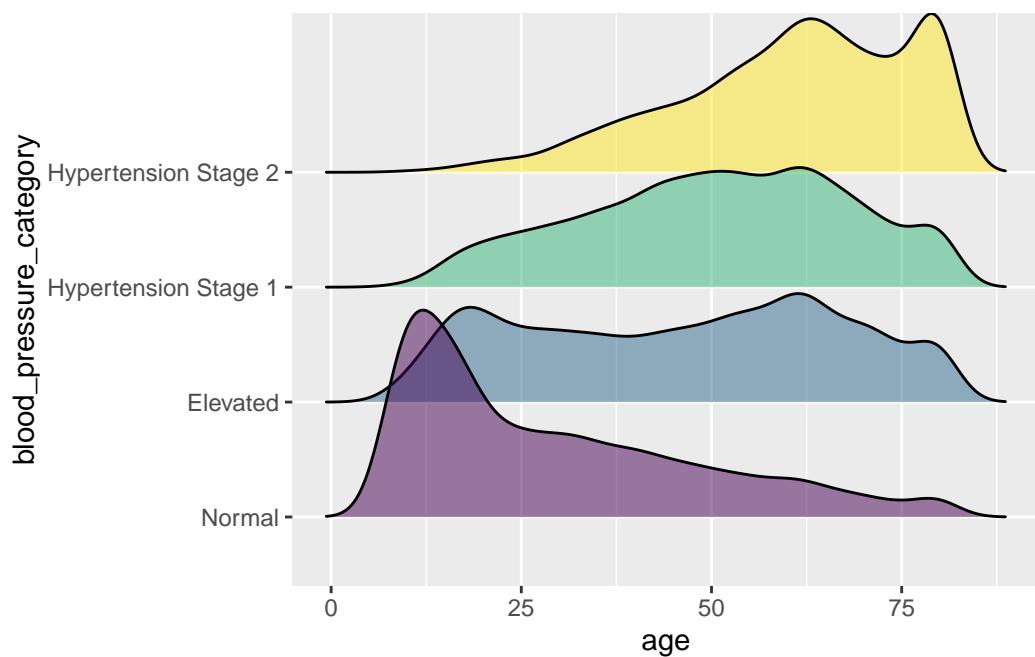
Який кореляційний зв'язок між вагою, віком та артеріальним тиском?

Були розраховані наступні коефіцієнти кореляції Спірмена:



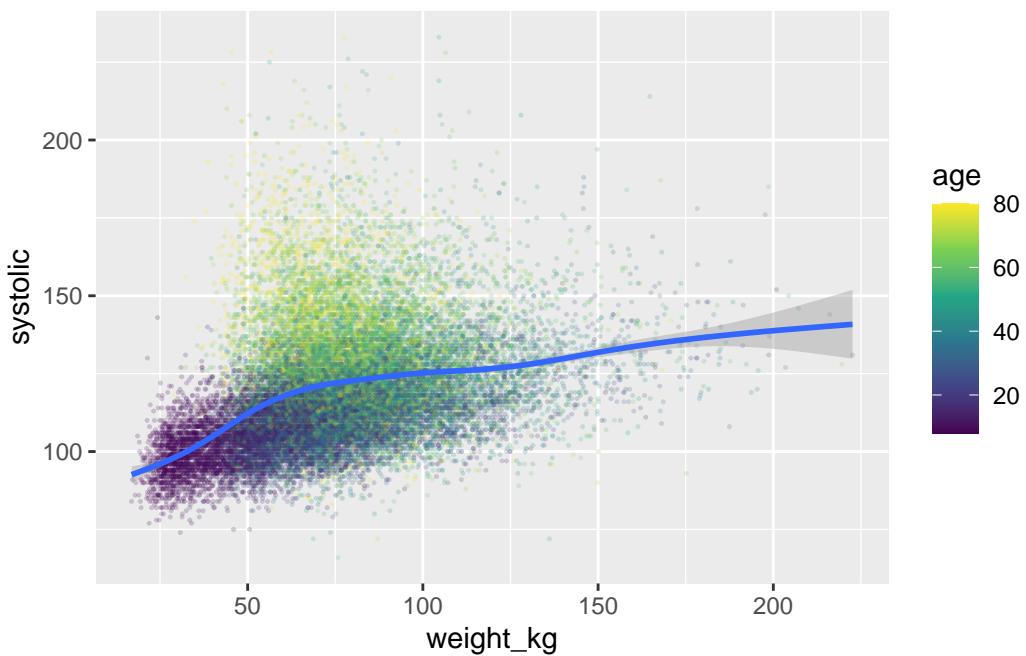
Чи мають люди старшого віку вищий артеріальний тиск?

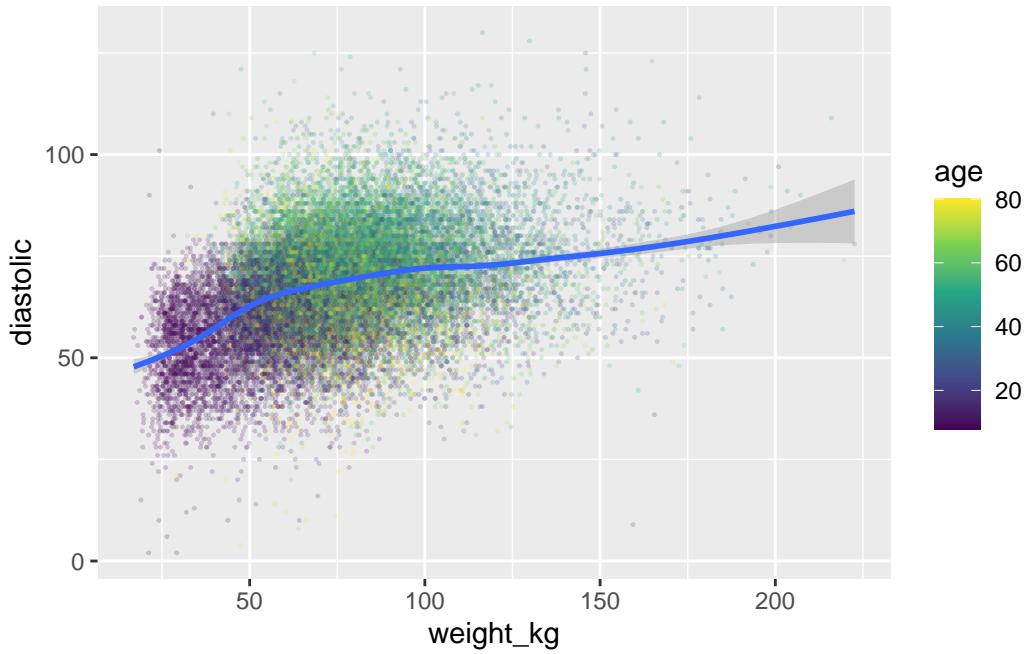




Вік має помітний вплив на артеріальний тиск, хоча вплив на систолічний та діастолічний тиск відрізняється.

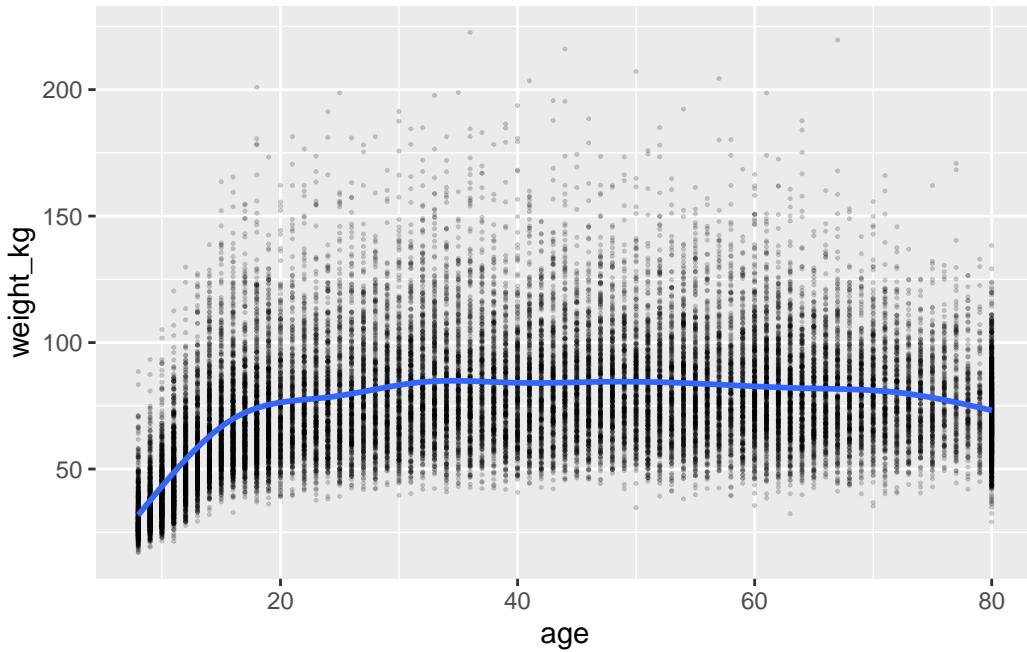
Чи мають люди з більшою вагою вищий артеріальний тиск?



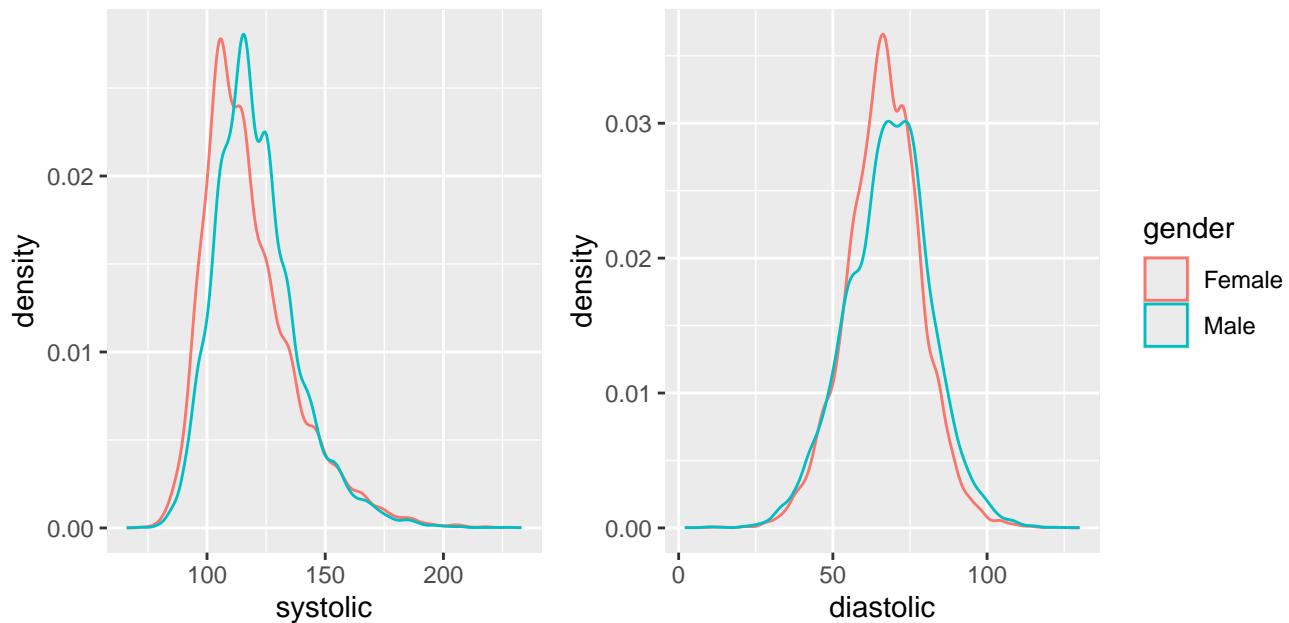


Різке зростання на початку графіку пояснюється швидким ростом молодих людей, у яких збільшується розмір кровоносних судин, що призводить до підвищення артеріального тиску. У дорослих також спостерігається підвищення артеріального тиску зі збільшенням ваги, хоча і не таке різке.

Дорослі не ростуть так інтенсивно, як діти:



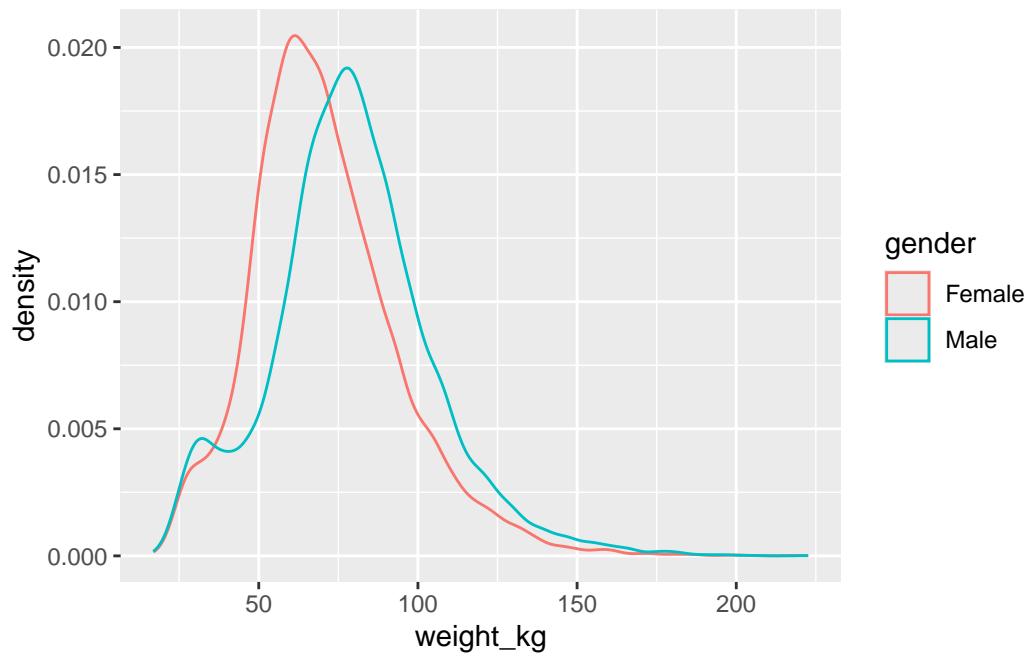
Чи впливає стать на артеріальний тиск?



