

Київський національний університет
імені Тараса Шевченка
Звіт
до лабораторної роботи 10,
дисципліни: “Технології аналізу та візуалізації даних”,
студента: Євчика Олексія,
групи Інформатика, Магістри

Тема:

Візуалізація даних з використанням пакету lattice у R

Мета роботи:

Метою роботи є ознайомлення з методами фасетування та багатовимірної візуалізації у R, зокрема за допомогою системи trellis graphics (пакет lattice). Студент має навчитися будувати dotplot, heatmap, contourplot та інші графіки, що дозволяють порівнювати показники між групами та виявляти структурні особливості даних.

Теоретичні відомості**Система trellis graphics**

Trellis graphics — це система багатопанельних графіків для аналізу багатовимірних даних, реалізована у пакеті lattice у R.

Основна ідея: графік автоматично розбивається на “решітку” (trellis) панелей, кожна з яких відповідає одній підгрупі даних.

Це дозволяє легко досліджувати залежності між змінними у різних категоріях, часових інтервалах чи рівнях факторів.

Особливості trellis graphics:

- автоматичне фасетування за умовами (),
- єдина шкала для всіх панелей (за замовчуванням),
- декларативний стиль: спочатку описуються що візуалізувати, а не як.

Основні функції пакету lattice

Функція	Призначення
xypplot()	Діаграма розсіювання, лінійні графіки
bwplot()	Boxplot у trellis-стилі
dotplot()	Dotplot — точкове порівняння категорій

densityplot()	Оцінка щільності розподілу
histogram()	Гістограма
levelplot()	Теплокарта (heatmap)
contourplot()	Контурні графіки
barchart()	Стовпчикові діаграми

Пакет використовує формулі типу:

$y \sim x | condition$

що означає «побудувати графік у від x окремо для кожного рівня $condition$ ».

Хід роботи

Завдання: Dotplot

Мета

Побудувати точкову діаграму для порівняння значень між категоріями.

Метод дозволяє ранжувати групи за значенням ознаки.

Код мовою R

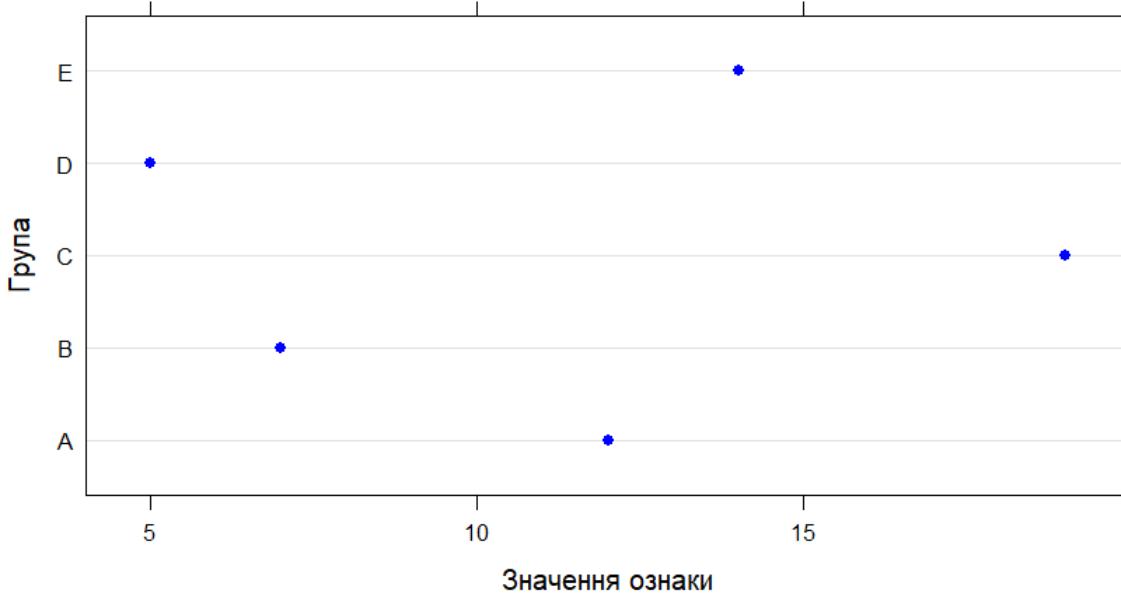
```
library(lattice)
library(tidyverse)

# Прикладові дані
df <- tibble(
  Group = c("A", "B", "C", "D", "E"),
  Feature = c(12, 7, 19, 5, 14)
)

# Dotplot
dotplot(Feature ~ Group, data = df,
        main = "Dotplot порівняння груп",
        xlab = "Значення ознаки",
        ylab = "Група",
        col = "blue",
        pch = 19)
```

Отримано графік

Dotplot порівняння груп



Коротка інтерпретація

Dotplot показує значення ознаки для кожної категорії у ранжованому вигляді. Група з найбільшим значенням («С») знаходиться праворуч, тоді як групи з меншими значеннями розташовані ліворуч. Такий тип діаграм дозволяє легко порівняти категорії між собою, виявити лідерів та аутсайдерів. Dotplot є ефективним інструментом для аналізу категоріальних даних з числовими показниками.

Висновки

У ході роботи я(як студент) навчився будувати багатопанельні графіки у стилі trellis graphics, застосовувати фасетування та працювати з категоріальними і числовими змінними у пакеті *lattice*. Розглянуті методи — dotplot, heatmap (levelplot), contourplot — дозволяють ефективно досліджувати структурні властивості даних, порівнювати групи та виявляти закономірності. Отримані навички є базовими для аналізу багатовимірних та часових даних.

Використана література

1. Sarkar D. *Lattice: Multivariate Data Visualization with R*. Springer.
2. Deepayan Sarkar. *lattice documentation* (CRAN).

3. Wickham H. *Elegant Graphics for Data Analysis*.
4. Venables W., Ripley B. *Modern Applied Statistics with S*.
5. Chang W. *R Graphics Cookbook*.
6. R Documentation: `help(lattice)`, `vignette("lattice")`.