

# Дослідницькі питання

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?
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?

# Опис даних

- NHANES 2011-2018
- 37399 рядків та 34 змінних

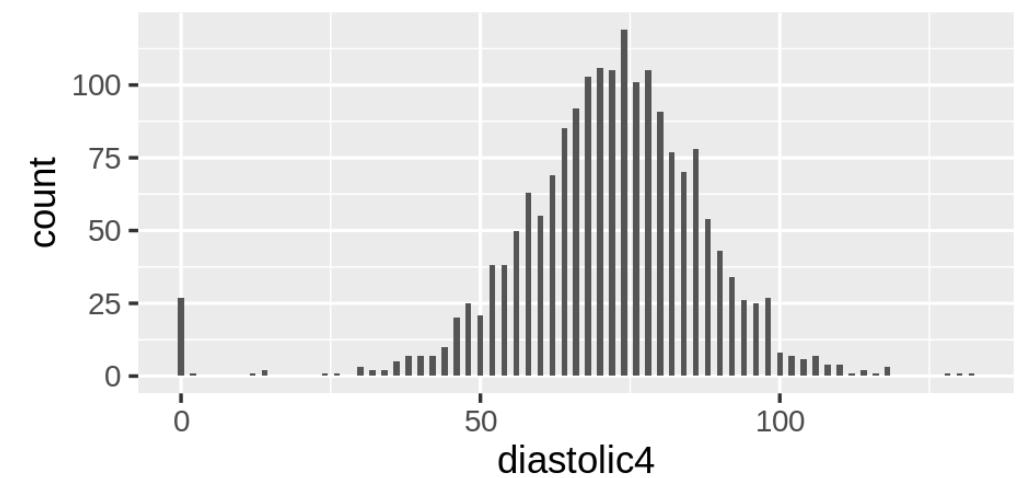
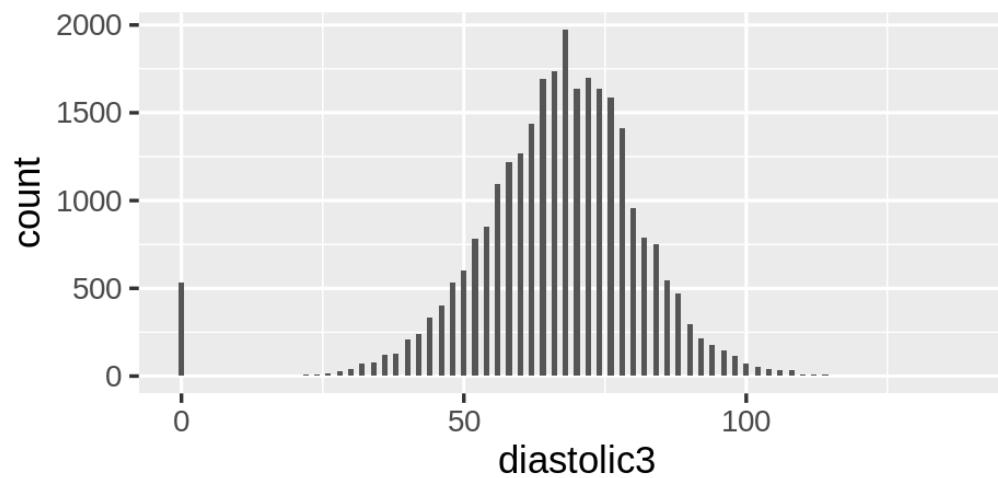
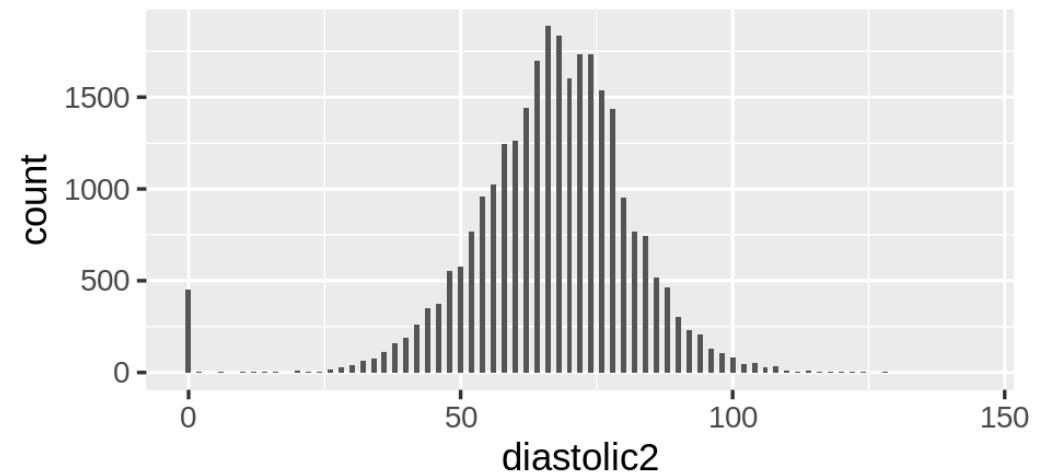
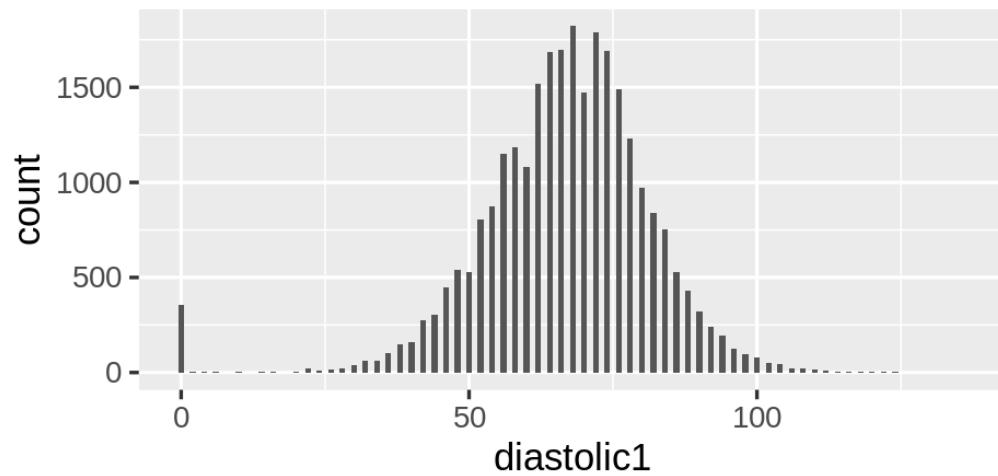
```
[1] "seqn"                      "diastolic1"  
[3] "diastolic2"                "diastolic3"  
[5] "diastolic4"                "systolic1"  
[7] "systolic2"                 "systolic3"  
[9] "systolic4"                 "pulse"  
[11] "weight_kg"                "waist_cm"  
[13] "subcutaneous_fat_g"       "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"
```

# Очищення даних

# Систолічний та діастолічний тиск

```
# A tibble: 8 × 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
```

- 28666 рядків з не пропущеним артеріальним тиском

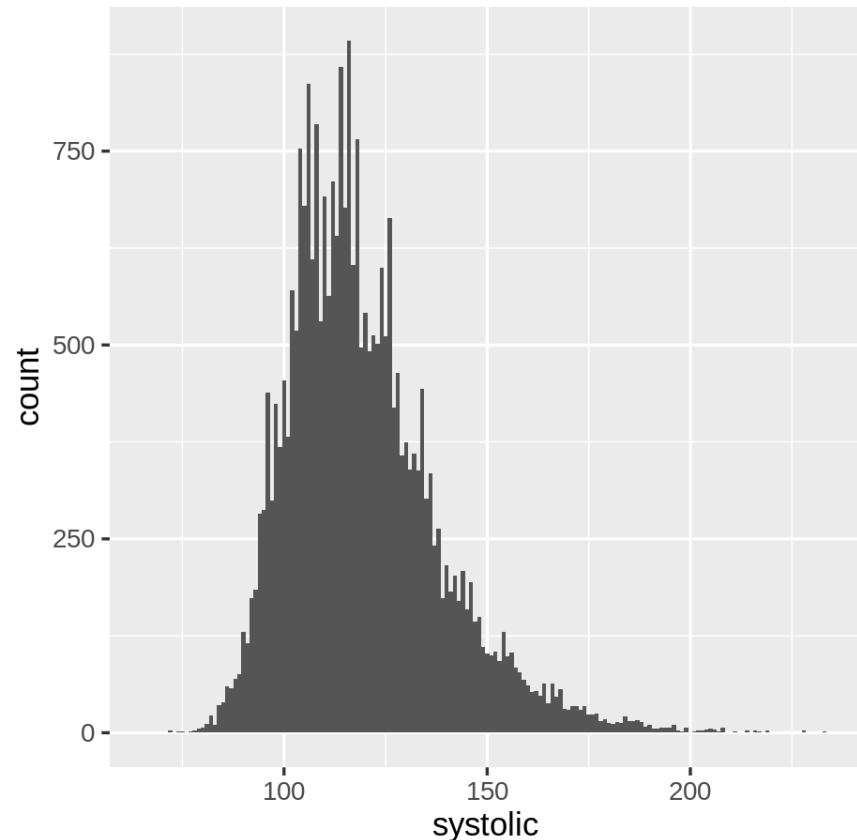
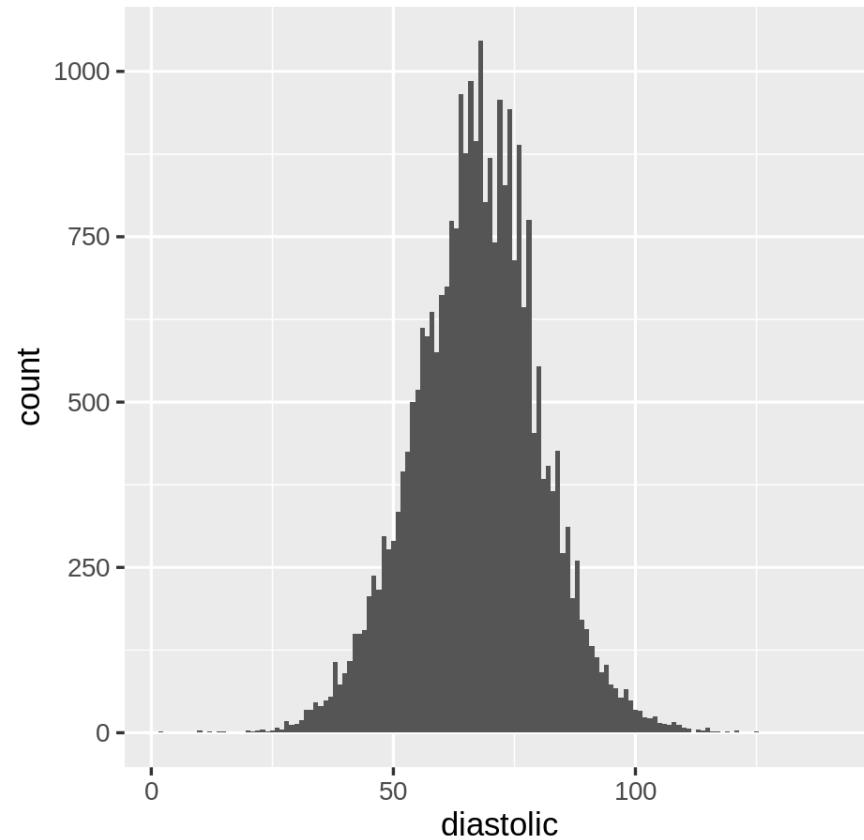


# Систолічний та діастолічний тиск

- 28666 рядків з не пропущеним артеріальним тиском
- 28479 рядків з не нульовим артеріальним тиском

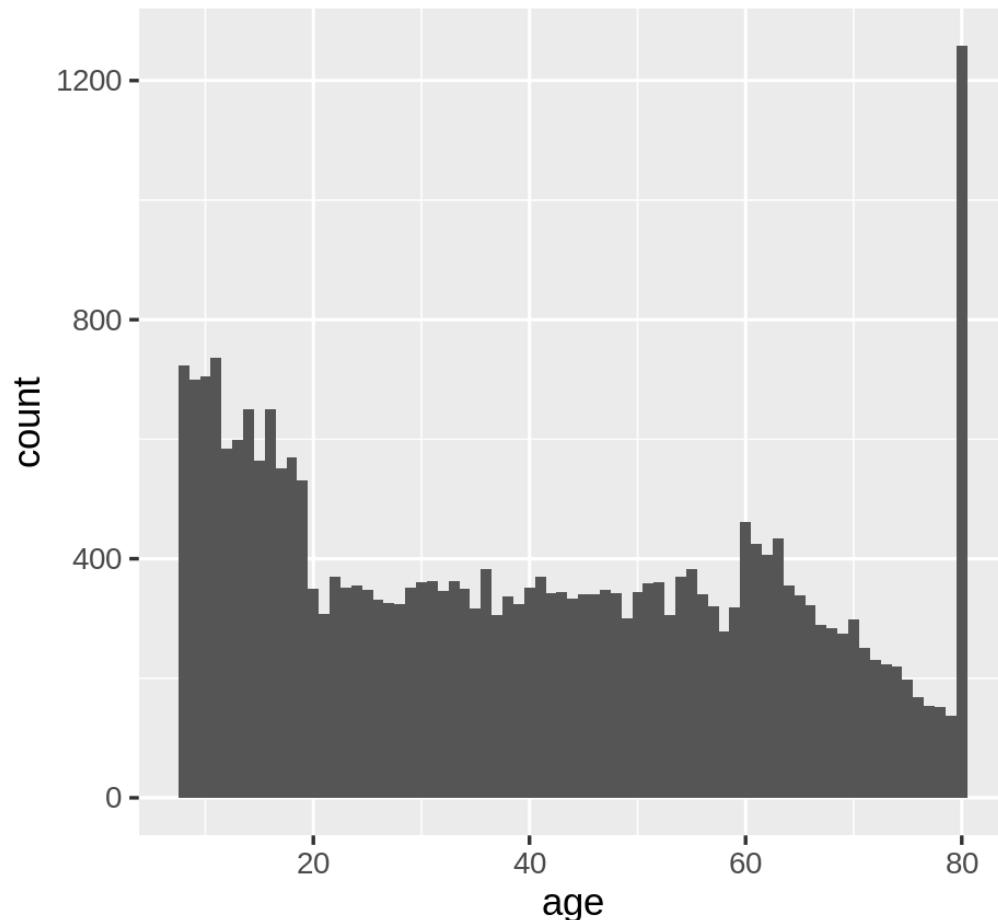
# Агрегація вимірювань

- Якщо два найближчі вимірювання відрізняються незначно (не більше ніж на 8), то беремо їх середнє значення
- Залишилось 28097 рядків

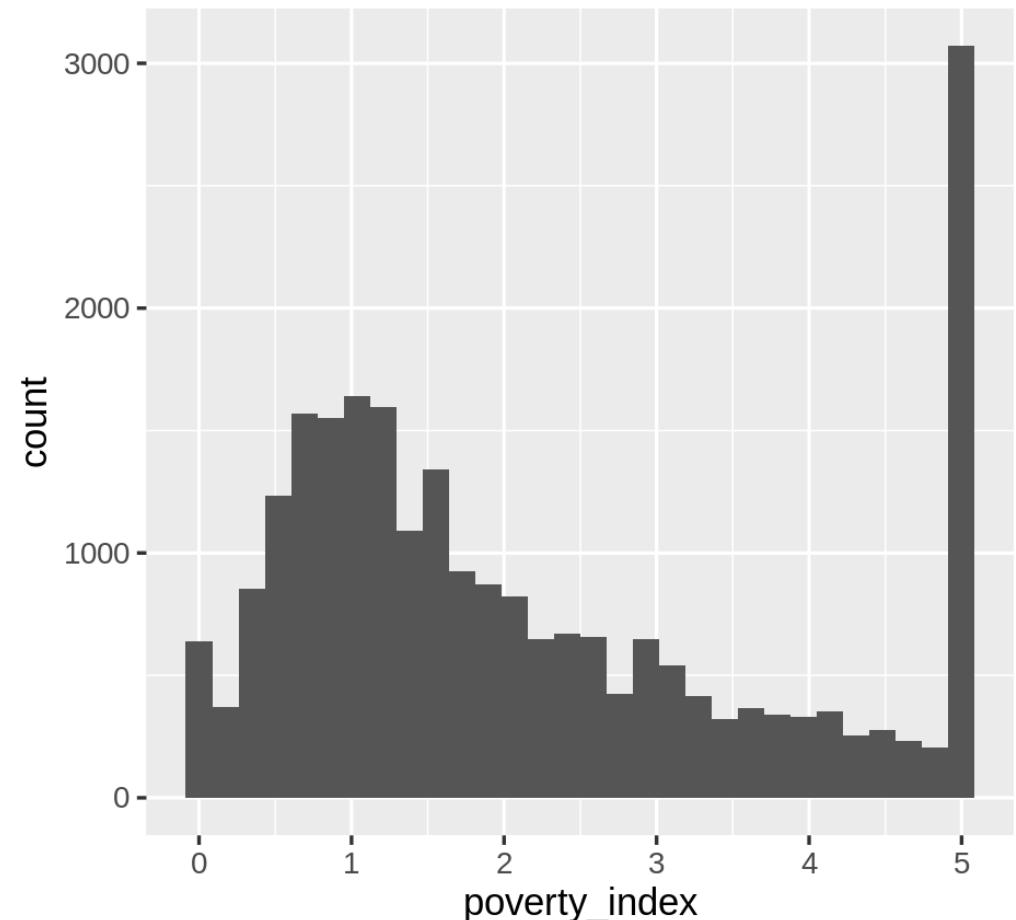


# Кодування інших змінних

Age  $\geq 80$  is coded as 80



Poverty index  $\geq 5$  is coded as 5



# Пропущені значення

# A tibble: 20 × 3			# A tibble: 12 × 3		
column	na_percent	na_count	column	na_percent	na_count
<chr>	<dbl>	<int>	<chr>	<dbl>	<int>
1 seqn	0	0	1 poverty_index	0.136	3827
2 diastolic	0	0	2 vigorous_phys_activity	0.133	3740
3 systolic	0	0	3 vigorous_phys_activity_days	0.133	3743
4 blood_pressure_category	0	0	4 moderate_phys_activity	0.133	3744
5 pulse	0.000107	3	5 moderate_phys_activity_days	0.134	3752
6 weight_kg	0.00883	248	6 sleep	0.191	5355
7 waist_cm	0.0447	1256	7 smoking_now	0.685	19233
8 subcutaneous_fat_g	0.343	9628	8 smoked_100_cigs	0.241	6758
9 visceral_fat_g	0.343	9630	9 diet_quality	0.188	5271
10 android_g	0.343	9640	10 android_fat_percent	0.343	9640
11 android_fat_g	0.343	9640	11 gynoid_fat_percent	0.343	9625
12 android_non_fat_g	0.343	9640	12 android_gynoid_ratio	0.344	9658
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			

