

виступати charges (витрати). Також будемо враховувати bmi і children. Логістичне модель дає змогу передбачити чи буде збільшення витрат.

Використавши predict() стало зрозуміло що витрати будуть тільки збільшуватися.

```
lm(formula = charges ~ age + bmi + children)
Residuals:
Min    1Q  Median    3Q   Max
-13884 -6994 -5092  7125 48627
Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) -6916.24   1757.48  -3.935 8.74e-05 ***
age          239.99    22.29  10.767 < 2e-16 ***
charges      542.86    258.24   2.102 0.0357 *
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 11370 on 1334 degrees of freedom
Multiple R-squared:  0.1201, Adjusted R-squared:  0.1181
F-statistic: 60.69 on 3 and 1334 DF, p-value: < 2.2e-16
```

Зі звіту можемо побачити тенденцію зростання величини страховки зі збільшенням віку клієнтів.

Отже, побудована логістична модель визначила залежність віку клієнта і показала, що надалі витрати будуть тільки збільшуватись. Зв'язок безпосередньо присутній. Автоматично зі збільшенням кількості похилих людей, витрати тільки будуть збільшуватись і витрати невідворотні. Також визначення страховки залежить і від інших параметрів, які були перераховані вище. За допомогою пакету R можна безпосередньо його прослідкувати.

#### **Список літератури**

1. Сайт визначення датасету <https://www.kaggle.com/goksunazlican/analysis-of-medical-cost-personal-datasets>
2. Джеймс Г., Уиттон Д., Тибишрани Р. Введение в статистическое обучение с примерами R. Изд. Второе, испр. Пер с англ. С.Э. Мاستицкого – М. ДМК Пресс, 2017. -456с.

**УДК 004.82:004:85**

*Крохмалюк В. В., студент 2 курсу спеціальності 122 «Комп'ютерні науки»  
Нескородєва Т. В., к.т.н., доцент, доцент кафедри інформаційних технологій*

## **АНАЛІЗ ДАНИХ ПРО СЕРДЕЧНІ ЗАХВОРЮВАННЯ МЕТОДАМИ СТАТИСТИЧНОГО НАВЧАННЯ**

*Донецький національний університет імені Василя Стуса, м. Вінниця*

Серцеві хвороби – це цілий ряд різних хвороб, які уражають серце та серцево-судинну систему. Вивченням хвороб серця, їх профілактикою та лікуванням займається кардіологія.

До захворювань серця відносяться, наприклад, хвороби кровеносних судин, такі, як захворювання коронарної артерії, аритмія, інфекційні захворювання серця, вроджені вади серця. Багато різновидів серцевих захворювань можна запобігти за допомогою здорового способу життя.

Щорічно вперше виявляється близько 2 млн хворих з цією патологією, з них кожний другий – працездатного віку. Смертність від хвороб серця та систем кровообігу в Україні займає перше місце, і у 2-4 рази вища, ніж у країнах ЄС та світу, причому в нашій країні вмирають від цих захворювань не тільки частіше, але й раніше.

Аналіз проведемо за допомогою пакету R. Для дослідження був використаний набір даних Heart. Ця база даних містить 76 атрибутів, але всі опубліковані експерименти відносяться до використання підмножини з 14 з них. Зокрема, база даних Клівленда є єдиною, до якої дослідники МЛ користувалися. Поле «мета» означає наявність у пацієнта хвороби серця. Ціле число оцінюється від 0 (немає присутності) до 4.

Інформація про атрибути:

- age – Вік
- sex – (1 = male; 0 = female)
- 3cp – Тип болю в грудях
- 4trestbps - Артеріальний тиск у спокої (у мм рт.ст. при надходженні в лікарню)
- chol – Сироватковий холестерин в мг / дл
- fbs – (цукор у крові натще > 120 мг / дл) (1 = вірно; 0 = помилково)
- restecg – Електрокардіографічні результати
- thalach – Досягнутий максимальний пульс
- exang – Стенокардія, викликана фізичними вправами (1 = так; 0 = ні)
- oidpeak – Депресія, викликана фізичними вправами відносно спокою
- slope – Нахил пікового вправи сегмента
- ca – Кількість основних судин (0-3), забарвлених флоросопією
- thal – 3 = нормальний; 6 = виправлений дефект; 7 = поворотний дефект
- target – 1 або 0

Створимо матрицю діаграм розсіювання за допомогою команд `attach(heart)` та `plot(heart)` для повного розуміння картини.

