EXPT NO:1	Implementation of data charts
DATE: 06.01.2026	

PRE-LAB QUESTIONS (PROVIDE BRIEF ANSWERS TO THE FOLLOWING QUESTIONS)

1. How can visualization help an academic institution improve student outcomes?
   Visualization transforms complex tables of marks into clear patterns. It helps institutions quickly identify at-risk students who need extra help, spot difficult subjects where the entire class is struggling, and track progress over time to see if teaching methods are working.

2. Which chart types are suitable for comparing subject-wise performance?
   - Bar Chart: Best for comparing average marks across different subjects side-by-side.
   - Box Plot: Excellent for showing the spread of marks (minimum, maximum, median) and identifying outliers in each subject.

3. What type of data scale is used for student marks?
   Student marks are measured on a ratio scale. This is because the data has a specific order (Example: 90 is better than 80), equal intervals between values, and a "true zero" point, it is possible to score 0, which represents a total absence of marks.

4. Why should raw academic data be cleaned before visualization?
   Raw data often contains errors like missing values, duplicates, or incorrect entries (Example: a mark of 105 out of 100). If not cleaned, these errors can skew averages, create misleading charts, and lead to incorrect conclusions about student performance.

5. How does visualization support evidence-based decision making?
   Visualization converts large, complex datasets into intuitive patterns, making it easier to objectively identify problems or successes. This allows to base the strategies on concrete visual proof rather than assumptions, ensuring resources are targeted exactly where the data shows they are needed.

IN-LAB EXERCISE:

OBJECTIVE:

To design appropriate data charts to analyze and compare academic performance indicators.

SCENARIO:

An autonomous engineering college wants to analyze internal assessment performance of first-year students across five subjects to identify difficult courses and improve teaching strategies.

IN-LAB TASKS (Using R Language)

- Load required R libraries (ggplot2, dplyr)
- Import dataset using read.csv()
- Perform basic data preprocessing
- Create bar chart for subject-wise average marks
- Generate line chart for performance trend across tests
- Plot pie chart for grade distribution

CODE:

```
[2]
✓ 1m
```

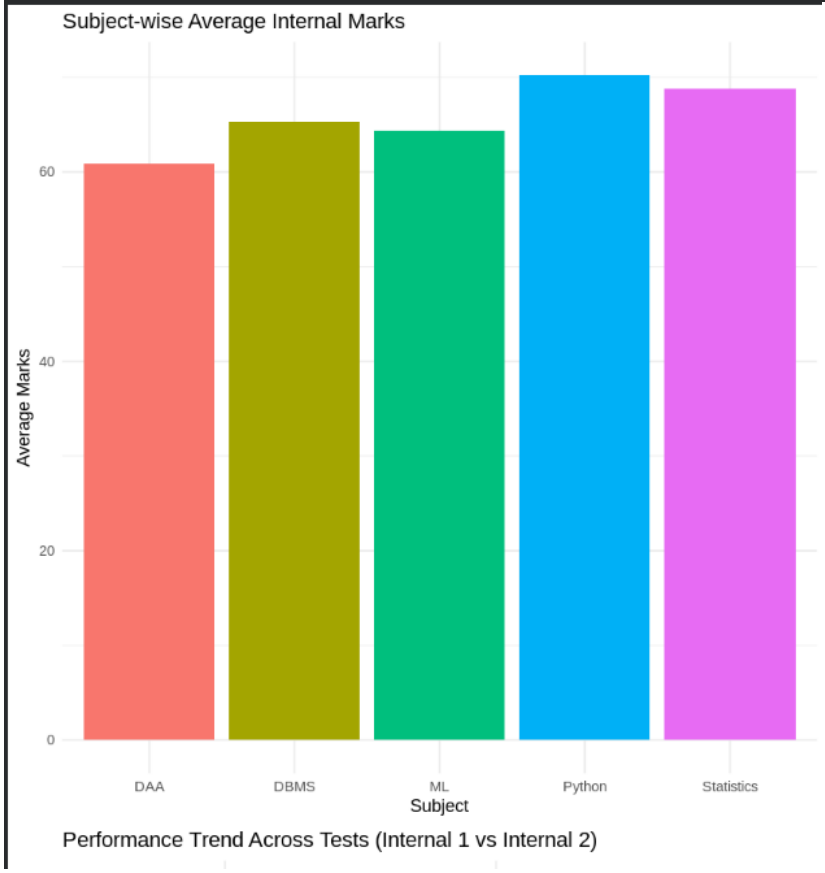
```
1
2 install.packages(c("ggplot2", "dplyr", "tidyr"))
3
```

