

LAB PROGRAM-DAY 5

```
store_data <- data.frame(
  StoreID = 1:48,
  Region = rep(c("North", "South", "East", "West"), 12),
  Month = rep(c("January", "February", "March", "April", "May", "June",
    "July", "August", "September", "October", "November", "December"),
    each = 4),
  TotalSales = c(
    10000, 15000, 12000, 13000,
    9000, 16000, 14000, 11000,
    11000, 17000, 13000, 15000,
    12000, 18000, 16000, 14000,
    13000, 19000, 15000, 16000,
    14000, 20000, 17000, 15000,
    15000, 21000, 16000, 14000,
    16000, 22000, 13000, 15000,
    12000, 18000, 16000, 14000,
    13000, 19000, 15000, 16000,
    14000, 20000, 17000, 15000,
    15000, 21000, 16000, 14000
  ),
  NumberOfCustomers = c(
    500, 700, 600, 650,
    450, 800, 700, 550,
    550, 850, 650, 750,
    600, 900, 800, 700,
    650, 950, 750, 800,
    700, 1000, 850, 750,
    750, 1050, 800, 700,
    800, 1100, 650, 750,
    600, 900, 800, 700,
    600, 900, 800, 700,
    600, 900, 800, 700
  )
)
```

```

650,950,750,800,
700,1000,850,750,
750,1050,800,700
),
AvgTransactionValue = c(20,21.43,20,20,20,20,20,20,20,20,20,
                        20,20,20,20,20,20,20,20,20,20,20,
                        20,20,20,20,20,20,20,20,20,20,20,
                        20,20,20,20,20,20,20,20,20,20,20)
)

```

```

library(dplyr)
library(ggplot2)
library(plotly)
library(tidyr)
library(lubridate)

# Check missing values
sum(is.na(store_data))

# Convert Month to ordered factor
store_data$Month <- factor(store_data$Month,
                            levels = month.name)

monthly_sales <- store_data %>%
  group_by(Month) %>%
  summarize(TotalSales = sum(TotalSales))

regional_sales <- store_data %>%
  group_by(Region) %>%
  summarize(TotalSales = sum(TotalSales),
            Customers = sum(NumberOfCustomers))

```

```
ggplot(regional_sales, aes(x=Region, y=TotalSales, fill=Region)) +  
  geom_bar(stat="identity") +  
  ggtitle("Total Sales by Region")
```

```
ggplot(monthly_sales, aes(Month, TotalSales, group=1)) +  
  geom_line(color="blue") +  
  geom_point(size=3) +  
  ggtitle("Monthly Sales Trend")
```

```
ggplot(store_data, aes(StoreID, TotalSales, color=Region)) +  
  geom_point(size=3) +  
  ggtitle("Store-wise Total Sales")
```

```
ggplot(store_data, aes(Month, NumberOfCustomers, fill=Region)) +  
  geom_bar(stat="identity", position="dodge") +  
  ggtitle("Customer Count per Month")
```

```
mean_ATV <- mean(store_data$AvgTransactionValue)  
median_ATV <- median(store_data$AvgTransactionValue)  
sd_ATV <- sd(store_data$AvgTransactionValue)
```

```
print(mean_ATV)  
print(median_ATV)  
print(sd_ATV)
```

```
ggplot(monthly_sales, aes(as.numeric(Month), TotalSales)) +  
  geom_point() +  
  geom_smooth(method="lm", color="red") +  
  ggtitle("Trendline of Monthly Sales")
```

```
ggplot(regional_sales, aes(TotalSales, Customers, color=Region)) +  
  geom_point(size=5) +
```

```

ggtitle("Region Sales vs Customers")

plot_ly(store_data,
       x = ~TotalSales,
       y = ~NumberOfCustomers,
       z = ~AvgTransactionValue,
       type = "scatter3d",
       mode = "markers",
       color = ~Region)

# 13. Boxplot by region
ggplot(store_data, aes(Region, TotalSales)) + geom_boxplot()

# 14. Histogram of sales
ggplot(store_data, aes(TotalSales)) + geom_histogram()

# 15. Density plot
ggplot(store_data, aes(TotalSales, fill=Region)) + geom_density(alpha=0.5)

# 16. Customers trendline
ggplot(store_data %>% group_by(Month) %>%
summarise(Cust=sum(NumberOfCustomers)),
       aes(Month, Cust, group=1)) +
       geom_line() + geom_point()

# 17. Scatter: Sales vs Customers
ggplot(store_data, aes(NumberOfCustomers, TotalSales, color=Region)) + geom_point()

# 18. Pie Chart of Regional Sales
ggplot(regional_sales, aes(x="", y=TotalSales, fill=Region)) +
       geom_bar(stat="identity", width=1) + coord_polar("y")

# 19. Facet by Region