# Підвибірки даних

- Було створено 3 підвибірки

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

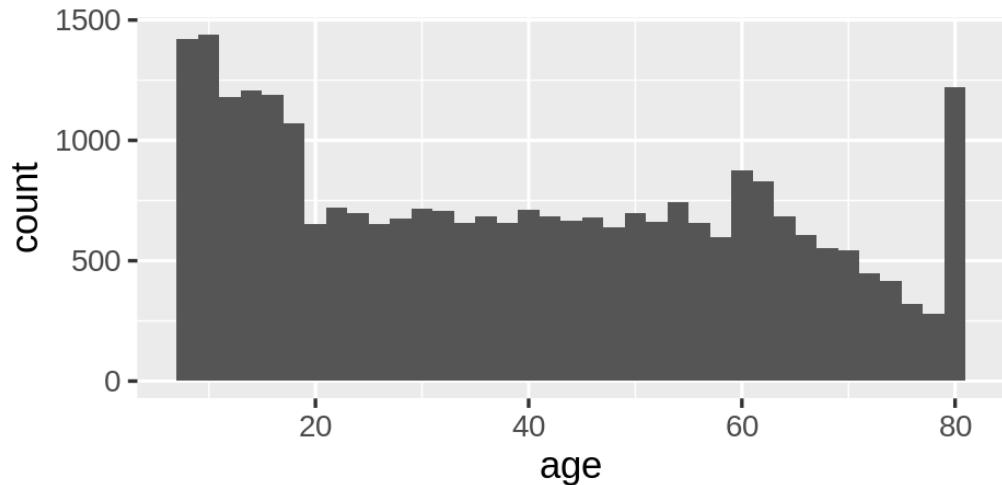
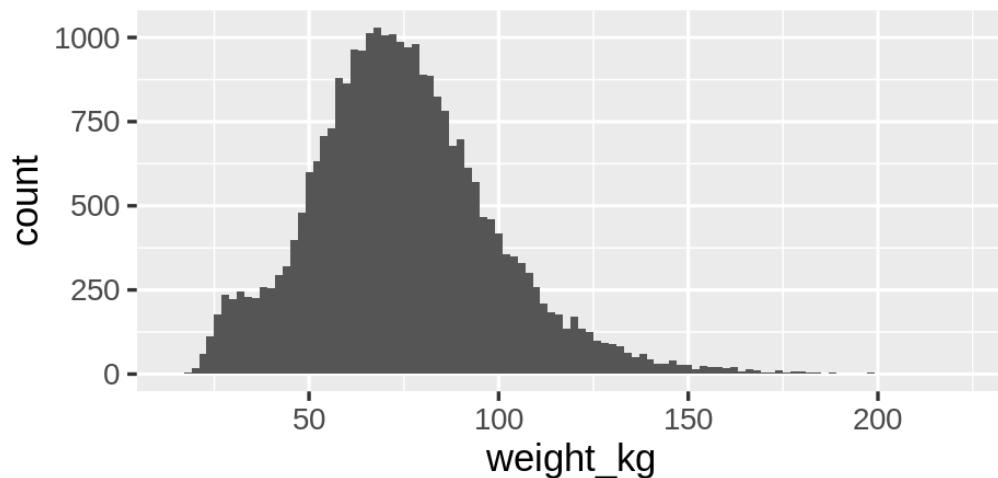
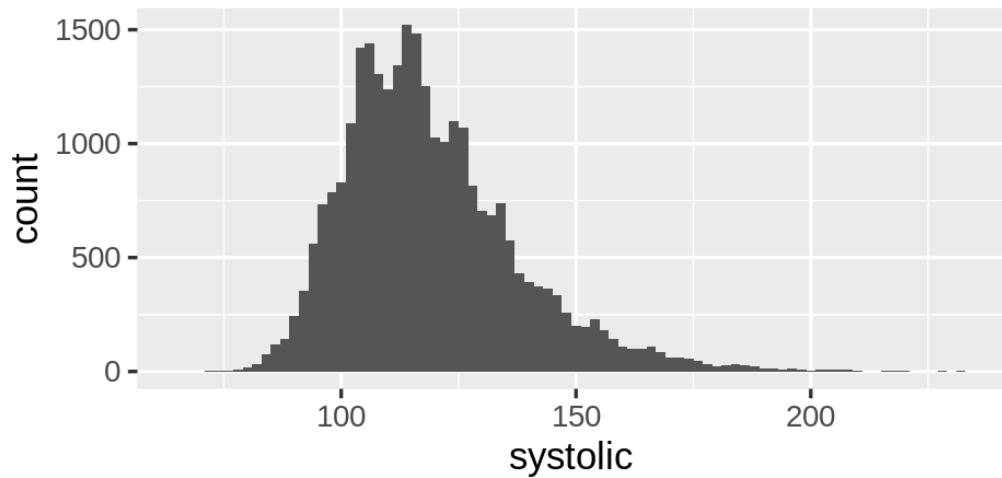
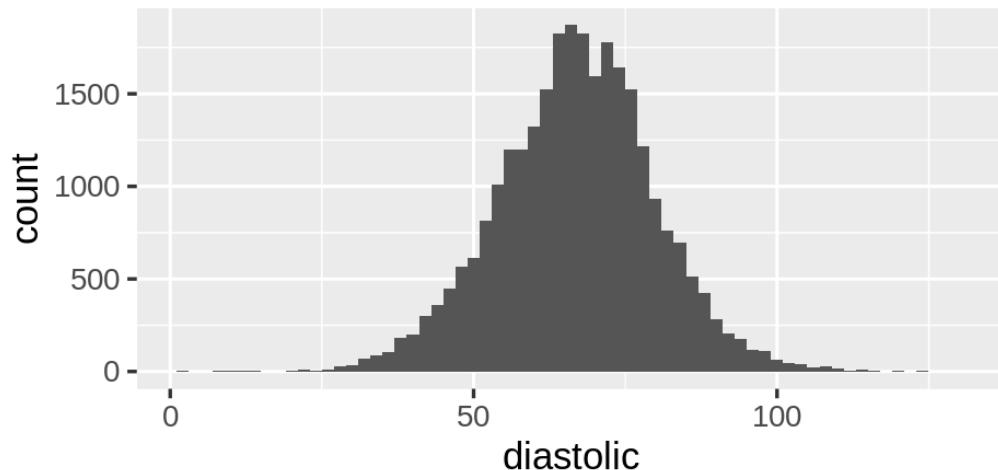
# **Резулътати EDA**

1. Як вік, вага, раса та стать впливають на артеріальний тиск?
2. Як фактори способу життя впливають на артеріальний тиск?
3. Як розподіл жирової тканини впливає на артеріальний тиск?

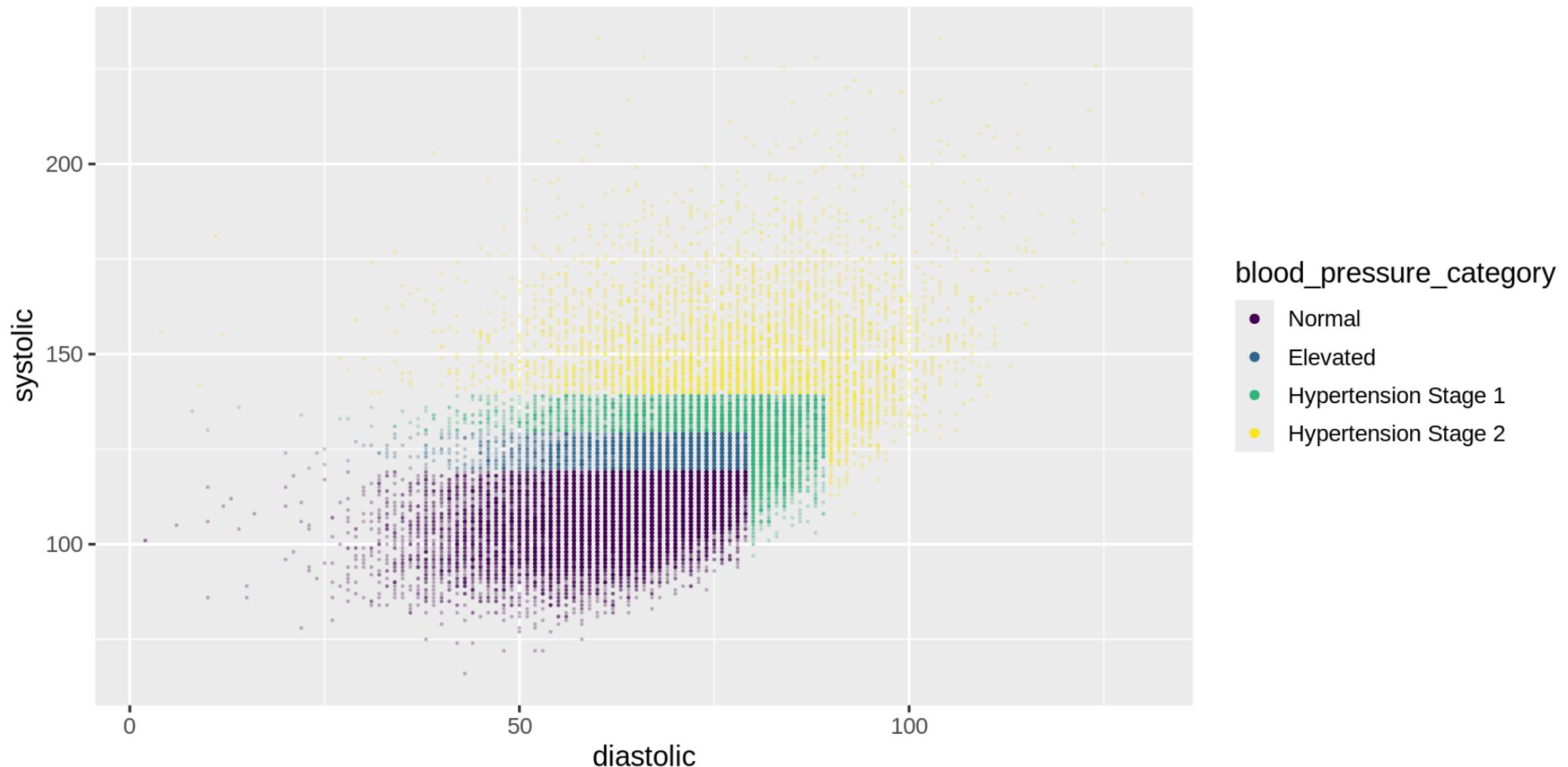
**Як вік, вага, раса та стать впливають на артеріальний тиск?**

# Розподіли числових змінних

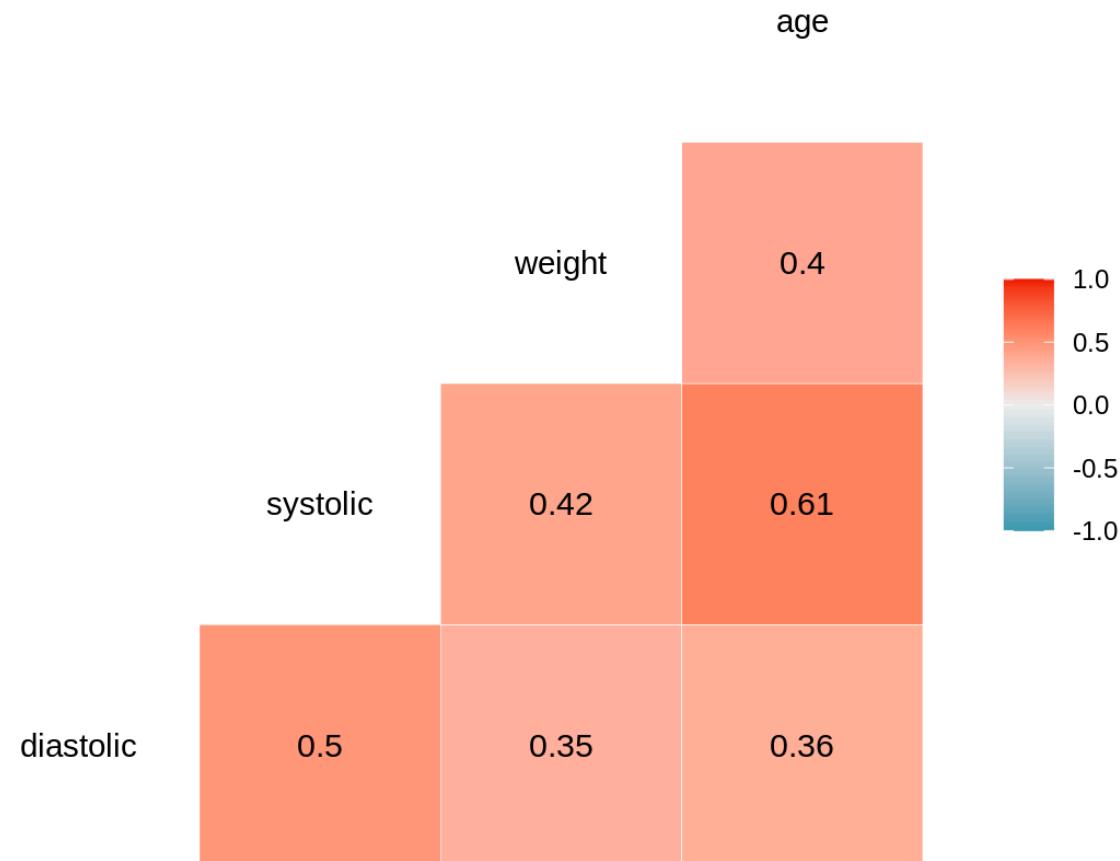
- Варто звернути увагу на розподіл віку



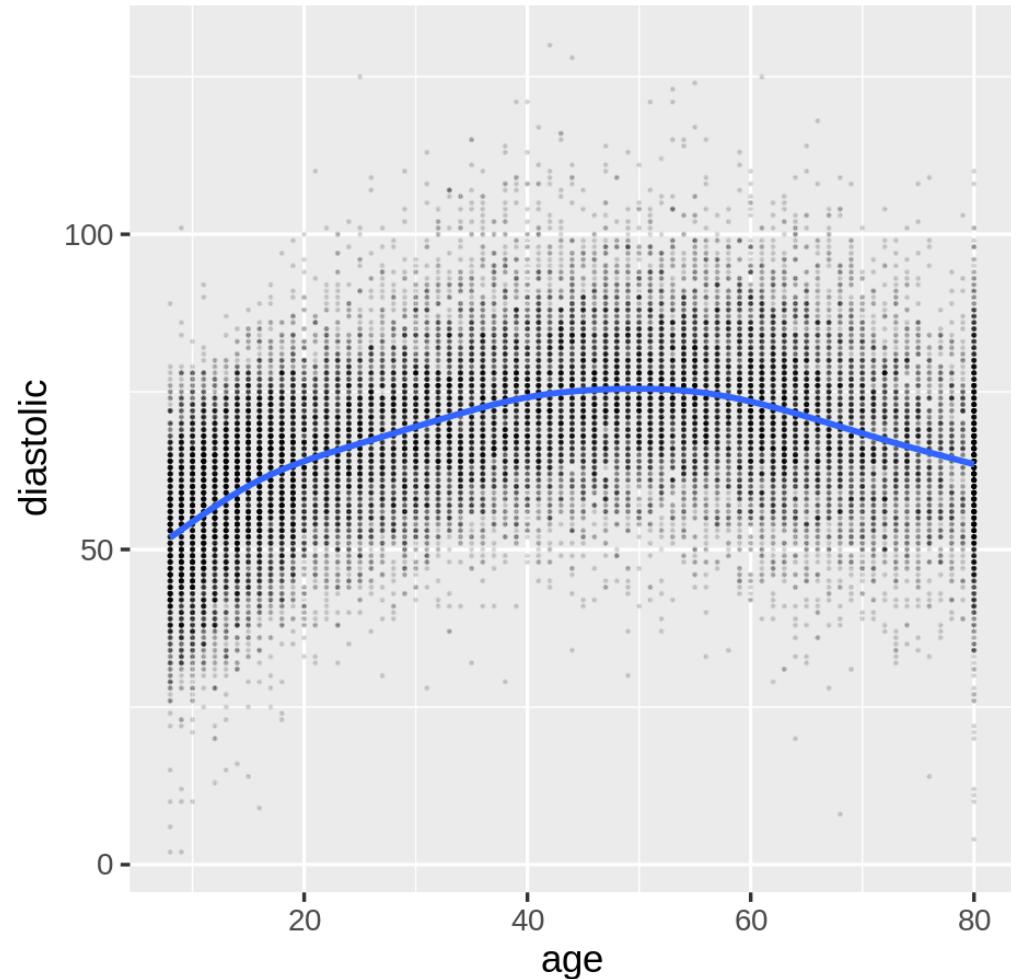
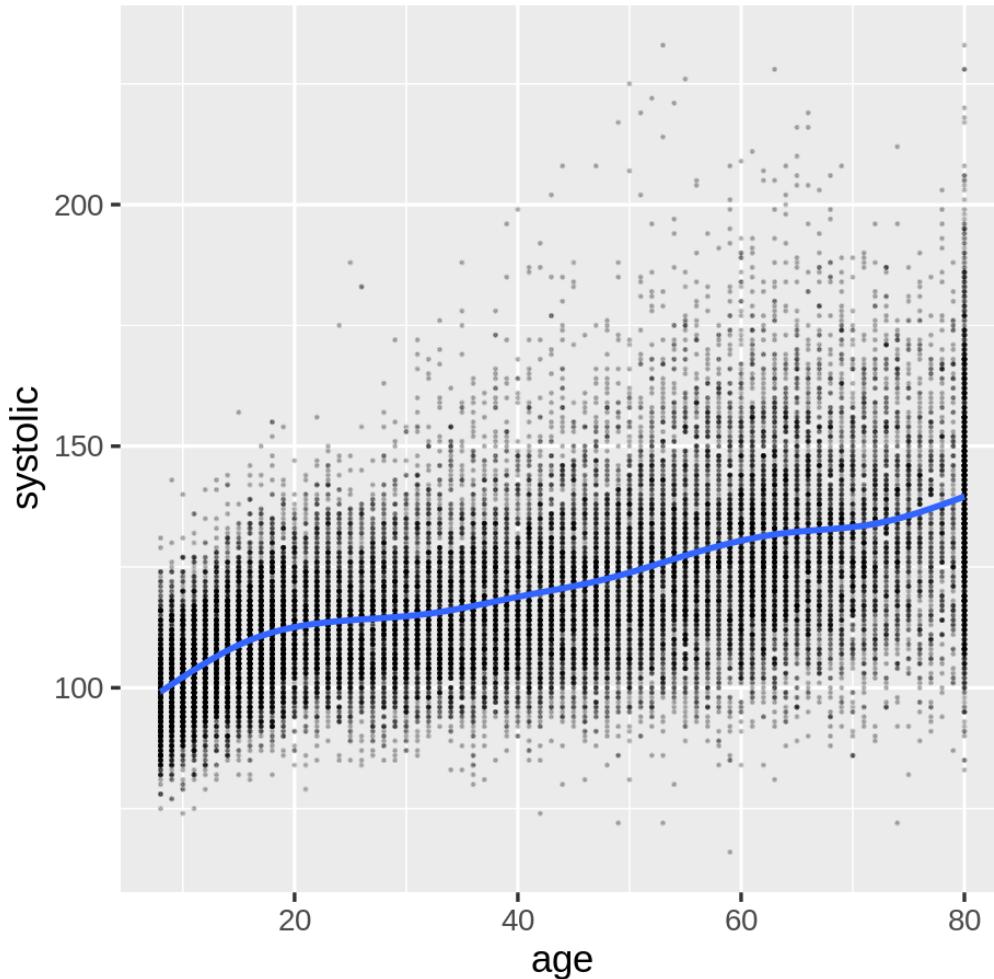
# Розподіл систолічного та діастолічного тиску

