ПВНЗ «Буковинський університет»

Факультет комп’ютерних систем і технологій

ЗВІТ

Про виконання лабораторної роботи №4

з дисципліни  “Інтелектуальний аналіз даних”

Виконав студент групи К-300  
Крячун Б.С.  
Прийняв викладач  
Артеменко О.І.

Лабораторна робота №4

Тема: Кореляційний та регресійний аналіз в RStudio.

Мета: Провести аналіз зв’язків між змінними. Розрахувати коефіцієнти кореляції Пірсона, Спірмена, Кендалла. Побудувати лінійну та поліноміальну регресію.

Варіант: №3

Посилання на GitHub: <https://github.com/kbs281103/IAD_KBS_K-300>

**Завдання 1**

Опис виконання:

Завантажено вбудований набір даних state.x77. Обрано перші 6 змінних. Обчислено коефіцієнти кореляції Пірсона та Спірмена. Додатково розраховано коваріацію та проведено тест на достовірність для зв'язку між доходом та освітою.

Код в R Studio:

states <- state.x77[, 1:6]

# Кореляції

cor(states)

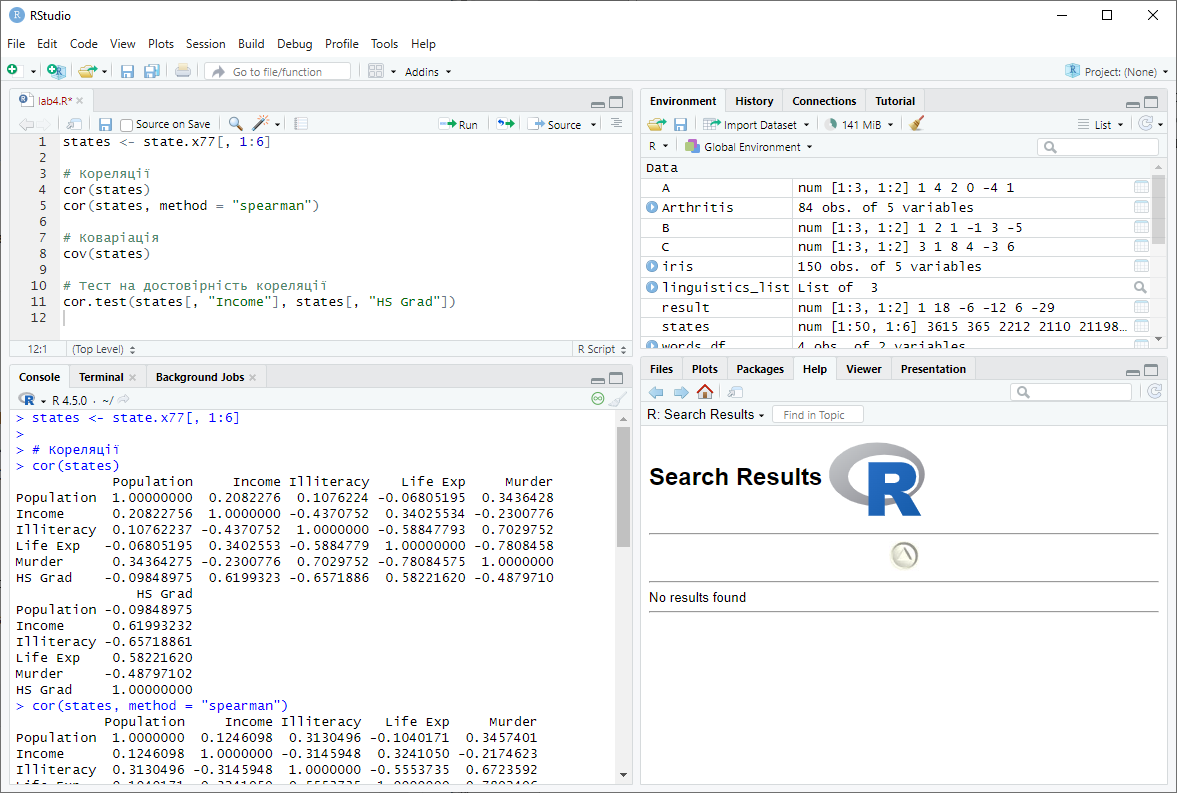
cor(states, method = "spearman")

