

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХЕРСОНСЬКИЙ ДЕРЖАВНИЙ УНІВЕРСИТЕТ
Факультет комп'ютерних наук, фізики та математики
Кафедра інформатики, програмної інженерії та економічної кібернетики

**ЗАСТОСУВАННЯ ФУНКЦІЙ ПАКЕТУ GGPLOT2 ПРОГРАМНОГО
СЕРЕДОВИЩА R ДЛЯ ВІЗУАЛІЗАЦІЇ ДАНИХ**

Кваліфікаційна робота (проект)
на здобуття ступеня вищої освіти «бакалавр»

Виконав: студент 4 курсу 431 групи

Спеціальності: 122 Комп'ютерні науки

Освітньо-професійної програми:

Комп'ютерні науки

Білоус Іван Валерійович

Керівник: доктор технічних наук, професор

Бабічев Сергій Анатолійович

Рецензент: Гончаренко Т.Л. кандидатка

педагогічних наук , доцентка

ЗМІСТ

РОЗДІЛ 1. Аналіз існуючих методів візуалізації даних

- 1.1 Ознайомлення з програмним середовищем R та візуалізацією..... 3
- 1.2 Аналіз методів візуалізації..... 3

РОЗДІЛ 2. Графічний пакет ggplot2 , структура та особливості застосування

- 2.1 Пакет ggplot2..... 12
- 2.2 Структура пакету ggplot2..... 13
- 2.3 Основні функції пакету ggplot2 та особливості їх застосування..... 13

РОЗДІЛ 3. Практичні приклади застосування функцій пакету ggplot2 для візуалізації даних

- 3.1 Підготовка до візуалізації..... 15
- 3.2 Наведення прикладів візуалізації даних..... 16

ВИСНОВКИ 28

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ..... 29

ДОДАТОК А.....30

РОЗДІЛ 1

АНАЛІЗ ІСНУЮЧИХ МЕТОДІВ ВІЗУАЛІЗАЦІЇ ДАНИХ

1 Ознайомлення з програмним середовищем R та візуалізацією

1.1 R - це мова і середовище для статистичних обчислень і графіки. R також надзвичайно гнучкий і простий у використанні, коли справа доходить до створення візуалізацій. Одна з його можливостей - створювати графіки хорошої якості з мінімальними кодами [1].

Особливість програми :

- 1) Підтримує широкий спектр даних - включаючи вектори, матриці, списки і фрейми.
- 2) Підтримує безліч статистичних методів для роботи з різними даними.
- 3) Володіє відмінною розширюваністю - до програми можна підключати більше 400 пакетів у вигляді бібліотек для різних цілей.
- 4) Дає можливість створювати якісну графіку, яка включає в тому числі і математичні символи

Перевага мови R :

- 1) Відкритий код програмного інтерфейсу
- 2) Можливості роботи на будь якій операційній системі
- 3) Взаємодія з іншими мовами
- 4) В R можна працювати з великим обсягом даних

Середовище в R це RStudio – середовище розробки . В даному середовищі є заздалегідь розділені області та додаткові модулі (робоча область , консоль , історія команд).

1.2 Аналіз методів візуалізації

Візуалізація – це створення умов для візуального спостереження. А також це інтерактивне вивчення візуального представлення абстрактних даних для

посилення людського пізнання. Абстрактні дані включають в себе як числові так і нечислові, такі як текст і географічна інформація.

1) Звичайне візуальне представлення кількісної інформації в схематичній формі. До цієї групи можна віднести всім відомі кругові та лінійні графіки [2].

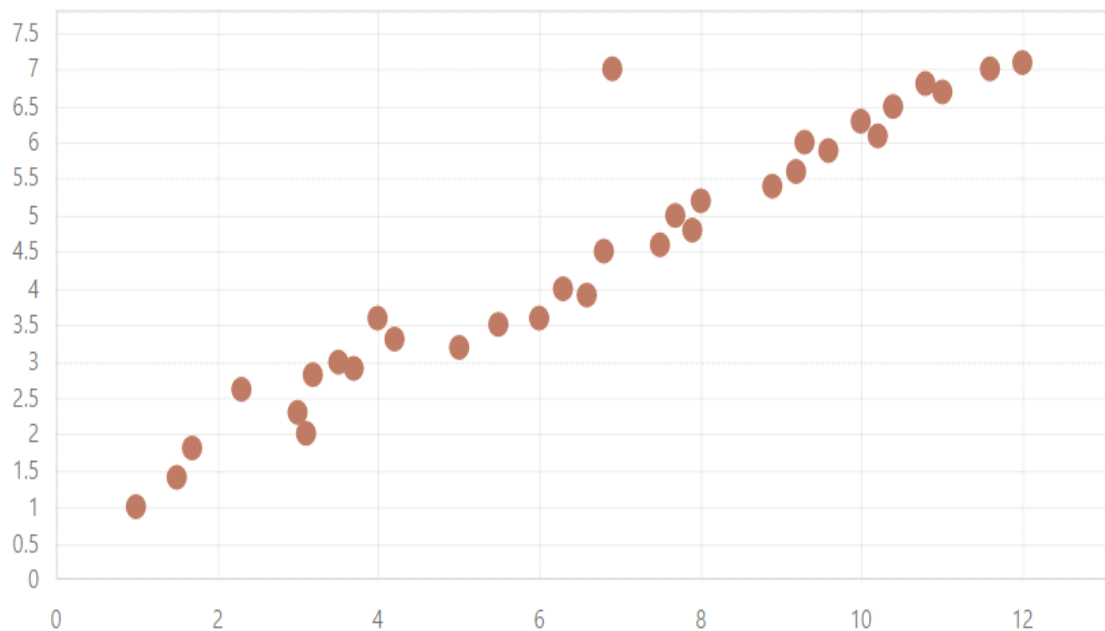


Рисунок 1.1- Точкова діаграма

Точкова діаграма показує взаємозв'язок між двома різними змінними або демонструє розподіляючі тенденції. Вона підходить, якщо у вас багато різних точкових даних, і ви хочете знайти загальне в наборі даних. Така візуалізація добре працює в пошуку винятків або закономірності розподілу даних.

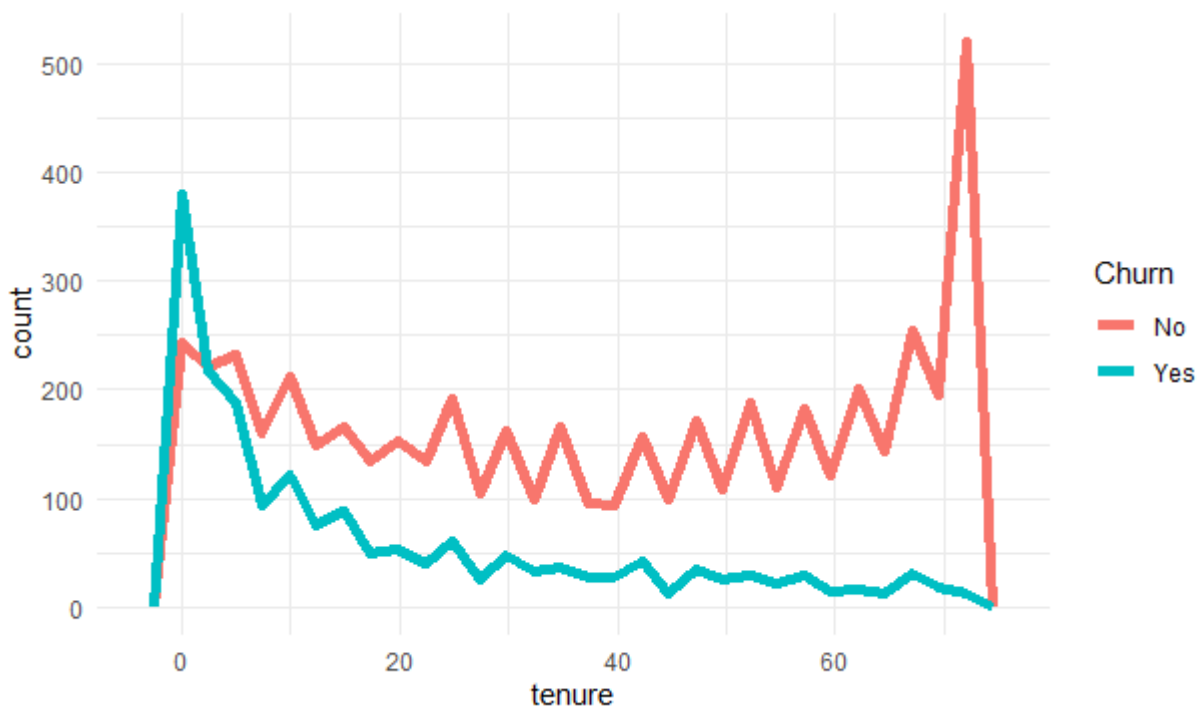


Рисунок 1.2 - Лінійний графік

Лінійний графік - це графік, який з'єднує ряд точок шляхом малювання відрізків між ними. Ці точки впорядковані за однією з їх координатних (зазвичай x-координатних) значень.

Лінійні графіки зазвичай використовуються для визначення тенденцій у даних [3].

2) Дані при візуалізації можуть бути перетворені у форму, що посилює сприйняття і аналіз цієї інформації. Наприклад, карта і полярний графік, тимчасова лінія і графік з паралельними осями, діаграма Ейлера [4].



Рисунок 1.3 - Діаграма Ейлера

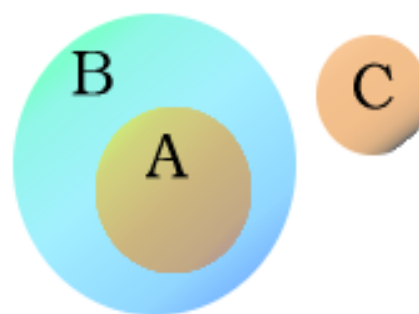


Рисунок 1.4 - Діаграма Ейлера

Діаграма Ейлера - геометрична схема, за допомогою якої можна зобразити відносини між підмножинами, для наочного уявлення.

