

Day 1 – Practical Session

Course Code: DSA0612

Course: Data Handling and Visualization

1. Visualize Scatter plot for given dataset.

Weight (kg)	Height (cm)
60	170
65	175
70	168
72	180
75	178

Sample Code:

```
# Load necessary library
```

```
library(ggplot2)
```

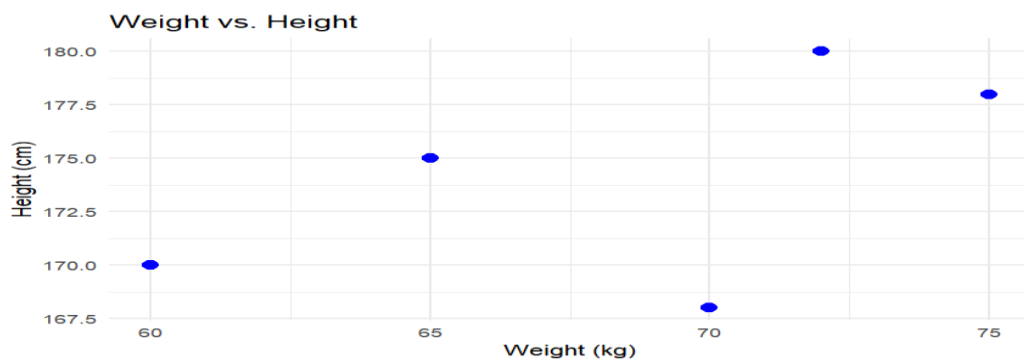
```
# Create a data frame
```

```
weight_height_data <- data.frame(  
  Weight = c(60, 65, 70, 72, 75),  
  Height = c(170, 175, 168, 180, 178)  
)
```

```
# Plot scatter plot
```

```
ggplot(weight_height_data, aes(x = Weight, y = Height)) +  
  geom_point(color = "blue", size = 3) +  
  labs(title = "Weight vs. Height",  
       x = "Weight (kg)",  
       y = "Height (cm)") +  
  theme_minimal()
```

OUTPUT:



2. Visualize a Line Plot for the Given Population Data

Year	Population (millions)
2010	1000
2011	1020
2012	1040
2013	1060
2014	1080

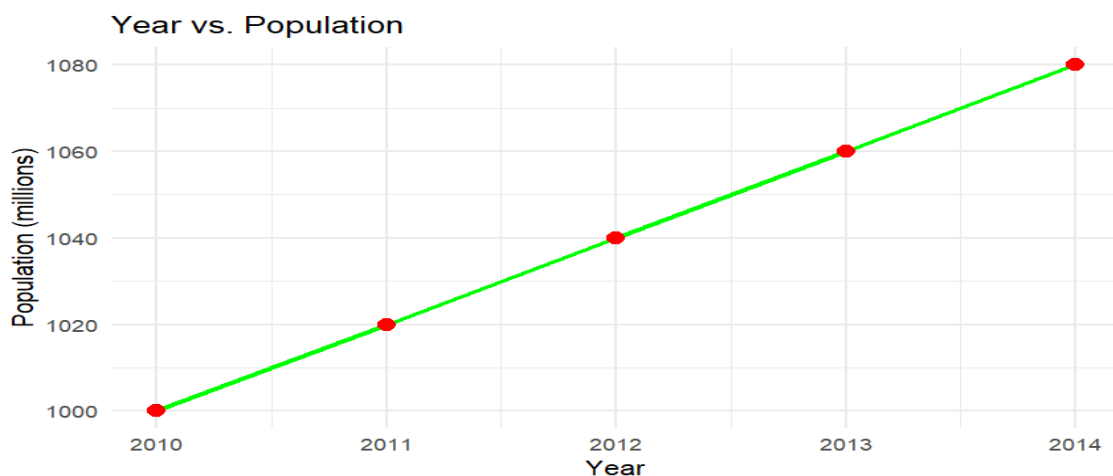
Create a data frame

```
year_population_data <- data.frame(  
  Year = 2010:2014,  
  Population = c(1000, 1020, 1040, 1060, 1080)  
)
```