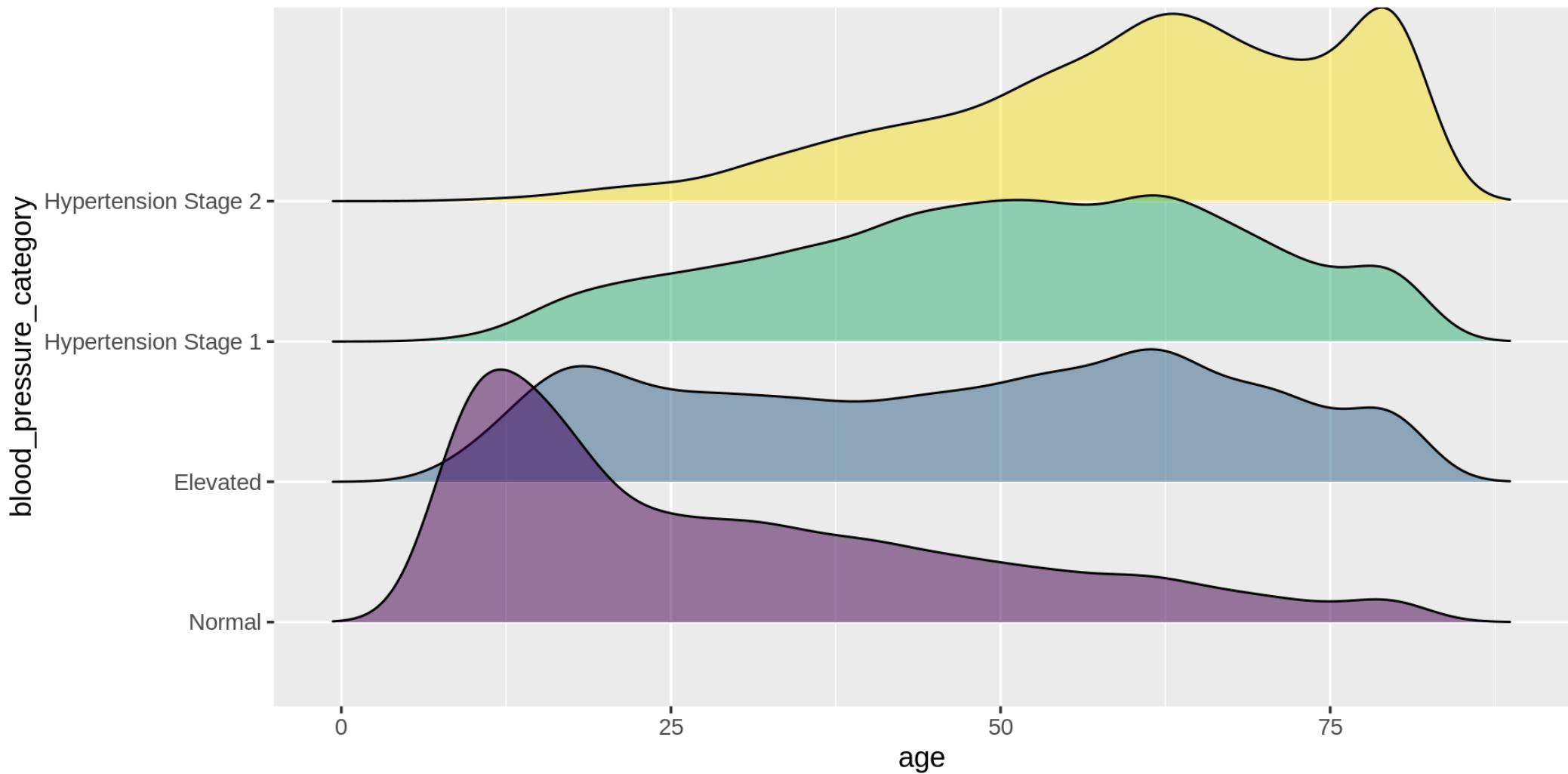


# Кореляційний зв'язок

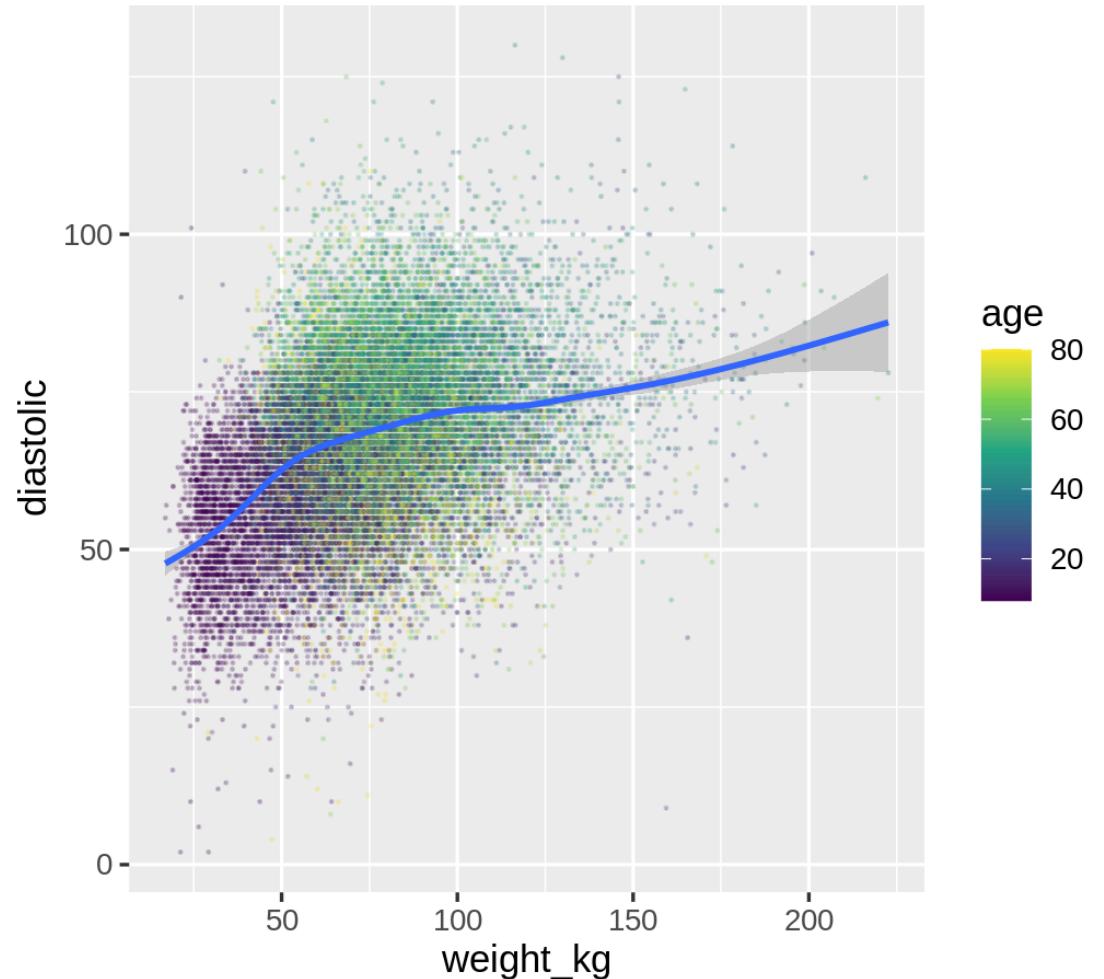
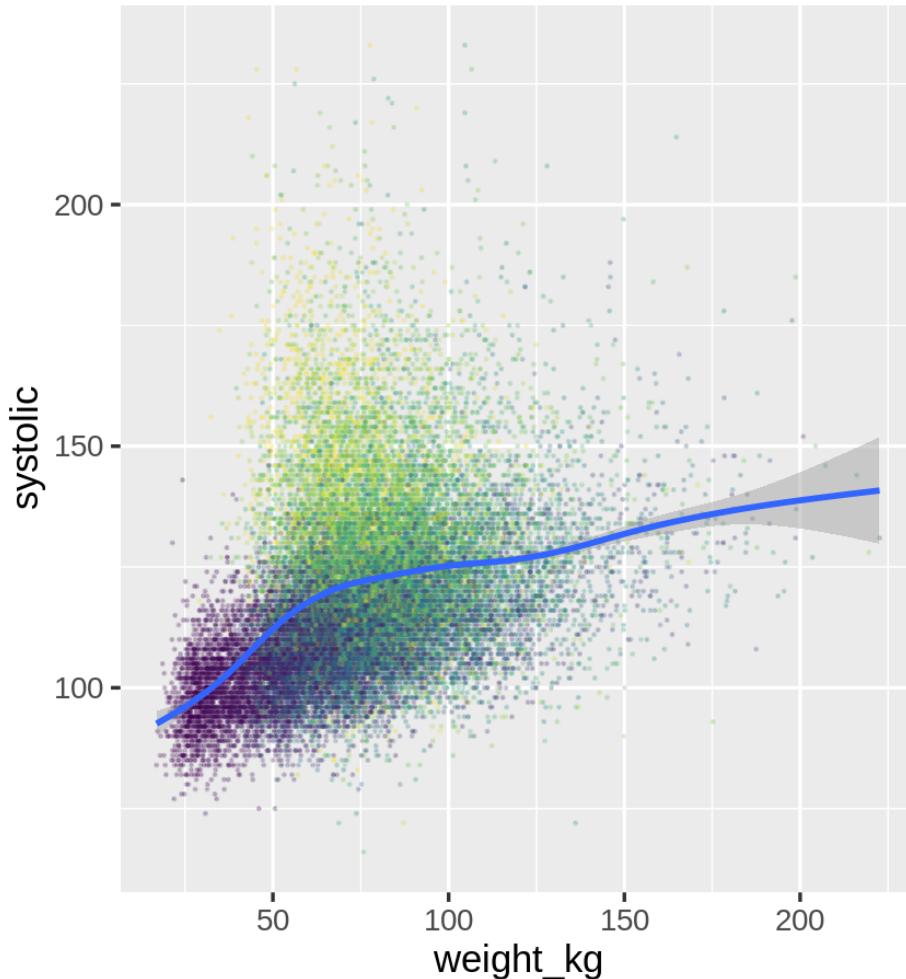


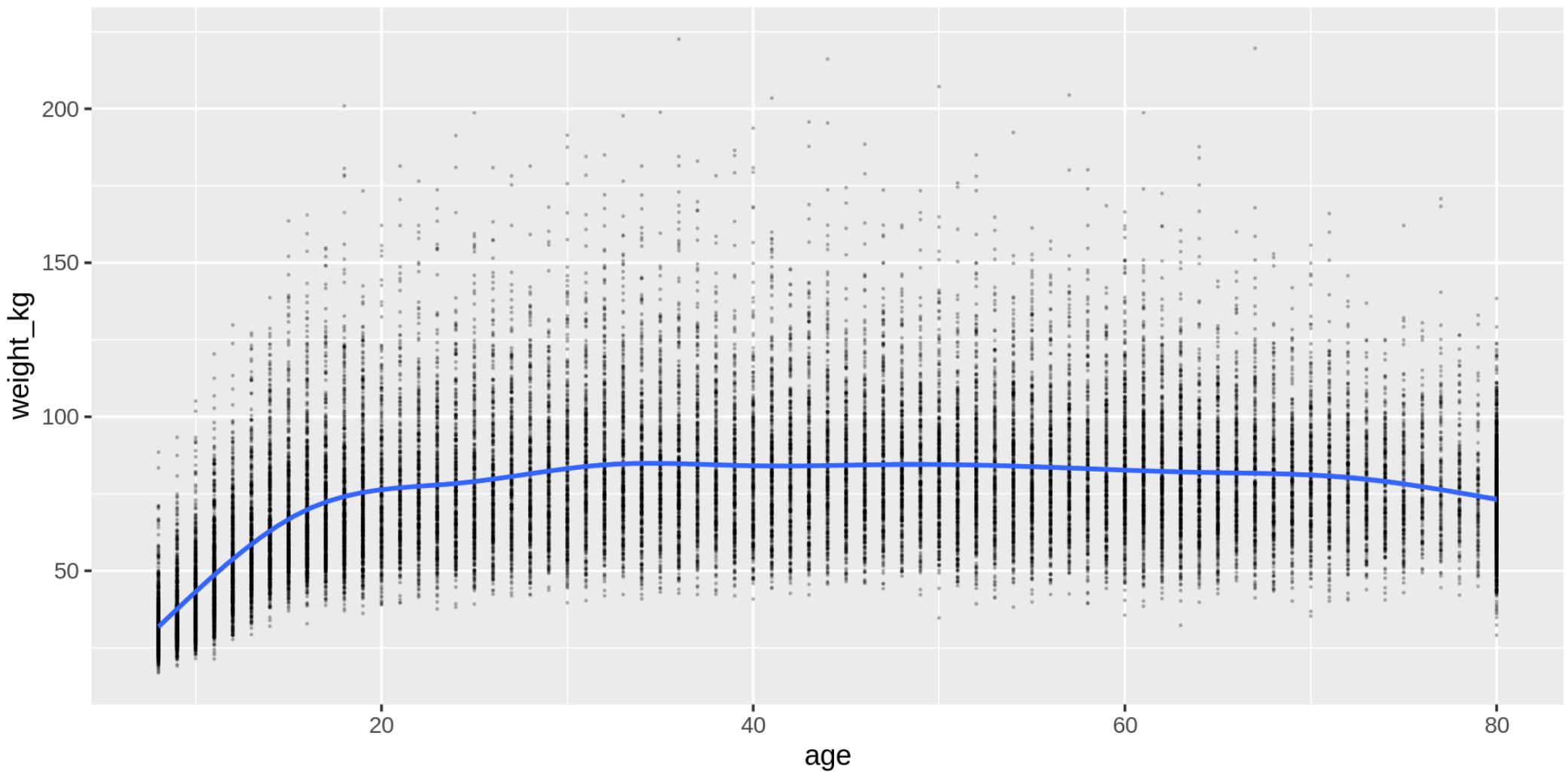
# Чи мають люди старшого віку вищий артеріальний тиск?



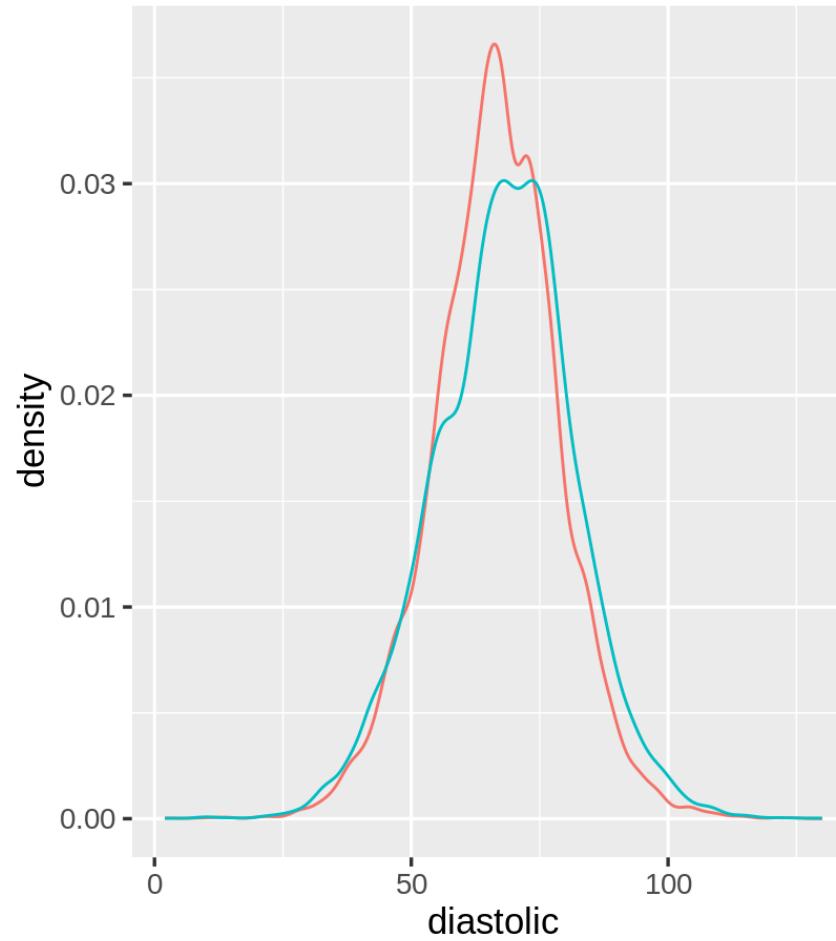
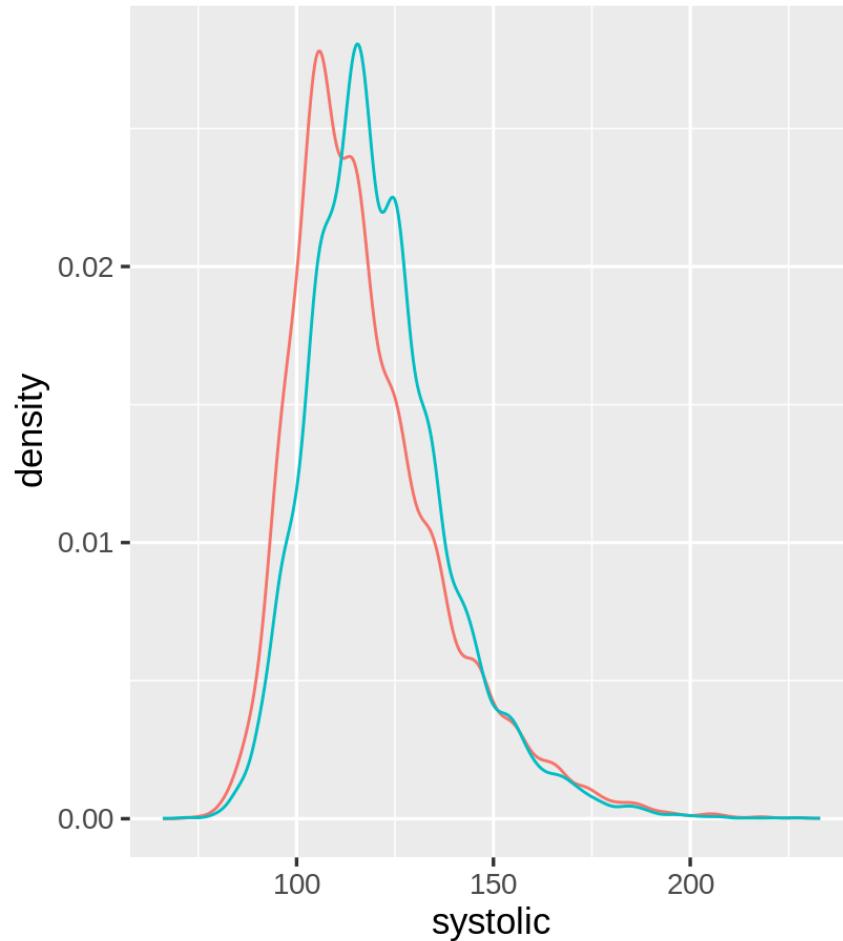