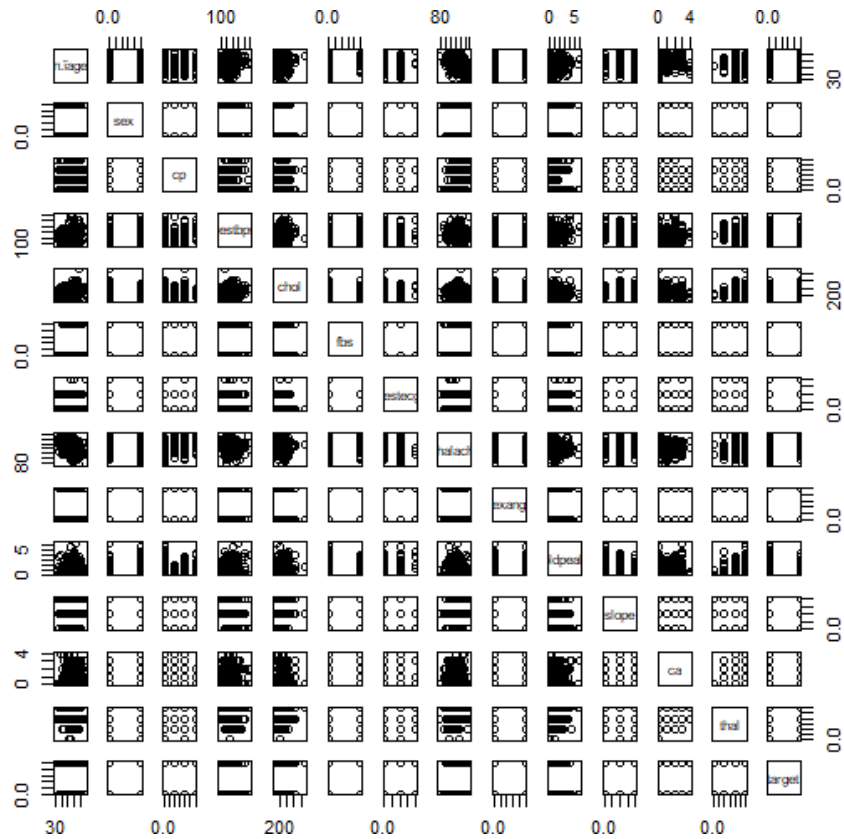


Рисунок 1 – Матриця діаграм розсіювання

Тепер скористаємось тими ж командами `attach()` та `plot()` для того, щоб дізнатися, в якому віці смертність є найвищою.