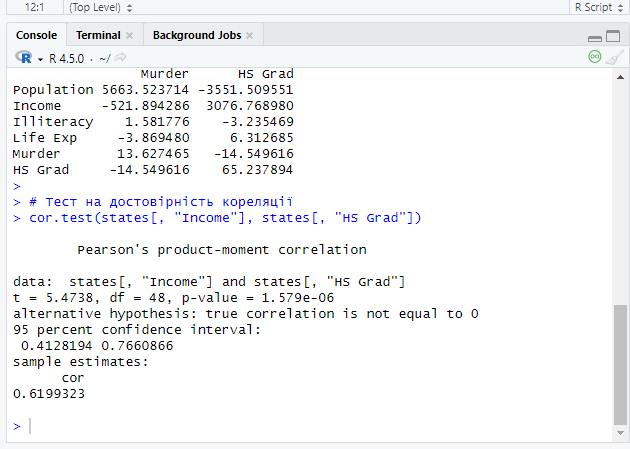
# Коваріація

cov(states)

# Тест на достовірність кореляції

cor.test(states[, "Income"], states[, "HS Grad"])





**Завдання 2**

Опис виконання:

У середовищі Microsoft Excel створено таблицю значень змінних X та Y. Побудовано точкову діаграму (scatter plot), додано лінію тренду (лінійну регресію) з відображенням рівняння. Обчислено коефіцієнт кореляції Пірсона.

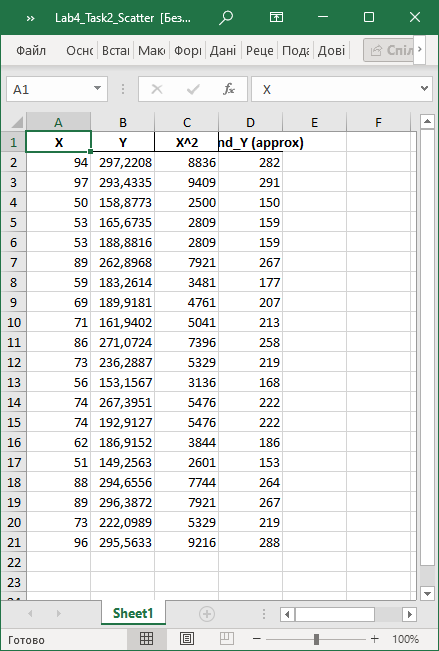
Алгоритм у Excel:

Вставити дані X і Y у дві колонки.

Виділити і вставити графік → “Вставка” → “Точкова”.

Додати трендлайн → “Параметри лінії тренду” → Показати рівняння та R².

Обчислити кореляцію:



Файл з даними: Lab4\_Task2\_Scatter.xlsx

**Завдання 3**

Опис виконання:

Використано набір women. Побудовано модель простої лінійної регресії (вага від зросту), згенеровано графік. Потім побудовано поліноміальну регресію другого ступеня (з додаванням height²), порівняно обидві моделі.

Код в R:

attach(women)

# Лінійна регресія

fit1 <- lm(weight ~ height, data = women)

summary(fit1)

plot(height, weight,

xlab = "Зріст (дюйми)",

ylab = "Вага (фунти)")

abline(fit1)

# Поліноміальна регресія

fit2 <- lm(weight ~ height + I(height^2), data = women)

summary(fit2)