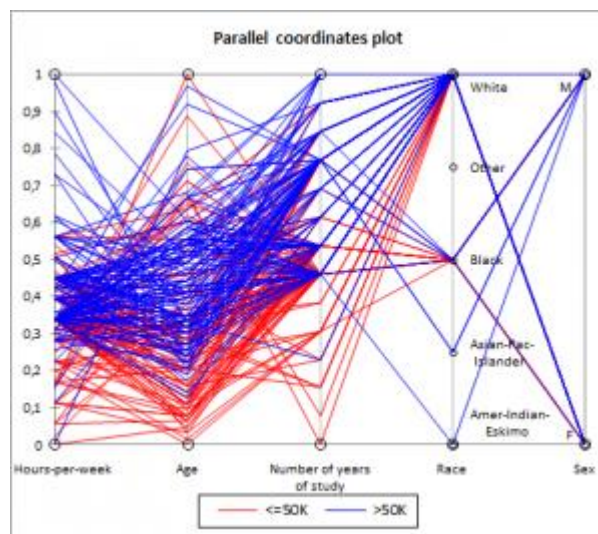
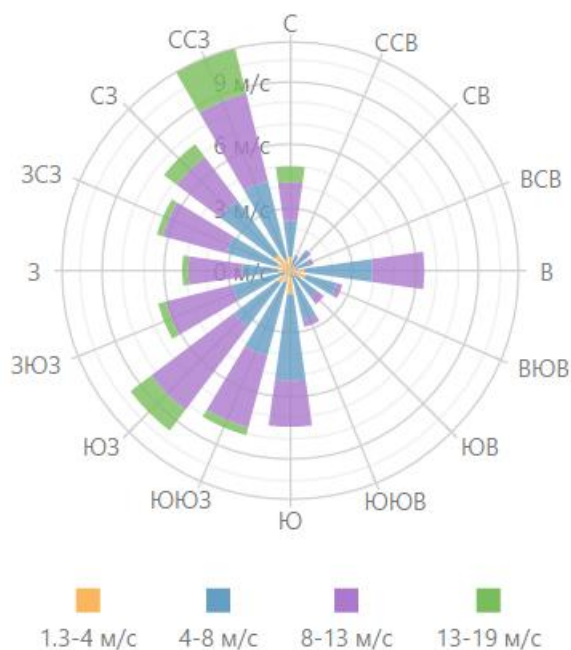


Рисунок 1.5- Полярна система

Рисунок1.6- Полярна система координат

Для наочного зображення множини елементів довільної природи представляють у вигляді множин точок, обмежених кривими. За допомогою таких діаграм зручно зображати відношення між множинами.

Полярна система координат — двовимірна система координат, в якій кожна точка на площині визначається двома числами — кутом та відстанню. Полярна система координат особливо корисна у випадках, коли відношення між точками найпростіше зобразити у вигляді відстаней та кутів

3) Концептуальна візуалізація дозволяє розробляти складні концепції, ідеї і плани за допомогою концептуальних карт, діаграм Ганта [5], графів з мінімальним шляхом та інших подібних видів діаграм.

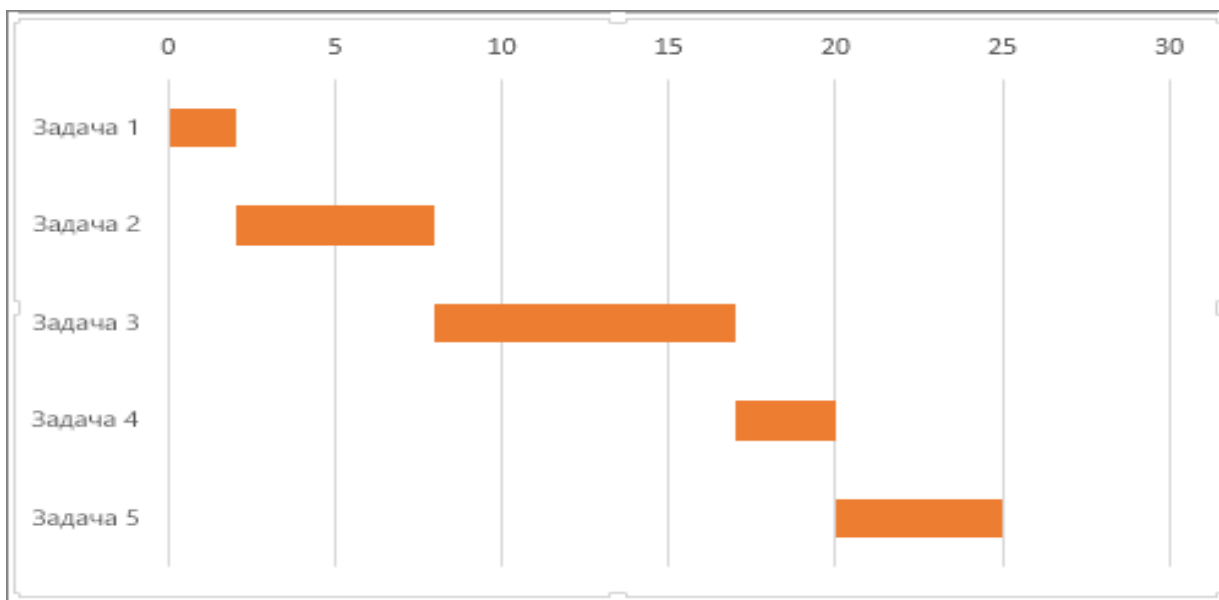


Рисунок 1.7 - Діограма Ганта

Діограма Ганта - це популярний тип стовпчастих діаграм (гістограм), який використовується для ілюстрації плану, графіка робіт з якого-небудь проекту.

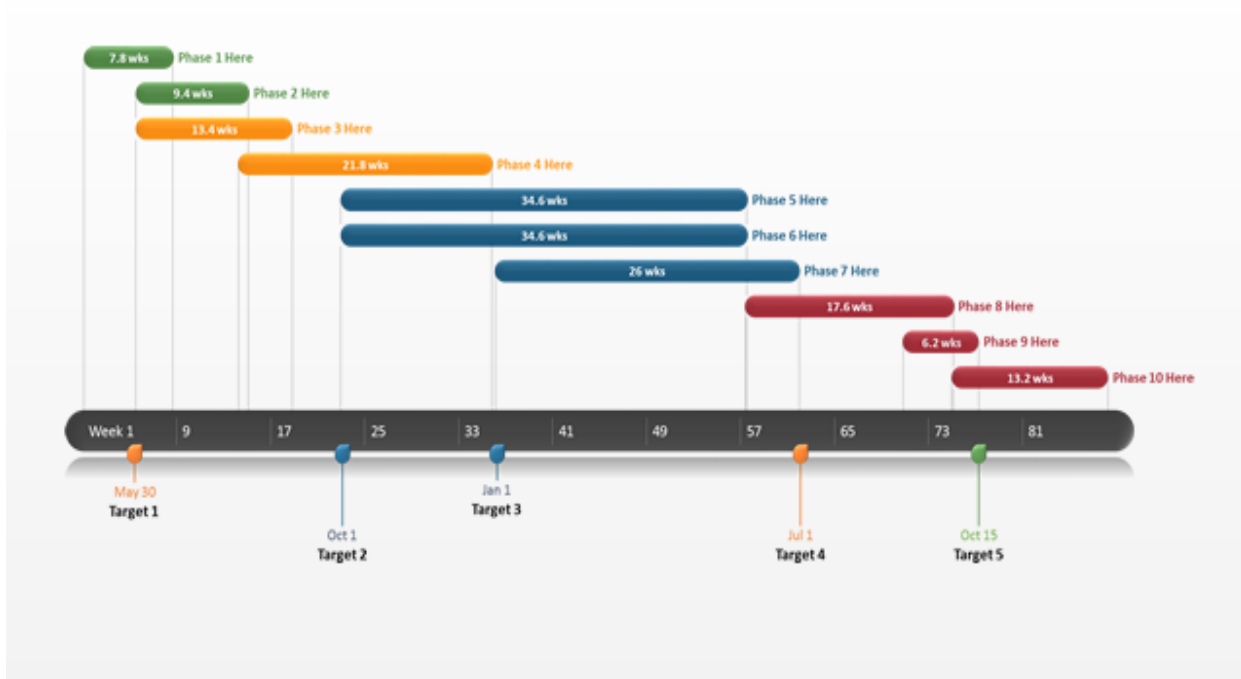


Рисунок 1.8 - Діограма Ганта

Діаграма Ганта являє собою відрізки (графічні плашки), розміщені на горизонтальній шкалі часу. Кожен відрізок відповідає окремому завданню або підзадачі. Завдання і підзадачі, складові плану, розміщуються по

вертикалі. Початок, кінець і довжина відрізка на шкалі часу відповідають початку, кінцю і тривалості завдання. На деяких діаграмах Ганта також показується залежність між завданнями. Є одним з методів планування проектів.

4) Стратегічна візуалізація переводить у візуальну форму різні дані про аспекти роботи організацій. Це всілякі графіки , життєвого циклу і графіки структур організацій.

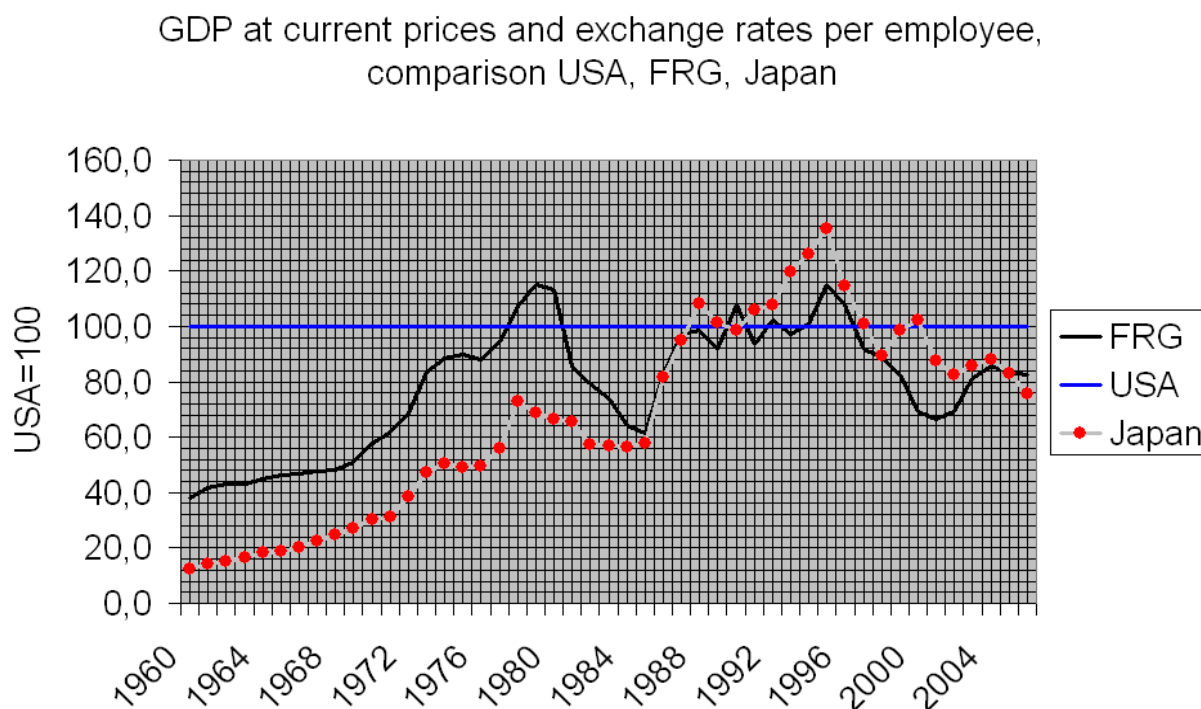


Рисунок 1.9 - Діаграма Ганта

Дана діаграма вказує на продуктивність роботи в той чи інший проміжок часу.