# Чи мають люди з більшою вагою вищий артеріальний тиск?





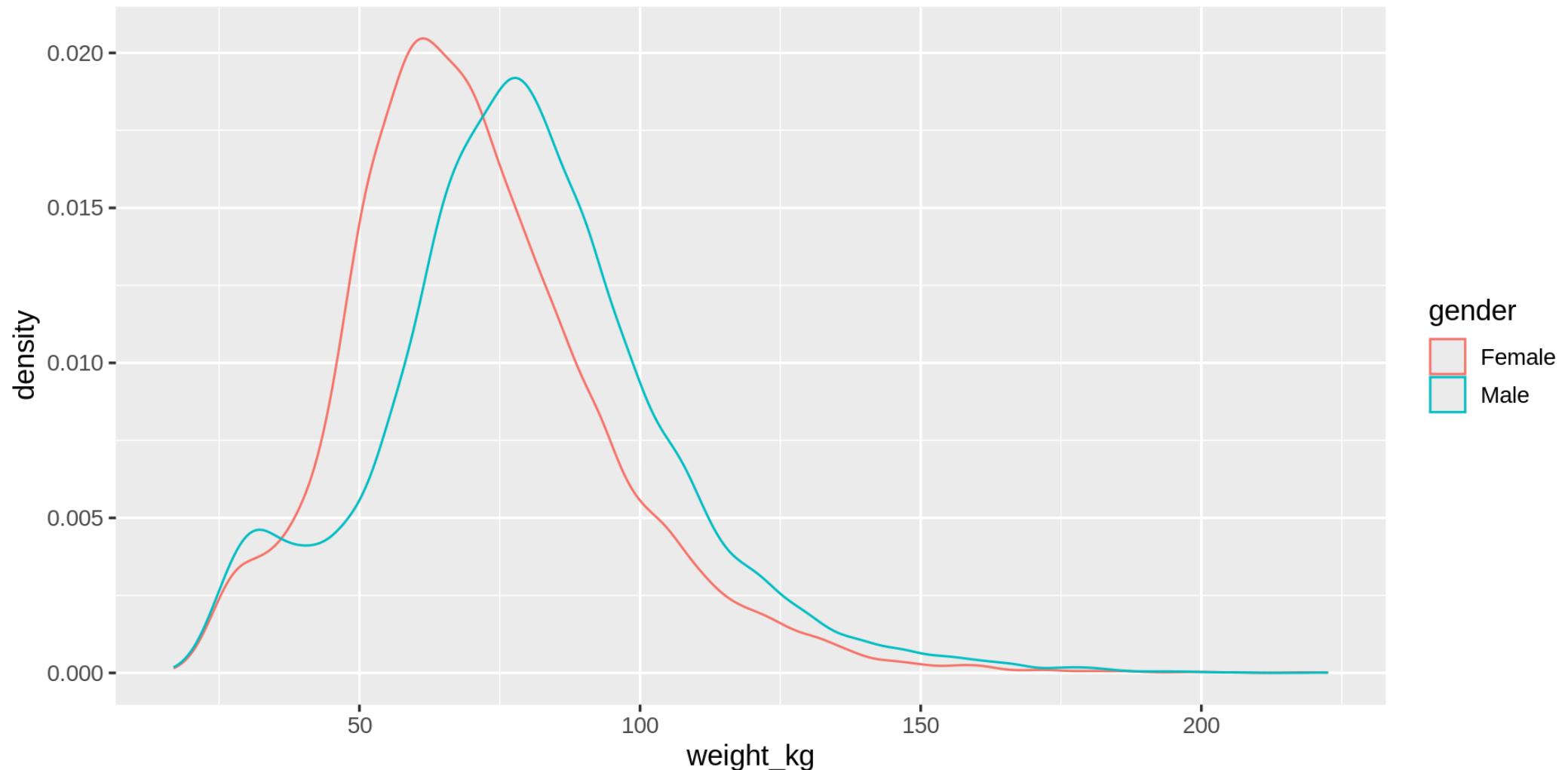
# Чи впливає стать на артеріальний тиск?



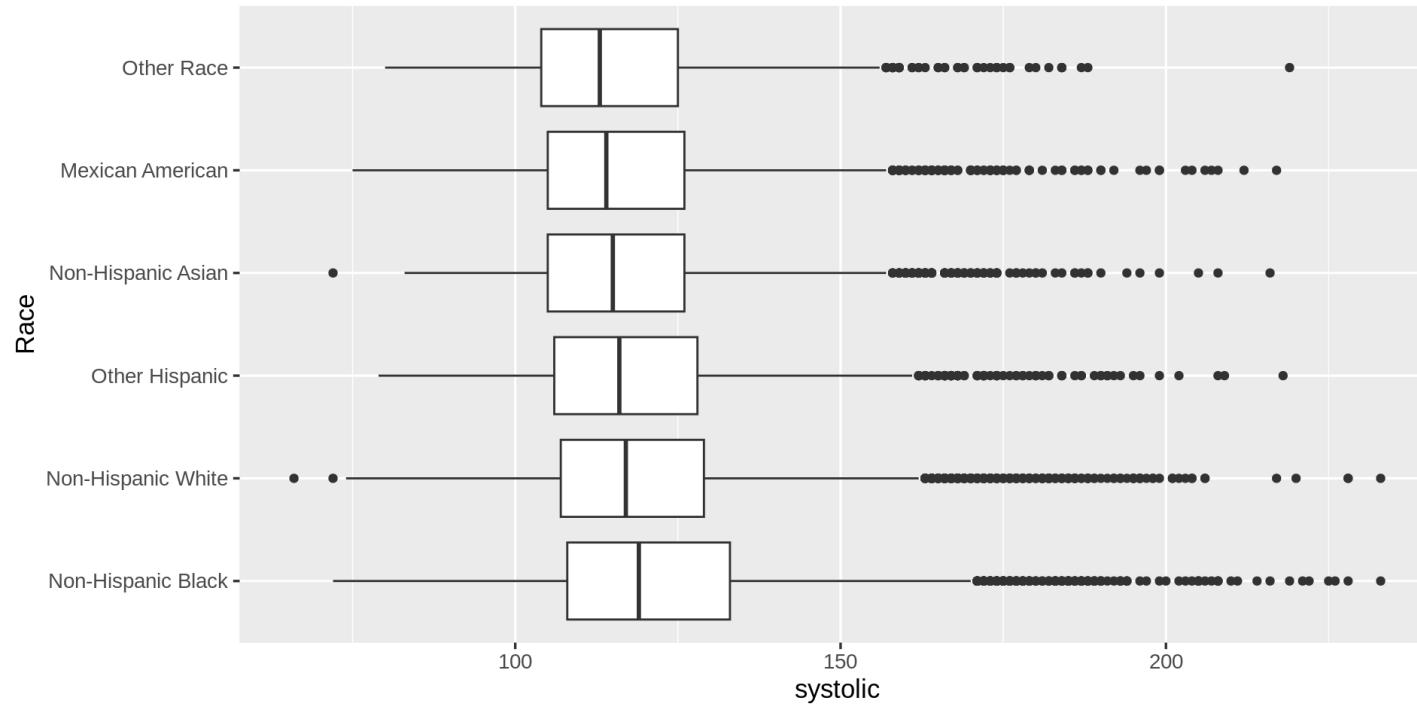
gender  
Female  
Male

```
# A tibble: 2 × 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
```

# Мабуть це тому що чоловіки більше за жінок?



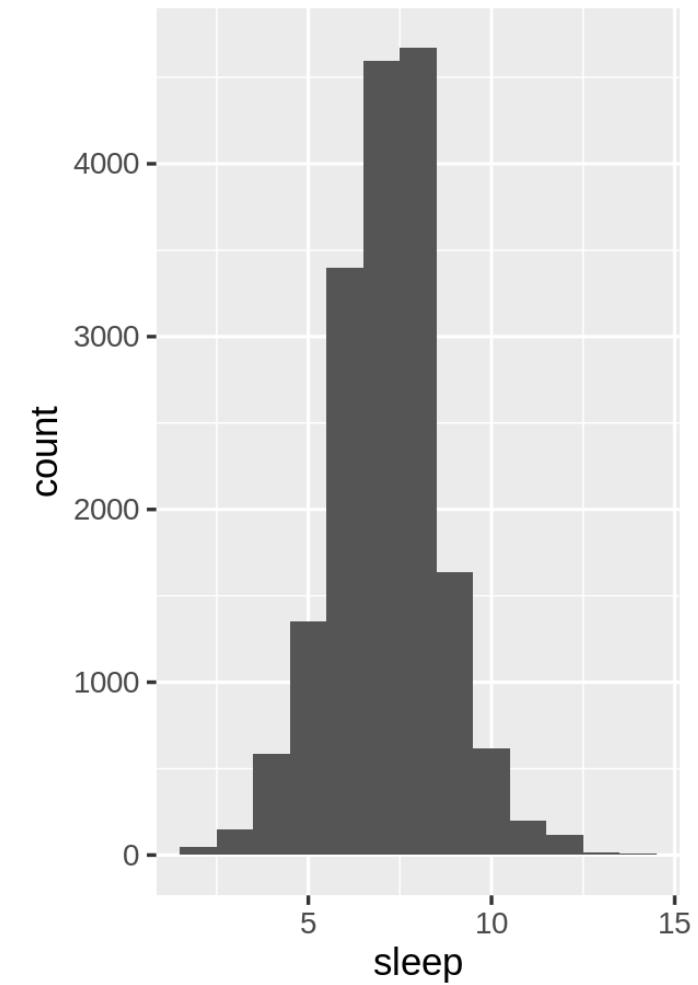
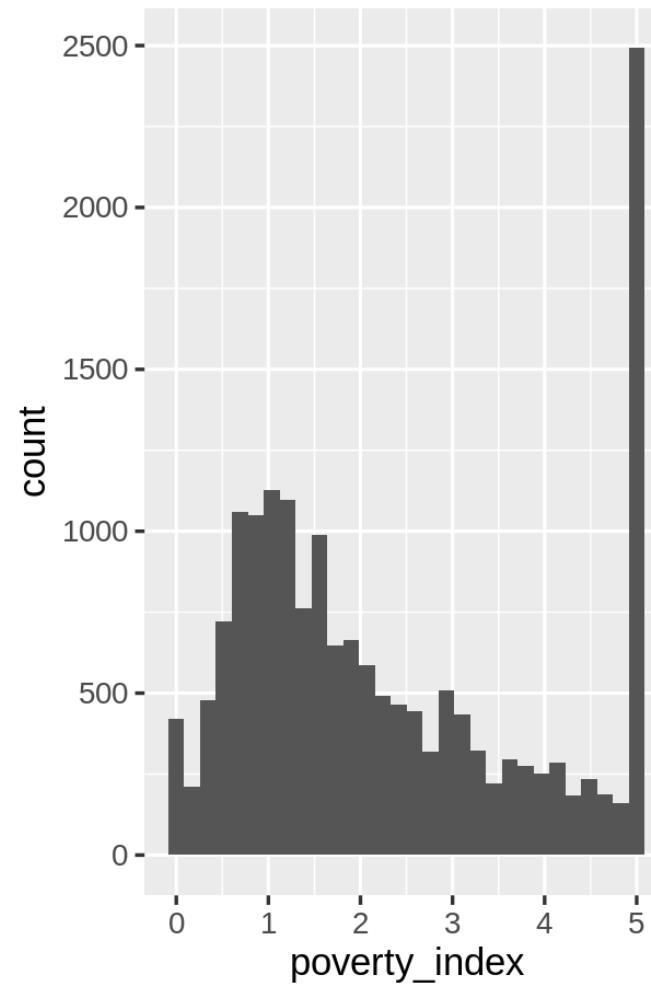
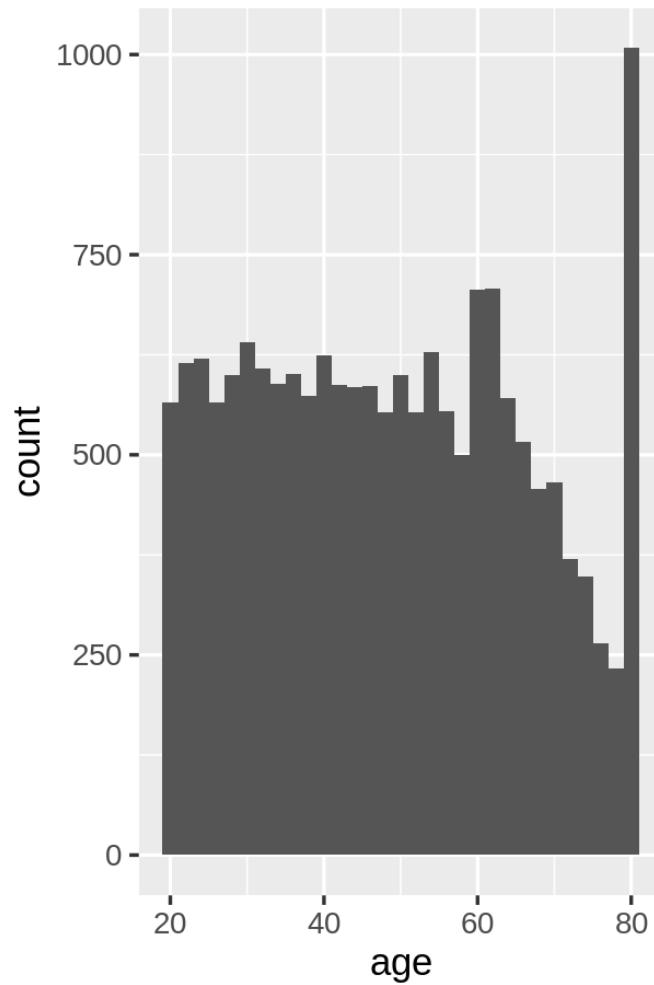
# Чи відрізняється артеріальний тиск залежно від раси?



```
# A tibble: 6 × 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
```