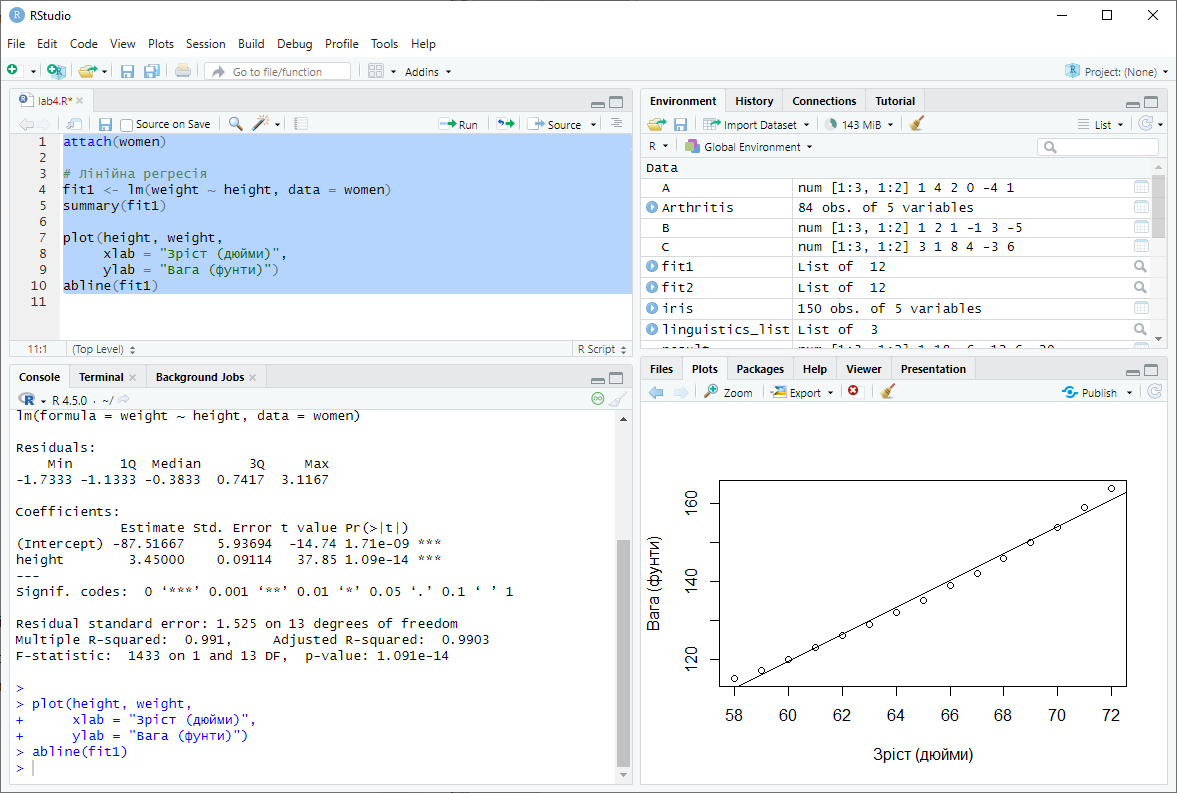
plot(height, weight,

xlab = "Зріст (дюйми)",

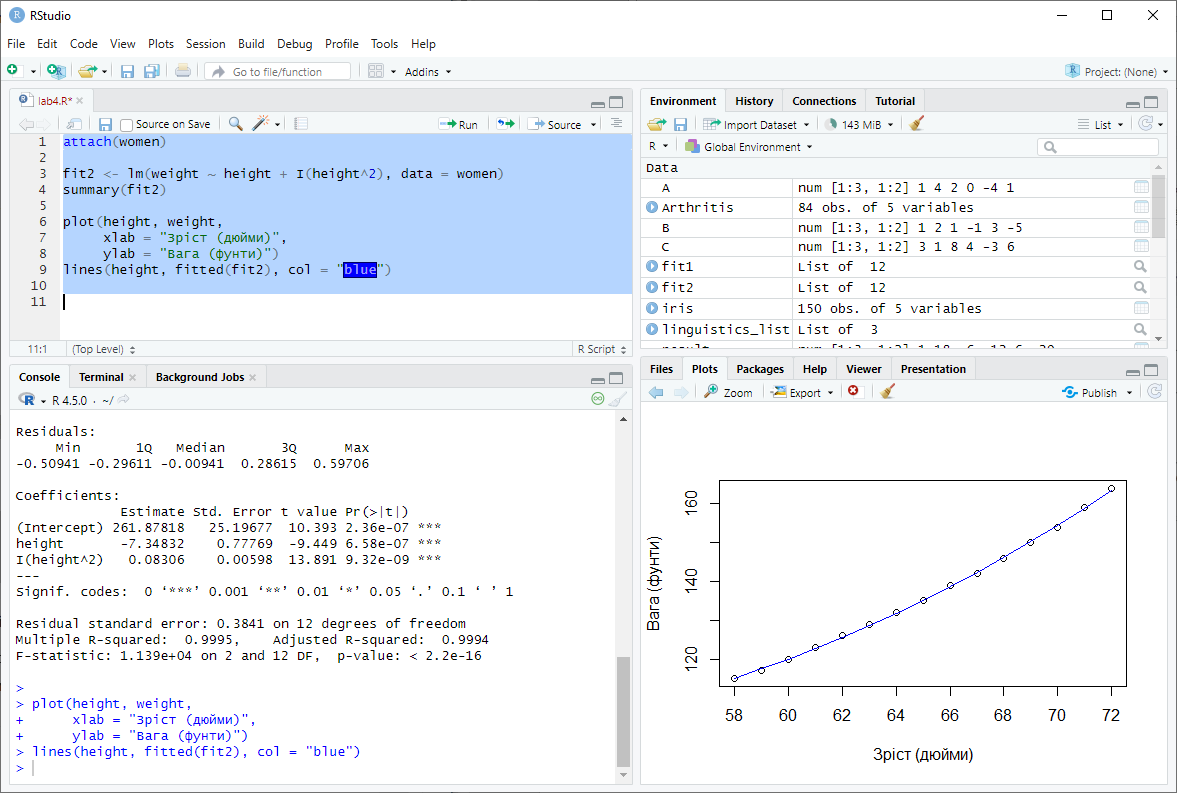
ylab = "Вага (фунти)")