5) Функцію `boxplot()` використовують для побудови діаграм типу "ящик з вусами" [6].

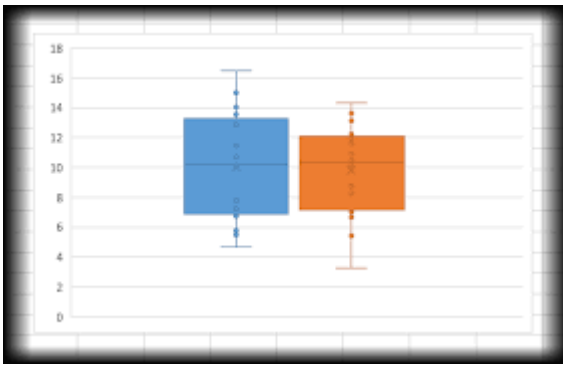


Рисунок 1.10 - Ящик з вусами

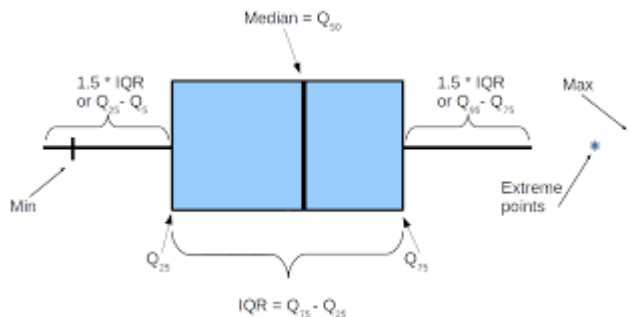


Рисунок 1.11 - Ящик з вусами

Такі діаграми відображають основні характеристики розподілу даних. Їх основними елементами є:

- 1) Прямокутник, дві протилежні сторони якого відповідають першому та третьому квантилям розподілу (ці квантілі відділяють першу та останню чверті даних у впорядкованому за зростанням масиві) якщо діаграма розташована вертикально, це будуть нижня та верхня сторони, а за її горизонтального розташування – ліва та права;
- 2) Відрізок у середині прямокутника, що відповідає медіані розподілу (це значення яке перебуває у центрі впорядкованого за зростанням масиву даних);
- 3) Вуси – два вертикальні (за вертикального розташування діаграми) або горизонтальні (за її горизонтального розташування) відрізки, що виходять з протилежних сторін прямокутника
- б) Скрипкові діаграми – це щось схоже на коробкову діаграму, але якщо потрібні більш точні данні використовують саме її.

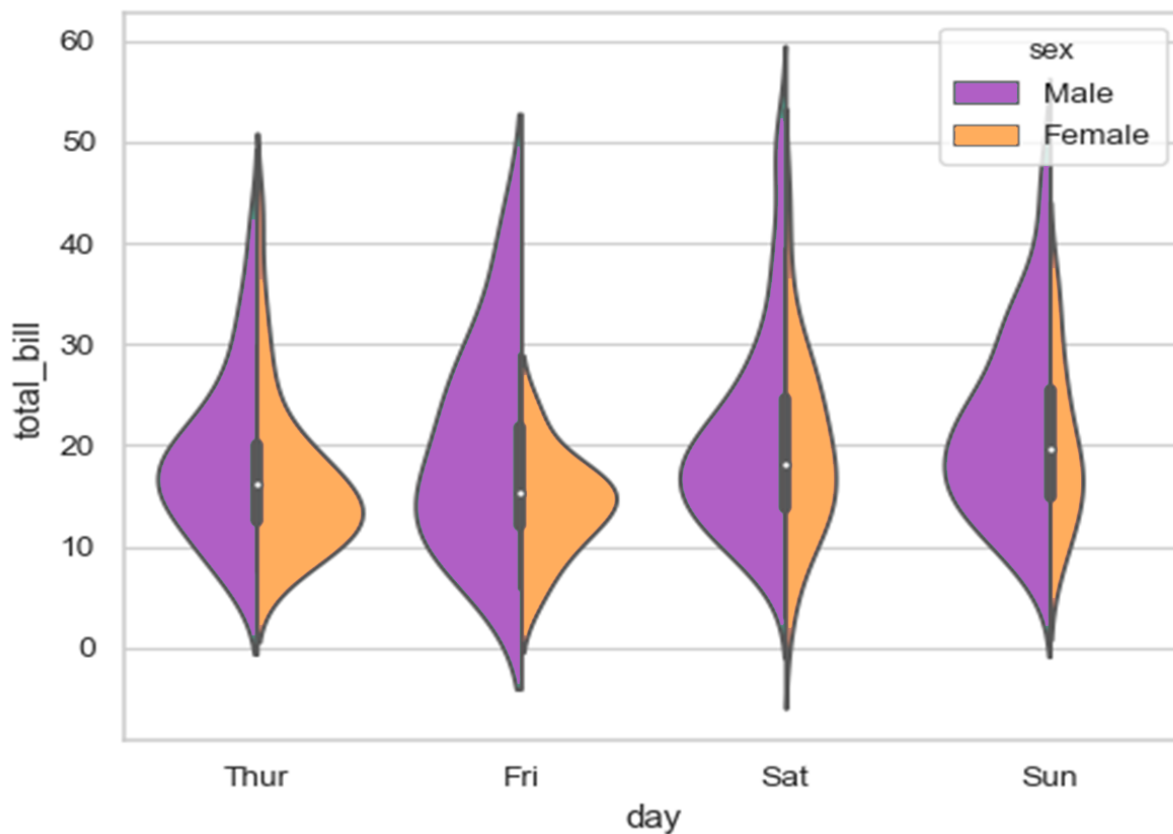


Рисунок 1.12 - Скрипкова діаграма

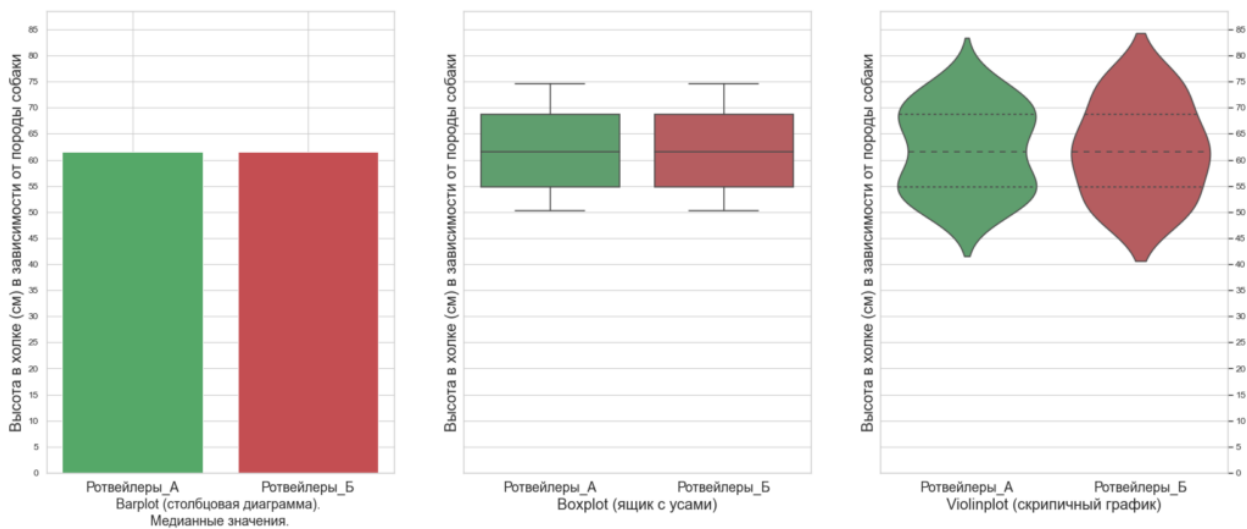


Рисунок 1.13 - Різноманітні діаграми

На даному графіку дуже гарно показується поєднання всіх діаграм або навіть як із однієї покроково виходить інша . Це так би мовити картинка яка може описати відразу 3 види діаграм .

Візуальна інформація краще сприймається і дозволяє швидко і ефективно донести до глядача власні думки та ідеї.

На графіку читач швидше знайде мінімальне і максимальне значення

1) 10% людина запам'ятовує з почутого, 20% – з прочитаного, і 80% – з побаченого .

2) на 323% краще людина виконує інструкцію, якщо вона містить ілюстрації. Очевидно, що людина схильна обробляти саме візуальну інформацію. Крім прекрасної обробки нашим мозком, візуалізація даних має кілька переваг:

- 1) Акцентування уваги на різних аспектах даних
- 2) Аналіз великого набору даних зі складною структурою
- 3) Зменшення інформаційного перевантаження людини і утримування його уваги
- 4) Однозначність і ясність виведених даних
- 5) Виділення взаємозв'язків і відносин, що містяться в інформації

Візуалізація даних відіграє важливу роль на всіх етапах статистичного аналізу - від первинного ознайомлення з властивостями даних до діагностики якості побудованих моделей і представлення отриманих результатів. З усього розмаїття статистичних програм виділяється R - інтенсивно розвивається і вільно поширювана система статистичних обчислень, в якій реалізовано безліч класичних і сучасних методів аналізу даних. Програмні реалізації алгоритмів, що входять в базову версію R, перевірені на практиці не одним поколінням користувачів і вчених.

Цілі візуалізації – це реалізація основної ідеї інформації, це те, заради чого потрібно показати вибрані дані, якого ефекту потрібно досягнути – виявлення взаємозв'язків в інформації, відображення розподілу даних, композиції або порівняння даних.

РОЗДІЛ 2

ГРАФІЧНИЙ ПАКЕТ GG PLOT2 , СТРУКТУРА ТА ОСОБЛИВОСТІ ЗАСТОСУВАННЯ

2.1 Пакет ggplot2

Ggplot2 - це розширення мови R, призначене для візуалізації даних. Для створення графіки пакет [7] використовує систему абстрактних понять: масив даних, візуальні засоби, геометричні об'єкти, зіставлення змінних з масиву візуальним засобам, статистичне перетворення змінних, системи координат. GG розшифровується як *grammar of graphics*. «Граматика графіки» - єдина причина, яка робить ggplot2 дуже потужним, оскільки розробник R не обмежений набором заздалегідь заданих графіки, яка використовується в інших пакетах. Під цим розуміється певна система правил, що дозволяють описувати і будувати графіки. Ggplot досить сильно відрізняється від стандартної графічної підсистеми R. Даний пакет призначений для побудови двовимірних графіків і діаграм.