```

```

ggplot(store_data, aes(Month, TotalSales)) +
  geom_line(aes(group=Region, color=Region)) +
  facet_wrap(~Region)

# 20. Heatmap: Sales by Region & Month
ggplot(store_data, aes(Region, Month, fill=TotalSales)) +
  geom_tile()

# 21. 3D bar: Month & Region
plot_ly(store_data, x=~Month, y=~Region, z=~TotalSales, type="bar3d")

# 22. Lollipop chart
ggplot(monthly_sales, aes(Month, TotalSales)) +
  geom_segment(aes(xend=Month, y=0, yend=TotalSales)) +
  geom_point(size=4)

# 23. Area chart
ggplot(monthly_sales, aes(Month, TotalSales)) + geom_area(fill="skyblue")

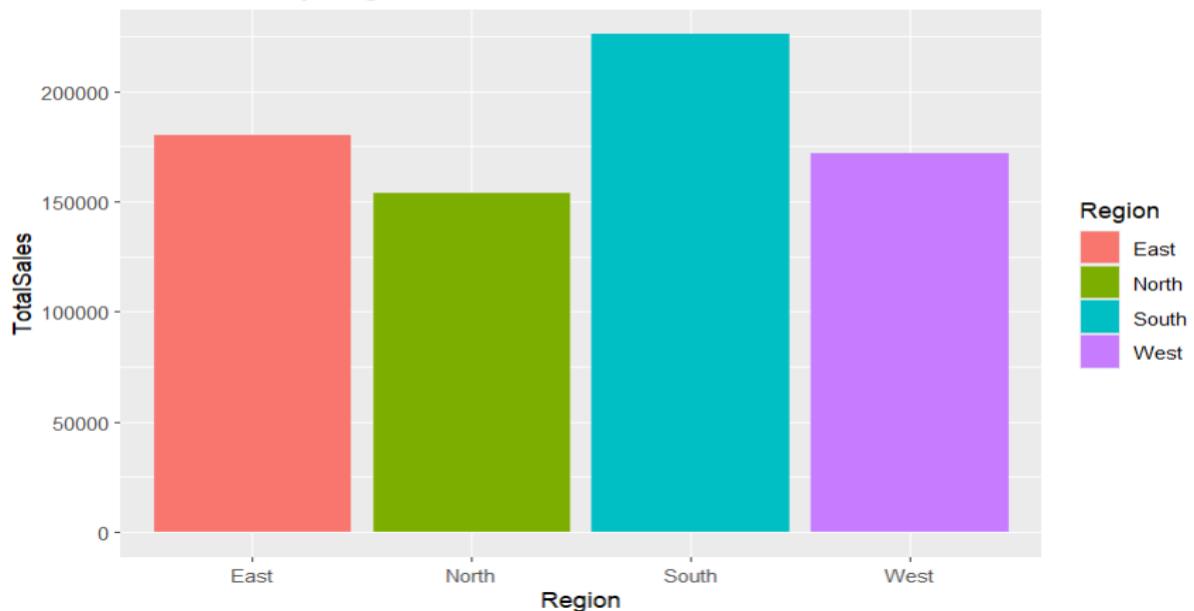
# 24. Violin: Customers by region
ggplot(store_data, aes(Region, NumberOfCustomers, fill=Region)) + geom_violin()

# 25. Correlation plot
cor_matrix <- store_data %>%
  select(TotalSales, NumberOfCustomers, AvgTransactionValue) %>% cor()
print(cor_matrix)

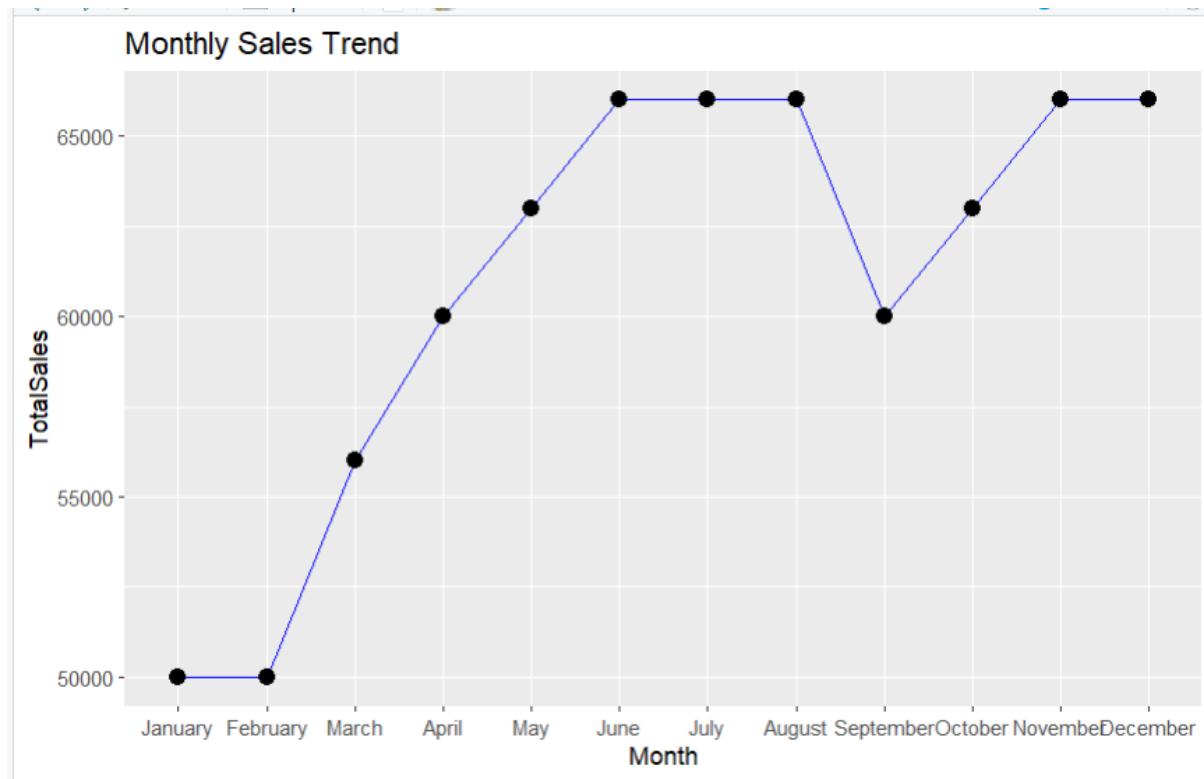
```