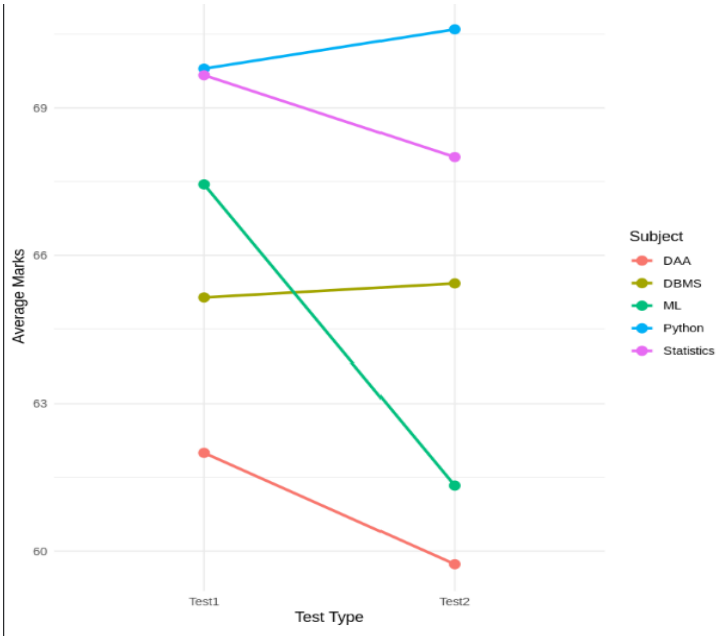
```
...  Installing packages into '/usr/local/lib/R/site-library'
     (as 'lib' is unspecified)
```

```r
# Print details
print("SAI VAISHNAVI R 23BAD094")

# Load libraries
library(ggplot2)
library(dplyr)
library(tidyr)

# Upload and read CSV file (File picker will open)
data <- read.csv("/content/1.student_performance.csv")

# View structure and summary
str(data)
summary(data)

# ---------------------------------
# Average of internal tests
# ---------------------------------
data <- data %>%
  mutate(
    Avg_Internal = (Internal_Test1 + Internal_Test2) / 2
  )

# ---------------------------------
# Subject-wise average
# ---------------------------------
subject_avg <- data %>%
  group_by(Subject) %>%
  summarise(
    Mean_Marks = mean(Avg_Internal, na.rm = TRUE)
  )

# Bar plot: Subject-wise average internal marks
ggplot(subject_avg, aes(x = Subject, y = Mean_Marks, fill = Subject)) +
  geom_bar(stat = "identity") +
  theme_minimal() +
  labs(
    title = "Subject-wise Average Internal Marks",
    x = "Subject",
    y = "Average Marks"
  ) +
  theme(legend.position = "none")

# ---------------------------------
# Trend across tests
# ---------------------------------
trend_data <- data %>%
  group_by(Subject) %>%
  summarise(
    Test1 = mean(Internal_Test1, na.rm = TRUE),
    Test2 = mean(Internal_Test2, na.rm = TRUE)
  ) %>%
  pivot_longer(
    cols = c(Test1, Test2),
    names_to = "Test",
    values_to = "Marks"
  )

# Line plot: Performance trend
ggplot(trend_data, aes(x = Test, y = Marks, group = Subject, color = Subject)) +
  geom_line(size = 1) +
  geom_point(size = 3) +
  theme_minimal() +
  labs(
    title = "Performance Trend Across Tests (Internal 1 vs Internal 2)",
    x = "Test Type",
    y = "Average Marks"
  )

# ---------------------------------
# Grade distribution
# ---------------------------------
grade_counts <- data %>%
  count(Final_Grade) %>%
  mutate(
    prop = n / sum(n) * 100,
    ypos = cumsum(prop) - 0.5 * prop
  )

# Pie chart: Grade distribution
ggplot(grade_counts, aes(x = "", y = prop, fill = Final_Grade)) +
  geom_bar(stat = "identity", width = 1, color = "white") +
  coord_polar("y", start = 0) +
  theme_void() +
  labs(title = "Final Grade Distribution") +
  geom_text(
    aes(y = ypos, label = paste0(round(prop, 1), "%")),
```

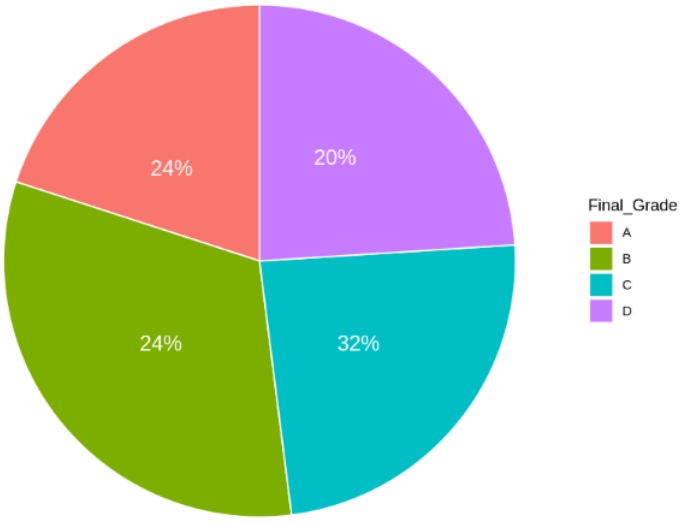
OUTPUT:

```
[1] "SAI VAISHNAVI R 23BAD094"
'data.frame':   50 obs. of  9 variables:
 $ Student_ID          : int  1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 ...
 $ Department          : chr  "IT" "AI&DS" "IT" "IT" ...
 $ Semester            : int  6 5 6 5 6 7 7 5 7 7 ...
 $ Subject             : chr  "Statistics" "DAA" "DAA" "ML" ...
 $ Internal_Test1      : int  81 84 45 67 67 83 83 59 69 50 ...
 $ Internal_Test2      : int  73 74 42 60 43 94 85 67 73 47 ...
 $ Assignment_Marks    : int  11 16 19 12 16 19 18 13 10 11 ...
 $ Attendance_Percentage: int  71 67 81 97 76 86 86 94 72 91 ...
 $ Final_Grade         : chr  "D" "A" "A" "C" ...
  Student_ID    Department          Semester      Subject
 Min.   :1001   Length:50          Min.   :5.00   Length:50
 1st Qu.:1013   Class :character   1st Qu.:5.00   Class :character
 Median :1026   Mode  :character   Median :6.00   Mode  :character
 Mean   :1026                      Mean   :5.86
 3rd Qu.:1038                      3rd Qu.:6.00
 Max.   :1050                      Max.   :7.00
 Internal_Test1  Internal_Test2  Assignment_Marks Attendance_Percentage
 Min.   :40.00   Min.   :42.00   Min.   :10.00    Min.   :66.00
 1st Qu.:52.50   1st Qu.:52.00   1st Qu.:12.00    1st Qu.:81.00
 Median :67.00   Median :64.00   Median :13.00    Median :88.50
 Mean   :66.36   Mean   :64.48   Mean   :14.06    Mean   :86.22
 3rd Qu.:79.50   3rd Qu.:74.00   3rd Qu.:17.00    3rd Qu.:94.75
 Max.   :88.00   Max.   :94.00   Max.   :19.00    Max.   :99.00
 Final_Grade
 Length:50
 Class :character
 Mode  :character
```





Final Grade Distribution



POST-LAB QUESTIONS (PROVIDE BRIEF ANSWERS TO THE FOLLOWING QUESTIONS)

1. Which subject shows consistently low performance and why?
   DAA shows consistently low performance. It has the lowest overall average (60.9) and performed poorly in both Internal Test 1 (62.0) and Internal Test 2 (59.7).

2. Why is a line chart suitable for trend analysis?
   A line chart is ideal because it connects individual data points with a continuous line. This visual continuity makes it instantly easy for the human eye to detect the direction of change (improvement or decline) and the rate of change (steep vs. gradual slope) between Test 1 and Test 2.

3. What limitations does a pie chart have in analytics?
   - Hard to Compare: It is difficult for the human eye to accurately compare the size of slices that are similar in percentage (e.g., 24% vs 26%).
   - No Trends: It cannot show changes over time.
   - Clutter: It becomes unreadable if there are too many categories (slices).

4. How can this analysis help curriculum planning?
   It helps to allocate resources more effectively. For example, knowing that DAA and ML in Test 2 are pain points, the curriculum can be adjusted to include more tutorial hours for these subjects, introduce prerequisite workshops, or split the heavy syllabus into two manageable semesters.

5. How can such visualizations be integrated into AI-driven academic analytics?
   These visualizations can serve as Early Warning Systems. An AI model can analyze this data in real-time to predict which students are likely to fail. The system can then automatically generate these charts on a teacher's dashboard to highlight at-risk students, allowing for intervention before the final exams.

ASSESSMENT

Description	Max Marks	Marks Awarded
Pre Lab Exercise	**5**	
In Lab Exercise	**10**	
Post Lab Exercise	**5**	
Viva	**10**	
Total	**30**	
	Faculty Signature	