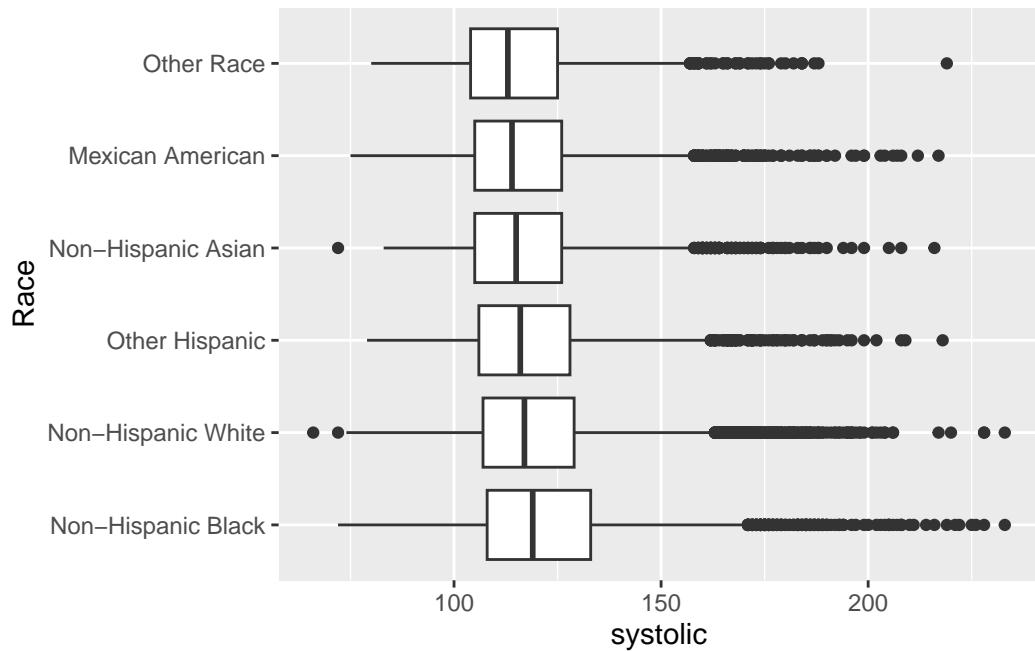
В середньому, у чоловіків спостерігається тенденція до дещо вищого артеріального тиску.

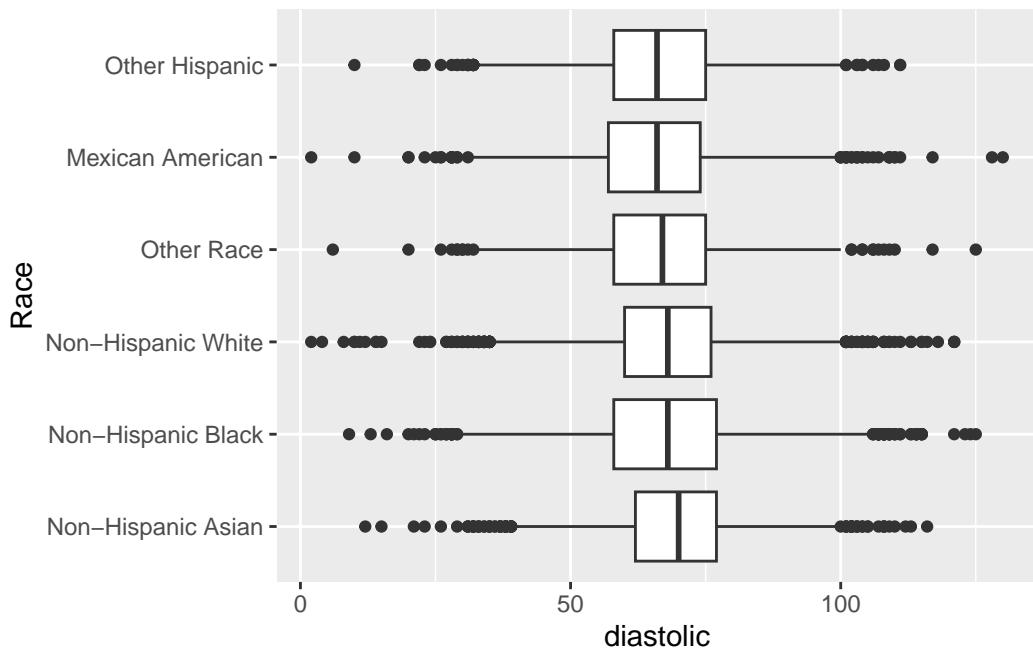
```
# A tibble: 2 x 5
  gender diastolic_mean diastolic_median systolic_mean systolic_median
  <chr>      <dbl>           <dbl>       <dbl>          <dbl>
1 Female       66.6            67        118.          114
2 Male         68.1            69        121.          118
```

Чоловіки та жінки також відрізняються за вагою.



Чи відрізняється артеріальний тиск залежно від раси?





```
# A tibble: 6 x 3
  race          systolic_median diastolic_median
  <chr>           <dbl>            <dbl>
1 Other Race      113              67
2 Mexican American 114              66
3 Non-Hispanic Asian 115             70
4 Other Hispanic   116              66
5 Non-Hispanic White 117              68
6 Non-Hispanic Black 119              68
```

Загалом, суттєвої різниці не спостерігається. Однак, у середньому, Non-Hispanic Black мають дещо вищий систолічний артеріальний тиск порівняно з іншими расами.

3.2 Як фактори способу життя впливають на артеріальний тиск?

Для відповіді на це питання використаємо вибірку `lifestyle_data`.

Дескриптивні статистики

Числові змінні

variable name	min	1st quartile	mean	median	3rd quartile	max
diastolic	8.0	64.00	70.86	71.00	78.00	130.0
systolic	66.0	111.00	124.10	121.00	134.00	233.0
age	20.0	34.00	49.15	49.00	63.00	80.0
weight_kg	32.3	66.40	82.02	78.60	93.60	222.6

variable name	min	1st quartile	mean	median	3rd quartile	max
poverty_index	0.0	1.01	2.30	1.84	3.54	5.0
sleep	2.0	6.00	7.23	7.00	8.00	14.5

Можна помітити, що середній систолічний та діастолічний артеріальний тиск вищий, ніж у вибірці `whole_data`. Це пояснюється тим, що ця вибірка містить лише людей віком 20 років і старше (див. мінімальне значення для віку).

Категорійні змінні

```
# A tibble: 4 x 3
  blood_pressure_category percentage     n
  <ord>                  <dbl> <int>
1 Normal                 0.555 15452
2 Elevated               0.139 3858
3 Hypertension Stage 1  0.160 4462
4 Hypertension Stage 2  0.146 4074

# A tibble: 2 x 3
  gender percentage     n
  <chr>      <dbl> <int>
1 Female     0.514  8942
2 Male       0.486  8451

# A tibble: 6 x 3
  race          percentage     n
  <chr>         <dbl> <int>
1 Mexican American 0.128  2218
2 Non-Hispanic Asian 0.125  2172
3 Non-Hispanic Black 0.221  3842
4 Non-Hispanic White 0.390  6779
5 Other Hispanic   0.101  1755
6 Other Race       0.0360 627

# A tibble: 5 x 3
  education          percentage     n
  <ord>              <dbl> <int>
1 Less than 9th grade 0.0848 1475
2 9-11th grade        0.122  2118
3 High school graduate 0.221  3842
4 Some college degree  0.317  5507
5 College graduate or above 0.256 4451
```

```

# A tibble: 8 x 3
  vigorous_phys_activity_days percentage     n
  <dbl>            <dbl> <int>
1 0        0.764 13292
2 1        0.0324 564
3 2        0.0475 827
4 3        0.0650 1131
5 4        0.0386 671
6 5        0.0306 533
7 6        0.0105 182
8 7        0.0111 193

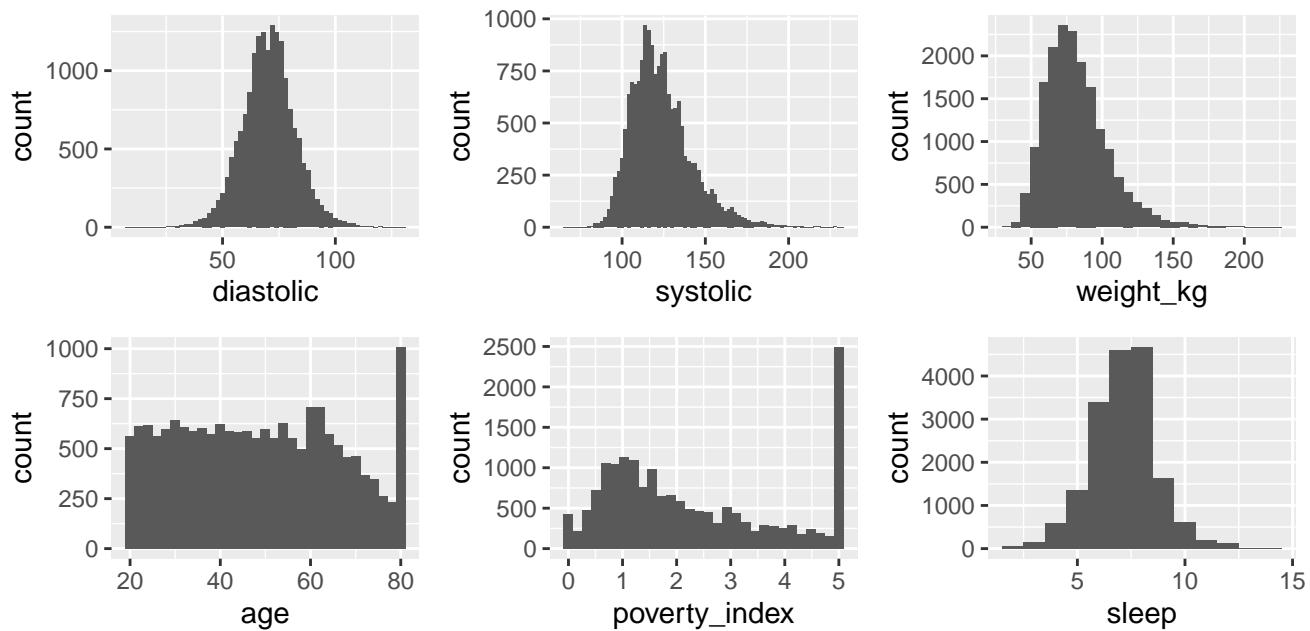
# A tibble: 8 x 3
  moderate_phys_activity_days percentage     n
  <dbl>            <dbl> <int>
1 0        0.581 10101
2 1        0.0554 964
3 2        0.0877 1525
4 3        0.101   1762
5 4        0.0498 866
6 5        0.0598 1040
7 6        0.0156 272
8 7        0.0496 863

# A tibble: 2 x 3
  smoked_100_cigs percentage     n
  <chr>            <dbl> <int>
1 No              0.567  9867
2 Yes             0.433  7526

# A tibble: 5 x 3
  diet_quality percentage     n
  <ord>            <dbl> <int>
1 Poor           0.0617 1073
2 Fair            0.239  4160
3 Good            0.412  7159
4 Very good      0.204  3545
5 Excellent       0.0837 1456

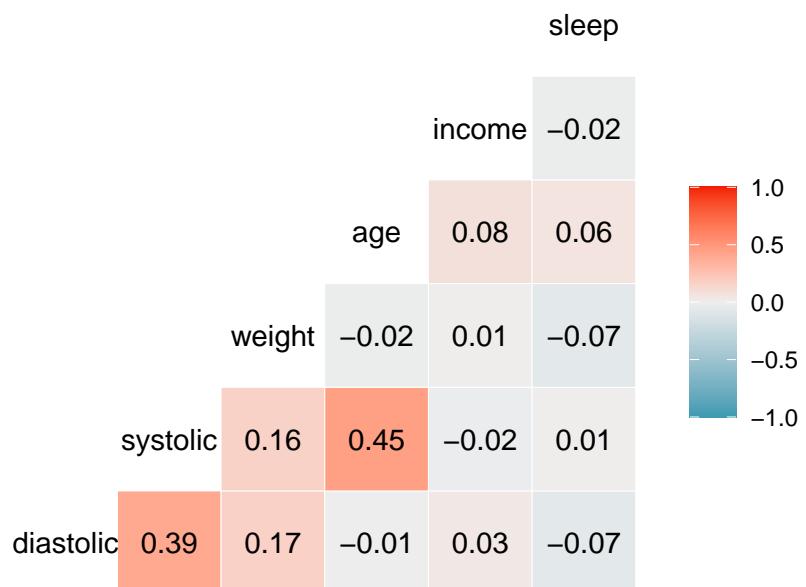
```