OUTPUT:

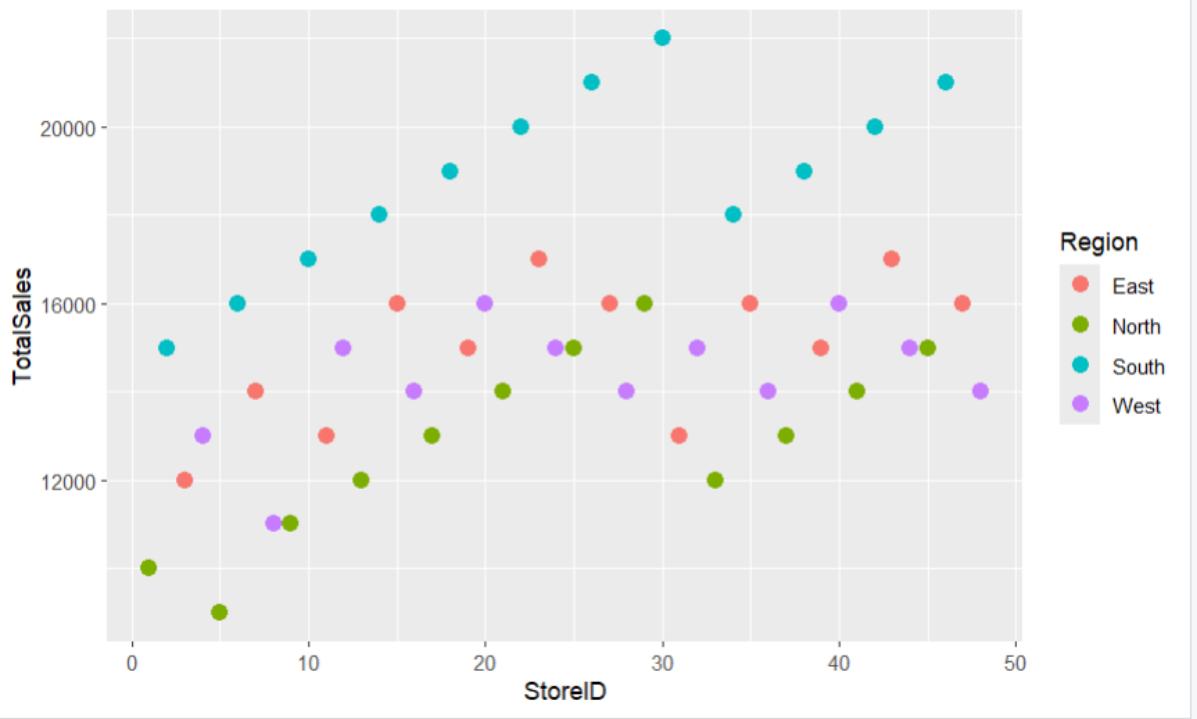
Total Sales by Region



Monthly Sales Trend



Store-wise Total Sales



Customer Count per Month