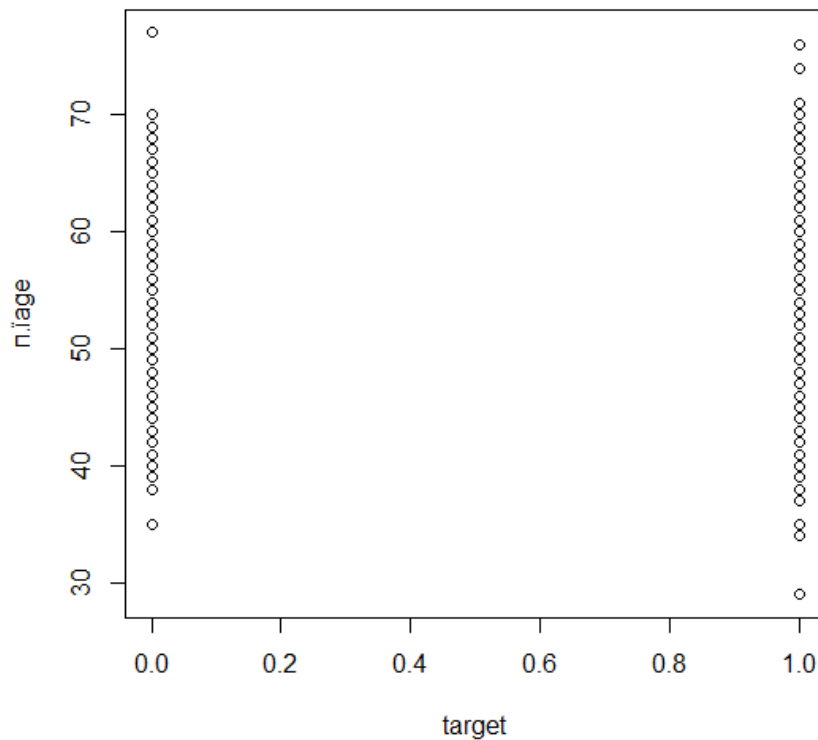
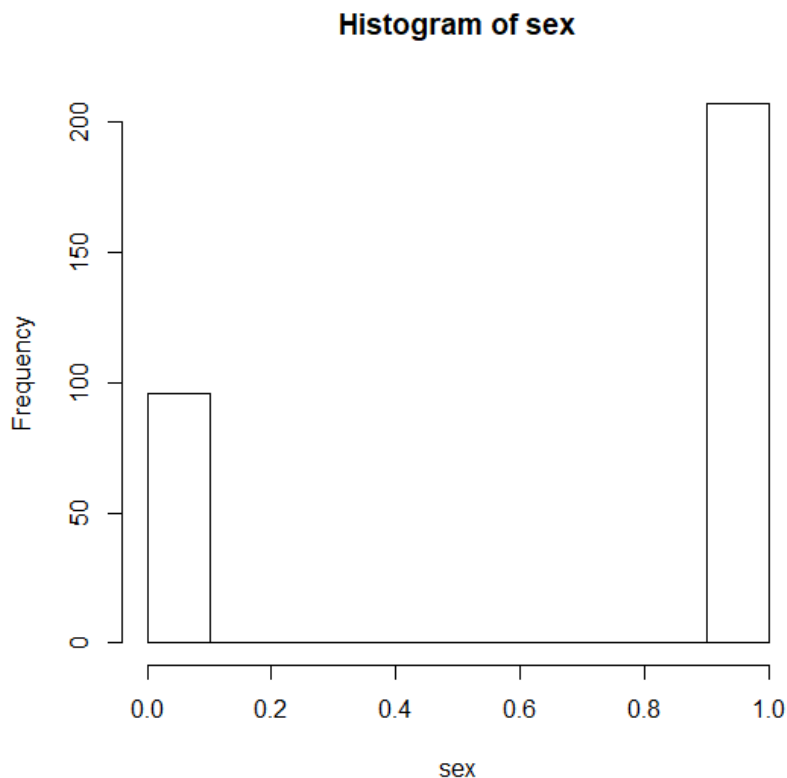


Рисунок 2 – Діаграма розсіювання

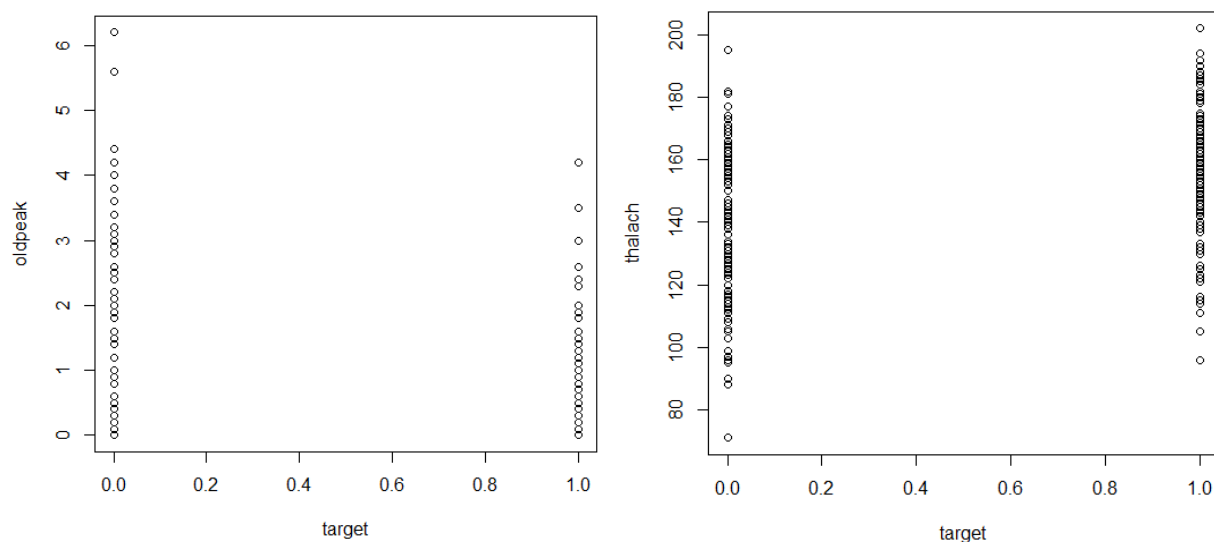
Зобразивши дані графічно, ми можемо бачити, що в основному, це люди у віці від 40 до 70 років. Та у більше ніж 50 % випадків життя врятувати не вдається, але також є рідкі випадки летального результату і у віці до 40 років.

