

# Лабораторна робота 4

Непараметрична регресія. Аналіз головних компонент

Піковець Артем КМ-22

- Питання: Як кількість вісцерального жиру впливає на систолічний тиск людини?

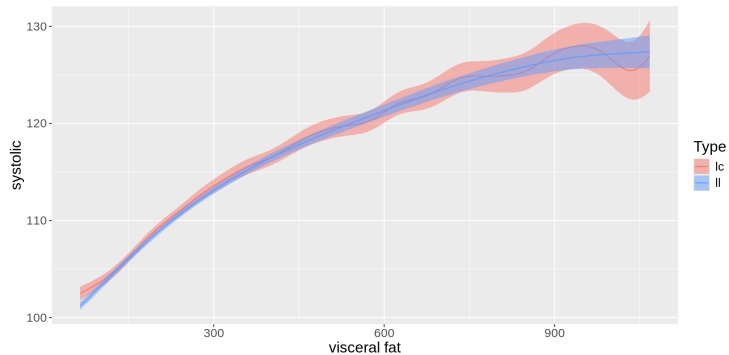
- Дані взяті з NHANES за роки 2011-2018.
- Були взяті базові демографічні змінні, та зміні що стосуються артеріального тиску та розподілу жирової й нежирової маси в тілі людини.
- Отримана вибірка містить 16788 людей.
- Вибірку розділено на тренувальну (50%) та тестову (50%).

# Модель 1

- $y = m(\text{visceral fat}) + \varepsilon$

# Модель 1

- $y = m(\text{visceral fat}) + \varepsilon$

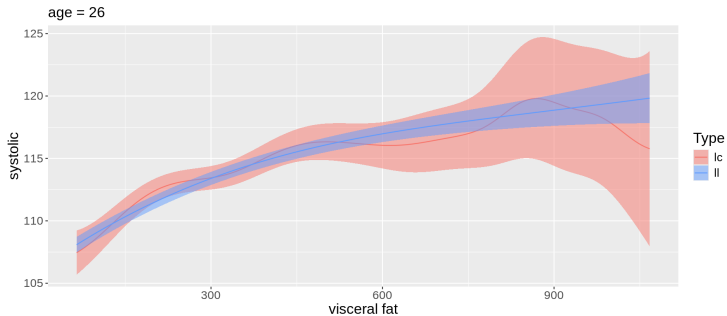


## Модель 2

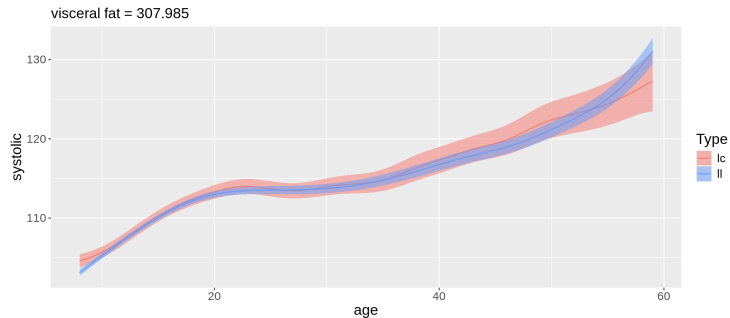
- $y = m(\text{visceral fat}, \text{age}) + \varepsilon$

# Модель 2

- $y = m(\text{visceral fat}, \text{age}) + \varepsilon$



# Модель 2





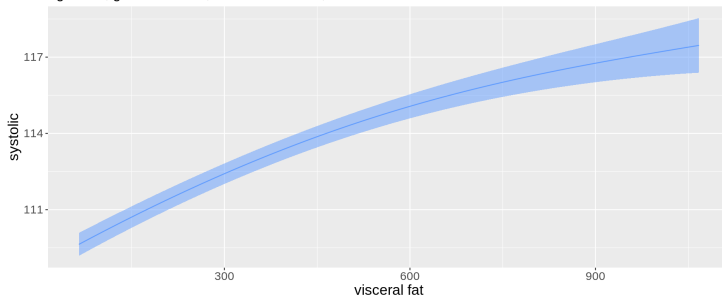
# Частково лінійна модель

- $y = m(\text{visceral fat}, \text{age}) + \beta_{is\ male} \cdot is\ male + \beta_{is\ black} \cdot is\ black + \beta_{lean\ mass} \cdot \ln(\text{lean mass}) + \varepsilon$