Розподіл числових змінних



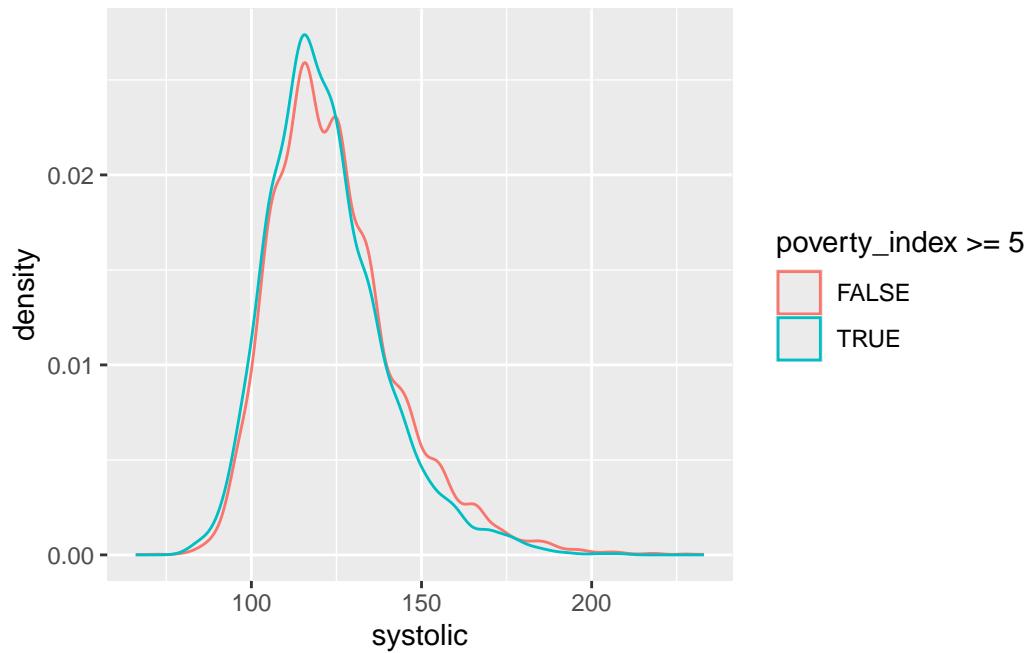
Який кореляційний зв'язок між доходом, кількістю сну та артеріальним тиском?

Були розраховані наступні коефіцієнти кореляції Спірмена:

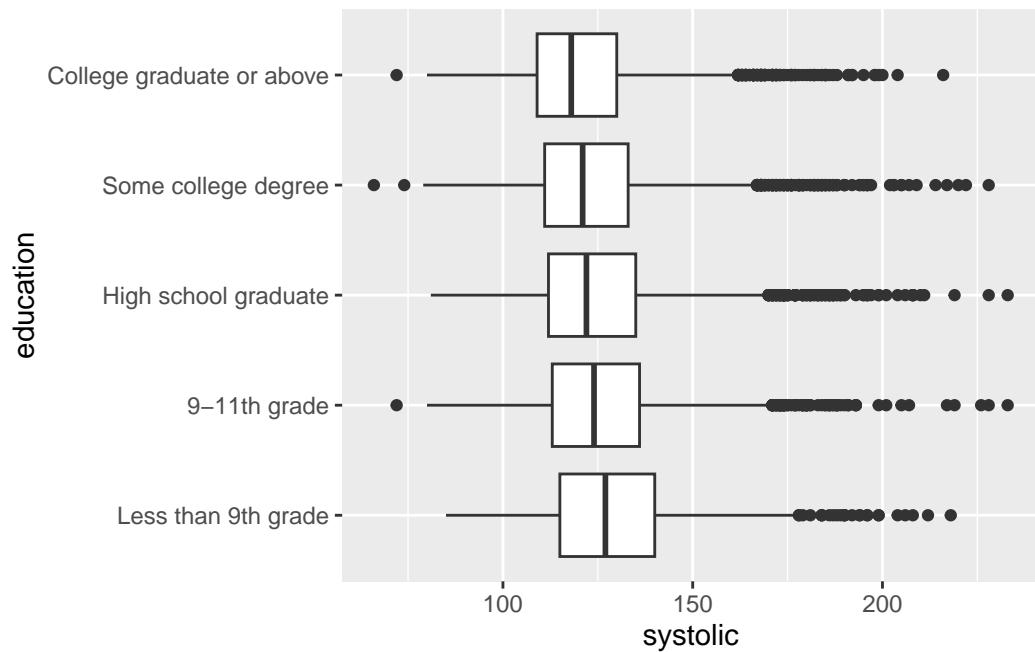


Сон та дохід не демонструють чіткого впливу на артеріальний тиск.

Навіть у заможних людей (з індексом бідності ≥ 5) розподіл артеріального тиску не відрізняється:

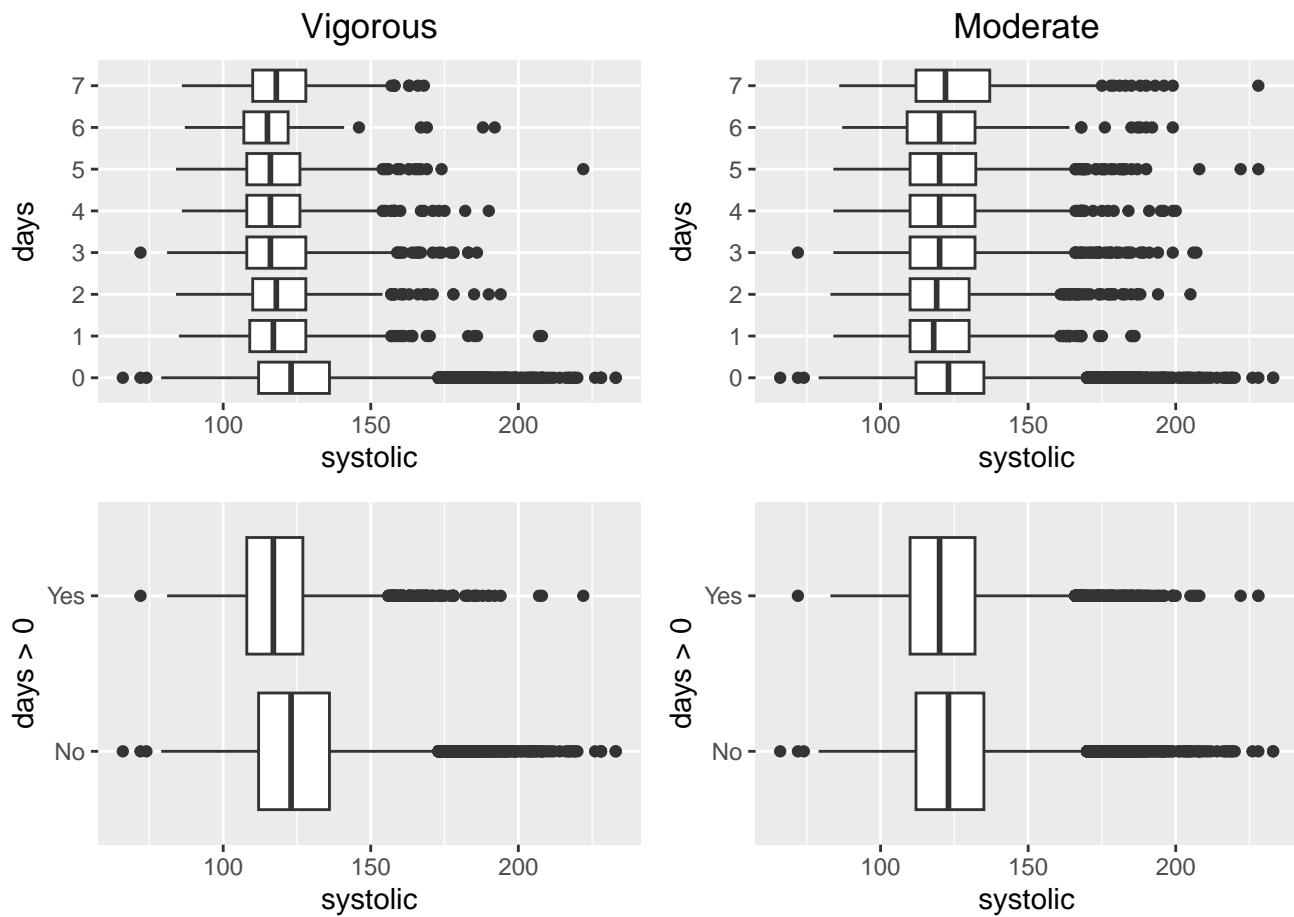


Чи мають люди з нижчим рівнем освіти менш здоровий артеріальний тиск?



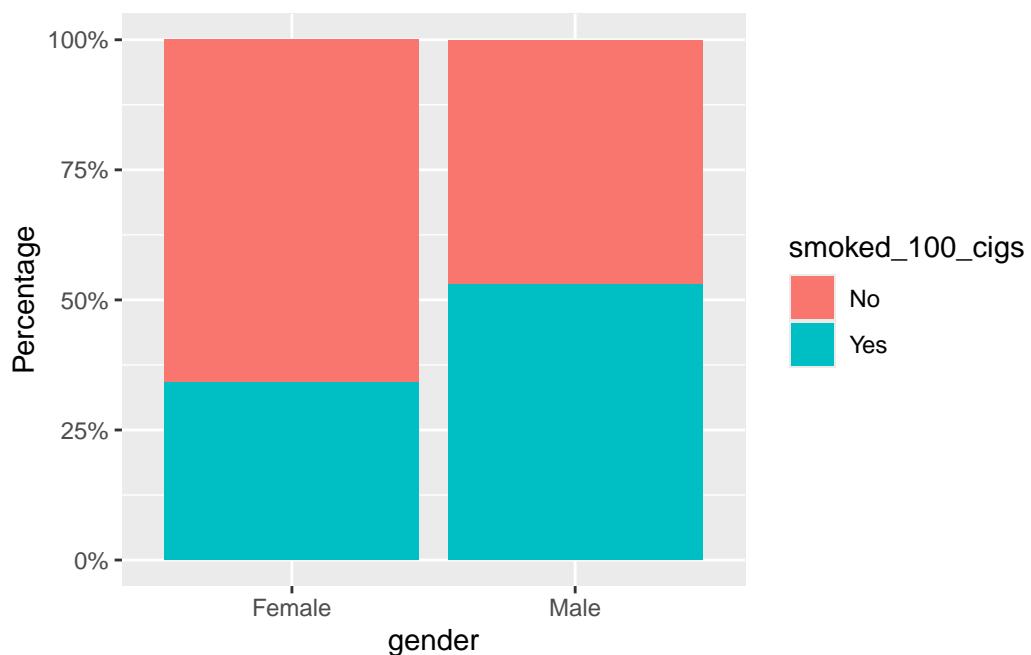
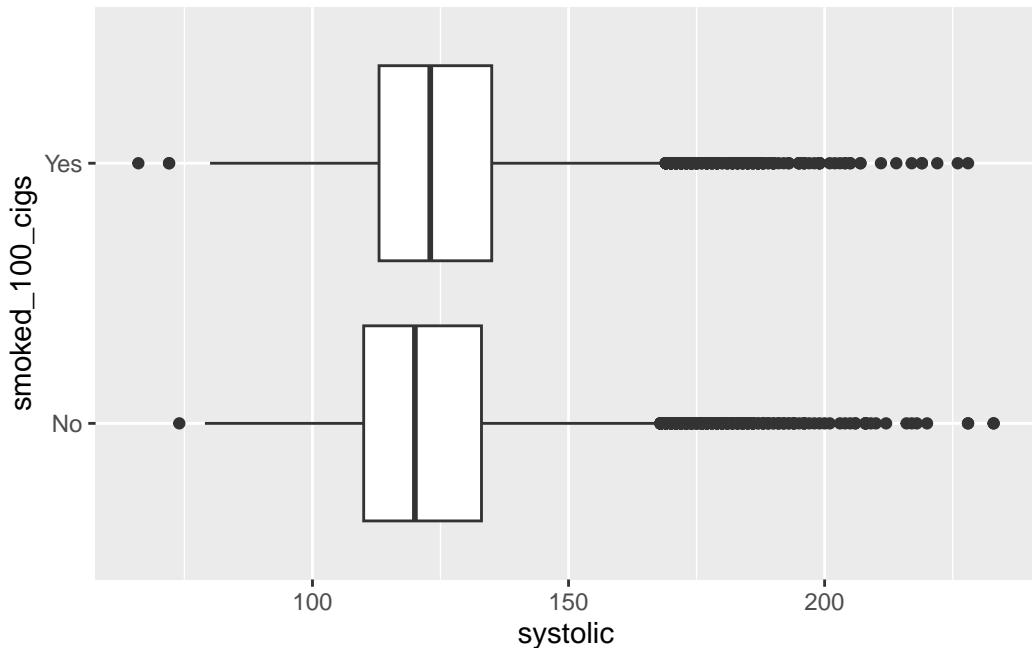
В середньому, спостерігається незначне підвищення артеріального тиску зі зменшенням освіти.

