# Лабораторна робота №3.

**Використання модуля Pandas.**

**Аналіз даних по серцево-судинних захворюваннях**

***Мета роботи:*** набути навичок роботи з модулем Pandas та провести первинний аналіз даних.

**Завдання**

Скільки чоловіків і жінок представлено в цьому наборі даних? Не було дано опису ознаки «стать» (якої статі відповідає 1, а якої - 2 в ознаці *gender*) - це можна визначити подивившись на зріст, при розумному припущенні в середньому чоловіки вище.

if df['height'].mean() > df[df['gender'] == 1].mean()['height']:

print('1 - women 2 - man')

else:

print('2 - women 1 - man')

Хто в середньому рідше вказує, що вживає алкоголь - чоловіки чи жінки?

if df[(df['alco'] == 1) & (df['gender'] == 1)]['gender'].value\_counts()[1] > df[(df['alco'] == 1) & (df['gender'] == 2)]['gender'].value\_counts()[2]:

print('man')

else:

print('women')

У скільки разів (округлити, *round*) відсоток курців серед чоловіків більше, ніж відсоток курців серед жінок (принаймні, за цими анкетними даними)?

percent\_smok\_man = (df[(df['alco'] == 1)]['gender'].value\_counts()[2] \* 100) / df['gender'].value\_counts()[2]

print("Smoke man %:" + repr(percent\_smok\_man))

percent\_smok\_women = (df[(df['alco'] == 1)]['gender'].value\_counts()[1] \* 100) / df['gender'].value\_counts()[1]

print("Smoke woman %:" + repr(percent\_smok\_women))

print('Smoke man %/Smoke woman %')

print(percent\_smok\_man/percent\_smok\_women)

Побудуйте нову ознаку - BMI (*Body Mass Index*). Для цього треба вагу у кілограмах поділити на квадрат зросту в метрах. Нормальними вважаються значення BMI від 18.5 до 25. Виберіть вірні твердження.

***Твердження:***

• Медіанний BMI по вибірці перевищує норму. вірно

• У жінок в середньому BMI нижче, ніж у чоловіків. не вірно

print (df['weight'].mean()/ ((df['height'].mean()/100) \* (df['height'].mean()/100)))

print (df[(df['gender']==1)]['weight'].mean() / ((df[(df['gender']==1)]['height'].mean()/100) \* (df[(df['gender']==1)]['height'].mean()/100)))

print (df[(df['gender']==2)]['weight'].mean() / ((df[(df['gender']==2)]['height'].mean()/100) \* (df[(df['gender']==2)]['height'].mean()/100)))