lines(height, fitted(fit2), col = "blue")

Лінійна регресія



Поліноміальна регресія



**Завдання 4**

Опис виконання:

Набір даних: attitude. Виконано аналіз вибраних змінних. Побудовано графіки, обчислено коефіцієнти кореляції Пірсона та Кендалла. Додатково створено графіки попарної кореляції та корелограми.

a) Графіки + регресія + кореляція:

data(attitude)

# Кореляції

cor(attitude$rating, attitude$complaints, method = "pearson")

cor(attitude$rating, attitude$complaints, method = "kendall")

# Графіки

plot(attitude$rating, attitude$complaints,

xlab = "Оцінка",

ylab = "Скарги")

abline(lm(complaints ~ rating, data = attitude), col = "blue")

# Поліноміальна регресія

fit\_poly <- lm(complaints ~ rating + I(rating^2), data = attitude)

lines(attitude$rating, fitted(fit\_poly), col = "red")

[Вставити графік з обома лініями]

b) Графік попарної кореляції:

pairs(attitude)

library(car)

scatterplotMatrix(attitude, spread = FALSE, lty.smooth = 2)

[Вставити обидва графіки: pairs() і scatterplotMatrix()]

c) Використання psych::corPlot:

library(psych)

corPlot(cor(attitude), numbers = TRUE)

[Вставити графік corPlot()]

d) Корелограма:

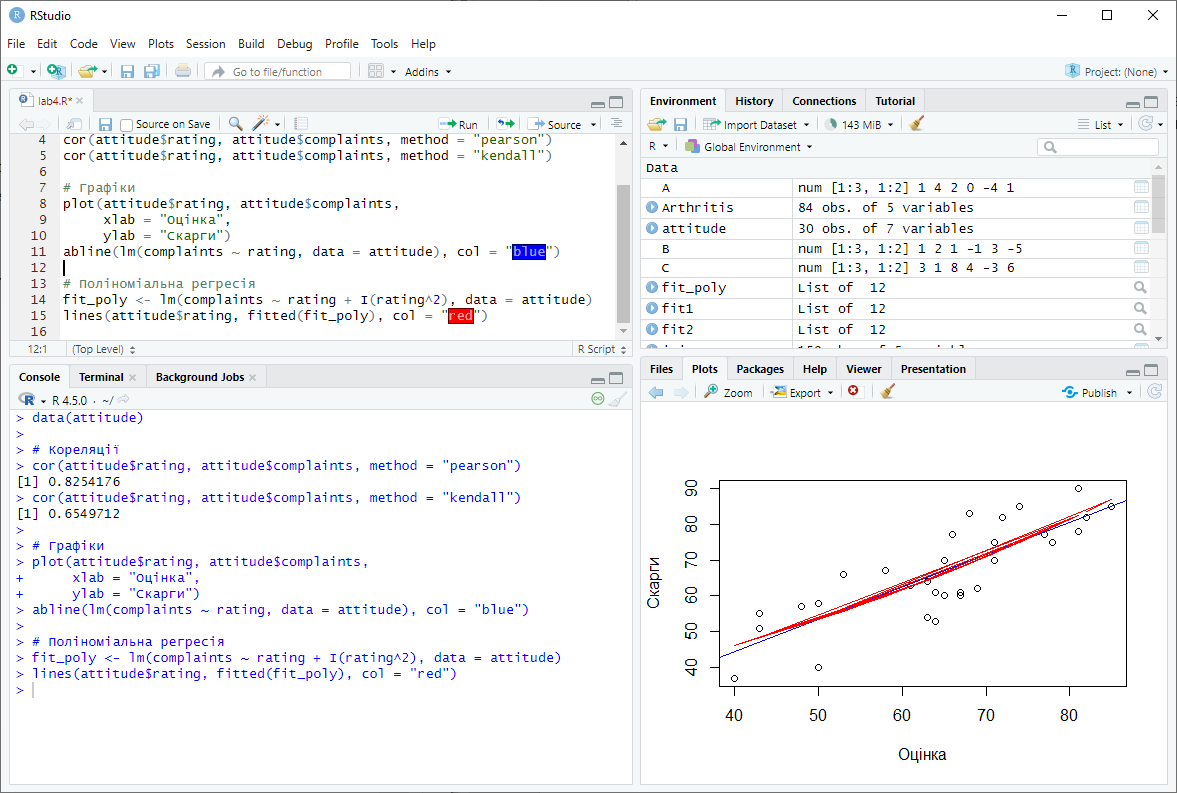
library(corrgram)

corrgram(attitude)

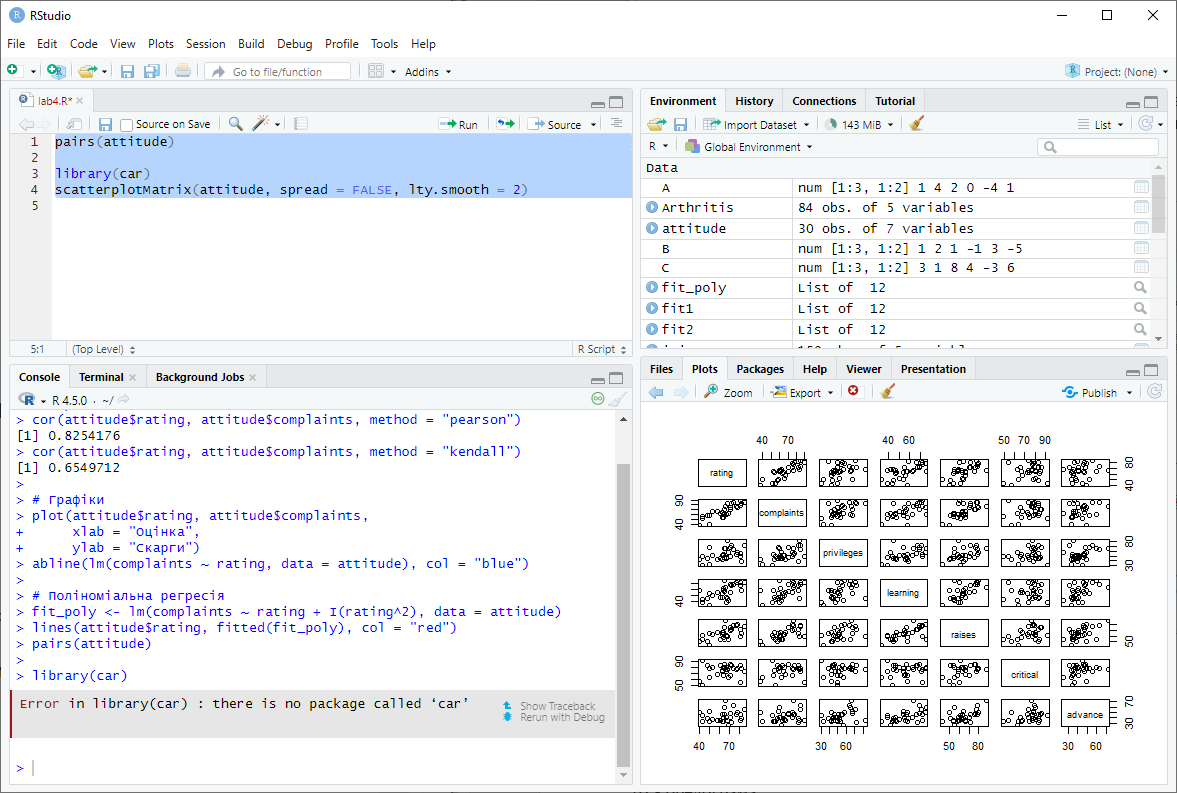
library(corrplot)