Чи мають фізично активні люди здоровіший артеріальний тиск?



В середньому, люди, які не є фізично активними, мають вищий систолічний артеріальний тиск.

Чи мають люди, які більше курять, гірший артеріальний тиск?

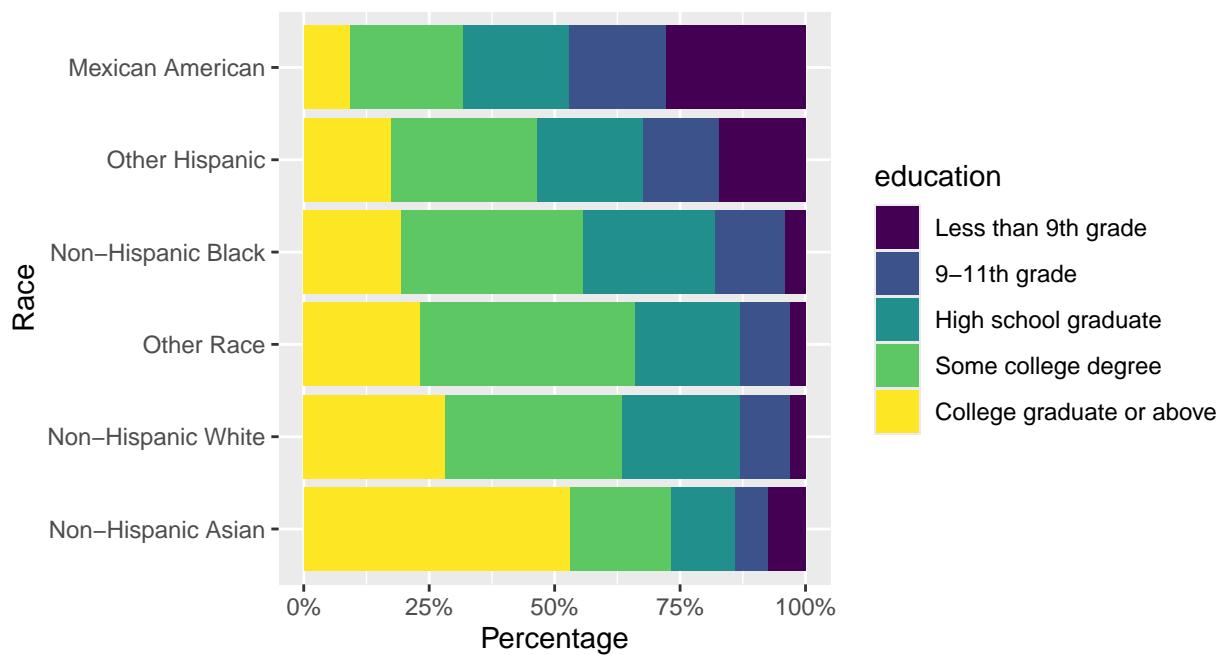
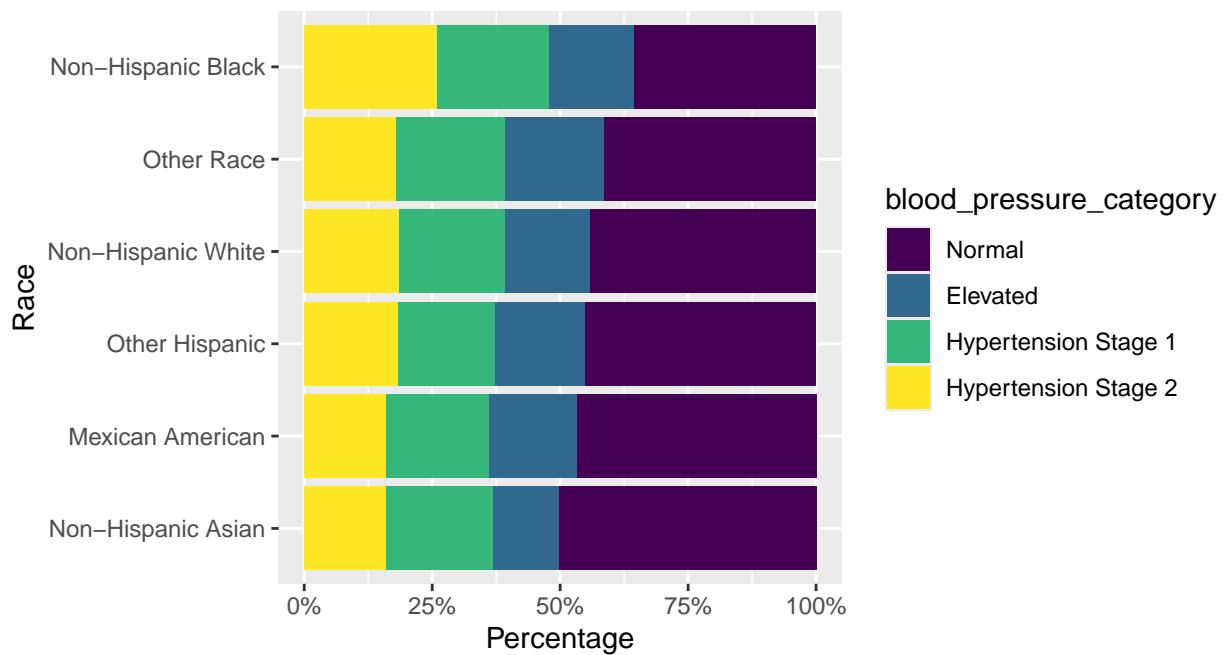


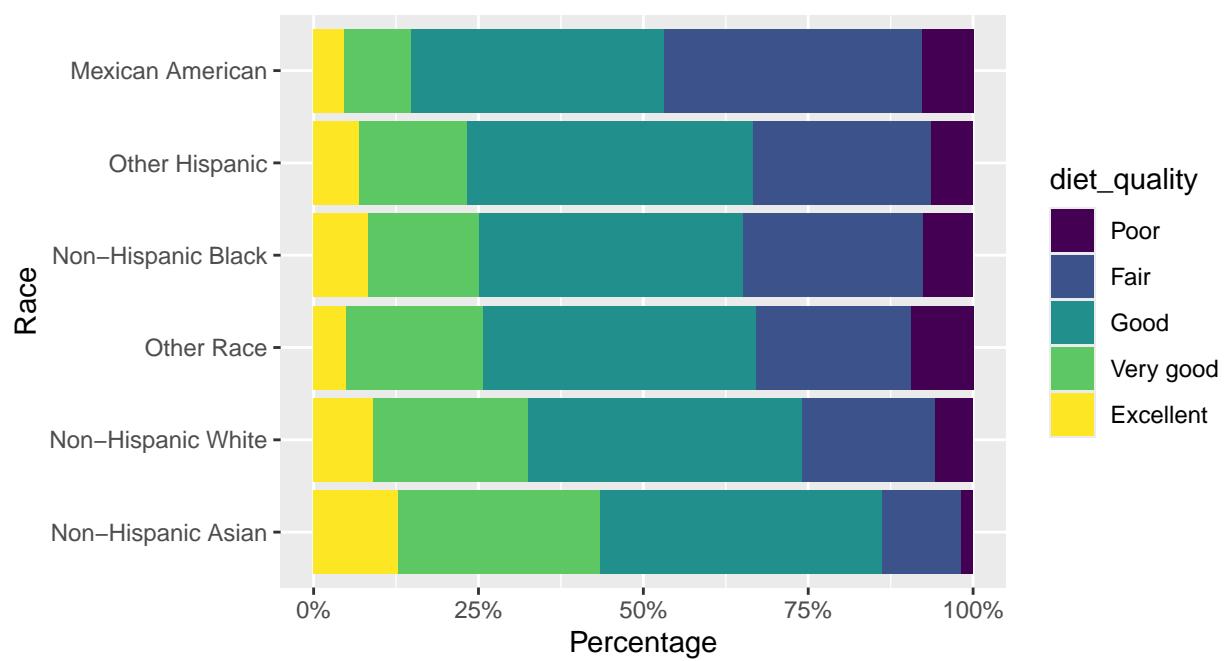
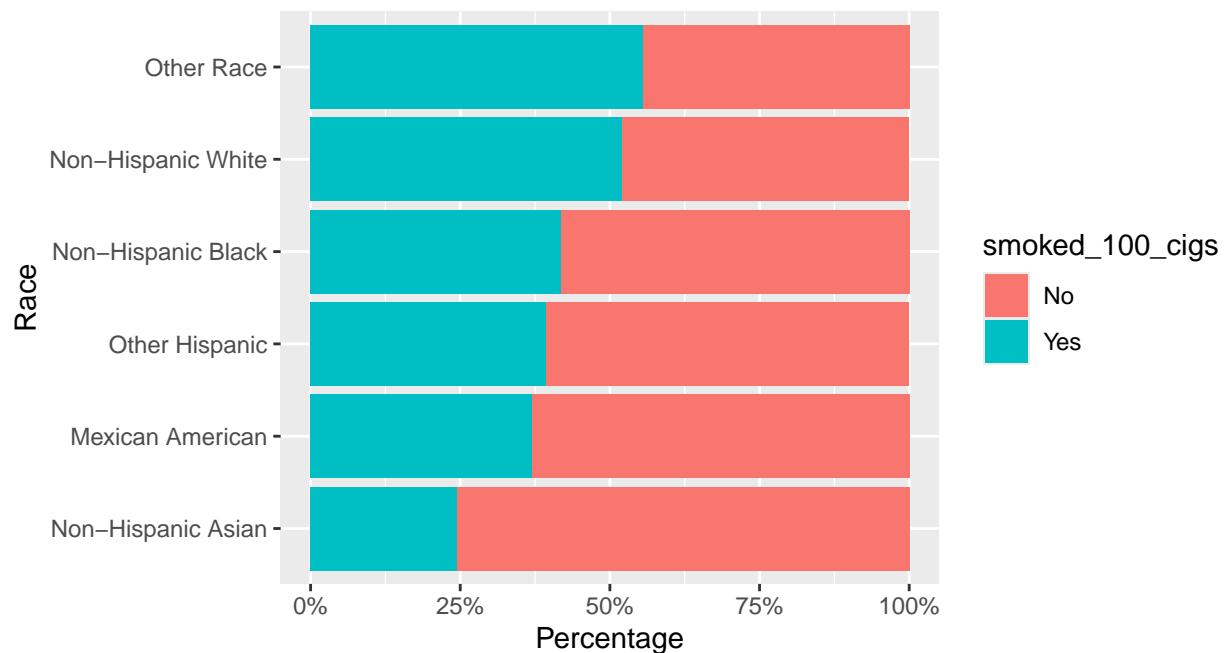
В середньому, у людей, які викурили щонайменше 100 цигарок, спостерігається вищий систолічний артеріальний тиск. Крім того, відсоток таких людей вищий серед чоловіків, ніж серед жінок.

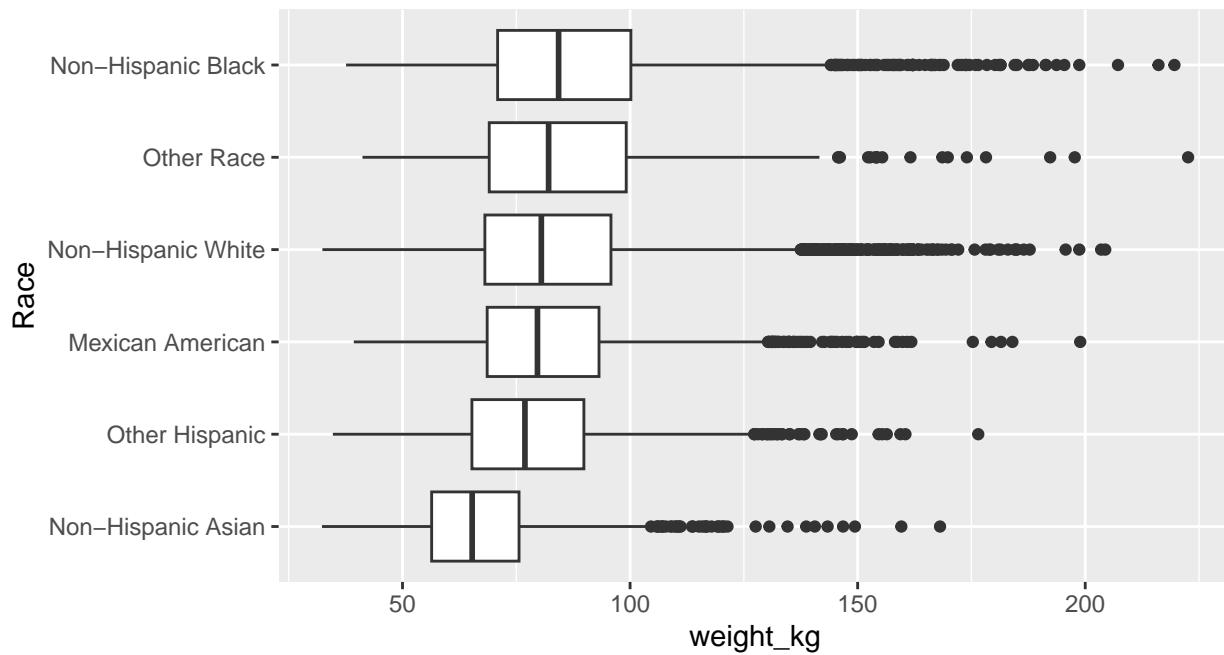
```
# A tibble: 2 x 2
  gender smoked_100_cigs_percent
```

<chr>	<dbl>
1 Female	0.341
2 Male	0.530

Як спосіб життя відрізняється залежно від раси?