**Як фактори способу життя впливають на артеріальний тиск?**

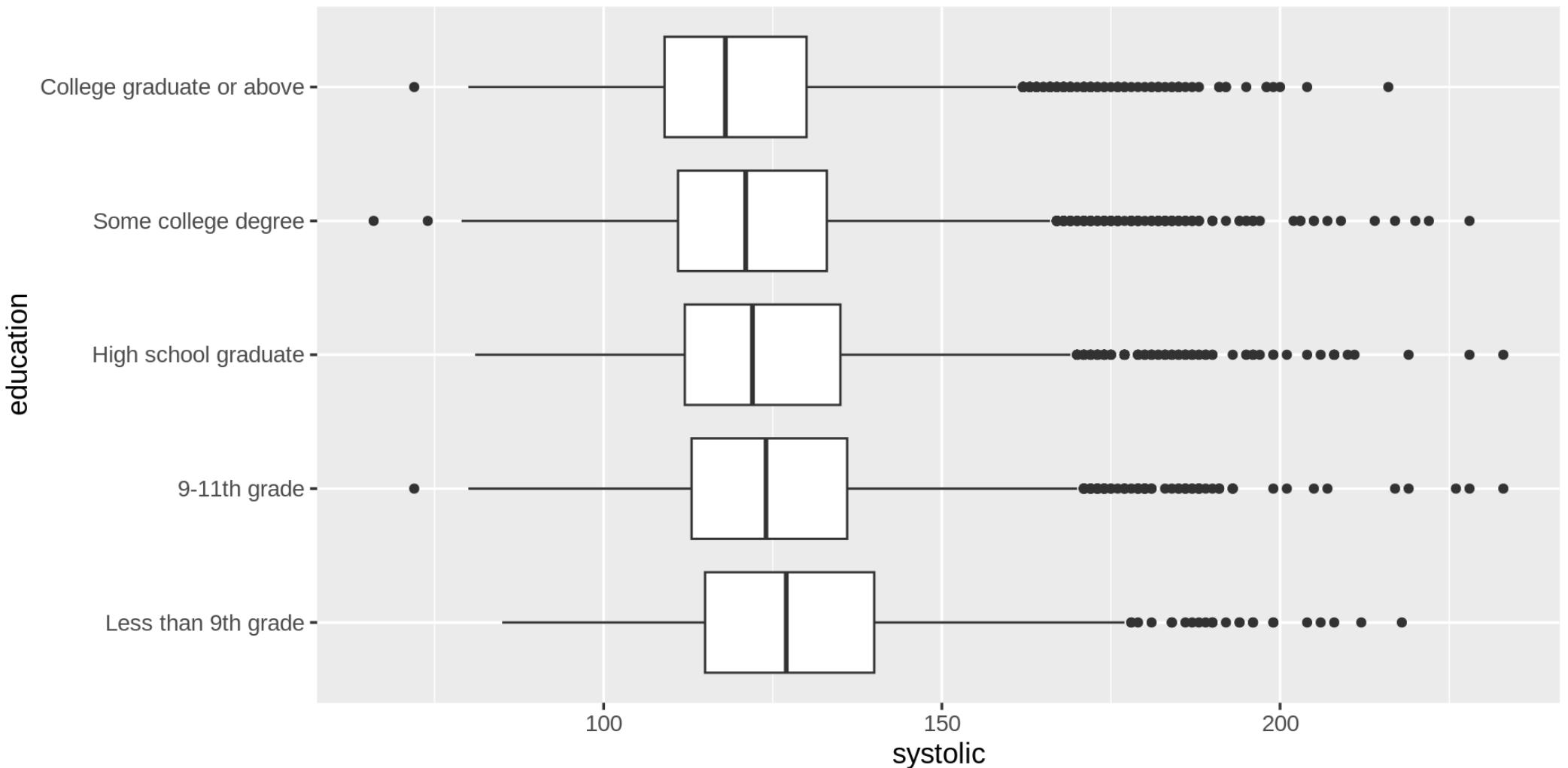
# Розподіл числових змінних



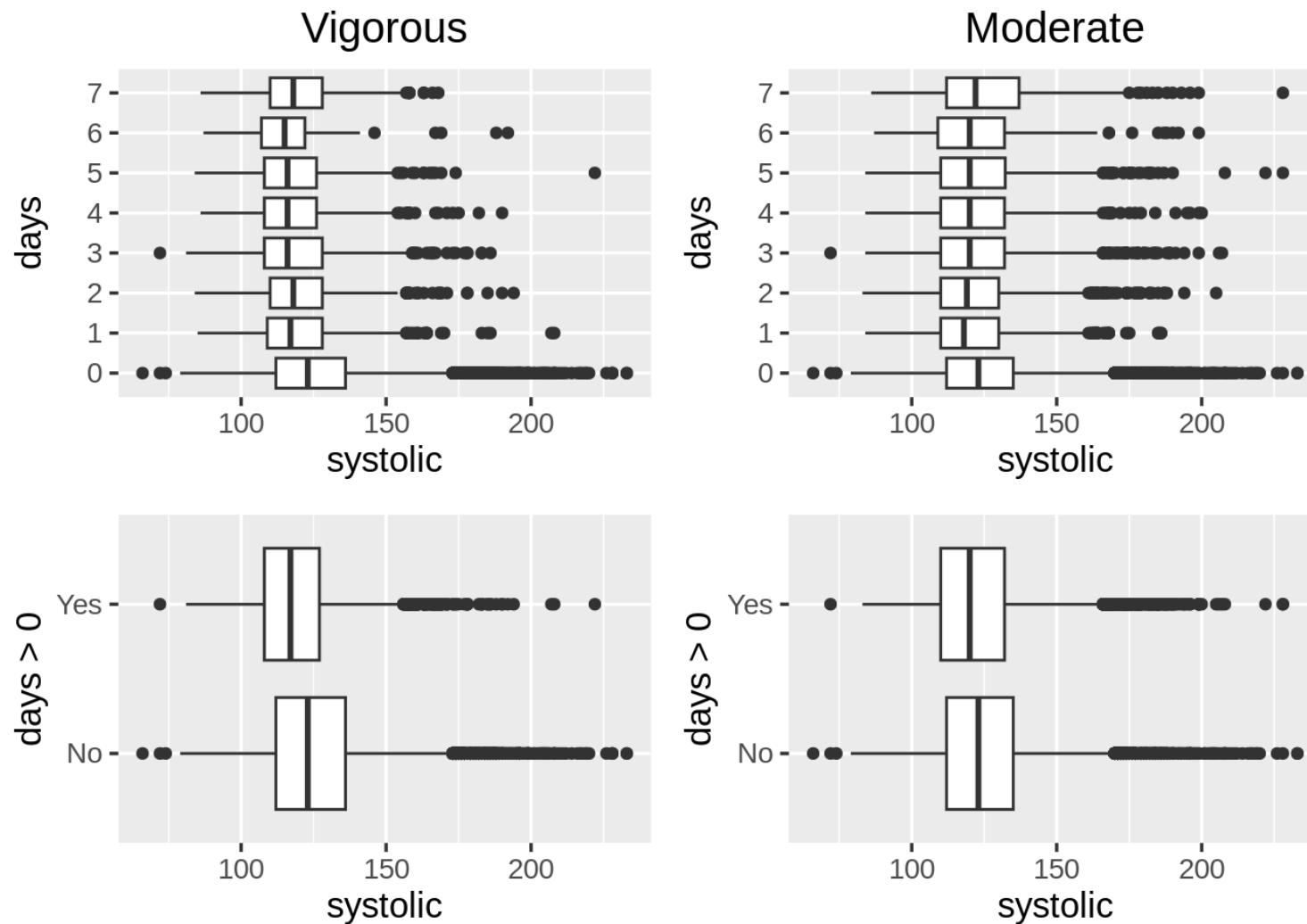
# Кореляційний зв'язок



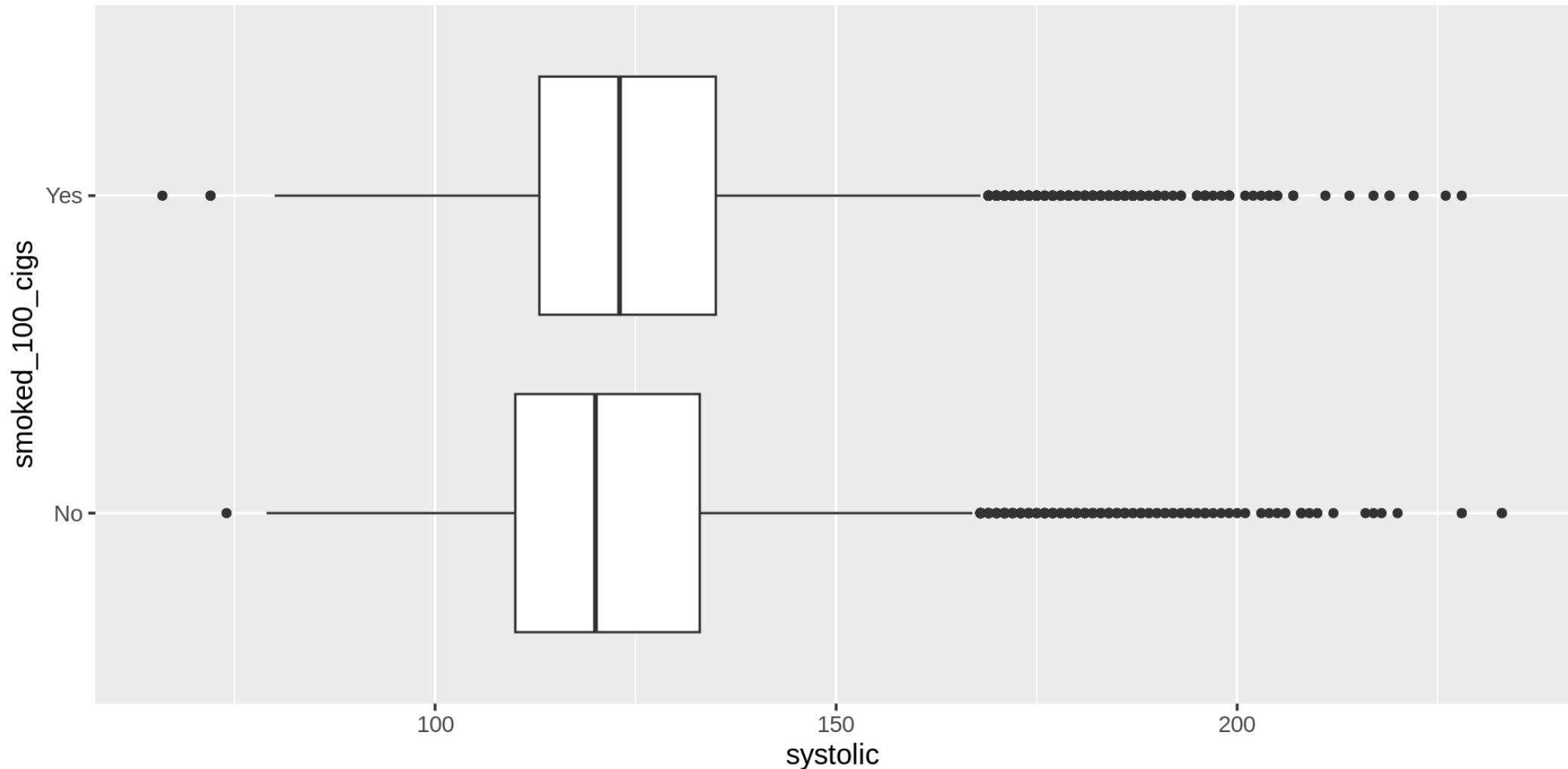
# Чи мають люди з нижчим рівнем освіти менш здоровий артеріальний тиск?

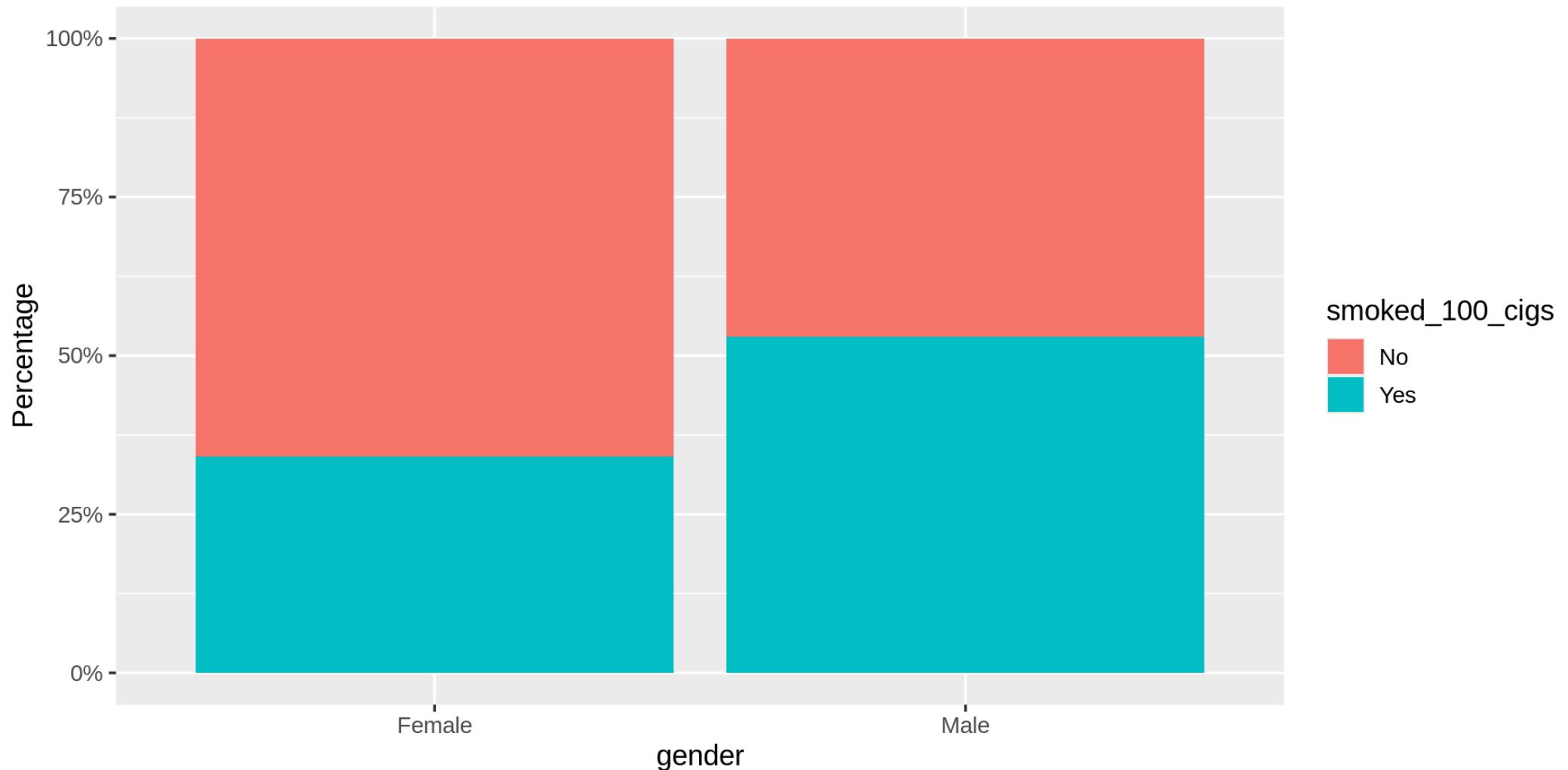


# Чи мають фізично активні люди здоровіший артеріальний тиск?



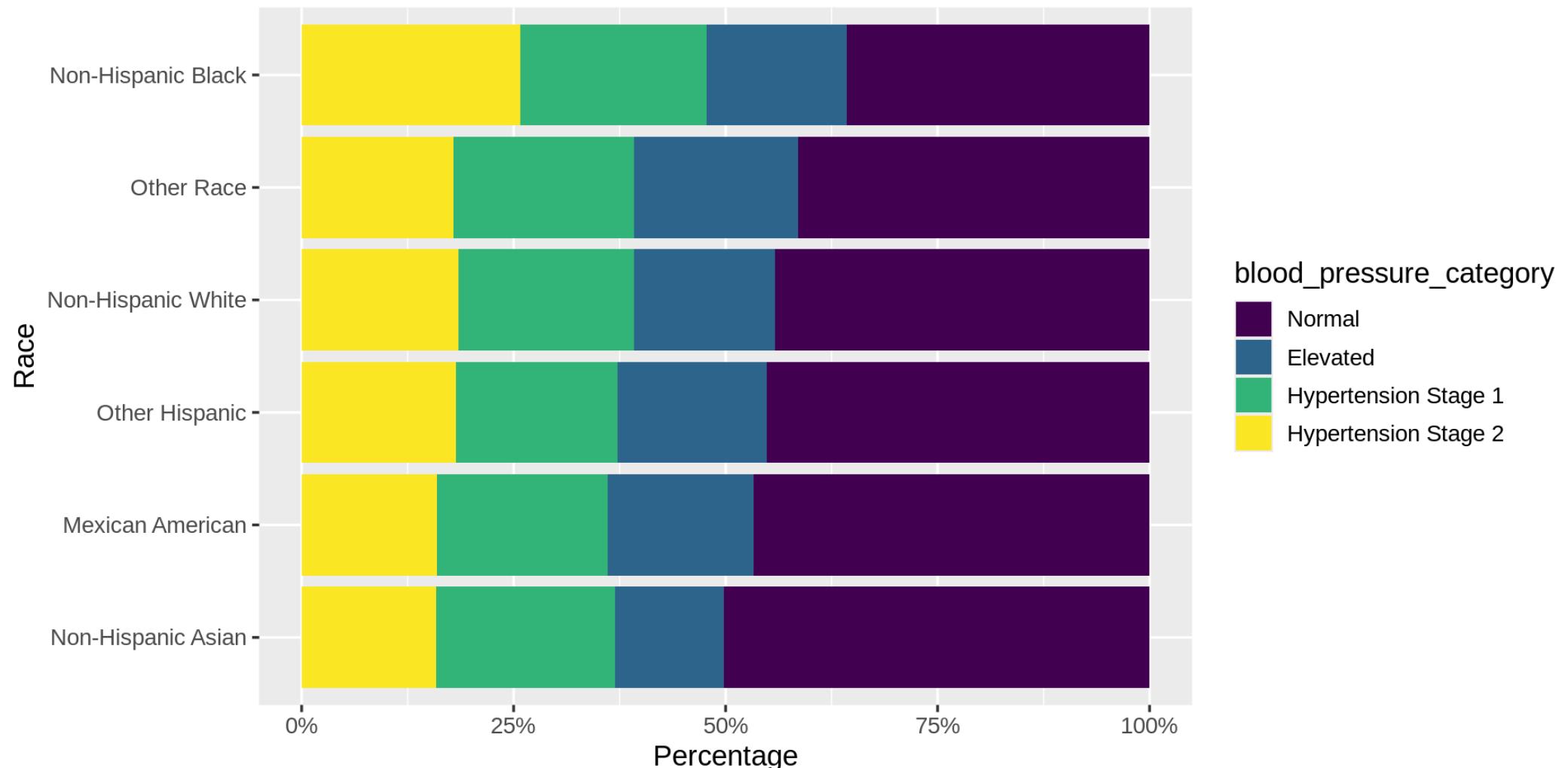
# Чи мають люди, які більше курять, гірший артеріальний тиск?