Графік в рамках пакету `ggplot2` представляється як конструктор, що складається з окремих частин, які можна з'єднати довільним чином, щоб досягти бажаного ефекту.

Пакет являє собою систему для побудови графіків в R, засновану на цілісній і логічній граматиці. Це дозволяє зробити графічні побудови послідовними, чого часто бракує R, і дає можливість користувачам створювати нові типи діаграм.

Найпростіший підхід до створення діаграм в пакеті `ggplot2` – це використання функції швидкого побудови діаграм `qplot()`.

`Ggplot2` - це пакет, розроблений спеціально для візуалізації даних і забезпечує кращий аналіз даних [8]. Це забезпечує гарні, безпроблемні сюжети, які піклуються про дрібні деталі, таких як малювання легенд і їх відображення. Популярність цього пакету обумовлена декількома причинами серед яких можна відзначити естетичну привабливість і додатне для публікації якість одержуваних з його допомогою графіків, можливість створювати призначені для користувача типи діаграм, а також великий набір опцій для тонкої настройки зовнішнього вигляду графіків.

2.2 Структура пакету `ggplot2`

Пакет `ggplot2` це набір інструментів та функцій для створення візуалізації, існує декілька основних інструментів без яких створення візуалізації неможливе.

Data – `data.frame`, що містить дані для подання.

Aes – задає зв'язок між даними і їх поданням, визначає які змінні є на осях, які змінні відповідають за колір і форму даних, представлених на графіку.

Geom – геометричний об'єкт, за допомогою якого проходить візуалізація даних. Визначає геометричні об'єкти, які задають тип діаграми. Цей параметр задається в форматі текстового вектора з одним або більше елементами: "point", "smooth", "Boxplot", "line", "histogram", "density", "Bar" і "jitter".

Stat - група функцій, що додають в графіки різних статистичних показників, наприклад, такі як, середнє значення за групами, лінійна апроксимація даних.

Coord - група функцій, яка задає систему координат для представлення даних на площині. Це можуть бути звичайні координати, полярні координати, або географічні координати для подання відповідних даних.

Facet - група функцій, які дозволяють групувати графіки по заданому параметру і представляти результати у вигляді набору графіків (сітки з графіків).

Theme – функції, що дозволяють змінювати оформлення графіка, наприклад розмір і колір шрифту координатних осей і поділів на них, фон графіка і всього малюнка.

2.3 Основні функції пакету ggplot2 та особливості їх застосування

В даному пакеті взагалі є дуже багато різноманітних функцій для побудови різноманітного виду візуалізації . Про деякі із функцій(базові) я на даному етапі роботи і розповім .

1) Одними з базових функцій є функції *ggplot*, *gplot* та *plot* - дані функції дозволяють саме створювати саму візуалізацію (зображення) , тобто вони створюють графік на якому будуть відображатися якісь дані . Ці функції використовуються завжди в даному пакеті візуалізації , для створення базових та не тільки діаграм та графіків. Але всі ці три функції мають як спільні так і відмінні характеристики але всі вони призначенні для побудови графіків , просто наприклад функція *ggplot* має кращій набір даних ніж функція *plot* , але і та і інша створенні для однієї задачі побудови графіків .

2) Друга основна (базова) функція це *geom* – ця функція нанесення того чи іншого типу малюнку (точок , ліній ...). За допомогою цієї функції можлива побудова базової , і не тільки , візуалізації. Так сказати це якийсь графічний елемент який наноситься на графік .

3) Функція *theme* – вона відповідає за тему діаграми . Як вже і так зрозуміло коли треба поміняти або поставити тему на графік треба використовувати дану функцію .

- 4) `ggtitle` – функція яка відповідає за написи на діаграмі , будь які слова , зміна назв осі або просто якесь слово .
- 5) `color` – зміна кольору рисунка.

РОЗДІЛ 3

ПРАКТИЧНІ ПРИКЛАДИ ЗАСТОСУВАННЯ ФУНКЦІЙ ПАКЕТУ GGPLOT2 ДЛЯ ВІЗУАЛІЗАЦІЇ ДАНИХ

3.1 Підготовка до візуалізації

Для початку створення візуалізації в R\Rstudio потрібно їх завантажити та встановити , після установки заходжу в середовище та підключаю пакети для роботи та створення візуалізації . Для цього мені потрібен пакет `ggplot2` , але його не має в даному програмному середовищі , його потрібно самому підключати , для цього є функція `install.packages` , та в дужках вводжу назву пакету якого мені потрібно завантажити . Після завантаження всіх потрібних пакетів треба їх викликати із файлу для початку роботи , для цього вводжу функцію `library` та також в дужках назву пакетів з якими я буду працювати . На цьому етапі я підключив два пакети один з яких це пакет `ggplot2` а другий пакет він був встановлений в базу R і при завантаженні я його просто визвав ,

цей пакет *MASS* , він мені потрібен тільки для того щоб брати з нього вже готові данні та по цим даним будувати різновиди графіків. Тепер можна переходити до самої візуалізації . Даний пакет є дуже великим він вміщає в себе 226 рядків та 10 стовбців з різними даними.

Основні данні яким я буду користуватися для створення візуалізації в вигляді діаграм

man 1 – a , b , c , ... g .

man 2 - -3 , -2 , ... 3 .

bwt – 703 , 1021 , ... 5000 . (значення осі y , параметр вага в грамах)

lwt - 158 , 160 , ... 250 . (значення осі x , вага в фунтах)

smoke – 0 , 1 , 0 , 1 (дане значення коливається від 0 до 1 тому що 0 не курят , 1 курять)

age – 14 , 15 ... 35 (вік)

rase – 1 , 2 , 1 , 2 ... (раса людини , чорний або білий)

Ці дані так називаємі дата фрейми якими є вектори кожен з яких повинен бути однакової довжини , мають вигляд двовимірної таблиці , як правило трактуються дата фрейми як спостереження (рядки) певного набору величин (стовбці) .

3.2 Наведення прикладів візуалізації даних

На початку наведу приклад простої можна навіть сказати базової візуалізації.

```
ggplot(data = man, aes(x = lwt, y = bwt)) + geom_line()
```


Даний приклад є простою лінією яка показує данні які були вказані в пакеті MASS . Такий графік використовується для простої візуалізації наприклад знаходження якихось даних в певний період віку чи чогось другого . А так як в моїх даних дуже багато факторів то можемо спостерігати що графік вийшов не дуже , на мою думку краще зробити точений графік для цих даних.

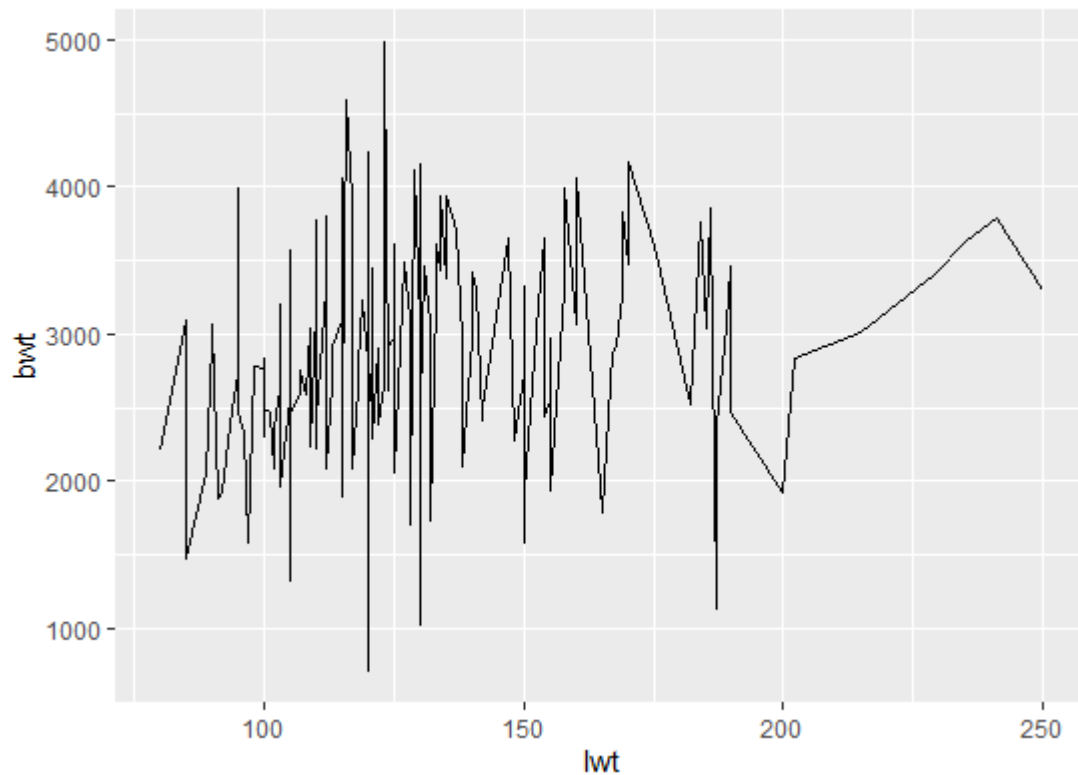


Рисунок 3.1 - Лінійний графік

Тепер я буду робити точечний графік треба всього лише змінити кінцеву назву графіка.

```
library(ggplot2)
```

```
ggplot(data = man , aes(x = lwt , y = bwt) ) + geom_point
```