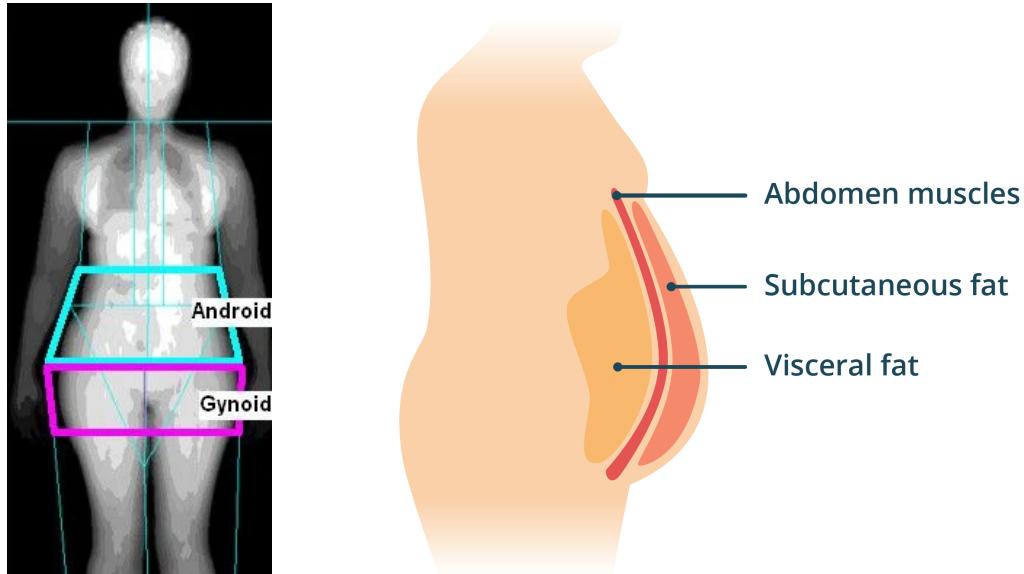


Цікаво, що Non-Hispanic Asian демонструють кращі показники в усіх аспектах.

3.3 Як розподіл жирової тканини впливає на артеріальний тиск?

Для відповіді на це питання використаємо вибірку fat_data.

Вибірка містить вимірювання кількості android (андроїдного) та gynoid (гіноїдного) жиру, а також вимірювання кількості visceral (вісцерального) та subcutaneous (підшкірного) жиру.



Дескриптивні статистики

Числові змінні

variable name	min	1st quartile	mean	median	3rd quartile	max
diastolic	2.00	58.00	66.75	67.00	75.00	128.00
systolic	74.00	104.00	114.59	113.00	123.00	233.00
age	8.00	15.00	29.68	28.00	44.00	59.00
weight_kg	16.90	55.80	72.67	71.00	87.60	203.50
waist_cm	46.30	76.00	90.06	88.70	102.30	176.00
subcutaneous_fat_g	85.99	749.96	1404.94	1277.92	1914.31	5494.25
visceral_fat_g	0.00	192.04	393.71	317.97	539.24	2003.75
android_fat_g	151.40	902.77	1979.14	1700.20	2696.08	10626.50
android_non_fat_g	636.70	2530.98	3521.91	3405.35	4400.88	10207.70
android_fat_percent	0.10	0.26	0.33	0.34	0.41	0.60
gynoid_fat_g	436.00	2641.00	4065.97	3745.30	5136.60	15904.50
gynoid_non_fat_g	1223.50	5591.98	7357.54	7275.60	9047.80	18544.40
gynoid_fat_percent	0.11	0.28	0.35	0.35	0.42	0.62

Варто зауважити що ця вибірка містить лише людей віком до 59 років (див. максимальне значення для віку).

Категорійні змінні

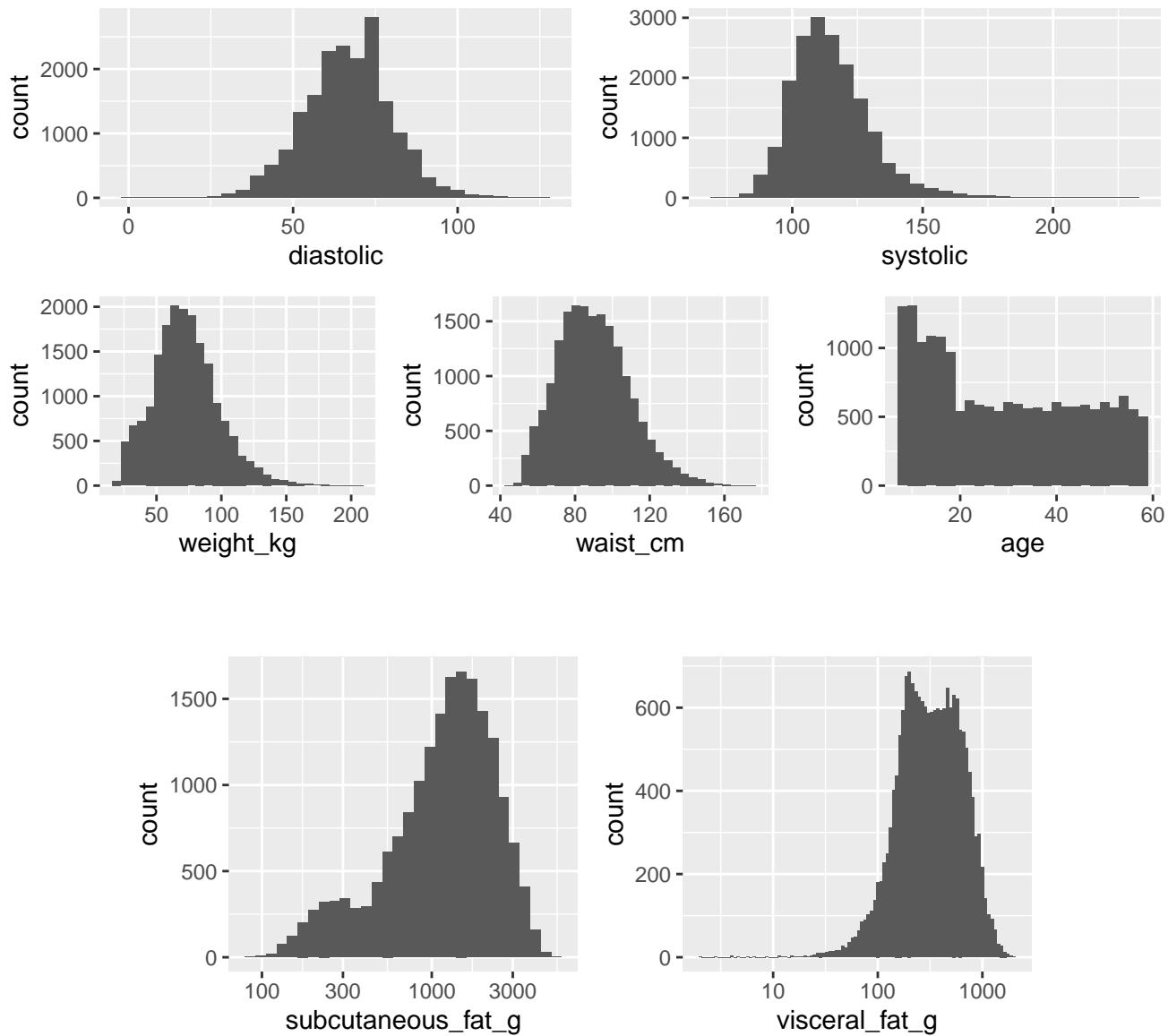
```
# A tibble: 4 x 3
  blood_pressure_category percentage     n
  <ord>                  <dbl> <int>
1 Normal                 0.555 15452
2 Elevated                0.139 3858
3 Hypertension Stage 1   0.160 4462
4 Hypertension Stage 2   0.146 4074

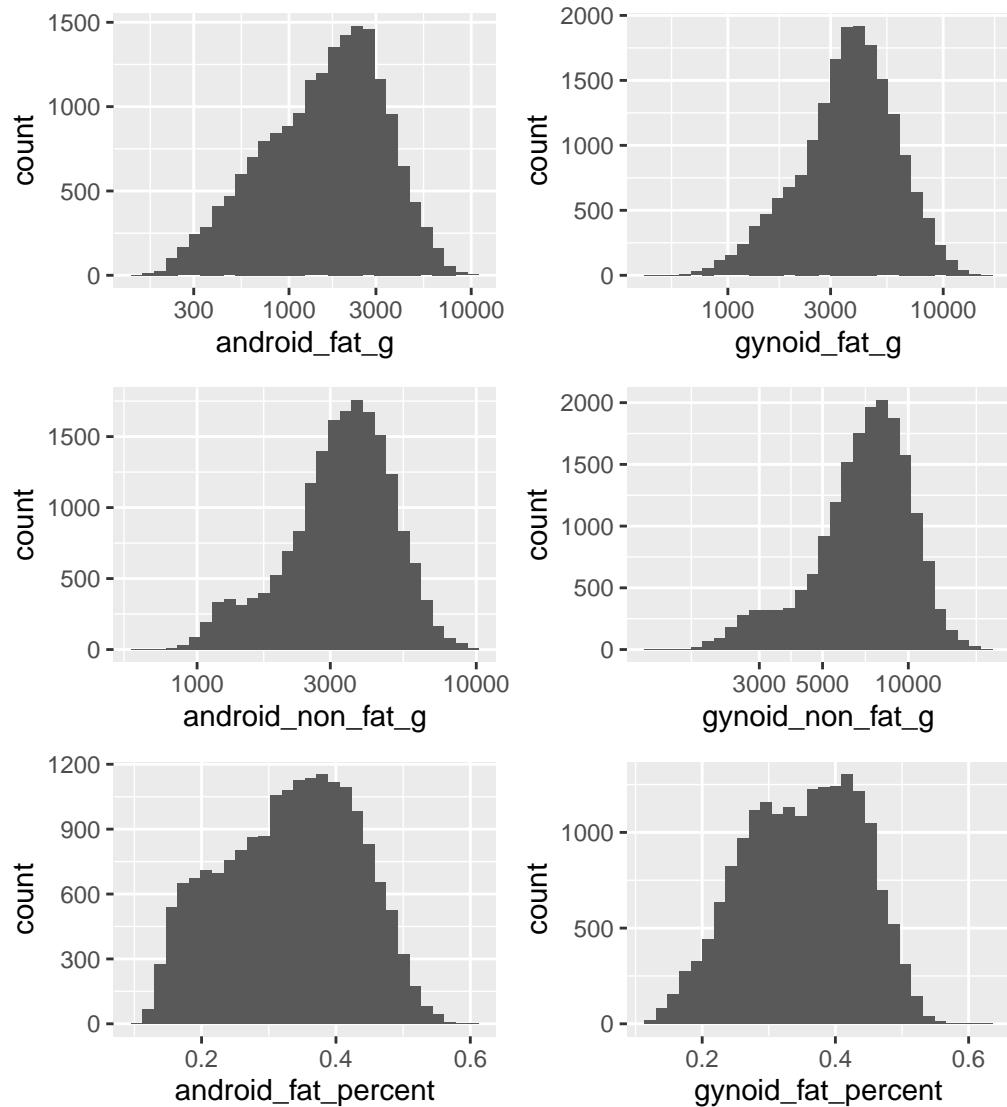
# A tibble: 2 x 3
  gender percentage     n
  <chr>      <dbl> <int>
1 Female     0.514  8942
2 Male       0.486  8451

# A tibble: 6 x 3
  race          percentage     n
  <chr>         <dbl> <int>
1 Mexican American 0.128  2218
2 Non-Hispanic Asian 0.125  2172
3 Non-Hispanic Black 0.221  3842
4 Non-Hispanic White 0.390  6779
5 Other Hispanic    0.101  1755
6 Other Race        0.0360 627
```