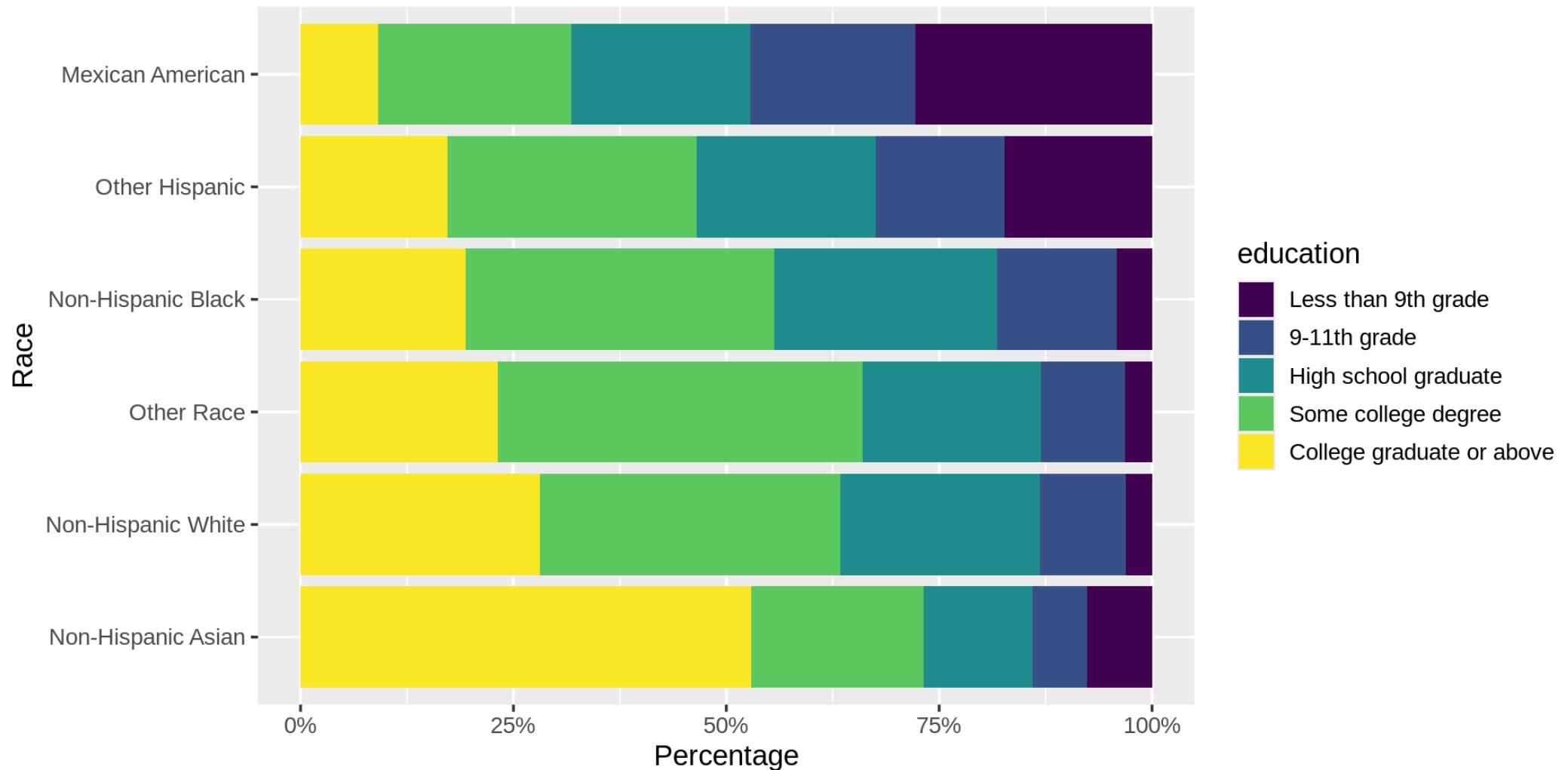


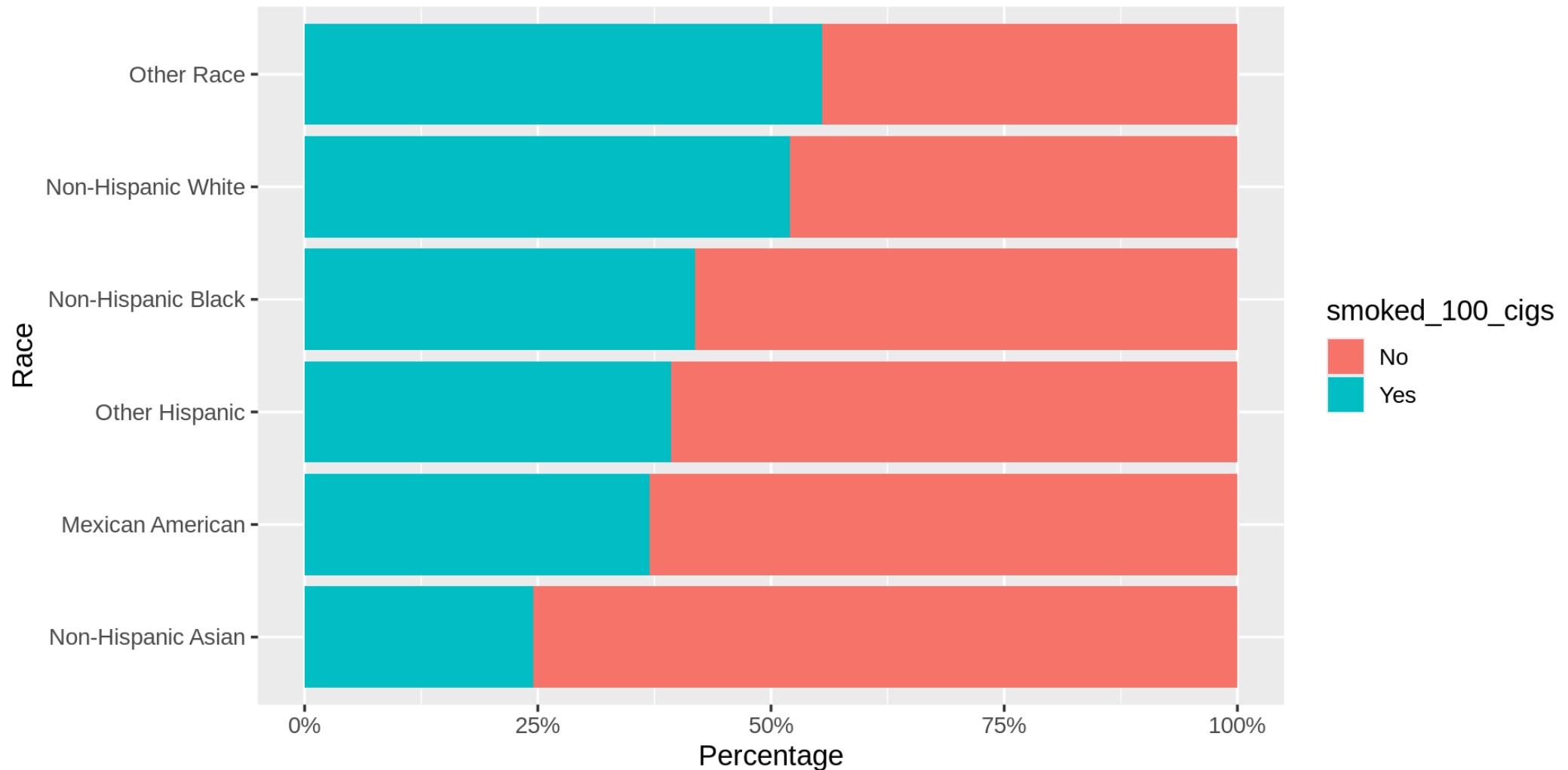


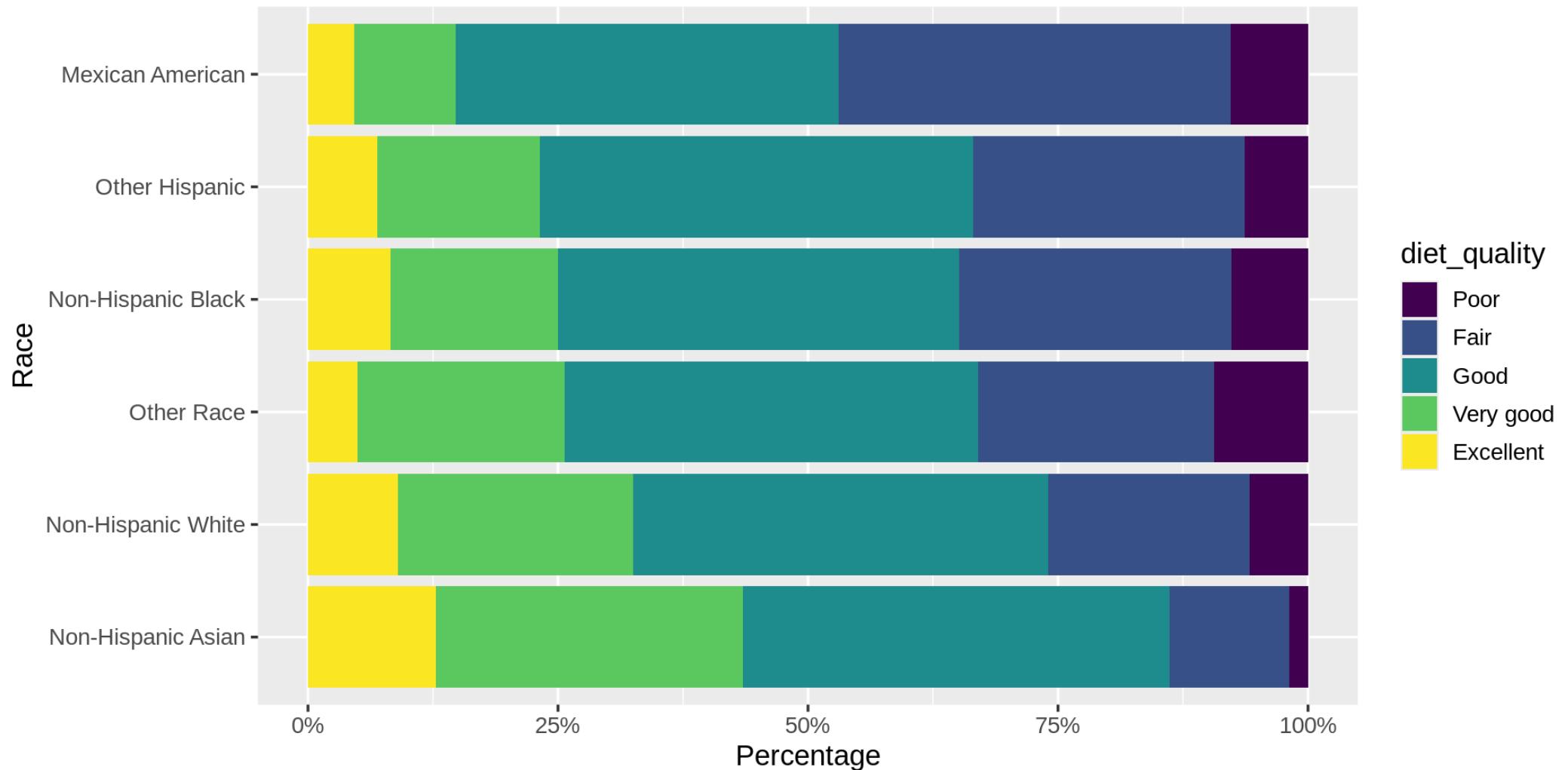
```
# A tibble: 2 × 2
  gender smoked_100_cigs_percent
  <chr>          <dbl>
1 Female         0.341
2 Male           0.530
```

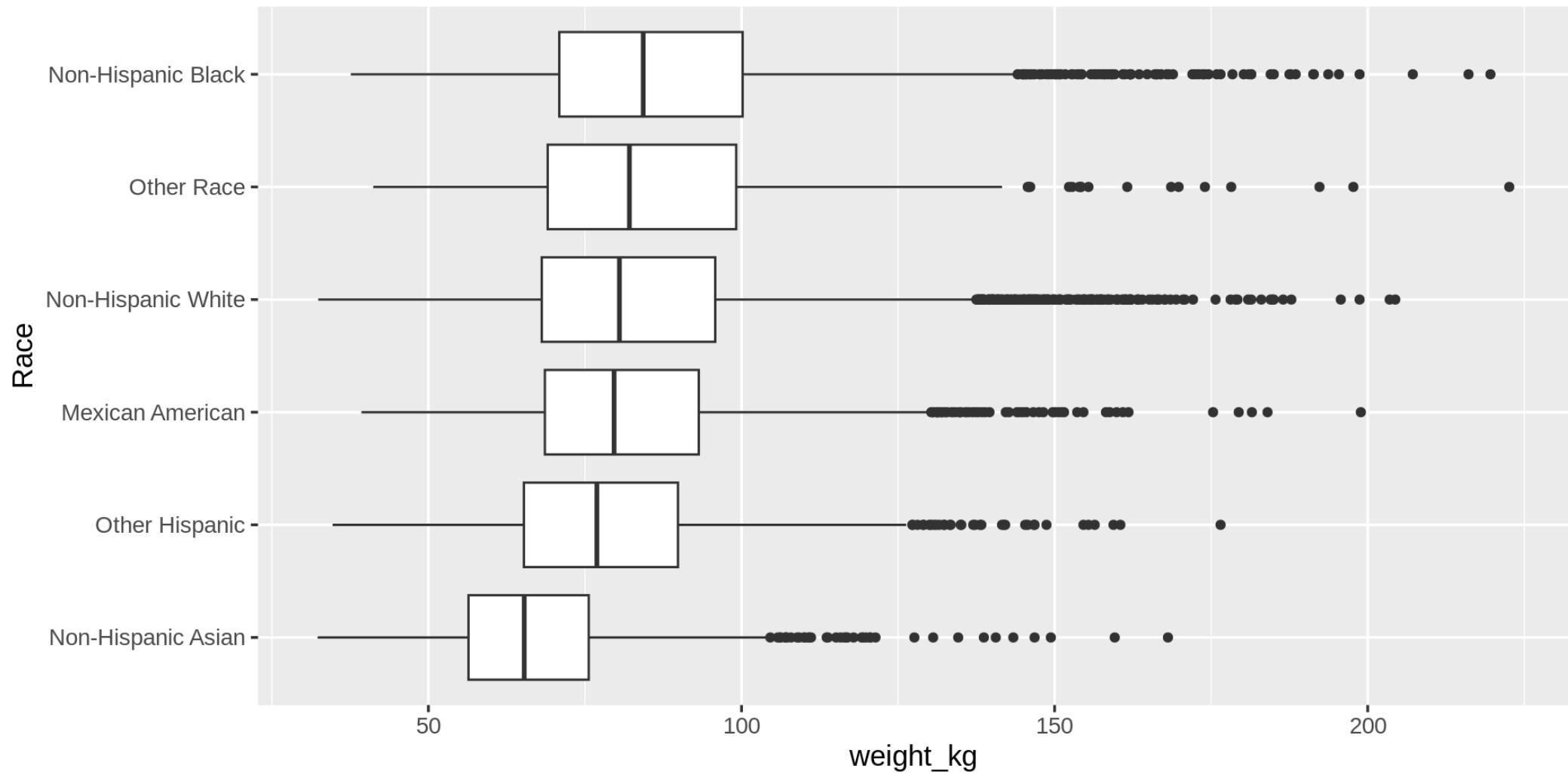
# Як спосіб життя відрізняється залежно від раси?





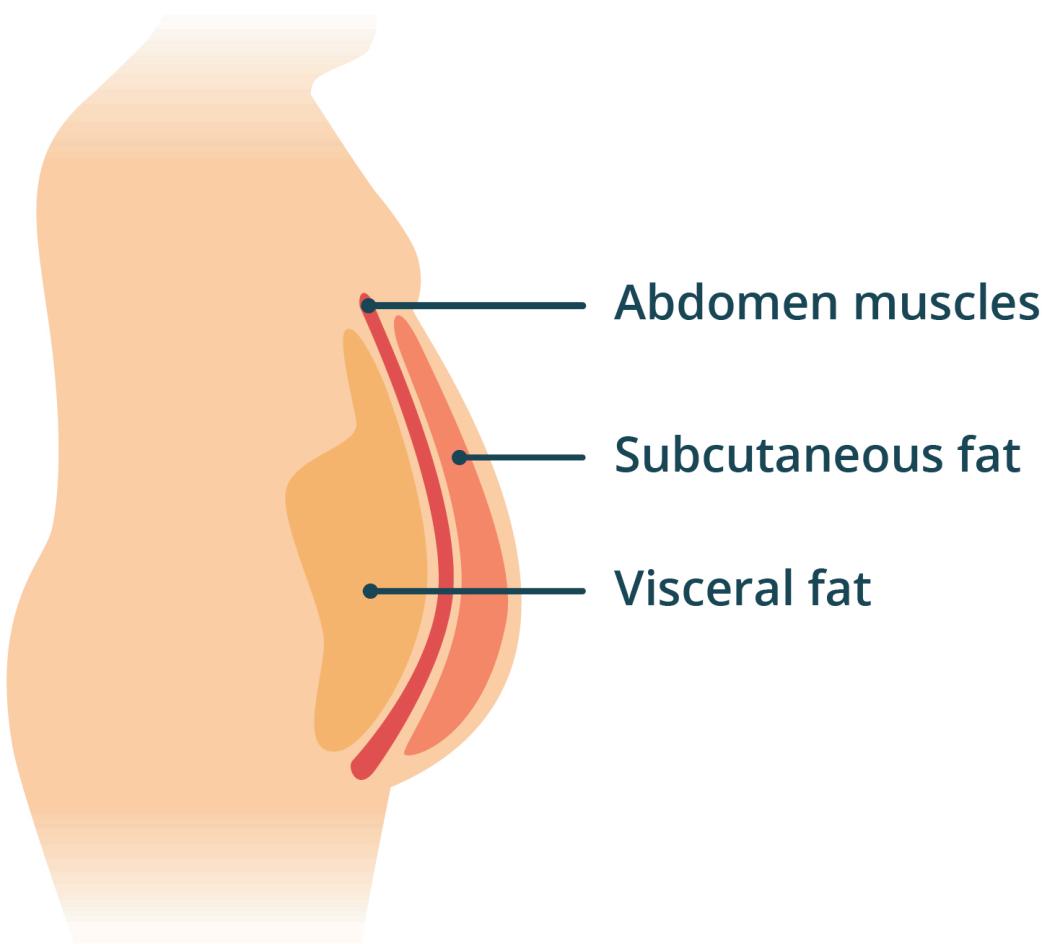
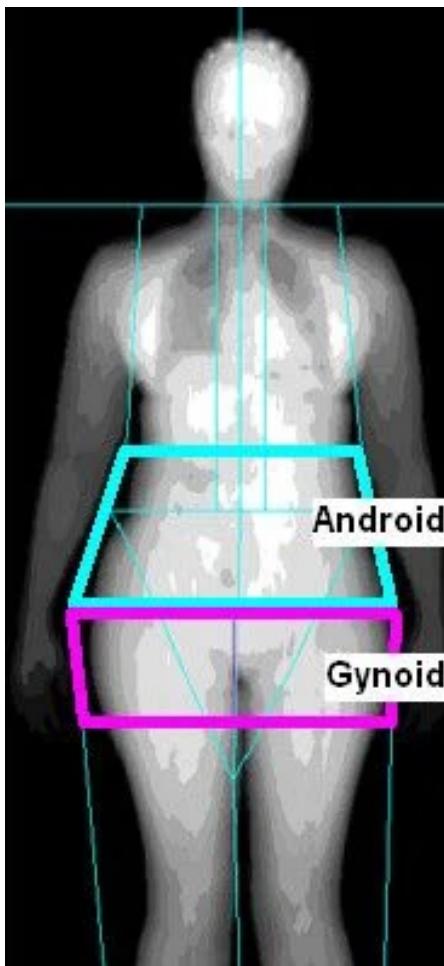