Plot line plot

```
ggplot(year_population_data, aes(x = Year, y = Population)) +  
  geom_line(color = "green", size = 1) +  
  geom_point(color = "red", size = 3) +  
  labs(title = "Year vs. Population",  
       x = "Year",  
       y = "Population (millions)") +  
  theme_minimal()
```

OUTPUT:



3. Bar Plot

Dataset: Sales by Product

Product	Sales
A	300
B	450
C	500
D	350
E	400

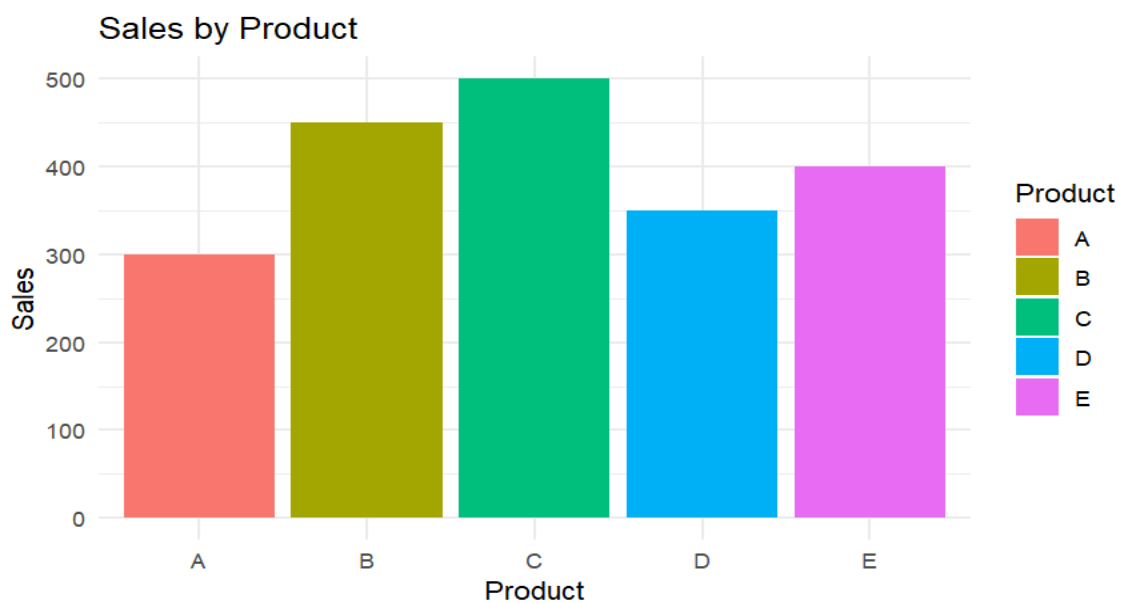
Create a data frame

```
sales_data <- data.frame(  
  Product = c("A", "B", "C", "D", "E"),  
  Sales = c(300, 450, 500, 350, 400)  
)
```

Plot bar plot

```
ggplot(sales_data, aes(x = Product, y = Sales, fill = Product)) +  
  geom_bar(stat = "identity") +  
  labs(title = "Sales by Product",  
       x = "Product",  
       y = "Sales") +  
  theme_minimal()
```

OUTPUT:



4. Histogram

Dataset: Age Distribution

Age
25
30
35
40
45
50
55
60
65
70

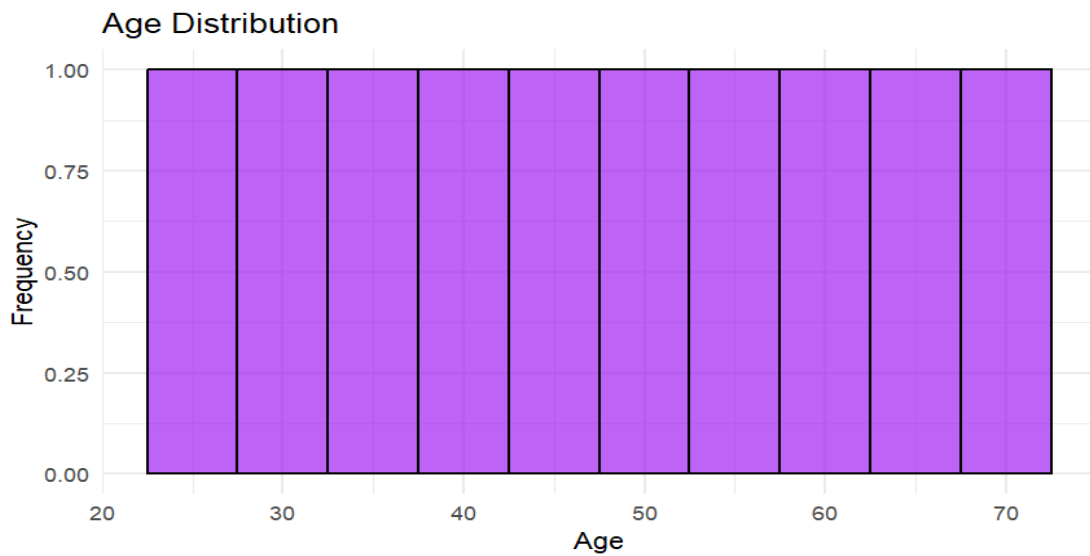
Create a data frame

```
age_data <- data.frame(  
  Age = c(25, 30, 35, 40, 45, 50, 55, 60, 65, 70)  
)
```

Plot histogram

```
ggplot(age_data, aes(x = Age)) +  
  geom_histogram(binwidth = 5, fill = "purple", color = "black", alpha = 0.7) +  
  labs(title = "Age Distribution",  
        x = "Age",  
        y = "Frequency") +  
  theme_minimal()
```

OUTPUT:



5. Pie Chart

Dataset: Market Share by Company

Company	Market Share
A	20%
B	30%
C	25%
D	25%

Create a data frame

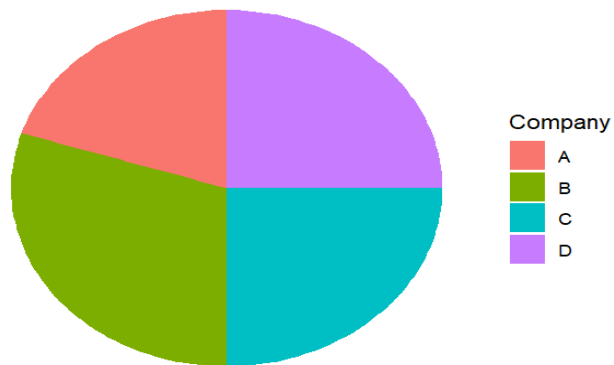
```
market_share_data <- data.frame(  
  Company = c("A", "B", "C", "D"),  
  Market_Share = c(20, 30, 25, 25)  
)
```

Plot pie chart

```
ggplot(market_share_data, aes(x = "", y = Market_Share, fill = Company)) +  
  geom_bar(stat = "identity", width = 1) +  
  coord_polar("y") +  
  labs(title = "Market Share by Company",  
    x = NULL,  
    y = NULL) +  
  theme_void() +  
  theme(legend.position = "right")
```

OUTPUT:

Market Share by Company



6. Box Plot

Dataset: Exam Scores by Class

Class	Scores
A	85
A	90
A	78
A	92
B	88
B	76
B	80
B	84

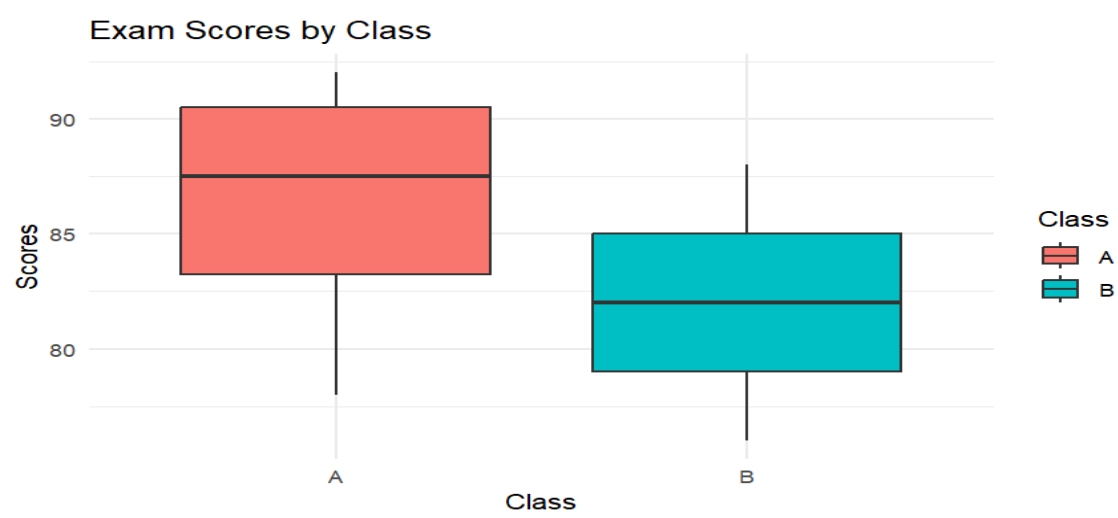
Create a data frame

```
scores_data <- data.frame(  
  Class = c("A", "A", "A", "A", "B", "B", "B", "B"),  
  Scores = c(85, 90, 78, 92, 88, 76, 80, 84)  
)
```

Plot box plot

```
ggplot(scores_data, aes(x = Class, y = Scores, fill = Class)) +  
  geom_boxplot() +  
  labs(title = "Exam Scores by Class",  
    x = "Class",  
    y = "Scores") +  
  theme_minimal()
```

OUTPUT:



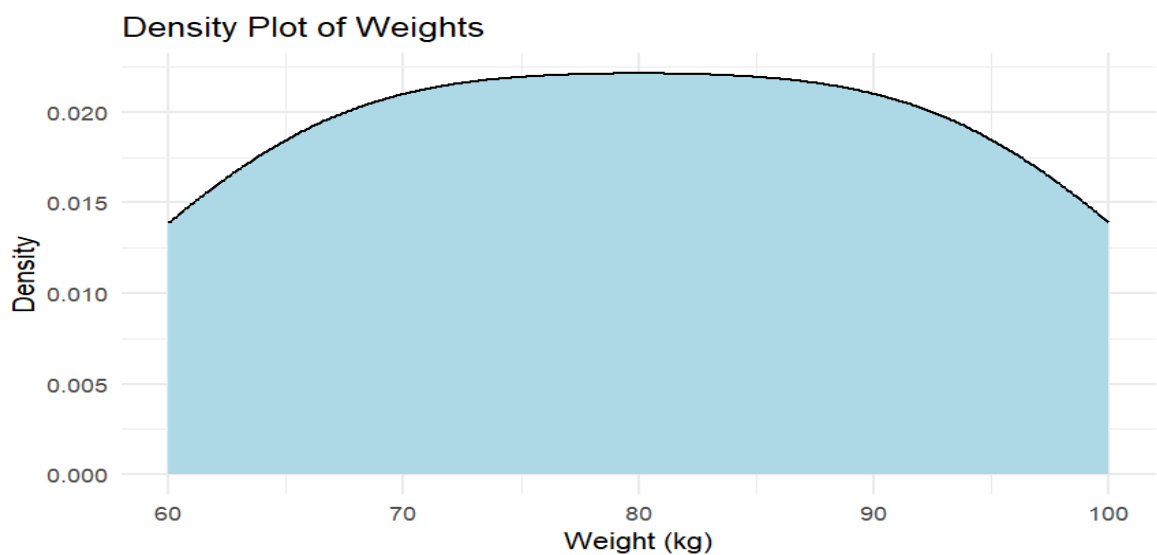
7. Density Plot

Dataset: Weights of Individuals

Weight (kg)
60
65
70
75
80
85
90
95
100

```
# Create a data frame
weights_data <- data.frame(
  Weight = c(60, 65, 70, 75, 80, 85, 90, 95, 100)
)
```

```
# Plot density plot
ggplot(weights_data, aes(x = Weight)) +
  geom_density(fill = "lightblue") +
  labs(title = "Density Plot of Weights",
       x = "Weight (kg)",
       y = "Density") +
  theme_minimal()
```



8. Violin Plot

Dataset: Scores by Group

Group	Score
X	80
X	85
X	78
X	92
Y	88
Y	76
Y	80
Y	84

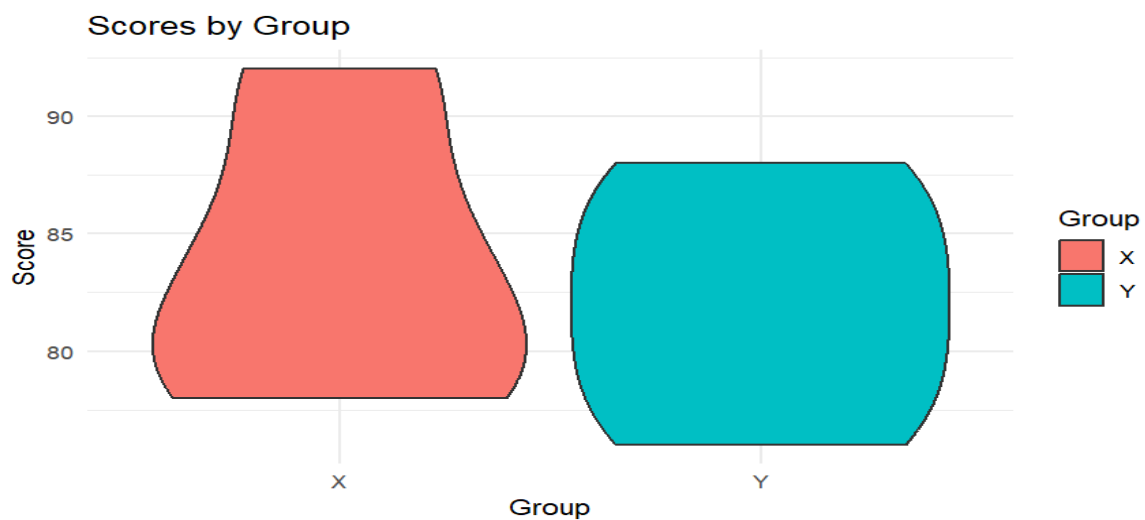

```
# Create a data frame
```

```
scores_group_data <- data.frame(  
  Group = c("X", "X", "X", "X", "Y", "Y", "Y", "Y"),  
  Score = c(80, 85, 78, 92, 88, 76, 80, 84)  
)
```

```
# Plot violin plot
```

```
ggplot(scores_group_data, aes(x = Group, y = Score, fill = Group)) +  
  geom_violin() +  
  labs(title = "Scores by Group",  
        x = "Group",  
        y = "Score") +  
  theme_minimal()
```

OUTPUT:



9. Heatmap

Dataset: Monthly Temperatures

Month	City A	City B	City C
Jan	5	10	15
Feb	6	11	16
Mar	7	12	17
Apr	8	13	18
May	9	14	19

```

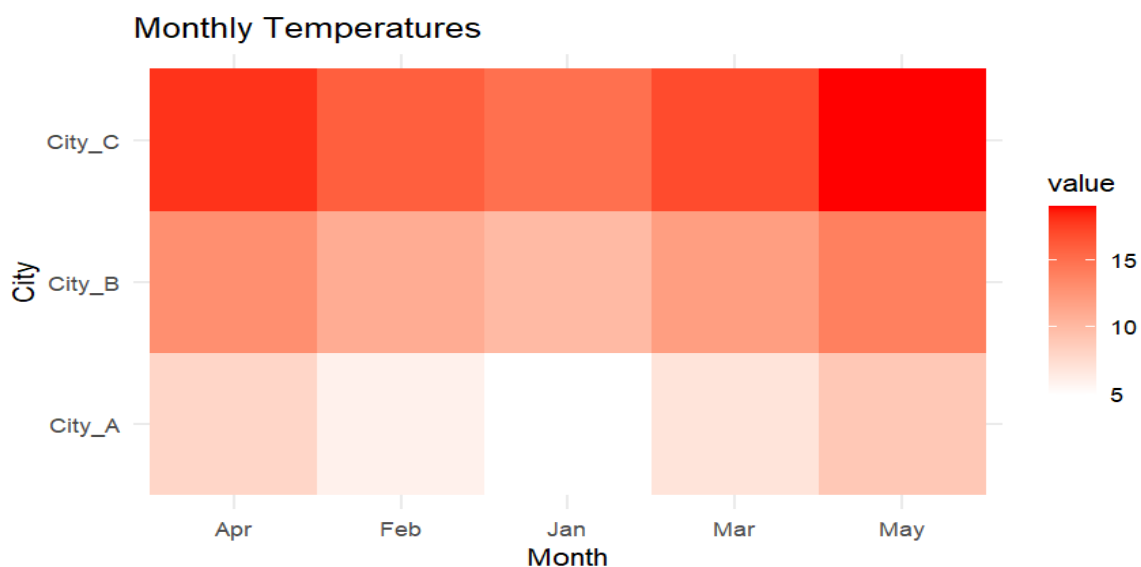
# Create a data frame
temperature_data <- data.frame(
  Month = c("Jan", "Feb", "Mar", "Apr", "May"),
  City_A = c(5, 6, 7, 8, 9),
  City_B = c(10, 11, 12, 13, 14),
  City_C = c(15, 16, 17, 18, 19)
)

# Reshape data for heatmap
library(reshape2)
melted_data <- melt(temperature_data, id.vars = "Month")

# Plot heatmap
ggplot(melted_data, aes(x = Month, y = variable, fill = value)) +
  geom_tile() +
  scale_fill_gradient(low = "white", high = "red") +
  labs(title = "Monthly Temperatures",
       x = "Month",
       y = "City") +
  theme_minimal()

```

OUTPUT:



10. Facet Plot

Dataset: Sales by Month and Region

Month	Region	Sales
Jan	East	200
Jan	West	150
Feb	East	220
Feb	West	170
Mar	East	210
Mar	West	160

Create a data frame

```
sales_region_data <- data.frame(  
  Month = c("Jan", "Jan", "Feb", "Feb", "Mar", "Mar"),  
  Region = c("East", "West", "East", "West", "East", "West"),  
  Sales = c(200, 150, 220, 170, 210, 160)  
)
```

Plot facet plot

```
ggplot(sales_region_data, aes(x = Month, y = Sales, fill = Region)) +  
  geom_bar(stat = "identity", position = "dodge") +  
  facet_wrap(~ Region) +  
  labs(title = "Sales by Month and Region",  
       x = "Month",  
       y = "Sales") +  
  theme_minimal()
```

OUTPUT:

