Lab_Sea_2304 — COVID-19 Data Visualization in R

This analysis uses **R**, **ggplot2**, **dplyr**, **tidyr**, and **scales** to visualize global COVID-19 data from the file `country_wise_latest.csv`.

We generate the following graphs:

- Top 10 countries by confirmed cases
- Regional comparison bar chart
- Pie chart (with percentage shown in legend)
- Stacked bar chart (Confirmed, Deaths, Recovered)

1️⃣ Loading Libraries and Dataset

```
# COVID-19 Data Analysis in R
This document explains the R script used to analyze and visualize **COVID-19
country-wise data** from `country_wise_latest.csv`.
It includes:
- A bar chart of top 10 countries
- A regional bar chart
- A pie chart with percentages in legends
- A stacked bar chart comparing key metrics

---

## 📌 1. Load Libraries and Read CSV

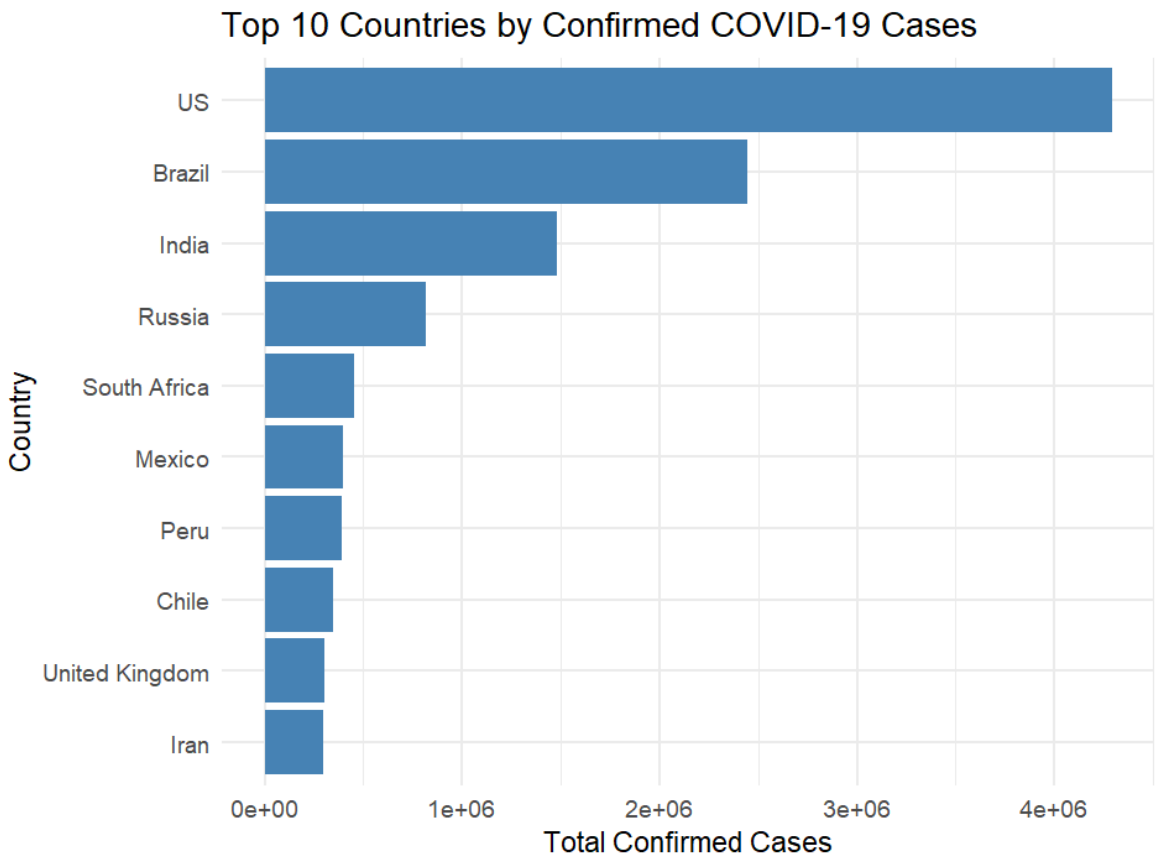
```r
library(ggplot2)
library(dplyr)
library(tidyr)
library(scales)

data <- read.csv("country_wise_latest.csv")
```

These libraries provide:

- **ggplot2** → for graphs
- **dplyr** → for data manipulation
- **tidyr** → for reshaping tables
- **scales** → formatting percentages

📌 2. Top 10 Countries by Confirmed Cases (Horizontal Bar)
```



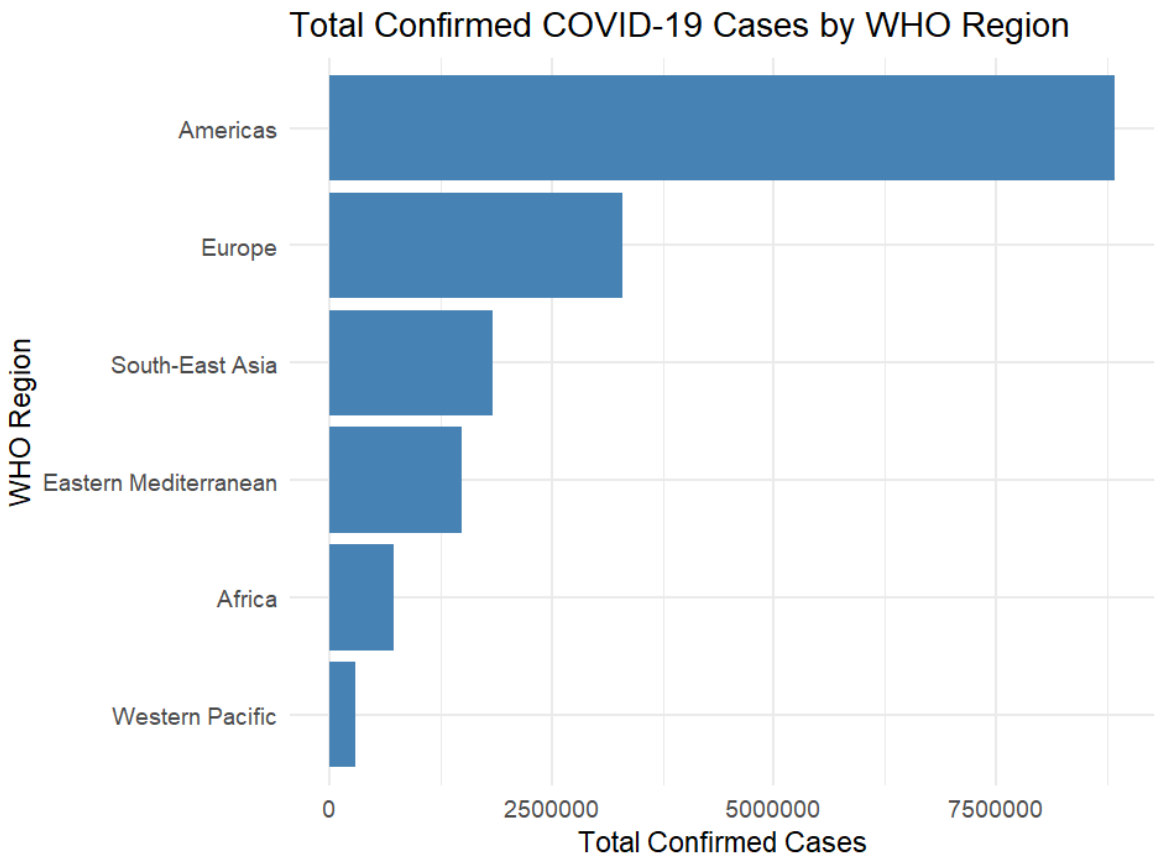

```r
top10 <- data %>%
  arrange(desc(Confirmed)) %>%
  head(10)

ggplot(top10, aes(x = reorder(Country.Region, Confirmed),
                  y = Confirmed)) +
  geom_bar(stat = "identity", fill = "steelblue") +
  coord_flip() +
  labs(title = "Top 10 Countries by Confirmed COVID-19 Cases",
       x = "Country",
       y = "Total Confirmed Cases") +
  theme_minimal(base_size = 14)
```

✔️ What this shows

A horizontal bar chart showing **which countries have the highest total cases**.

📌 3. Total Confirmed Cases by WHO Region (Bar Chart)



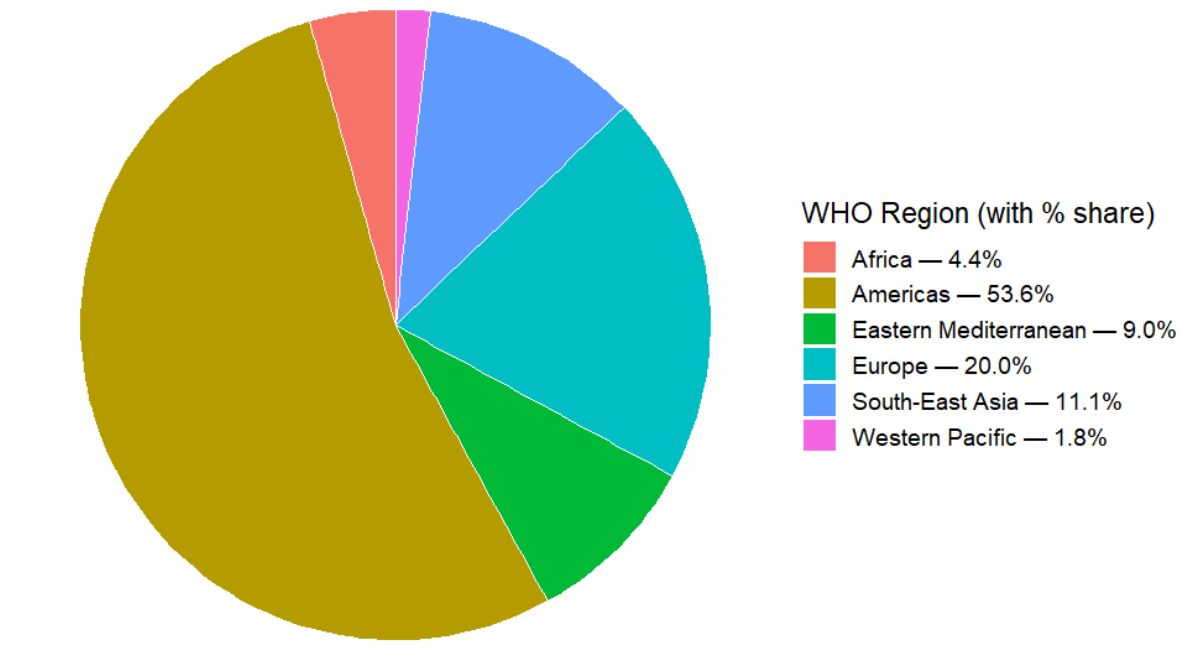

```r
region_data <- data %>%
  group_by(WHO.Region) %>%
  summarise(TotalConfirmed = sum(Confirmed, na.rm = TRUE)) %>%
  arrange(desc(TotalConfirmed))

ggplot(region_data, aes(x = reorder(WHO.Region, TotalConfirmed),
                        y = TotalConfirmed)) +
  geom_bar(stat = "identity", fill = "steelblue") +
  coord_flip() +
  labs(title = "Total Confirmed COVID-19 Cases by WHO Region",
       x = "WHO Region",
       y = "Total Confirmed Cases") +
  theme_minimal(base_size = 14)
```

✔️ What this shows

This graph compares **regions instead of individual countries**.

📌 4. Pie Chart of Region-wise Case Share (Percentages in Legend)