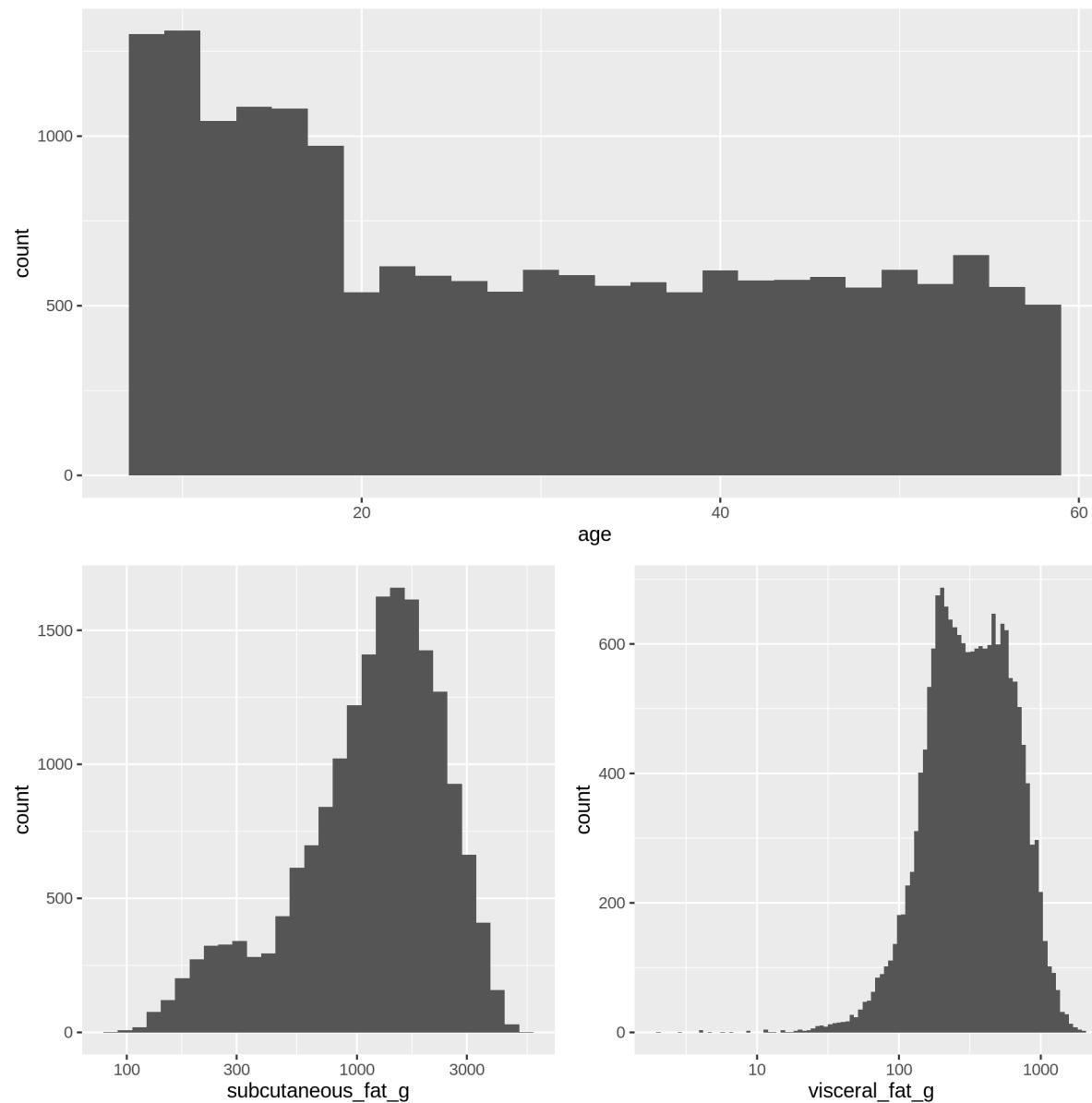


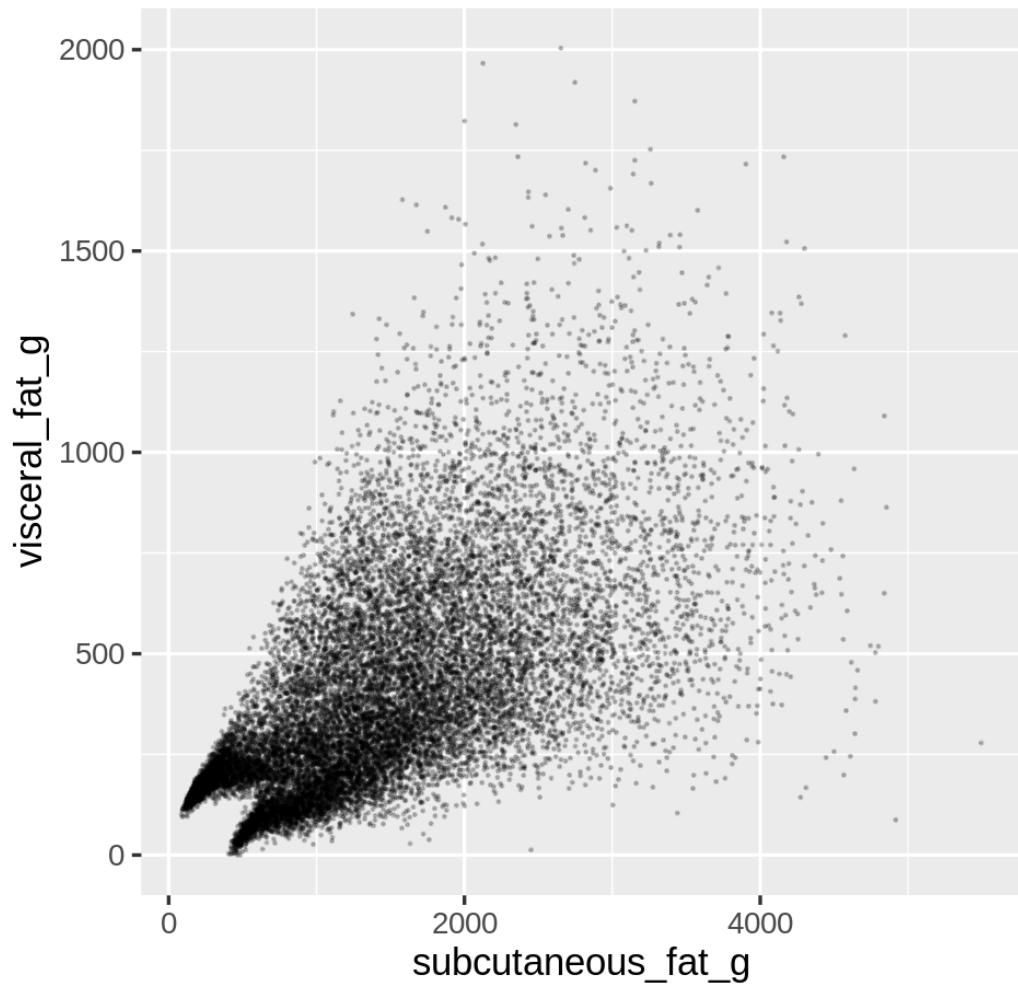
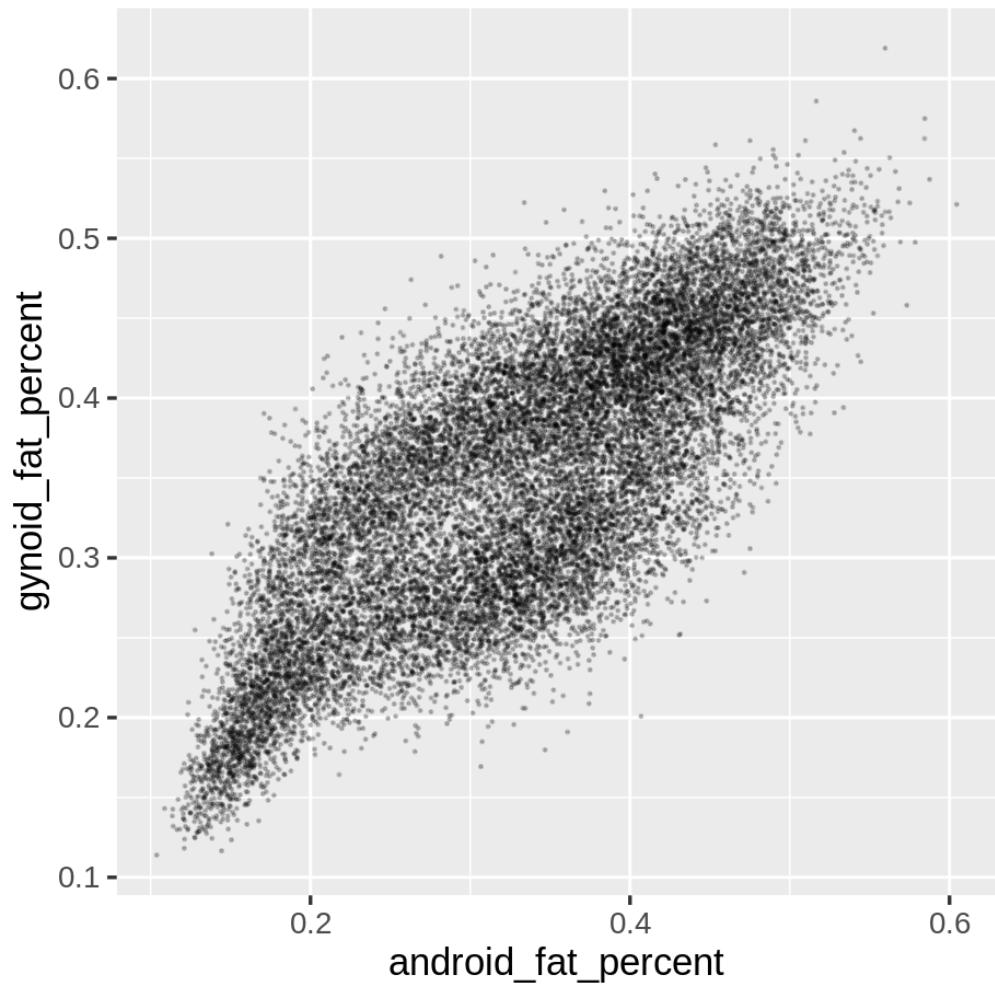
**Як розподіл жирової тканини впливає на артеріальний тиск?**

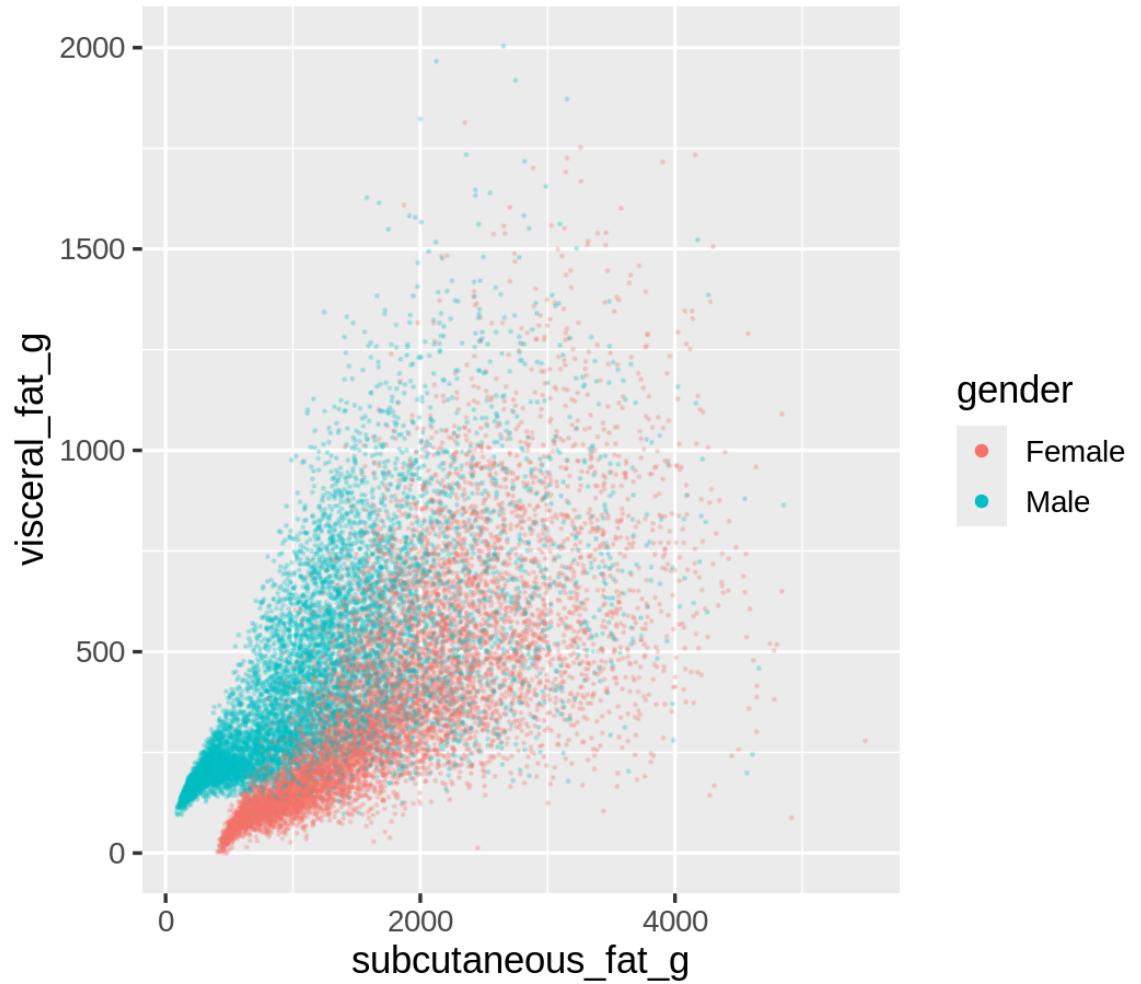
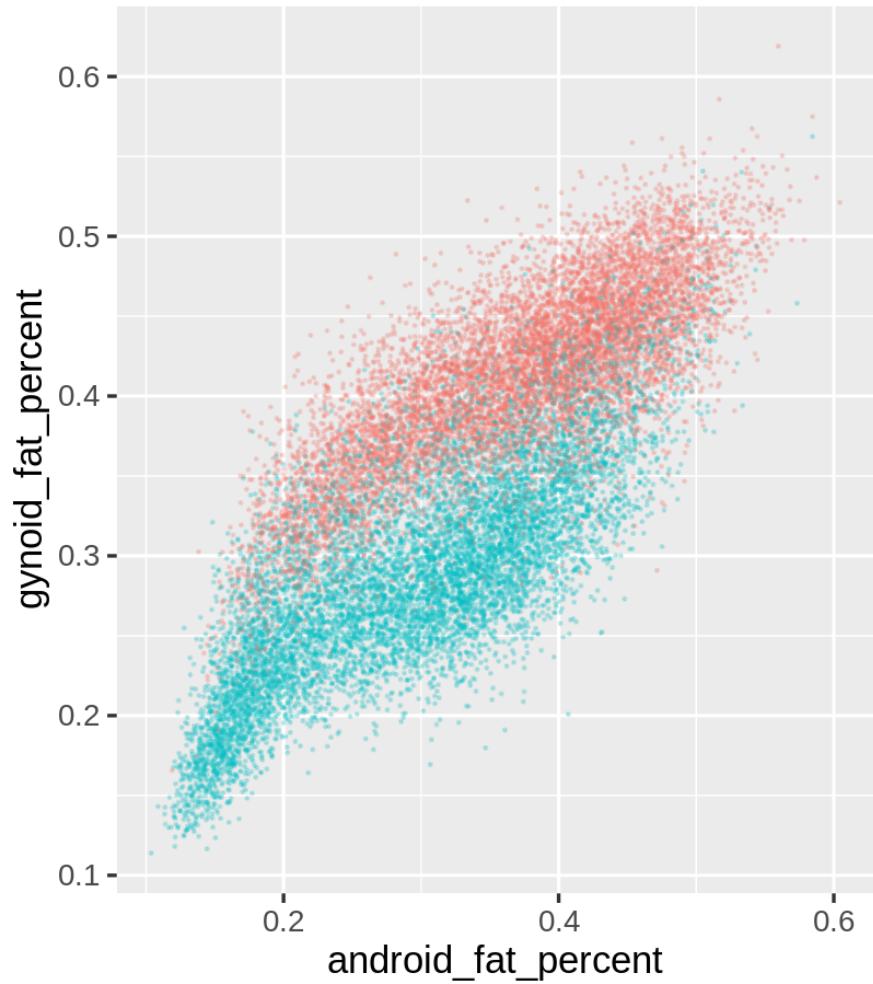
# Види жиру