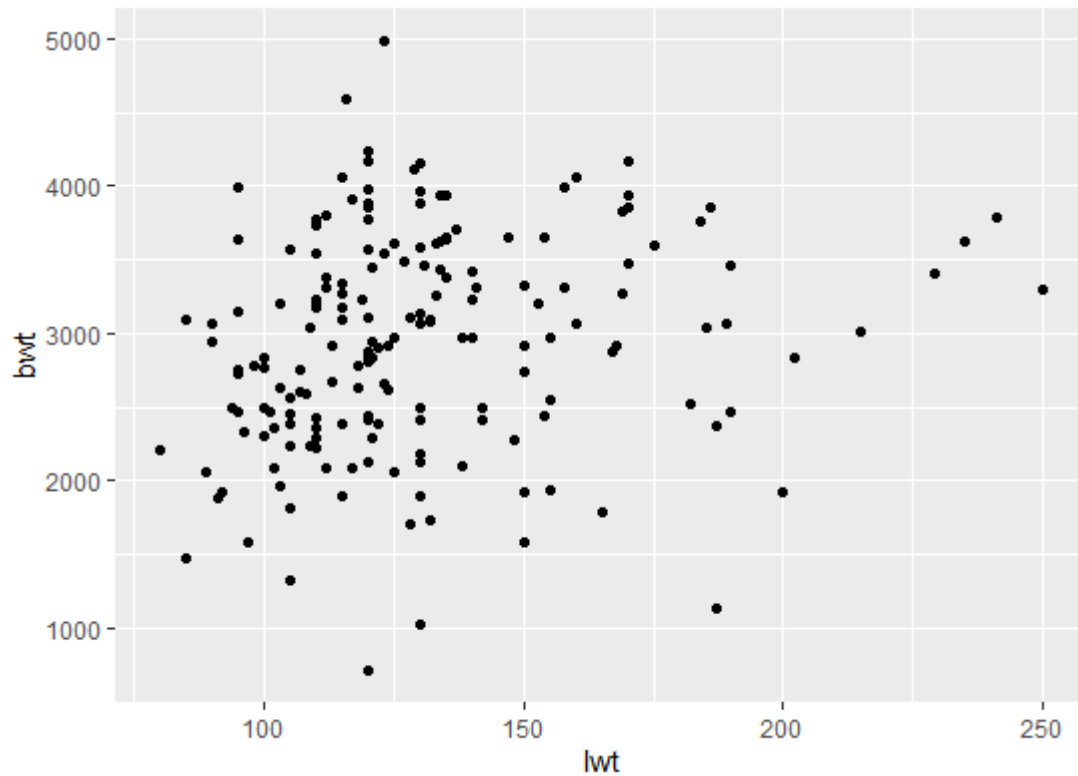


Рисунок 3.2 - Діограма розсіювання

Тепер можна з'єднати ці два види графіків та зробити один .

```
ggplot(data = man, aes(x = lwt, y = bwt)) +  
  geom_line() + geom_point()
```

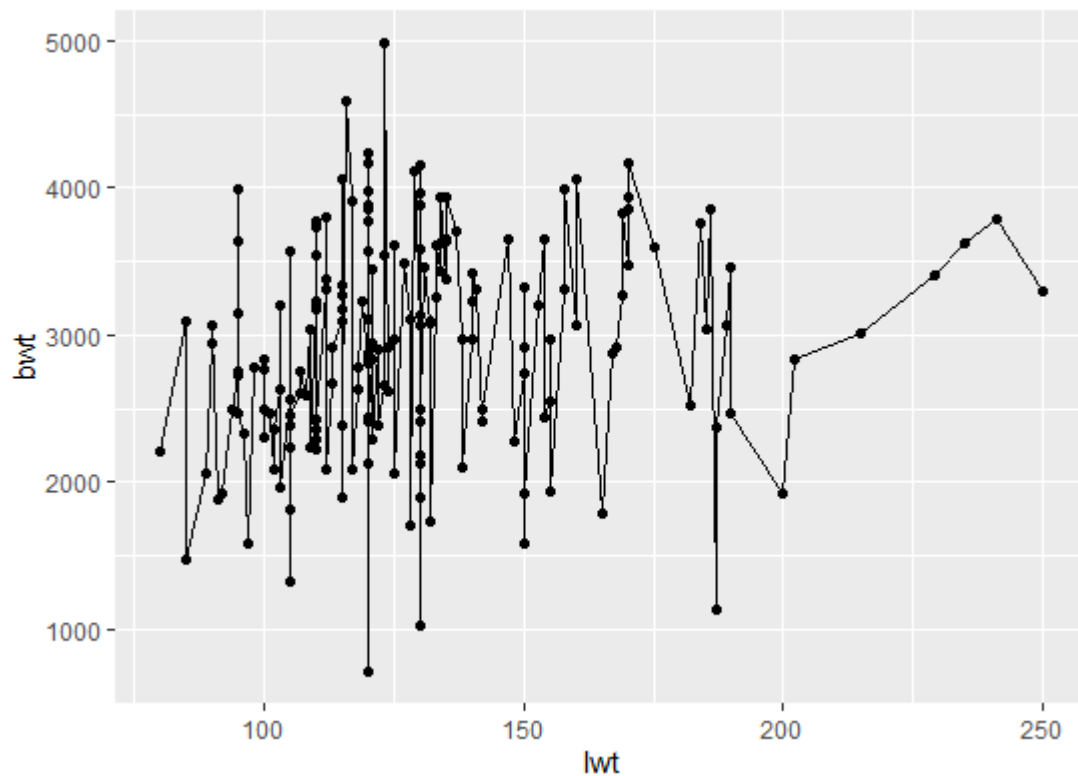


Рисунок 3.3 - Лінійний графік з'єднаний точками

Як показано вище прикладі для з'єднання графіків даного виду потрібно просто їх об'єднати .

З наведеного вище коду вийшов такий собі базовий графік де візуалізація в основі розміру , кольору була задана за змовчуванням або як в моєму випадку ще не задана .

На даному етапі я тільки задав що це буде графік розсіювання та назви осей . Далі я буду видозмінювати колір та розмір кульок . Для цього вводжу наступну функцію :

```
library(ggplot2)
ggplot(data = man , aes(x = lwt , y = bwt) )
+ geom_point(color = 'green' , size = 5)
```

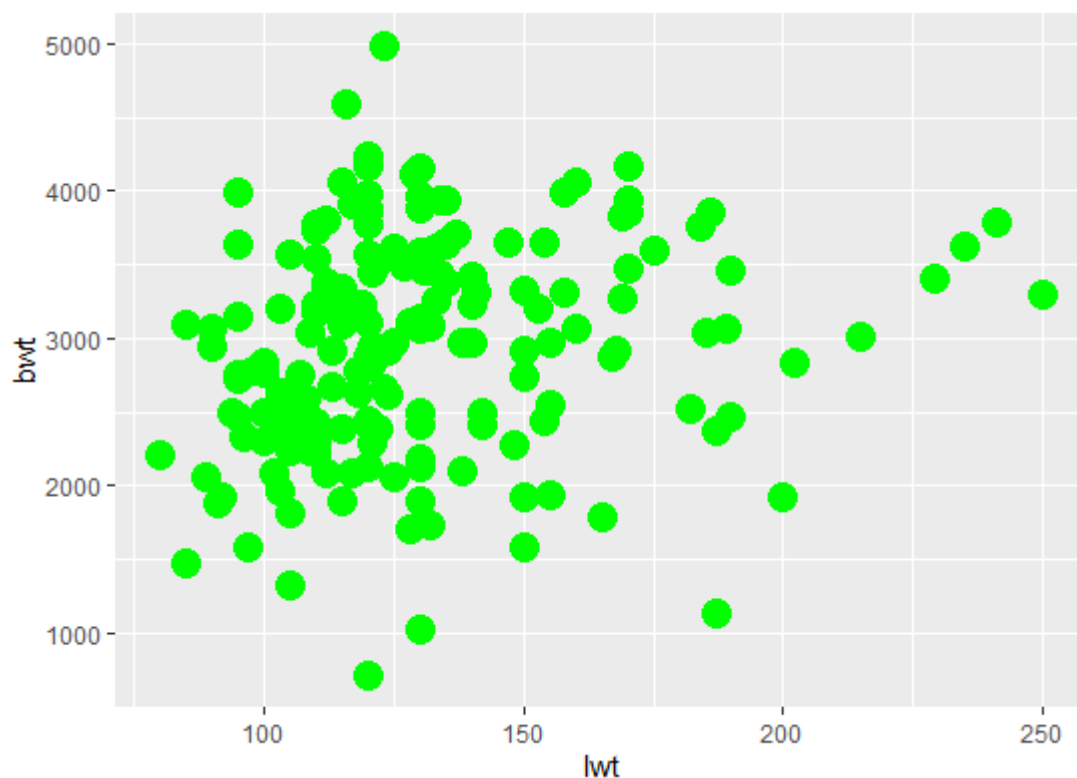


Рисунок 3.4 - Діограма розсіювання

Ось як бачите я змінив колір кульок за допомогою функції `color` та розмір за допомогою `size`.

Можна також видозмінювати коли графіки об'єднані за допомогою цих же функцій але не за змовчуванням а задати колір та розмір.

```
ggplot(data = man, aes(x = lwt, y = bwt)) +
  geom_line(color="red") +
  geom_point(size= 1.5 )
```

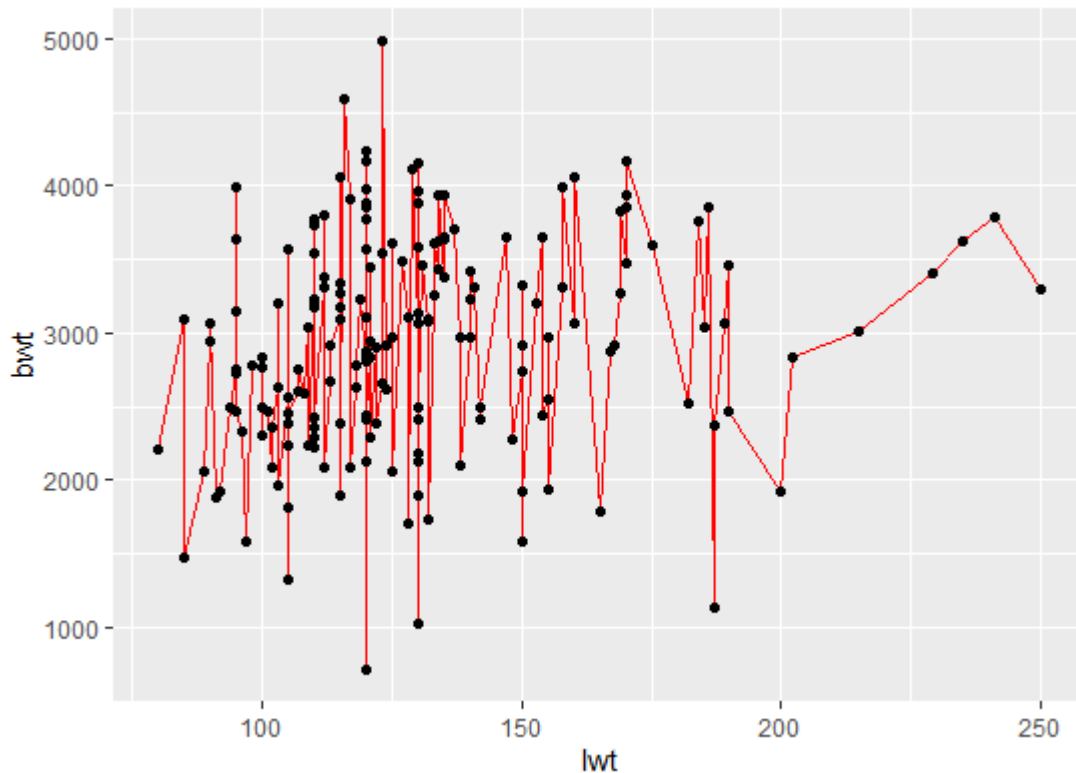


Рисунок 3.5 - Лінійний графік з'єднаний точками

Ось як видно на картинці та в функції все дуже легко та просто можна змінити колір та розмір графіка.