corrplot(cor(attitude), method = "circle")

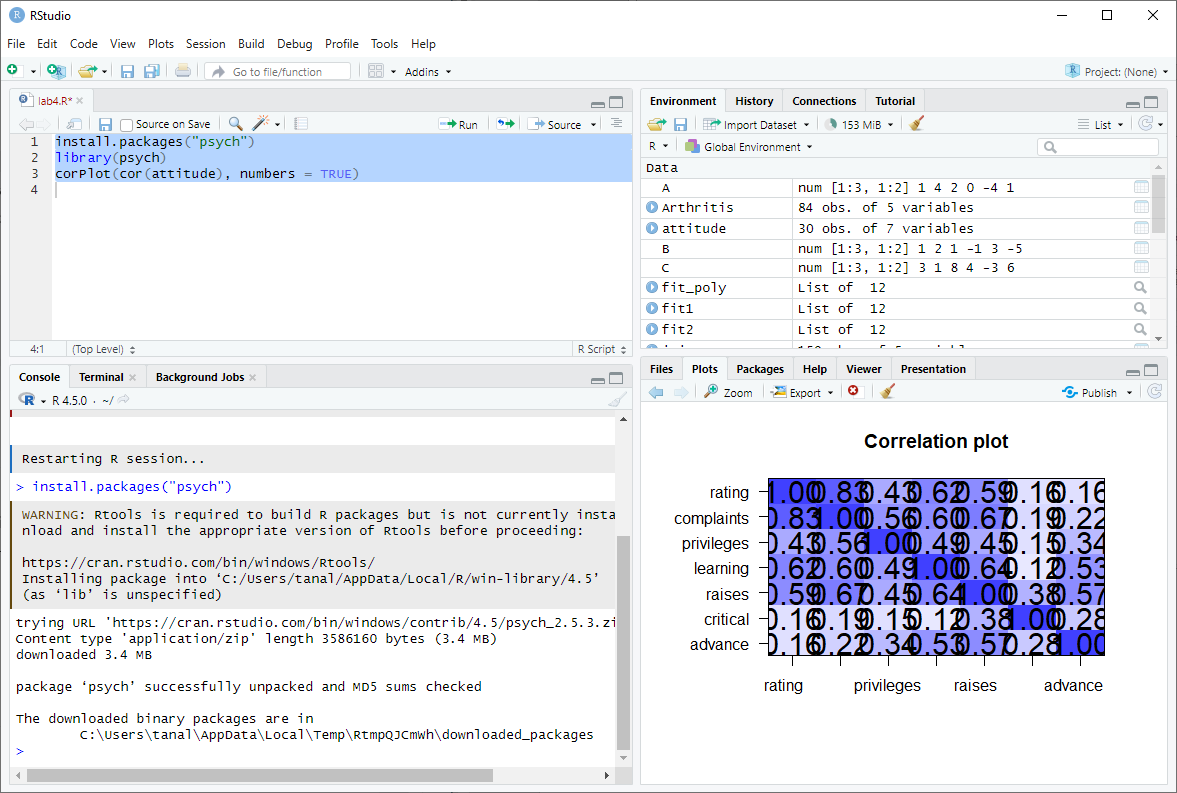
a) Графіки + регресія + кореляція



b) Графік попарної кореляції



c) Використання psych::corPlot



d) Корелограма