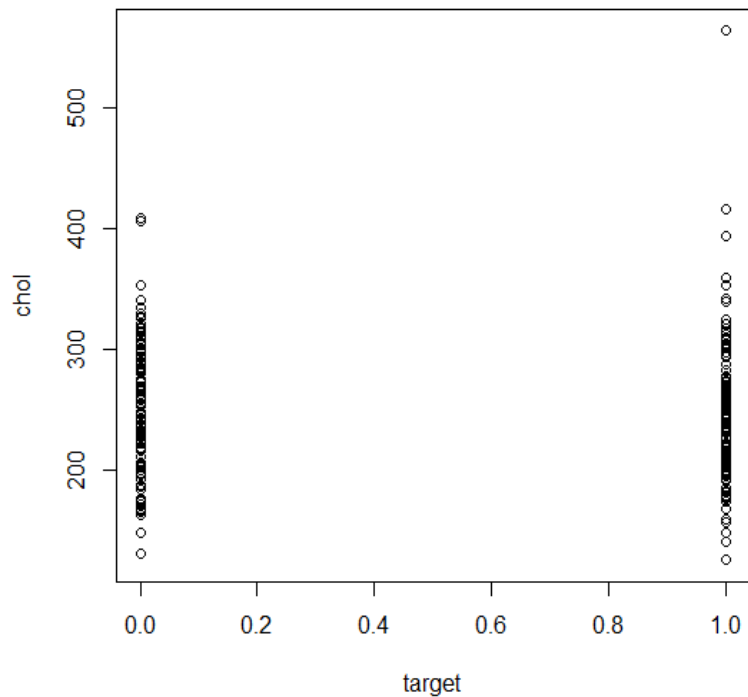
Спробуємо в'яснити, яка стать частіше стикається з цим недугом (рис.3)



*Рисунок 3 – Гістограма*

Уважно подивившись на (рис.3) ми можемо точно сказати, що чоловіки частіше мають проблеми з серцево-судинною системою.





Також, якщо звернути увагу на депресію, викликану фізичними вправами відносного спокою (oidpeak), максимальним пульсом(thalach) та рівнем холестерина (chol), то з цих трьох атрибутів, рівень холестерину є основною причиною серцево-судинних захворювань. З чого можна зробити висновок: для того, щоб знизити шанс захворювання серця, потрібно слідкувати за своїм здоров'ям, вести здоровий образ життя та правильно харчуватися, хоча і за своє психічне здоров'я забувати не слід

#### **Список використаної літератури**

1. Wikiwand [Електронний ресурс]. Режим доступу – [https://www.wikiwand.com/uk/%D0%A1%D0%B5%D1%80%D1%86%D0%B5%D0%B2%D1%96\\_%D1%85%D0%B2%D0%BE%D1%80%D0%BE%D0%B1%D0%B8](https://www.wikiwand.com/uk/%D0%A1%D0%B5%D1%80%D1%86%D0%B5%D0%B2%D1%96_%D1%85%D0%B2%D0%BE%D1%80%D0%BE%D0%B1%D0%B8)
2. Kaggle [Електронний ресурс]. Режим доступу - <https://www.kaggle.com/ronitf/heart-disease-uci>

**УДК 004.891.3**

*Литвинюк В. С., студент 3 курсу спеціальності 122 «Комп'ютерні науки»  
Антонов Ю. С., к.ф.-м. наук, доцент,  
доцент кафедри інформаційних технологій*

## **РОЗРОБКА СИСТЕМИ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ ДЛЯ УСУНЕННЯ ПРОБЛЕМ З ТРАНСПОРТНИМ ЗАСОБОМ**

*Донецький національний університет імені Василя Стуса, м. Вінниця*