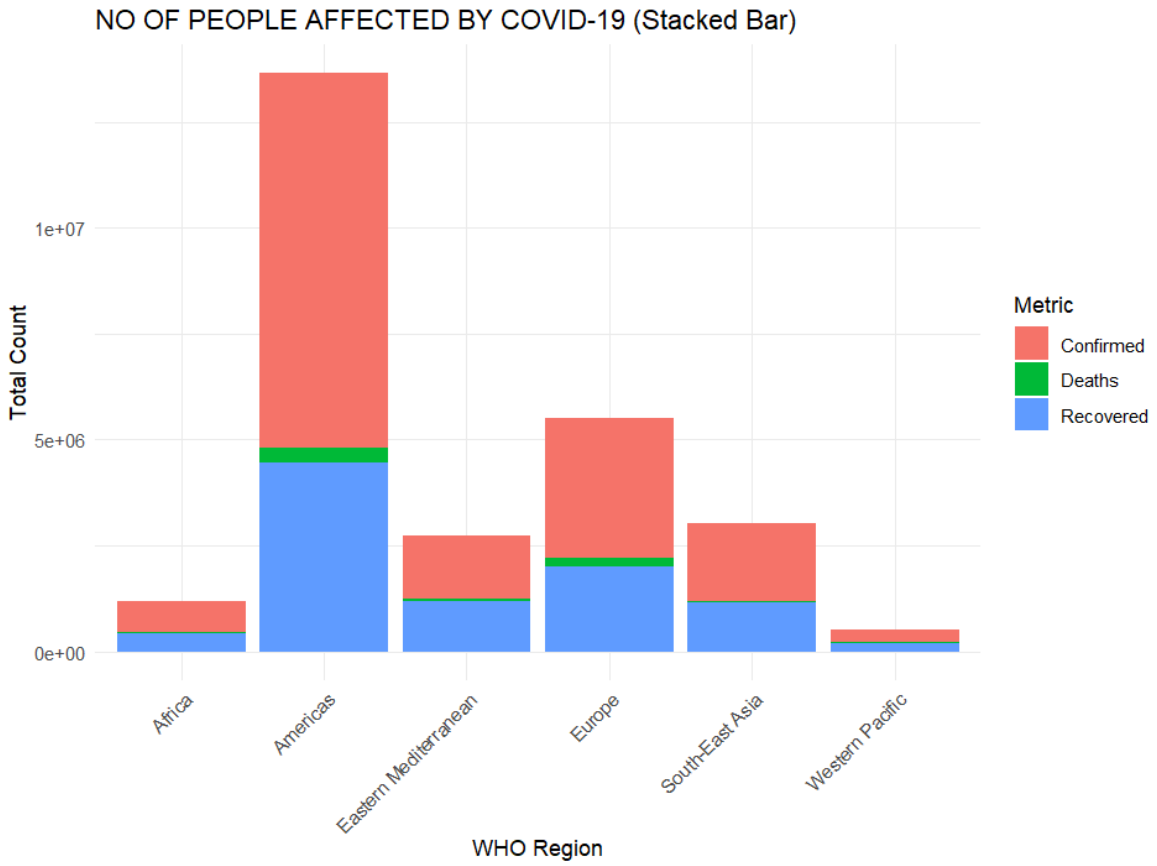

```r
region_data <- data %>%
  group_by(WHO.Region) %>%
  summarise(TotalConfirmed = sum(Confirmed, na.rm = TRUE)) %>%
  mutate(Percent = TotalConfirmed / sum(TotalConfirmed),
         LegendLabel = paste0(WHO.Region, " — ", percent(Percent)))

ggplot(region_data, aes(x = "", y = TotalConfirmed, fill = LegendLabel)) +
  geom_bar(width = 1, stat = "identity", color = "white") +
  coord_polar("y") +
  labs(title = "Share of Confirmed COVID-19 Cases by WHO Region") +
  theme_void(base_size = 14) +
  guides(fill = guide_legend(title = "WHO Region (with % share)"))
```

✔️ What this shows

A clean pie chart where **percent values appear in the legend** for better readability.

📌 5. Stacked Bar Chart (Confirmed vs Deaths vs Recovered)




```r
region_stack <- data %>%
  group_by(WHO.Region) %>%
  summarise(
    Confirmed = sum(Confirmed, na.rm = TRUE),
    Deaths = sum(Deaths, na.rm = TRUE),
    Recovered = sum(Recovered, na.rm = TRUE)
  )

region_long <- region_stack %>%
  pivot_longer(cols = c(Confirmed, Deaths, Recovered),
               names_to = "Metric",
               values_to = "Value")

ggplot(region_long, aes(x = WHO.Region, y = Value, fill = Metric)) +
  geom_bar(stat = "identity") +
  labs(title = "NO OF PEOPLE AFFECTED BY COVID-19 (Stacked Bar)",
       x = "WHO Region",
       y = "Total Count",
       fill = "Metric") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

✔ What this shows

A **stacked comparison** of Confirmed, Deaths, and Recovered across regions.

✅ Summary

This analysis gives:

- **Country-level** and **region-level** insights
- A mix of **bar, stacked, and pie charts**
- Clean, readable ggplot visualizations
- Proper grouping, percentages, and labeling

If you want this turned into a **PDF report**, **PowerPoint**, or **HTML dashboard**, tell me bab.