Зараз наведу приклад естетики , порівняня за тими ж даними що були й раніше , порівняня теми курять та не курять . Але для цього беруться тільки кульки і далі вони видозмінюються , для цього в пакеті ggplot2 є багато наведених прикладів в яку форму можна замінити кульки як в мене.

```
man$smoke <- factor(man$smoke , levels = (1 , 2)
  labels = c('smoker' , 'no smoker')
```

```

rase : men$rase <- factor (man$rase , levels = c (1,2),
labels = c ('white' , 'black'))
ggplot(man , aes (x = lwt , y = bwt , color = smoke ,))
+geom_point(siz = 5 ,shape=17)

```

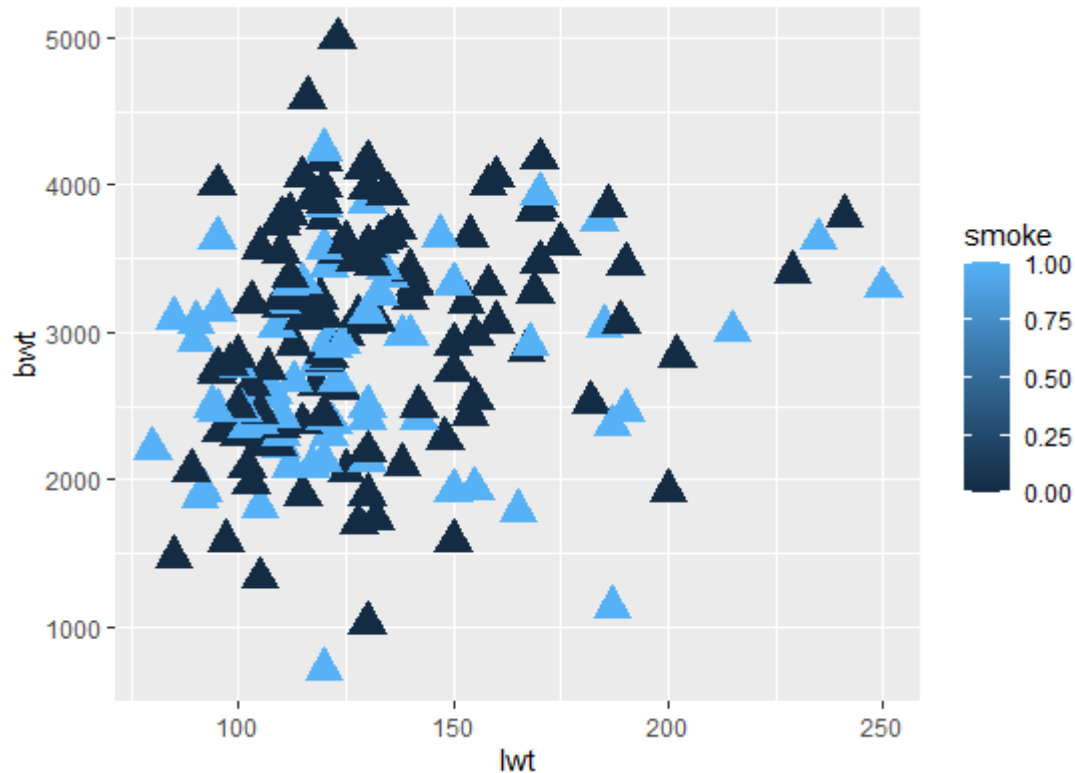


Рисунок 3.6 - Естетики

На цьому етапі я показав кількість вірогідних курців та присвоїв їм значення. На даному прикладі я показав яка кількість курців де значення 0 = взагалі не курили ніколи 1 = курці які курять цигарки щоденно.

Дану візуалізацію можна проводити з будь якими даними де вам потрібно визначити певний діапазон від x до y , де x = 0 а y = 1 .

Далі наведу приклад побудови гістограми.

```

library(ggplot2)
ggplot(man , aes (x=bwt)) + geom_histogram(binwidth=50)

```

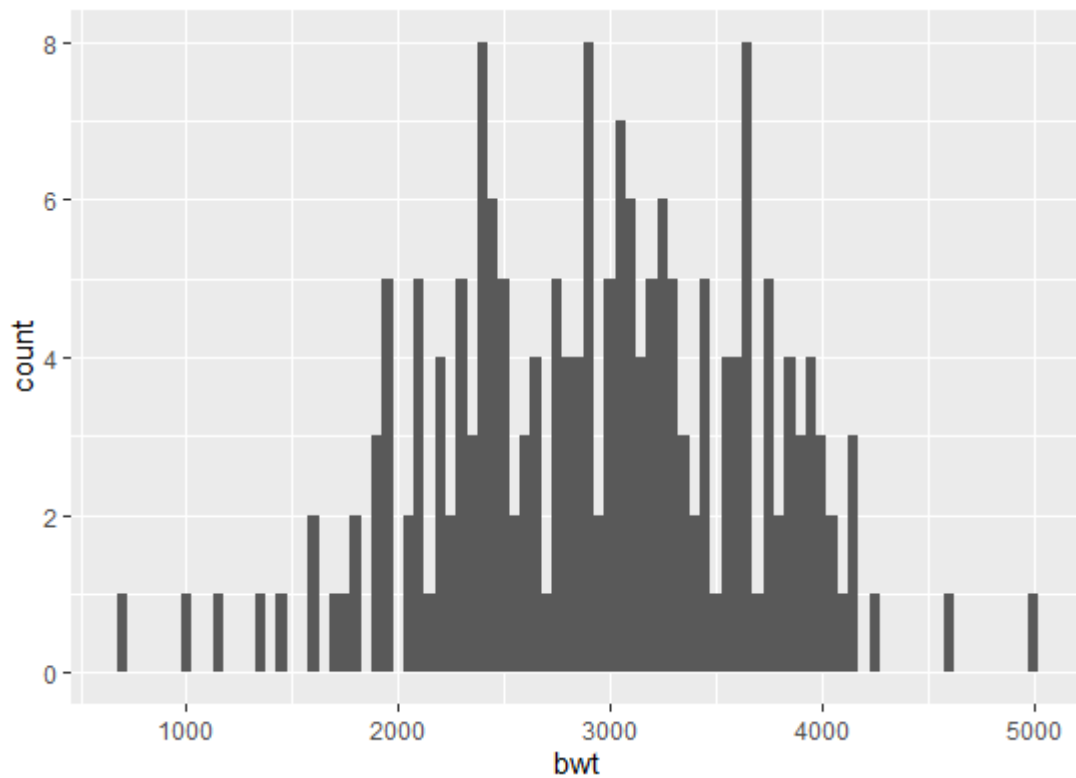


Рисунок 3.7 - Гістограма

На даному прикладі я побудував гістограму . Вона була побудована за тими ж даними що і точечна діаграма але на даному прикладі все ж я вказав деякі данні це $binwidth=50$ – ці данні визначають ширину стовбця , в цьому прикладі також можна змінити ширину стовбця на більшу або колір та розмір , але це за бажанням автора .

На даному етапі координатна система по змовчуванню – декартова але наприклад якщо я хочу змінити площину то як це зробити , думав я поки не знайшов відповідь , це можна зробити за допомогою функції *coord_polar*.

Зараз я наведу приклад коду для зміни координатної площини :

```
library(ggplot2)
```

```
ggplot(data = man, aes(x = lwt)) + geom_bar() + coord_polar()
```

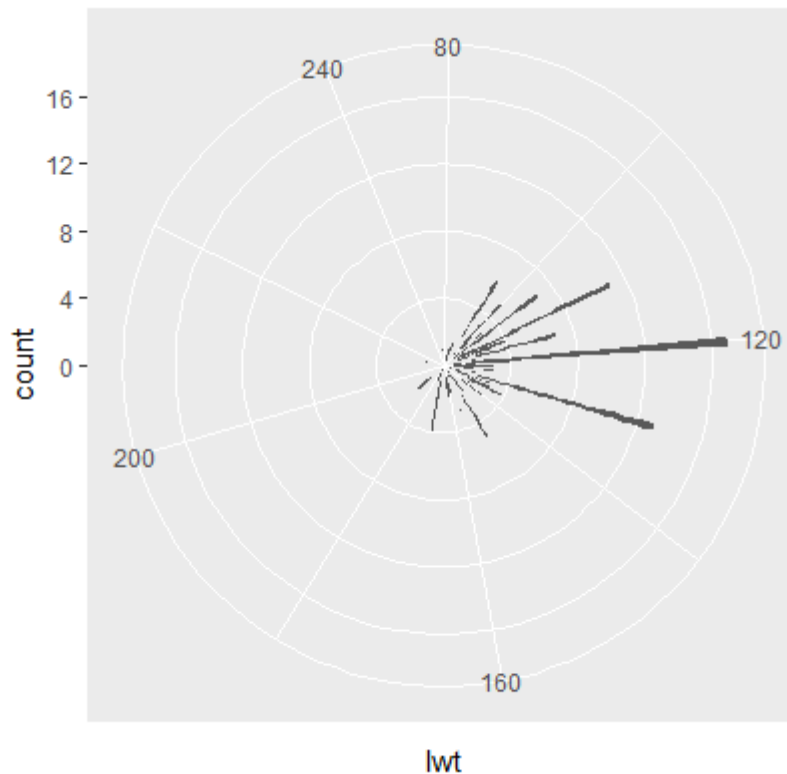


Рисунок 3.8 - Полярна система координат

Побудова `box_plot` (коробчатої діаграми)

Функцію `boxplot()` використовують для побудови діаграм типу "ящик з вусами". Такі діаграми відображають основні характеристики розподілу даних. Їх основними елементами є.

- 1) Прямокутник, дві протилежні сторони якого відповідають першому та третьому квантилям розподілу (ці квантілі відділяють першу та останню чверті даних у впорядкованому за зростанням масиві); якщо діаграма розташована вертикально, це будуть нижня та верхня сторони, а за її горизонтального розташування – ліва та права;
- 2) Відрізок у середині прямокутника, що відповідає медіані розподілу (це значення яке перебуває у центрі впорядкованого за зростанням масиву даних);
- 3) Вуси – два вертикальні (за вертикального розташування діаграми) або горизонтальні (за її горизонтального розташування) відрізки, що виходять з протилежних сторін прямокутника;
- 4) Точки, що виходять поза межі діапазону, який показують вуси.

```
ggplot(data = man, aes(y = bwt)) +
  geom_boxplot()
```

Думаю тут нічого пояснювати не треба і так все зрозуміло , але якщо будувати не таку примітивну коробку а може навіть декілька потрібно дійсно попрацювати с функціями .

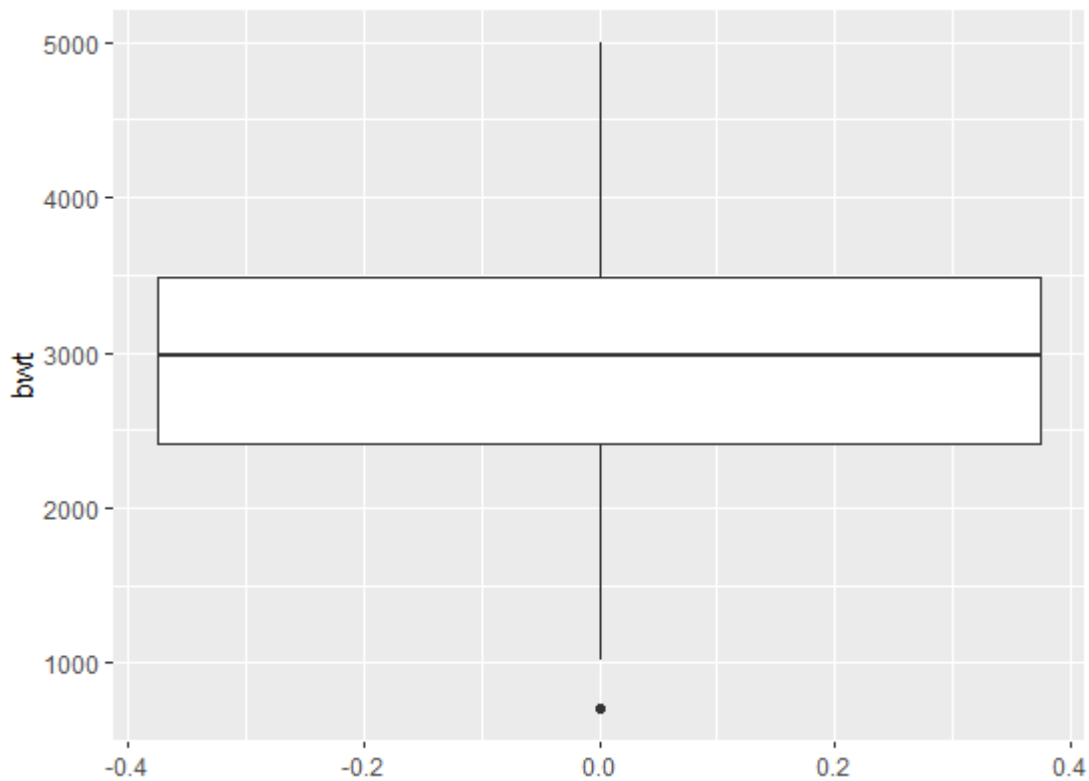


Рисунок 3.9 - Ящик з вусами

Зараз побудую графік що наведений вище по групам , тобто розбив на 2 об'єкти , та надавши колір за змовчуванням.

```
ggplot(data = man, aes(x = "", y = bwt, group = temp, fill = active)) +
  geom_boxplot()
```

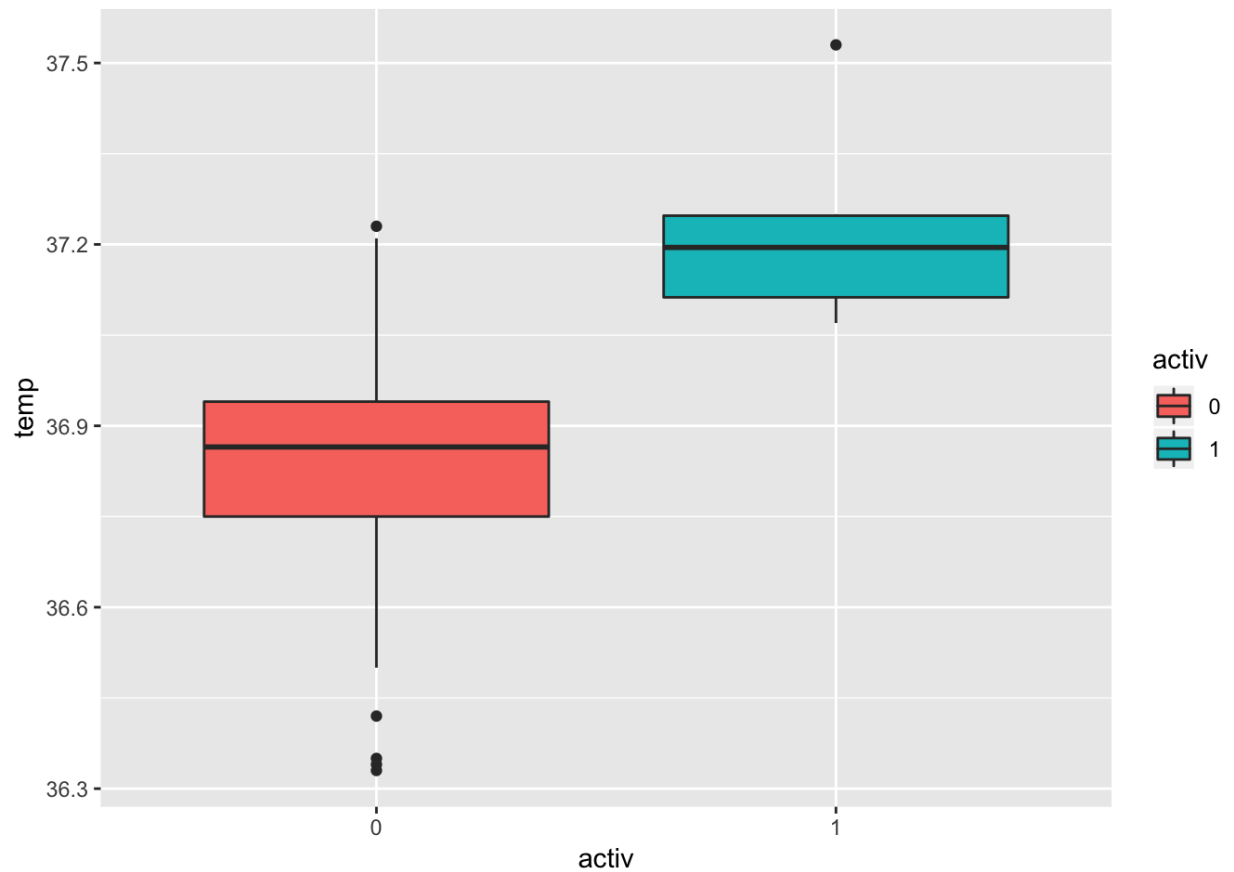



Рисунок 3.10 - Ящик з вусами

Приклад побудови `bar_plot`.

Функція `barplot()` використовується для створення стовпчастих (лінійчастих) діаграм.

```
df <- data.frame(
  man1 = c("a", "b", "c", "d", "e", "f", "g"),
  man2 = c(-3, -2, -1, 0, 1, 2, 3))
ggplot(df, aes(x=man1, y=man2, fill=man1)) +
  geom_bar(stat="identity", position="dodge") +
  guides(fill=FALSE)
```

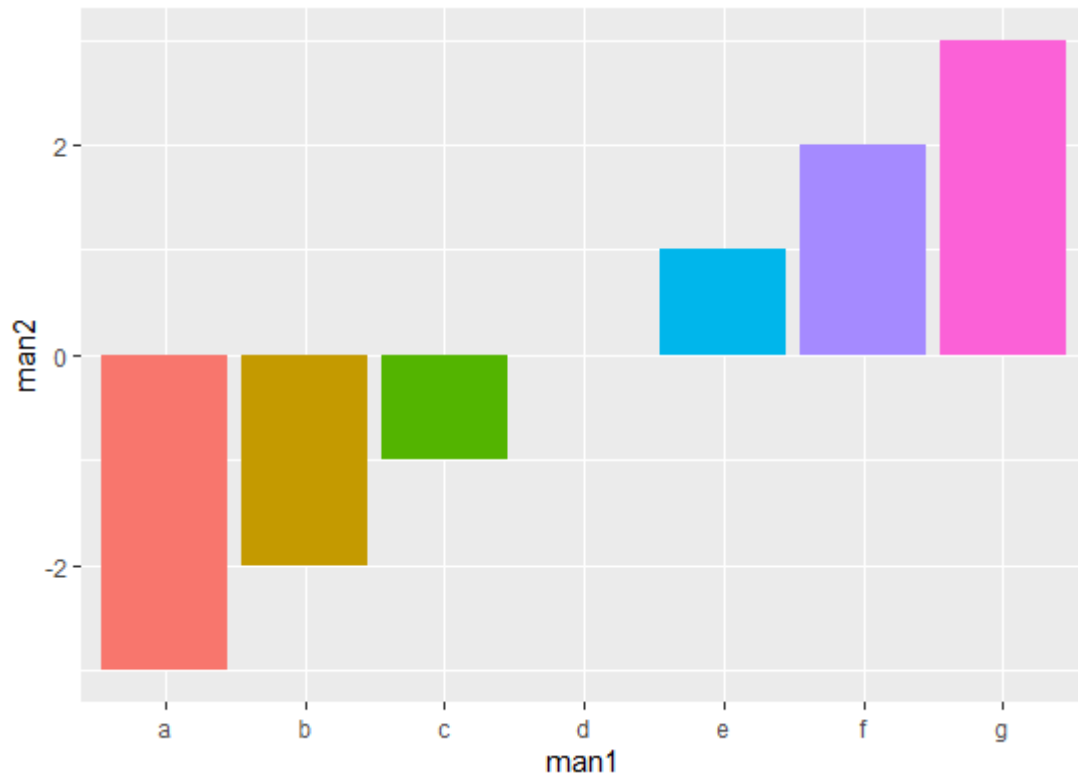


Рисунок 3.11 - Стовбчаста діаграма

Функція `barplot()` використовується для створення стовпчастих (лінійчастих) діаграм. До її основних аргументів належать такі:

- 1) `height` – числовий вектор або матриця, що задають висоти стовпчиків.
- 2) `beside` – використовується, якщо `height` є матрицею і може набувати значень `TRUE` або `FALSE`, у першому випадку стовпчики діаграми групуються згідно із стовпчиками матриці, а у другому – будується діаграма із накопиченням.
- 3) `width` – числовий вектор, що задає ширину стовпчиків; це не обов'язків параметр, за його відсутності всі стовпчики будуть мати однакову ширину.
- 4) `angle` – кут нахилу штрихів у градусах.