Розподіл числових змінних

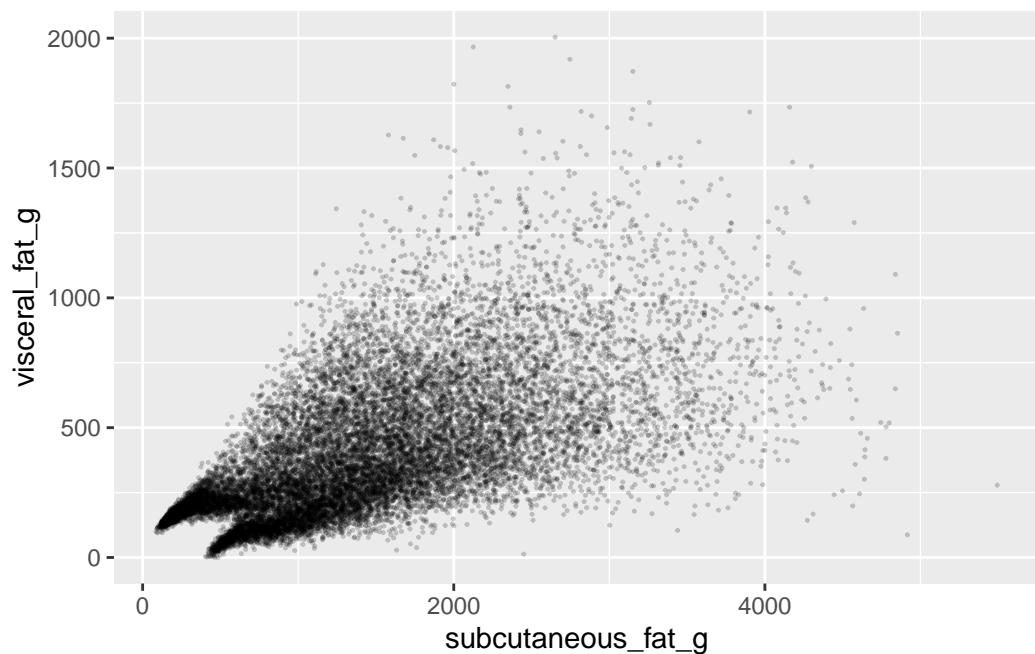
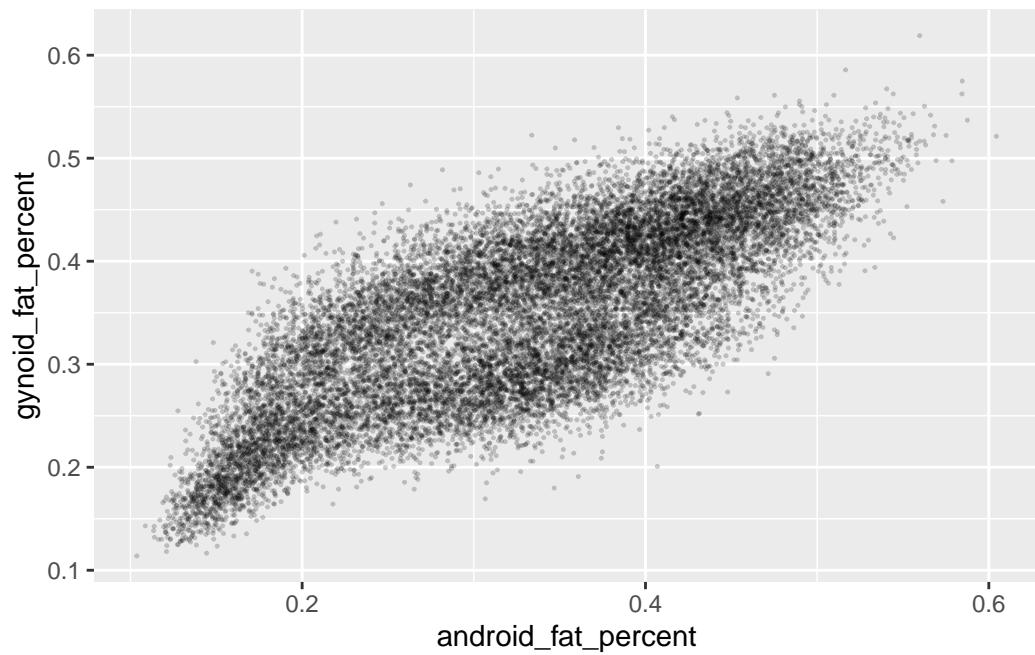
Змінні, що представляють кількість жиру, відображені в логарифмічному масштабі.



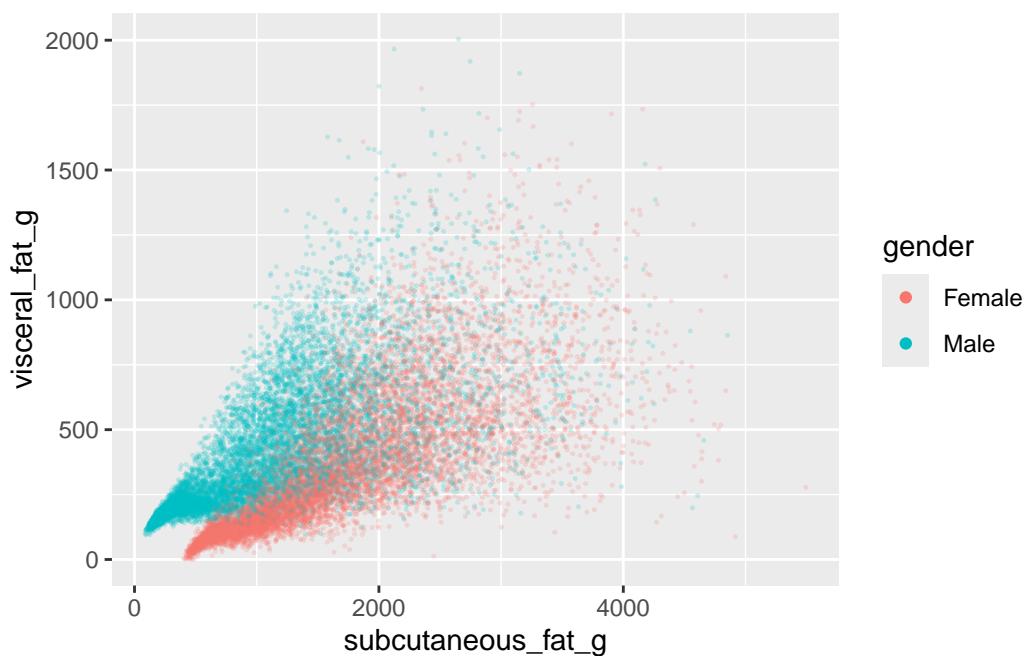
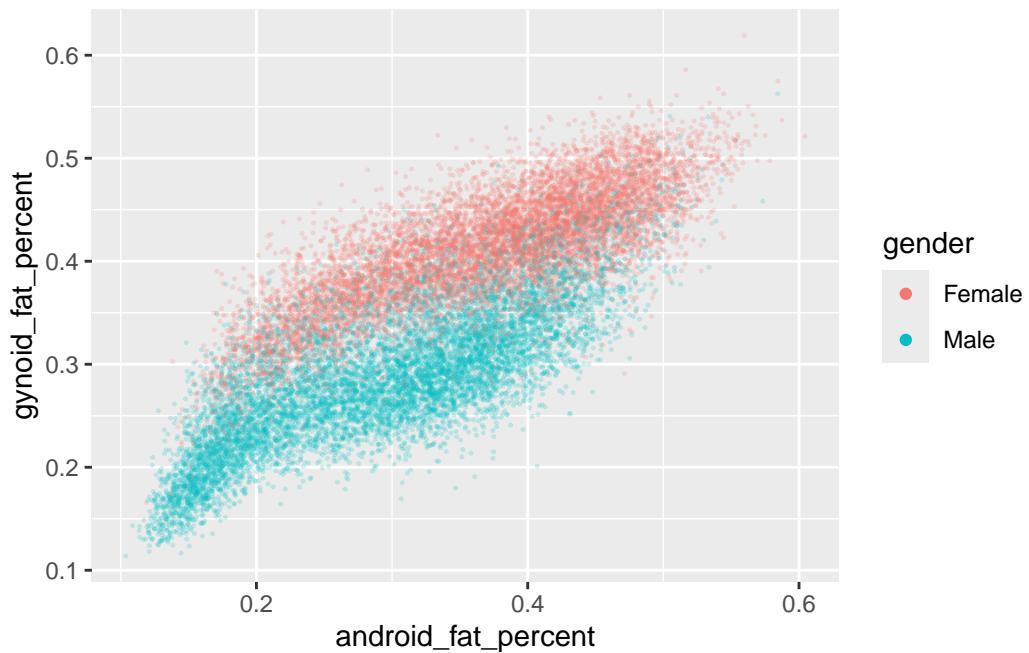


Відмінності між чоловіками та жінками

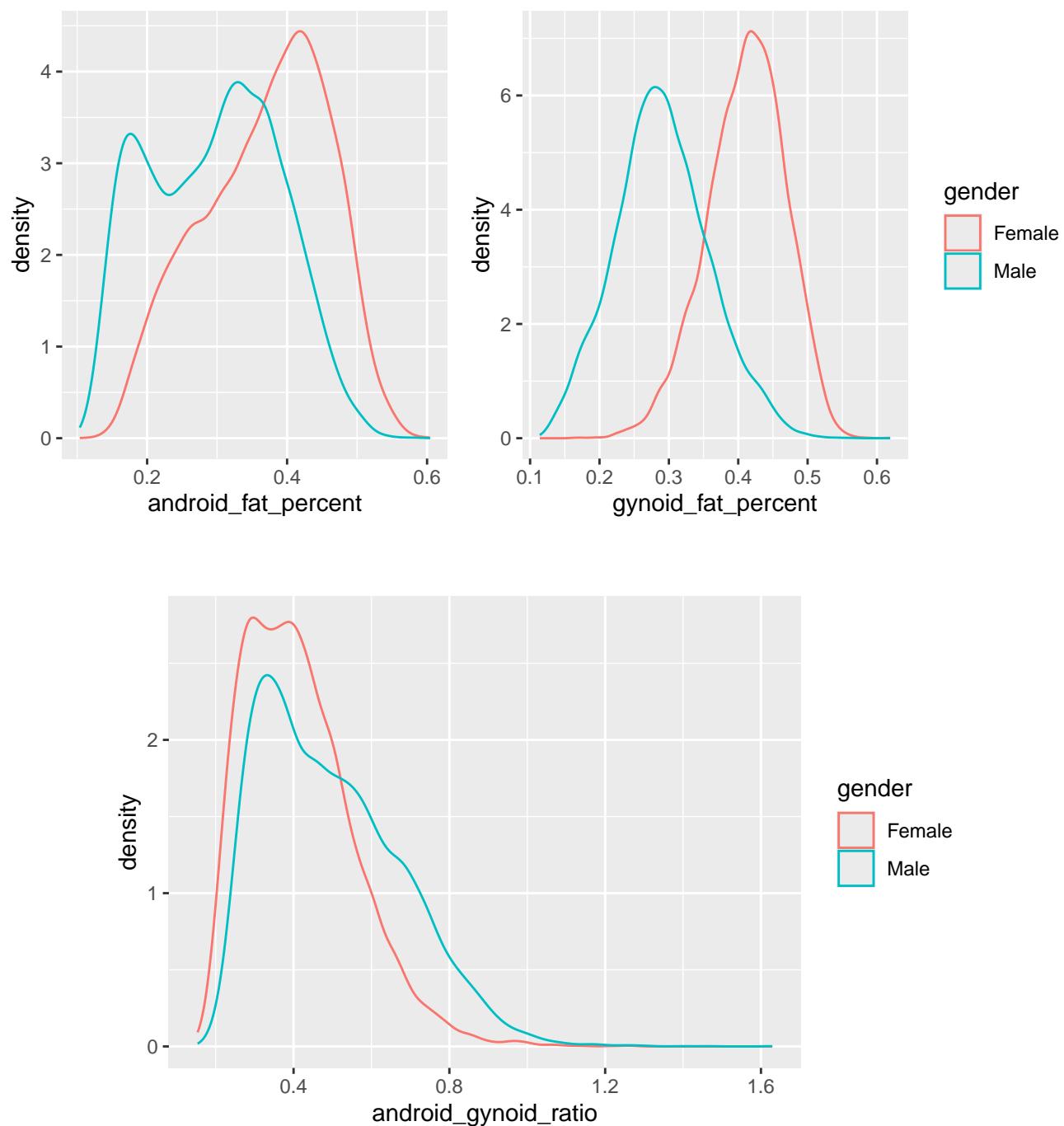
Розглянемо наступні два графіки.



Можна помітити два кластери. Їх пояснення полягає в тому, що у чоловіків та жінок різний розподіл жирової тканини.



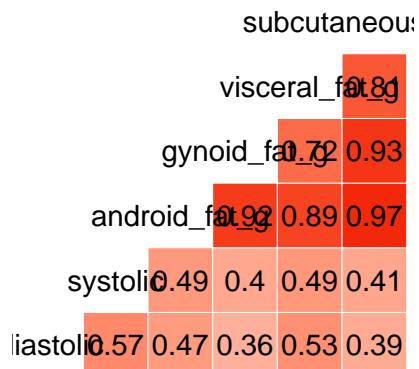
У жінок відсоток жирової тканини вищий, ніж у чоловіків, та жінки більш склонні до накопичення жиру в гіноїдній області, ніж чоловіки.