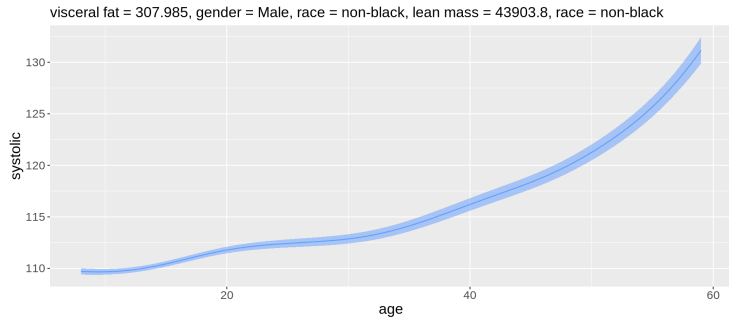
# Частково лінійна модель

- $$y = m(\text{visceral fat}, \text{age}) + \beta_{is\ male} \cdot is\ male + \beta_{is\ black} \cdot is\ black + \beta_{lean\ mass} \cdot \ln(\text{lean mass}) + \varepsilon$$

age = 26, gender = Male, race = non-black, lean mass = 43903.8



# Частково лінійна модель

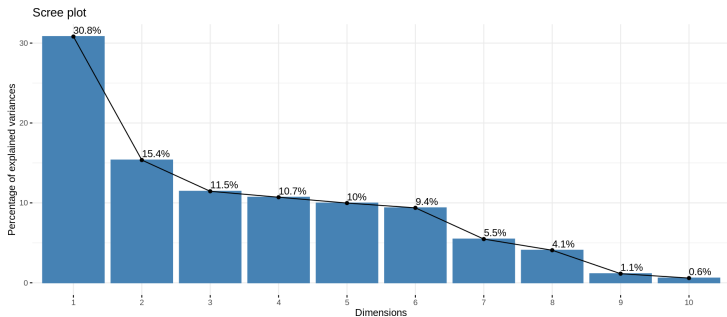


# Порівняння частково лінійної моделі з лінійною

	(1)	(2)
(Intercept)	-4.557 (5.193)	
genderMale	1.471*** (0.219)	1.475*** (0.337)
raceNon-Hispanic Black	3.611*** (0.233)	3.746*** (0.328)
log(lean_mass_g)	10.582*** (0.491)	10.752*** (0.775)
log(visceral_fat_g)	-1.162+ (0.675)	
age	-1.132*** (0.270)	
I(age^2)	0.006 (0.005)	
I(age^3)	0.000*** (0.000)	
log(visceral_fat_g) × age	0.332*** (0.053)	
log(visceral_fat_g) × I(age^2)	-0.005*** (0.001)	
Num.Obs.	16788	8394

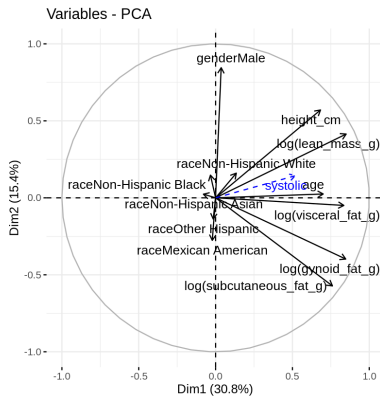
- Для PCA було взято 12 змінних.

- Для PCA було взято 12 змінних.

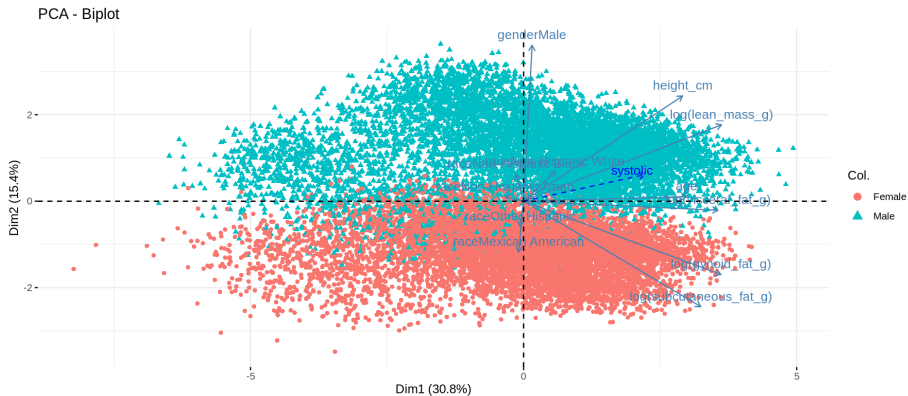


- Перші 6 компонент пояснюють 87.8% дисперсії.