Скрипкові діаграми

Вони зображають відображення числових даних. Це схоже на графік коробки, з додаванням оберненого графіку щільності ядра з кожного боку.

```
ggplot(data = man, aes(x = "", y = temp, group = activ, fill = activ)) +
  geom_violin()
```

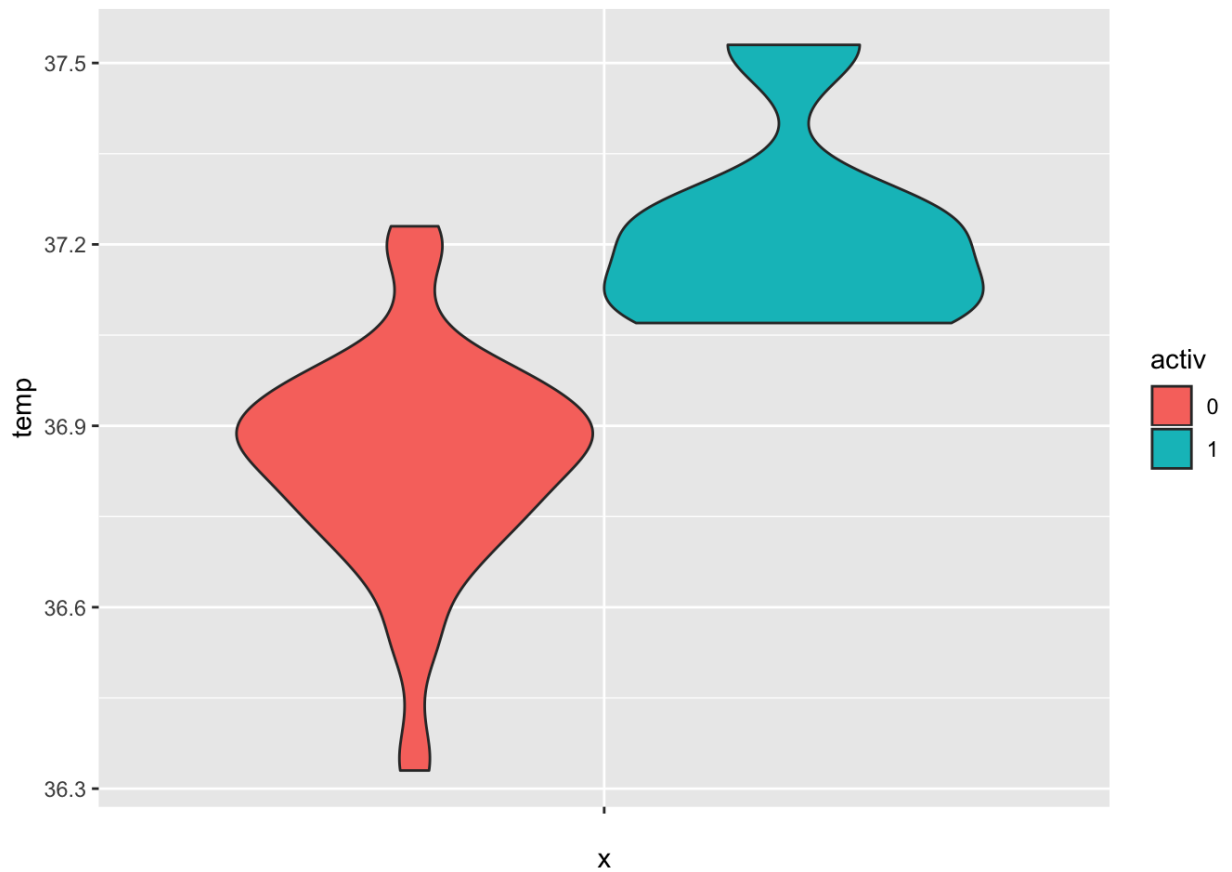


Рисунок 3.12 - Скрипкова діаграма

Скрипкова діаграма є більш інформативною, ніж звичайний графік. У той час як коробковий графік відображає лише зведені статистичні дані, такі як середнє, середнє та міжквартильні діапазони, скрипкова діаграма показує повний розподіл даних.

ВИСНОВКИ

На мою думку будувати візуалізацію це дуже захоплюючий та клопітливий момент. Клопітливий тому, що якщо треба будувати якусь не складну так би мовити базову візуалізацію та як я на власному прикладі показав, програмний код в R дуже зручний в цьому плані, але є одне але, для побудови завжди потрібна якась шпаргалка з тими чи іншими функціями, бо всі відразу в голові не втримаєш. При побудові візуалізації треба збирати її як конструктор (по камінцях), але це все одно дуже цікаво хоча і складно. Що до пакету ggplot2 мої висновки такі – цей пакет був створений для посилення візуалізації в R і він свою місію виконав, цей пакет володіє великим набором різноманітних функцій для створення візуалізації, тому він і став самим відомим пакетом для візуалізації даних в програмному середовищі R. На даний момент його можливості здаються майже безмежними, тобто всі види які я бачив та будував в програмі це все можна і було побудоване за допомогою цього пакету, але я навів тільки базові функції цього пакету а їх ще багато. Взагалі візуалізація це кращий та простіший спосіб прийняття інформації, як довели вчені мозок в 3 рази швидше зрозуміє данні по картинці ніж так.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- 1) <https://cran.r-project.org>
- 2) <https://coderlessons.com/tutorials/kompiuternoje-programmirovanie/learn-r-programmirovanie/r-lineinye-grafiki>
- 3) <https://uk.photo-555.com/9221395-line-graph-in-r>
- 4) <https://uk.skylineropescourse.com/245-venn-diagram-complete-explanation-and-examples-of-use>
- 5) <https://nachasi.com/creative/2020/09/03/gantt-chart/>
- 6) https://datavizcatalogue.com/RU/metody/diagramma_razmaha.html
- 7) <https://matrunich.com/blog/2014/02/27/what-is-ggplot2-grammar-of-graphics/>
- 8) <https://tsamsonov.github.io/r-geo-course/advgraphics.html>
- 9) <http://www.qsar4u.com/files/rintro/06.html>
- 10) <http://docs.ggplot2.org/current/>
- 11) <https://ru.linkedin.com>
- 12) <https://coderlessons.com/tutorial-ggplot2>
- 13) <https://tvimc.jimdofree.com>
- 14) <https://www.prostir.ua/?kb=andrij-hazin-robot-a-z-danymy-v-seredovyschi-r>
- 15) <https://www.google.com.ua>
- 16) <https://conf.ztu.edu.ua/wp-content/uploads/2017/06/151-2.pdf>
- 17) <https://coderoad.ru/11583948>
- 18) <https://answer-id.com/ru/67874548>
- 19) <https://habr.com/ru/post/267123/>
- 20) <https://gis-lab.info/forum/viewtopic.php?t=18790>
- 21) <https://www.machinelearningmastery.ru>

ДОДАТОК А

**КОДЕКС АКАДЕМІЧНОЇ ДОБРОЧЕСНОСТІ
ЗДОБУВАЧА ВИЩОЇ ОСВІТИ ХЕРСОНСЬКОГО
ДЕРЖАВНОГО УНІВЕРСИТЕТУ**

Я, Білоус Іван Валерійович
учасник освітнього процесу Херсонського державного університету,
УСВІДОМЛЮЮ, що академічна доброчесність – це фундаментальна етична
цінність усієї академічної спільноти світу.

ЗАЯВЛЯЮ, що у своїй освітній і науковій діяльності
ЗОБОВ'ЯЗУЮСЯ:

- дотримуватися:
 - вимог законодавства України та внутрішніх нормативних документів університету, зокрема Статуту Університету;
 - принципів та правил академічної доброчесності;
 - нульової толерантності до академічного плагіату;
 - моральних норм та правил етичної поведінки;
 - толерантного ставлення до інших;
 - дотримуватися високого рівня культури спілкування;
- надавати згоду на:
 - безпосередню перевірку курсових, кваліфікаційних робіт тощо на ознаки наявності академічного плагіату за допомогою спеціалізованих програмних продуктів;
 - оброблення, збереження й розміщення кваліфікаційних робіт у відкритому доступі в інституційному репозитарії;
 - використання робіт для перевірки на ознаки наявності академічного плагіату в інших роботах виключно з метою виявлення можливих ознак академічного плагіату;
- самостійно виконувати навчальні завдання, завдання поточного й підсумкового контролю результатів навчання;
- надавати достовірну інформацію щодо результатів власної навчальної (наукової, творчої) діяльності, використаних методик досліджень та джерел інформації;
- не використовувати результати досліджень інших авторів без використання покликань на їхню роботу;
- своєю діяльністю сприяти збереженню та примноженню традицій університету, формуванню його позитивного іміджу;
- не чинити правопорушень і не сприяти їхньому скоєнню іншими особами;
- підтримувати атмосферу довіри, взаємної відповідальності та співпраці в освітньому середовищі;

- поважати честь, гідність та особисту недоторканність особи, незважаючи на її стать, вік, матеріальний стан, соціальне становище, расову належність, релігійні й політичні переконання;
- не дискримінувати людей на підставі академічного статусу, а також за національною, расовою, статевою чи іншою належністю;
- відповідально ставитися до своїх обов'язків, вчасно та сумлінно виконувати необхідні навчальні та науково-дослідницькі завдання;
- запобігати виникненню у своїй діяльності конфлікту інтересів, зокрема не використовувати службових і родинних зв'язків з метою отримання нечесної переваги в навчальній, науковій і трудовій діяльності;
- не брати участі в будь-якій діяльності, пов'язаній із обманом, нечесністю, списуванням, фабрикацією;
- не підроблювати документи;
- не поширювати неправдиву та компрометуючу інформацію про інших здобувачів вищої освіти, викладачів і співробітників;
- не отримувати і не пропонувати винагород за несправедливе отримання будь-яких переваг або здійснення впливу на зміну отриманої академічної оцінки;
- не залякувати й не проявляти агресії та насильства проти інших, сексуальні домагання;
- не завдавати шкоди матеріальним цінностям, матеріально-технічній базі університету та особистій власності інших студентів та/або працівників;
- не використовувати без дозволу ректорату (деканату) символіки університету в заходах, не пов'язаних з діяльністю університету;
- не здійснювати і не заохочувати будь-яких спроб, спрямованих на те, щоб за допомогою нечесних і негідних методів досягати власних корисних цілей;
- не завдавати загрози власному здоров'ю або безпеці іншим студентам та/або працівникам.

УСВІДОМЛЮЮ, що відповідно до чинного законодавства у разі недотримання Кодексу академічної доброчесності буду нести академічну та/або інші види відповідальності й до мене можуть бути застосовані заходи дисциплінарного характеру за порушення принципів академічної доброчесності.