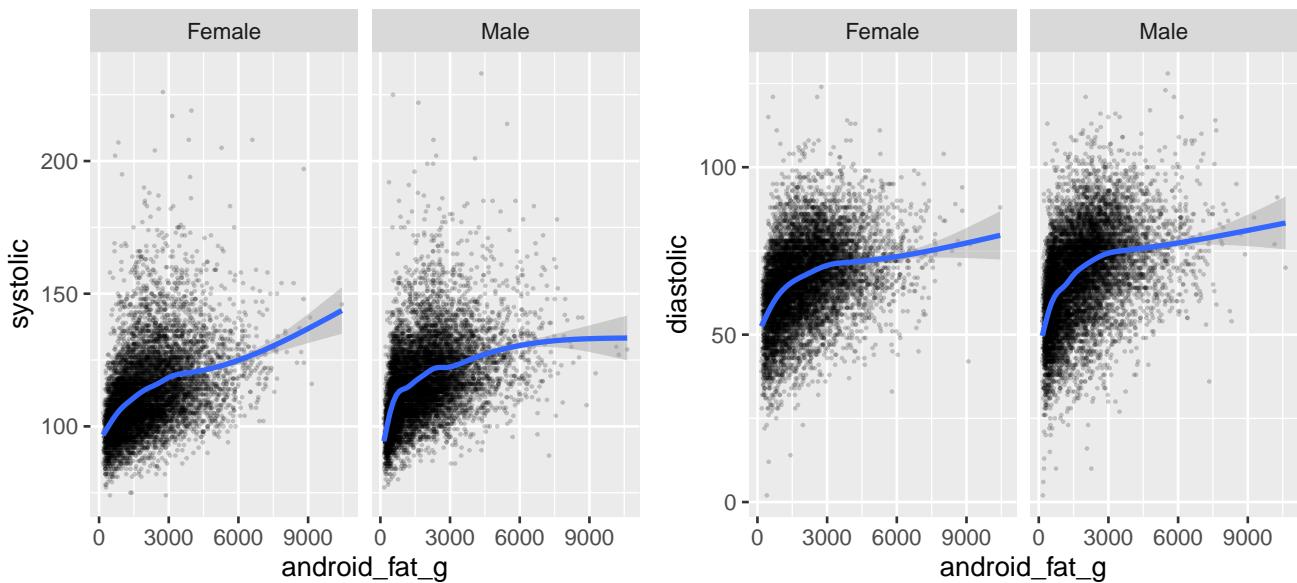
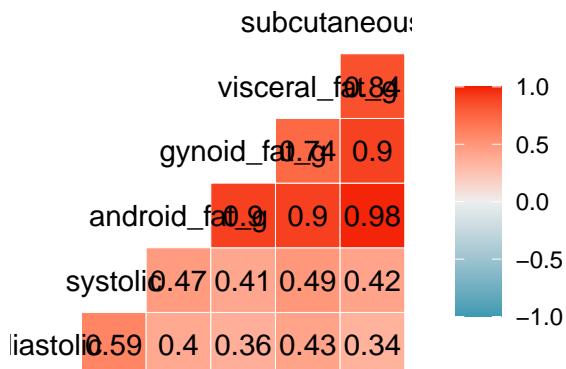
Тому, при дослідженні впливу жиру на артеріальний тиск, ми будемо розглядати чоловіків та жінок окремо.

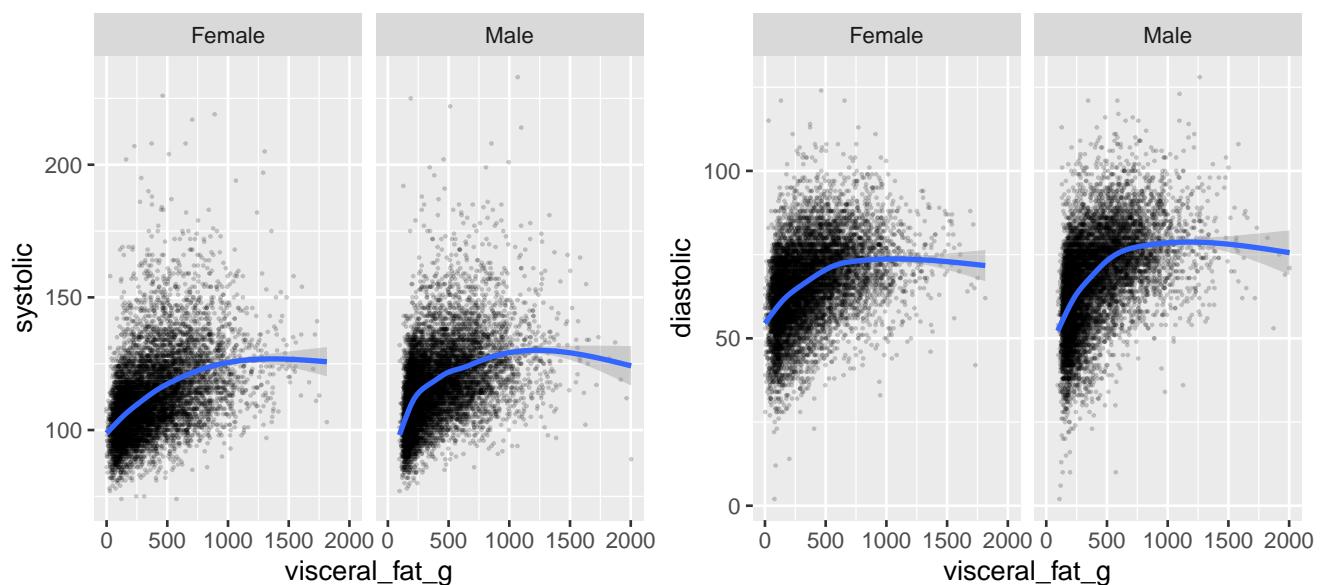
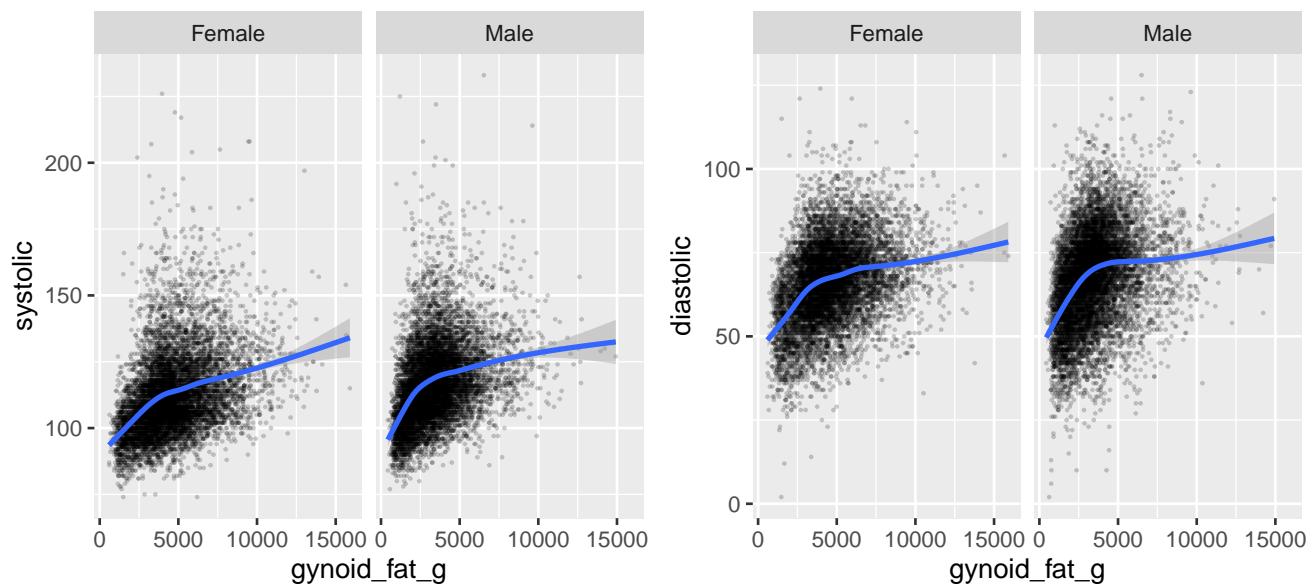
Чи свідчить більша кількість жиру про гірший артеріальний тиск?

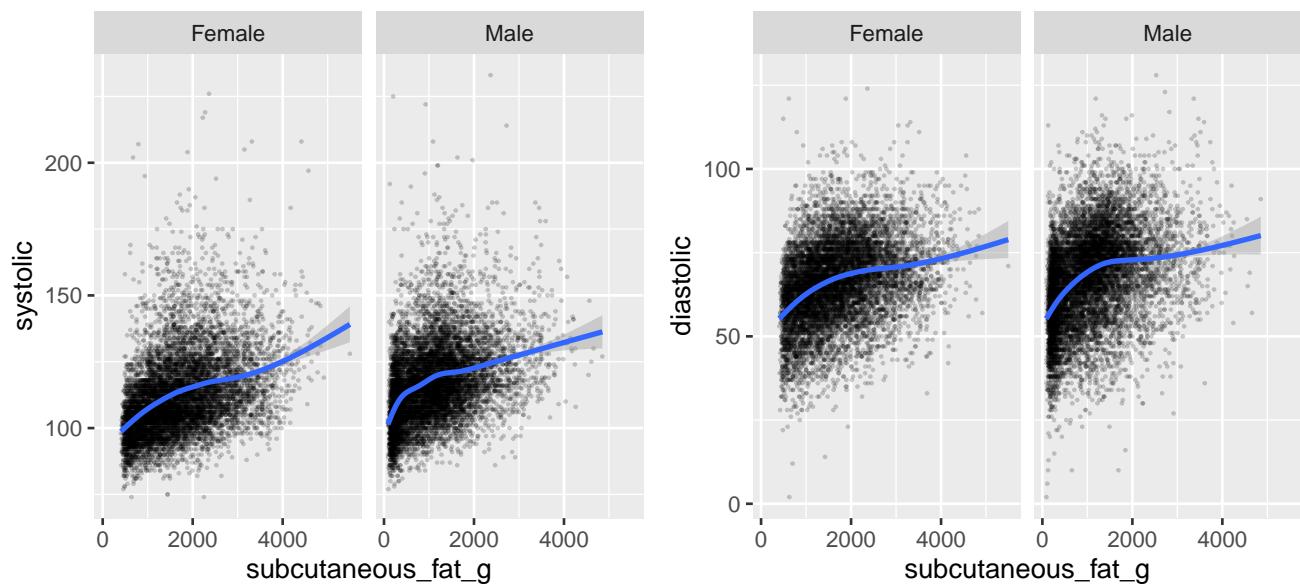
Male



Female

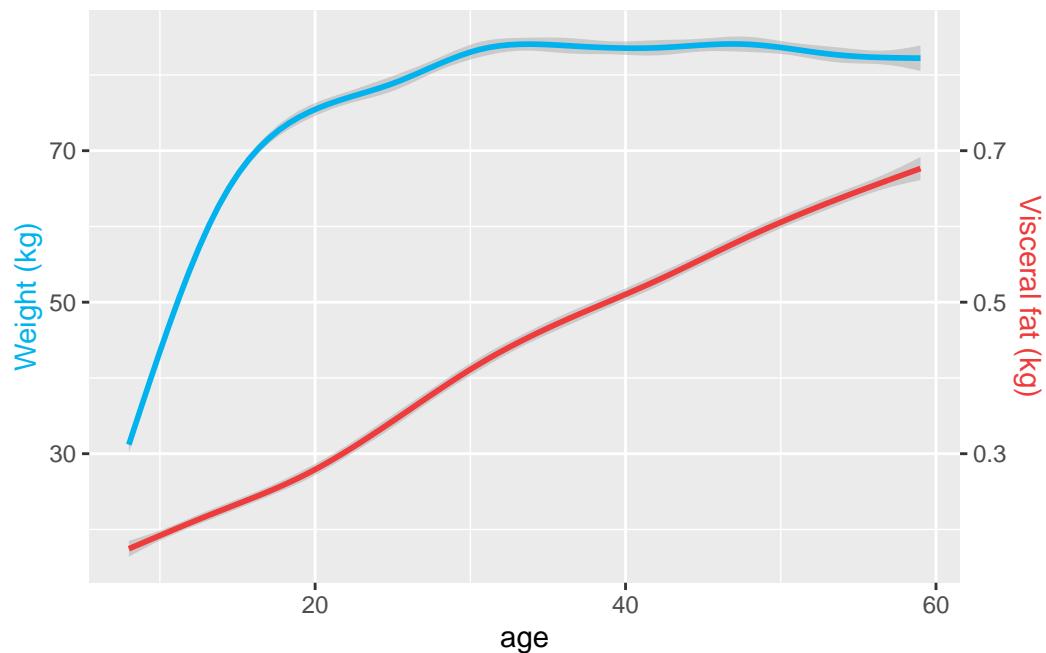


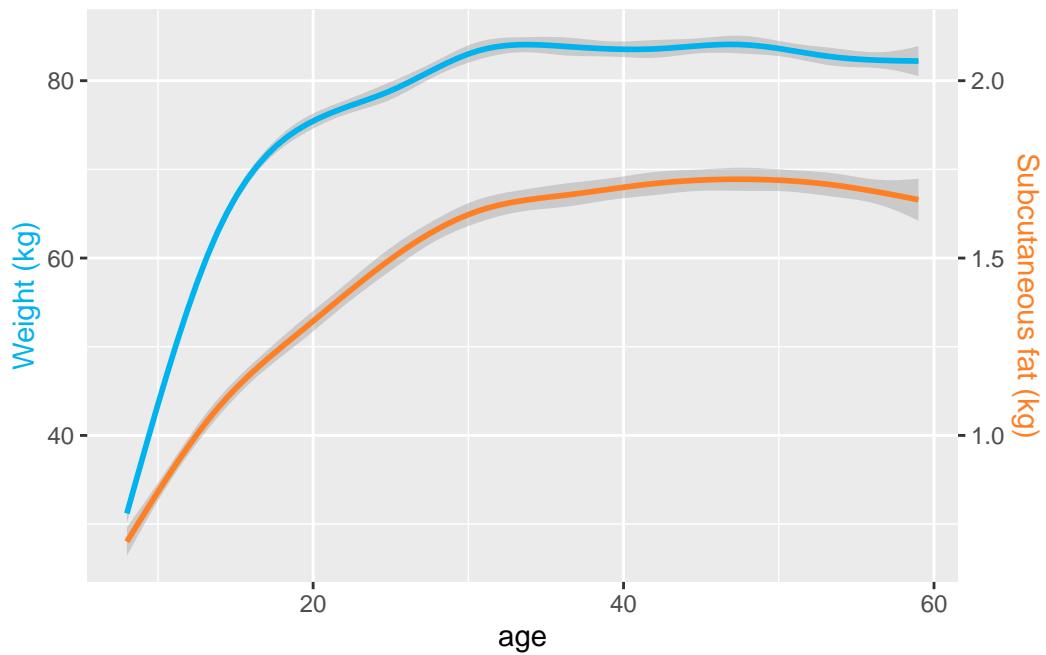




В середньому, більша кількість жиру пов'язана з вищим артеріальним тиском. Visceral жир, здається, має найсильніший вплив.