# PCA

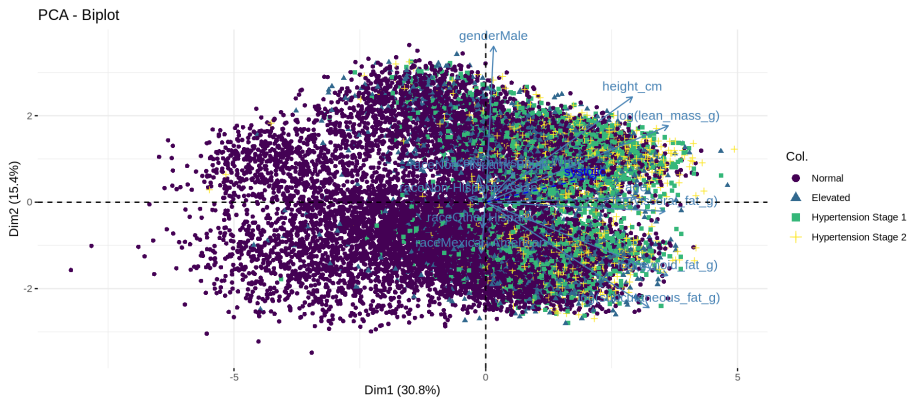


# PCA

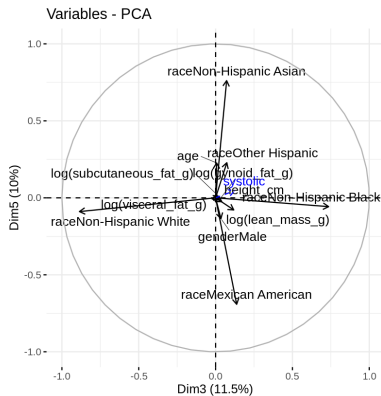




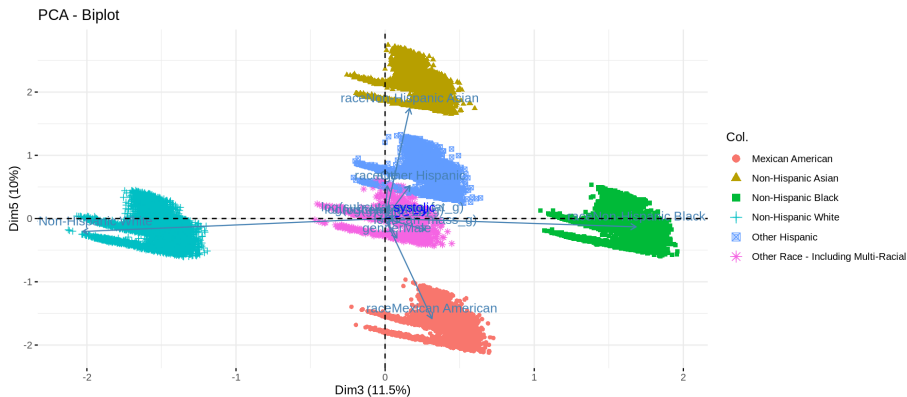
# PCA



# PCA



# PCA



- Непараметричні моделі дали дещо схожий результат з попередньою лабораторною роботою: зі збільшенням кількості вісцерального жиру підвищується артеріальний тиск, але сам вплив незначний. Побудовані непараметричні моделі не враховували деякі важливі фактори взаємодії (наприклад, між кількість жиру та статтю).
- Непараметричні моделі дали змогу побачити нелінійні зв'язки (такі як між віком та тиском).
- Значного зменшення розмірності досягти за допомогою PCA не вдалось, але можна використати деякі компоненти як змінні які характеризують певні характеристики людини (такі як її розмір).