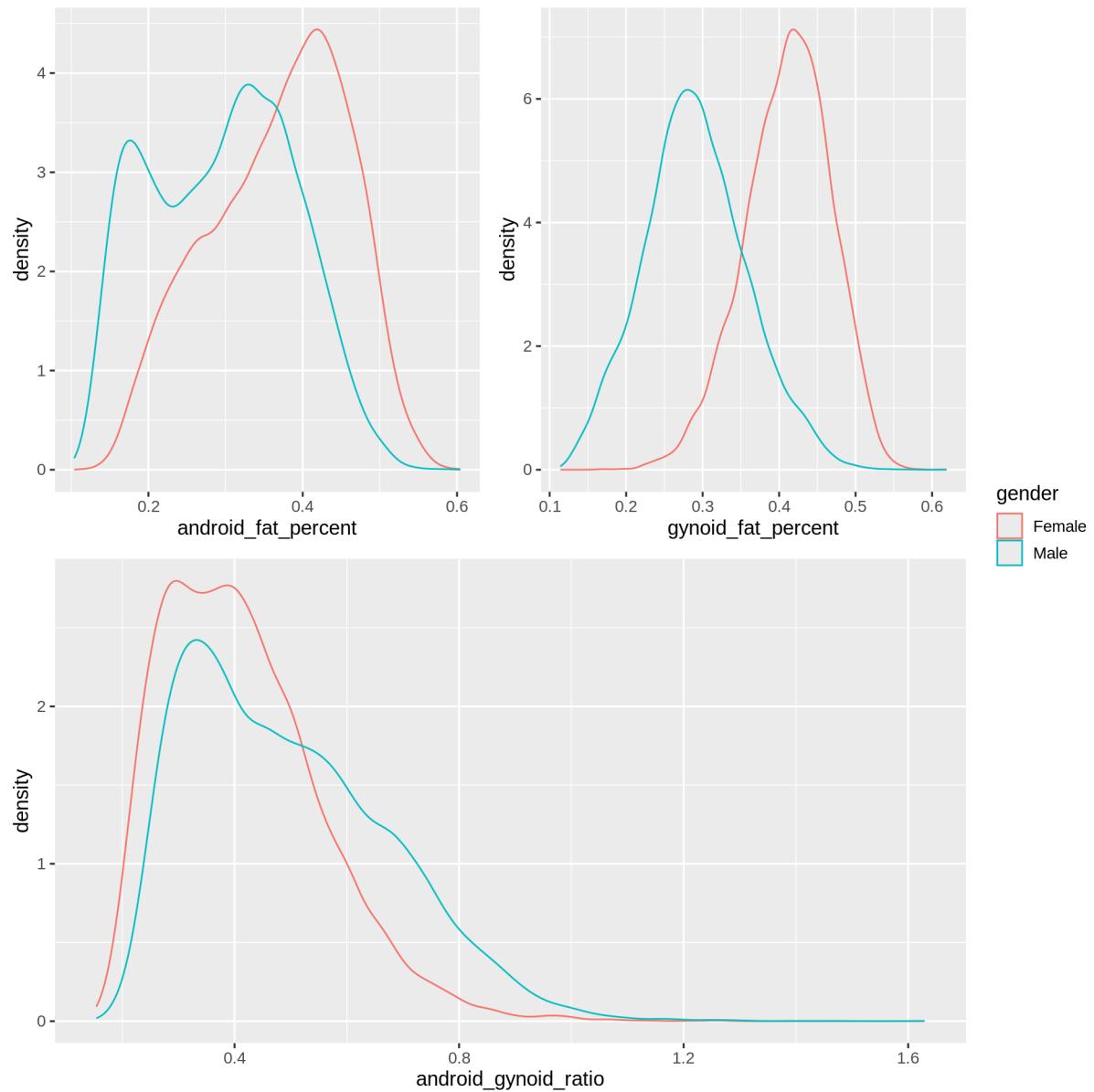
# Розподіл числових змінних





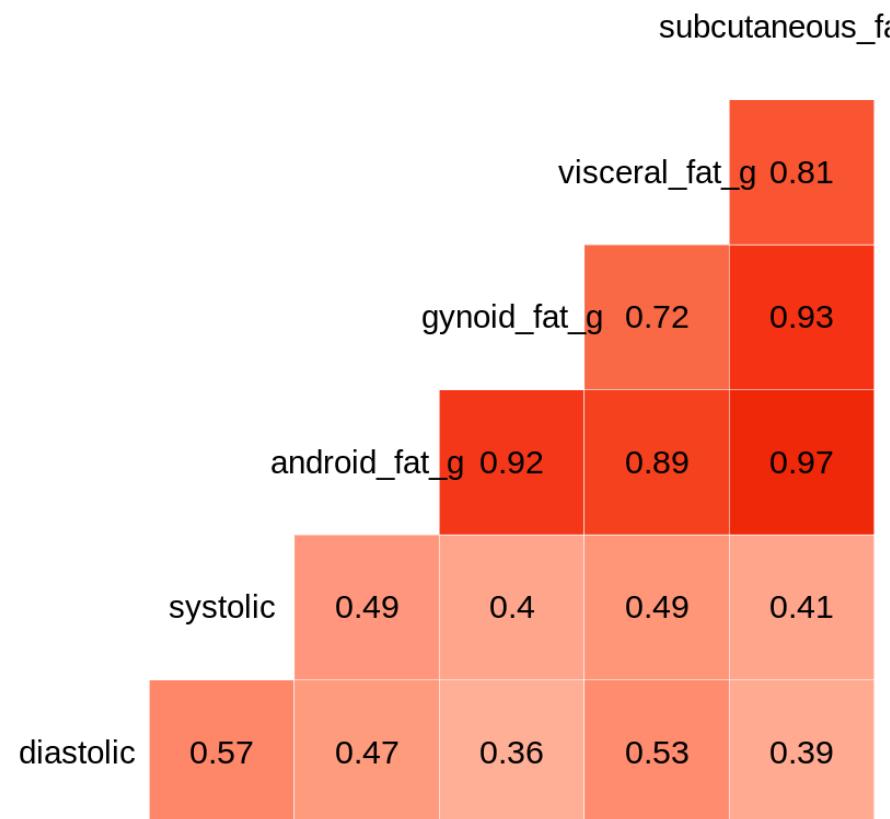


# Відмінності між чоловіками та жінками

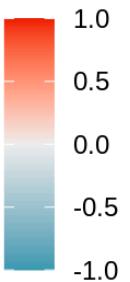
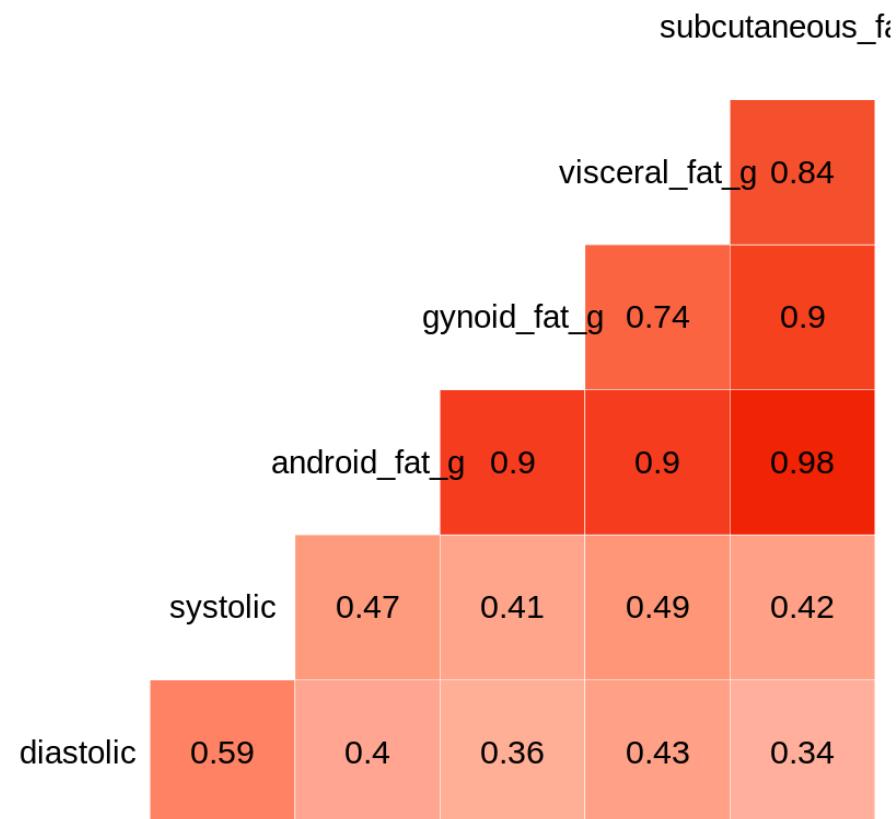


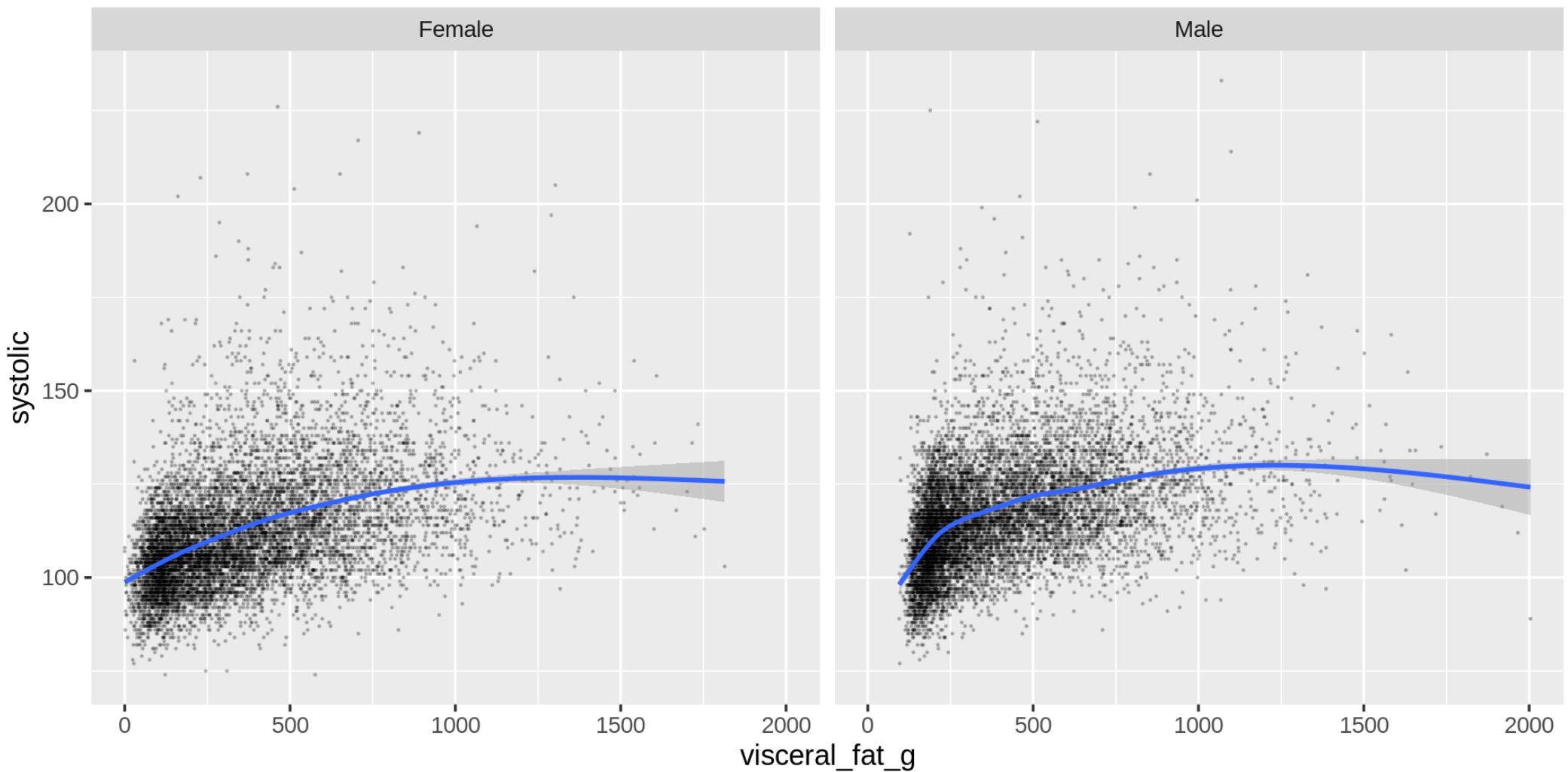
# Чи свідчить більша кількість жиру про гірший артеріальний тиск?

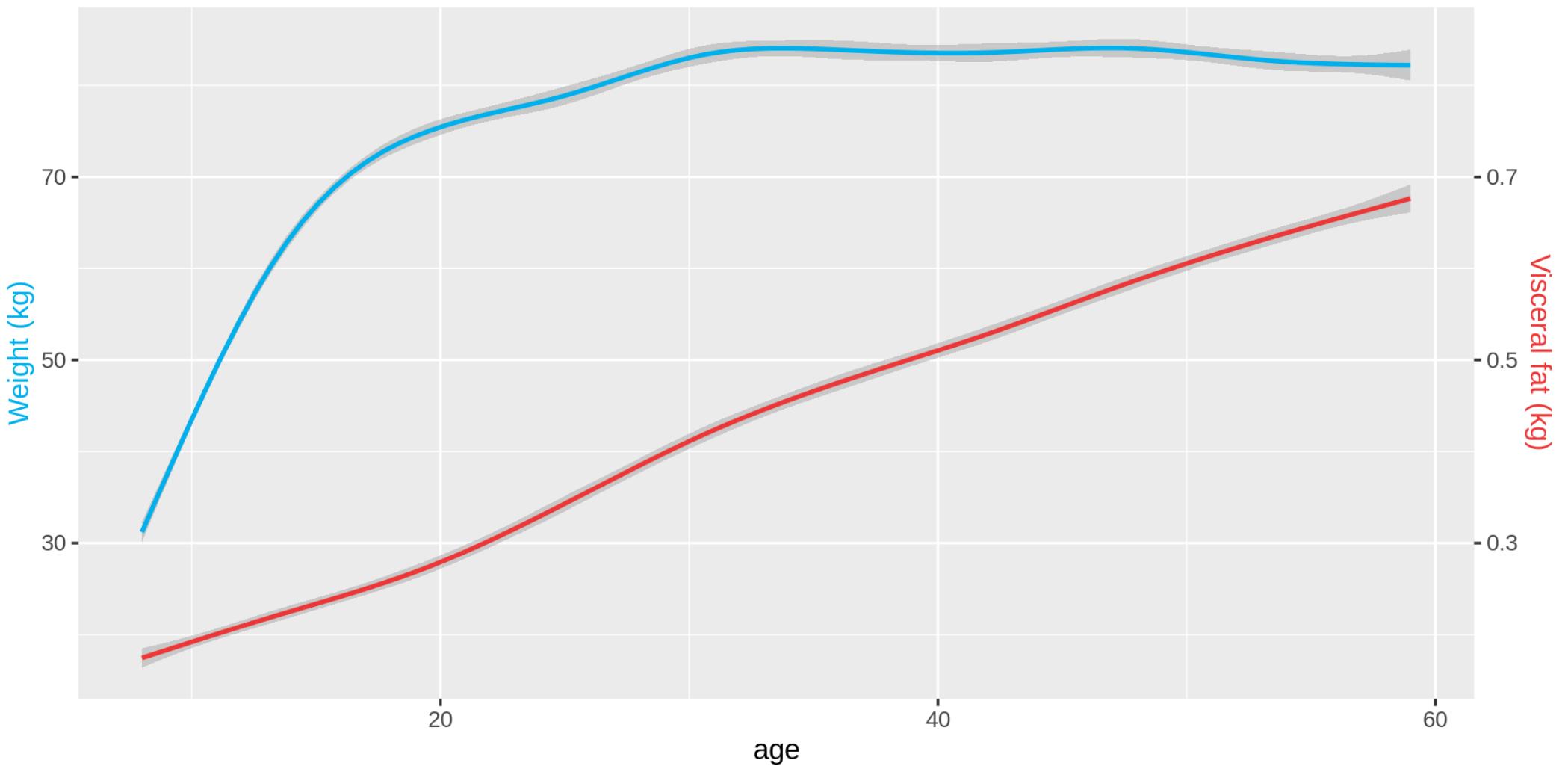
Male



Female



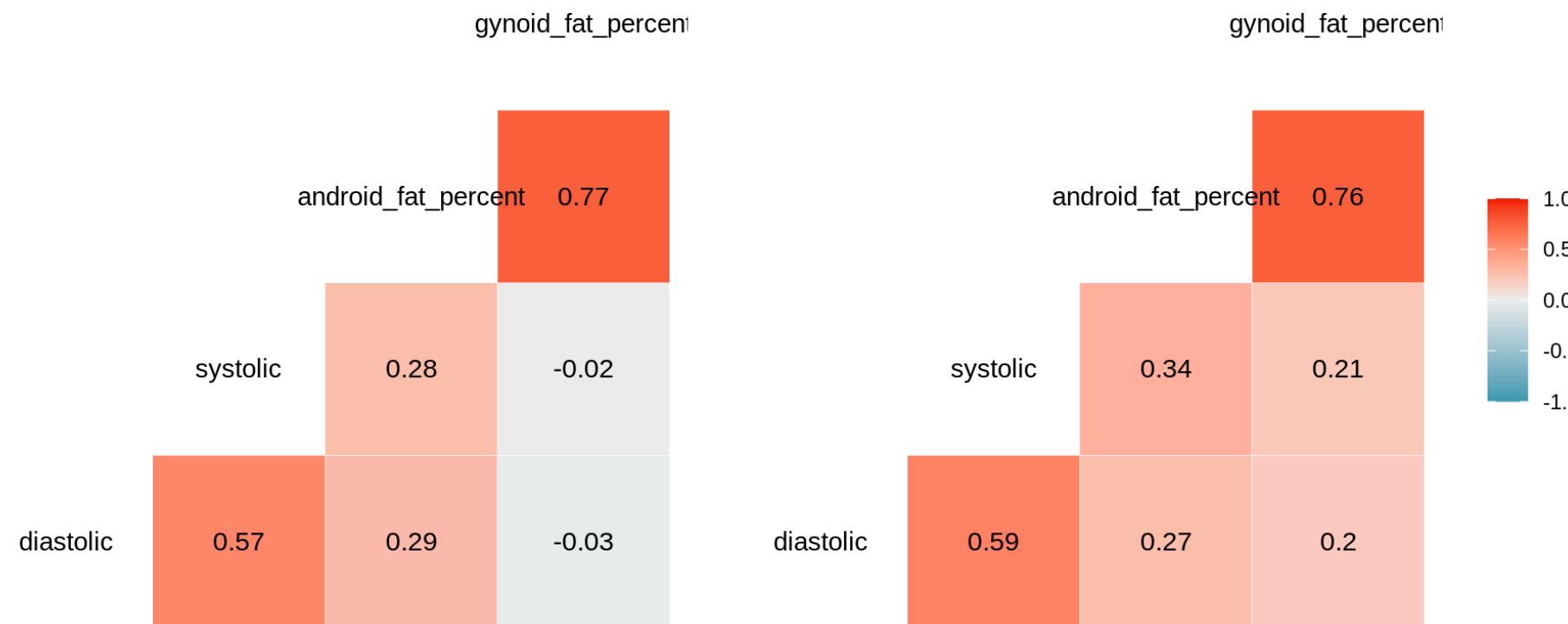




# Що більше впливає на підвищення артеріального тиску: жирова чи нежирова маса?

Male

Female



# Чи андроїдний жир шкідливіший за гіноїдний?

Male

Female

android\_gynoid\_ratio

systolic

0.49

diastolic

0.57

0.53

android\_gynoid\_ratio

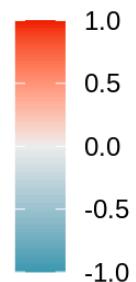
systolic

0.41

diastolic

0.59

0.32



# **Висновки**

- Найбільш помітний зв'язок з артеріальним тиском спостерігався у таких факторів, як вік, вага та кількість жиру
- Фактори способу життя мали певний, але незначний вплив
- Чоловіки та жінки відрізнялися за розподілом жиру
- Різні типи жиру мали різний вплив на артеріальний тиск