Також, visceral жир в середньому збільшується з віком, та продовжує збільшуватися у пізні роки на відміну від ваги та підшкірного жиру.

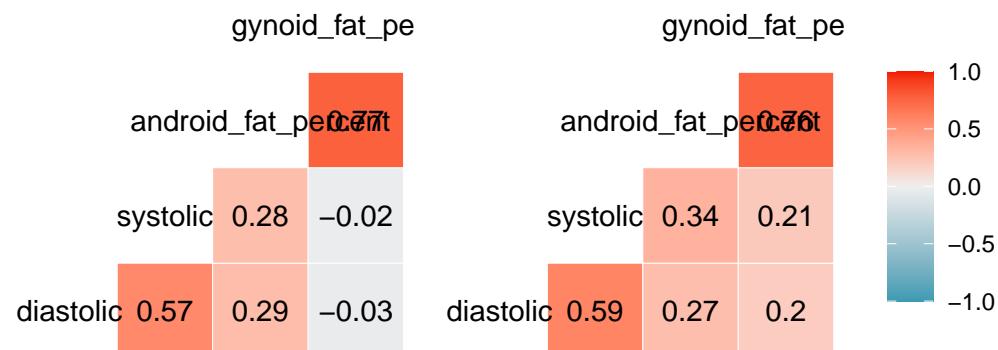


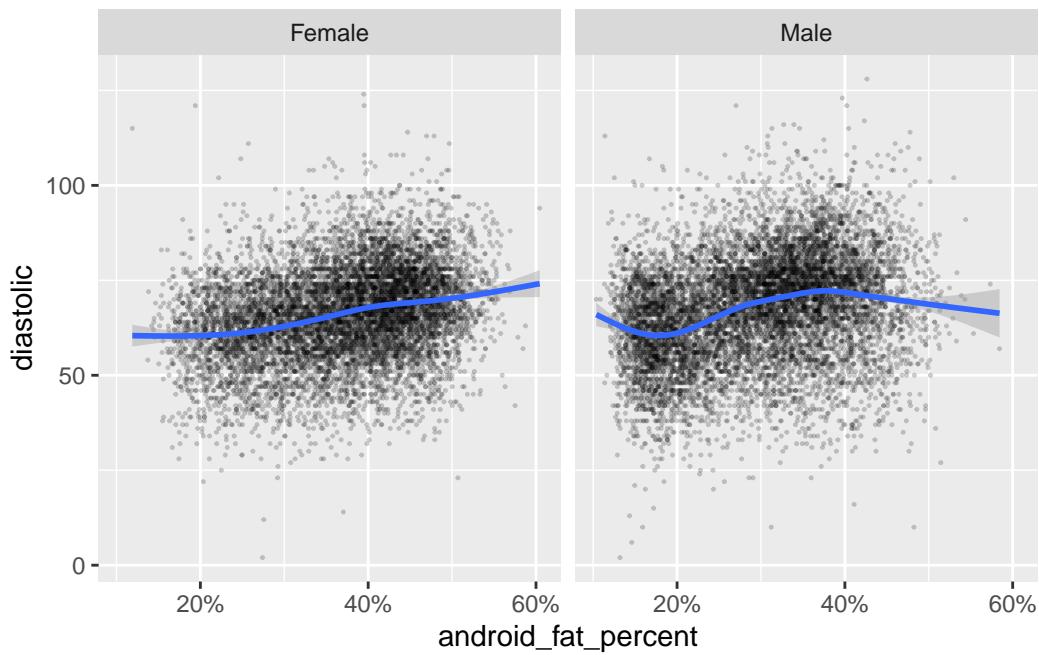
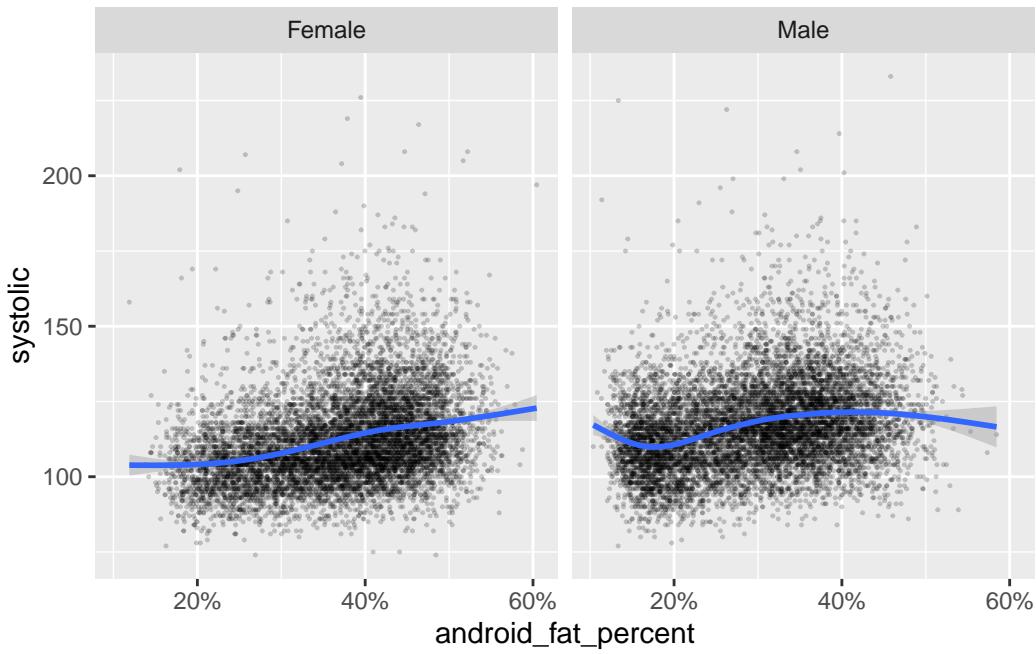


Що більше впливає на підвищення артеріального тиску: жирова чи нежирова маса?

Male

Female



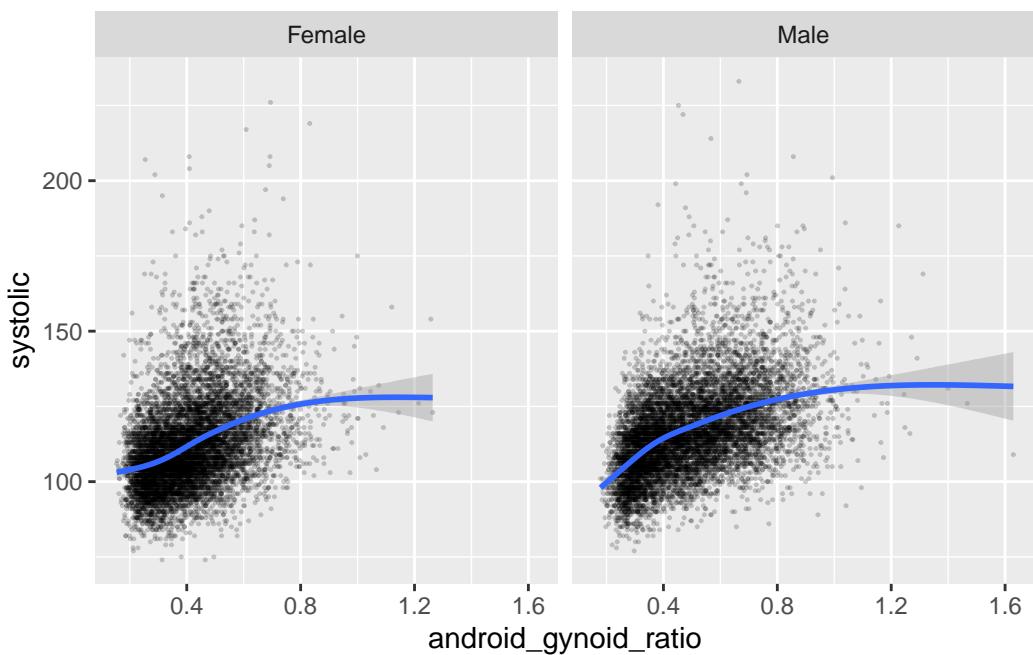
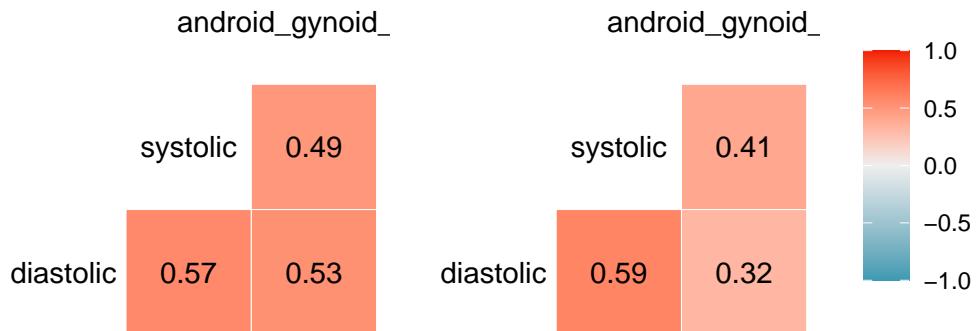


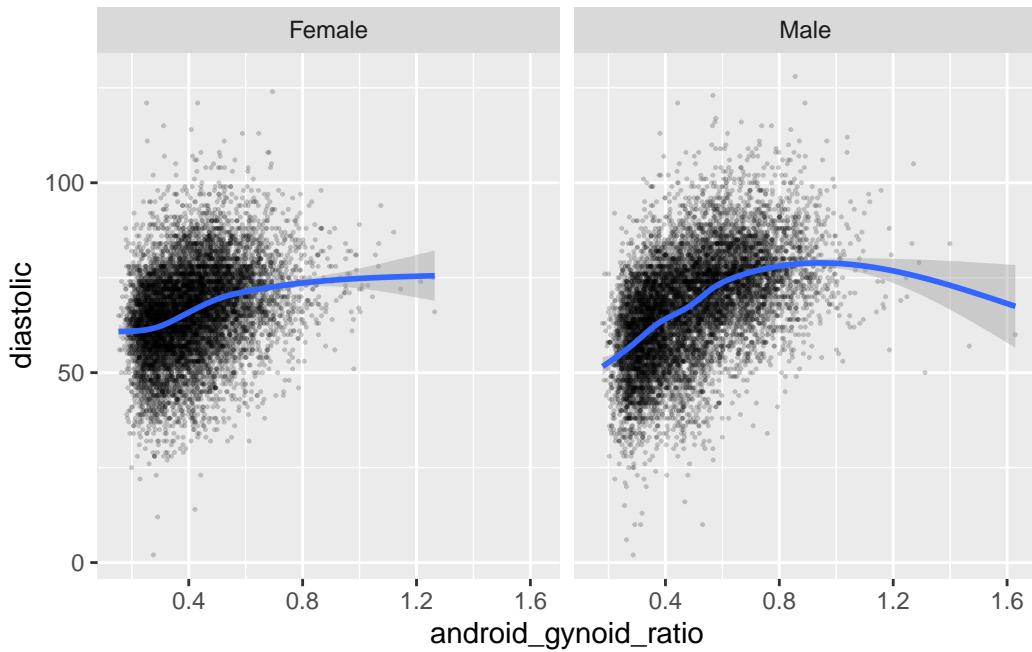
Відсоток андроїдного жиру має позитивну кореляцію з артеріальним тиском. В середньому, вище співвідношення жирової та нежирової маси в андроїдній області свідчить про вищий артеріальний тиск.

Чи андроїдний жир шкідливіший за гіноїдний?

Male

Female





Люди з вищим співвідношенням андроїдного/гіноїдного жиру в середньому мають вищий артеріальний тиск. Цей зв'язок сильніший у чоловіків, ніж у жінок.

4 Висновки

Найбільш помітний зв'язок з артеріальним тиском спостерігався у таких факторів, як вік, вага та кількість жиру. Хотілося б дослідити ефект ваги та жиру, контролюючи вік. Фактори способу життя мали певний, але незначний вплив, можливо, через недостатньо якісні змінні, що їх відображали. Жирова маса, здається, відіграла більшу роль, ніж нежирова маса, але це питання потребує подальшого вивчення. Як і передбачалося, різні типи жиру мали різний вплив на артеріальний тиск, наприклад, андроїдний жир виявився шкідливішим за гіноїдний.

5 Використані джерела

- [1] Golemund, G., & Wickham, H. (2017). R for Data Science. O'Reilly Media.