COVID\_19 Report

Monone Damilola

2024-04-10

# Introduction

This report provides an analysis of COVID-19 data, focusing on different continents’ population, total cases, and death cases. It is a personal project to practice R studio, in which the data set used is an open data from COVID\_19 website.

## Data cleaning and manipulation

cleaned\_data <- na.omit(covid\_data)  
print(cleaned\_data)

## # A tibble: 69,468 × 7  
## iso\_code continent location date population total\_cases  
## <chr> <chr> <chr> <dttm> <dbl> <dbl>  
## 1 AFG Asia Afghanistan 2020-03-22 00:00:00 38928341 34  
## 2 AFG Asia Afghanistan 2020-03-23 00:00:00 38928341 41  
## 3 AFG Asia Afghanistan 2020-03-24 00:00:00 38928341 43  
## 4 AFG Asia Afghanistan 2020-03-25 00:00:00 38928341 76  
## 5 AFG Asia Afghanistan 2020-03-26 00:00:00 38928341 80  
## 6 AFG Asia Afghanistan 2020-03-27 00:00:00 38928341 91  
## 7 AFG Asia Afghanistan 2020-03-28 00:00:00 38928341 107  
## 8 AFG Asia Afghanistan 2020-03-29 00:00:00 38928341 118  
## 9 AFG Asia Afghanistan 2020-03-30 00:00:00 38928341 146  
## 10 AFG Asia Afghanistan 2020-03-31 00:00:00 38928341 175  
## # ℹ 69,458 more rows  
## # ℹ 1 more variable: total\_deaths <dbl>

## Summary Statistics

# Overview of the Dataset

print(summary)

## function (object, ...)   
## UseMethod("summary")  
## <bytecode: 0x00000224f634ccb8>  
## <environment: namespace:base>

# Calculate total population, total cases and death cases per continent

population\_table <- cleaned\_data %>%  
 group\_by(continent) %>%  
 summarise(total\_population = sum(population))  
  
deaths\_table <- cleaned\_data %>%  
 group\_by(continent) %>%  
 summarise(total\_deaths = sum(total\_deaths))  
  
cases\_table <- cleaned\_data %>%  
 group\_by(continent) %>%  
 summarise(total\_cases = sum(total\_cases))

# Tables

print(population\_table)

## # A tibble: 6 × 2  
## continent total\_population  
## <chr> <dbl>  
## 1 Africa 525928577499  
## 2 Asia 1959500225160  
## 3 Europe 313849887963  
## 4 North America 245949704166  
## 5 Oceania 15500274648  
## 6 South America 176001182162

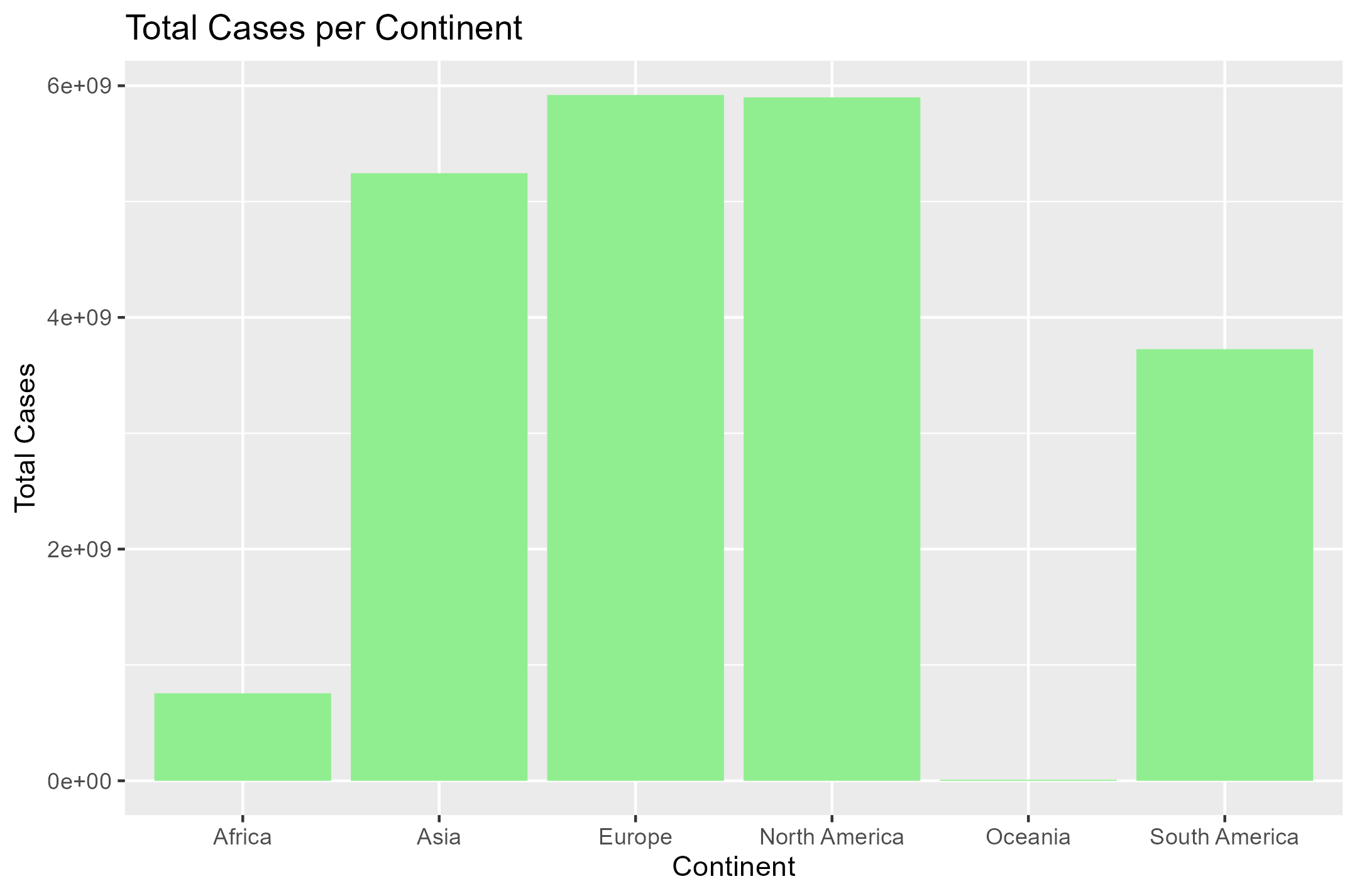
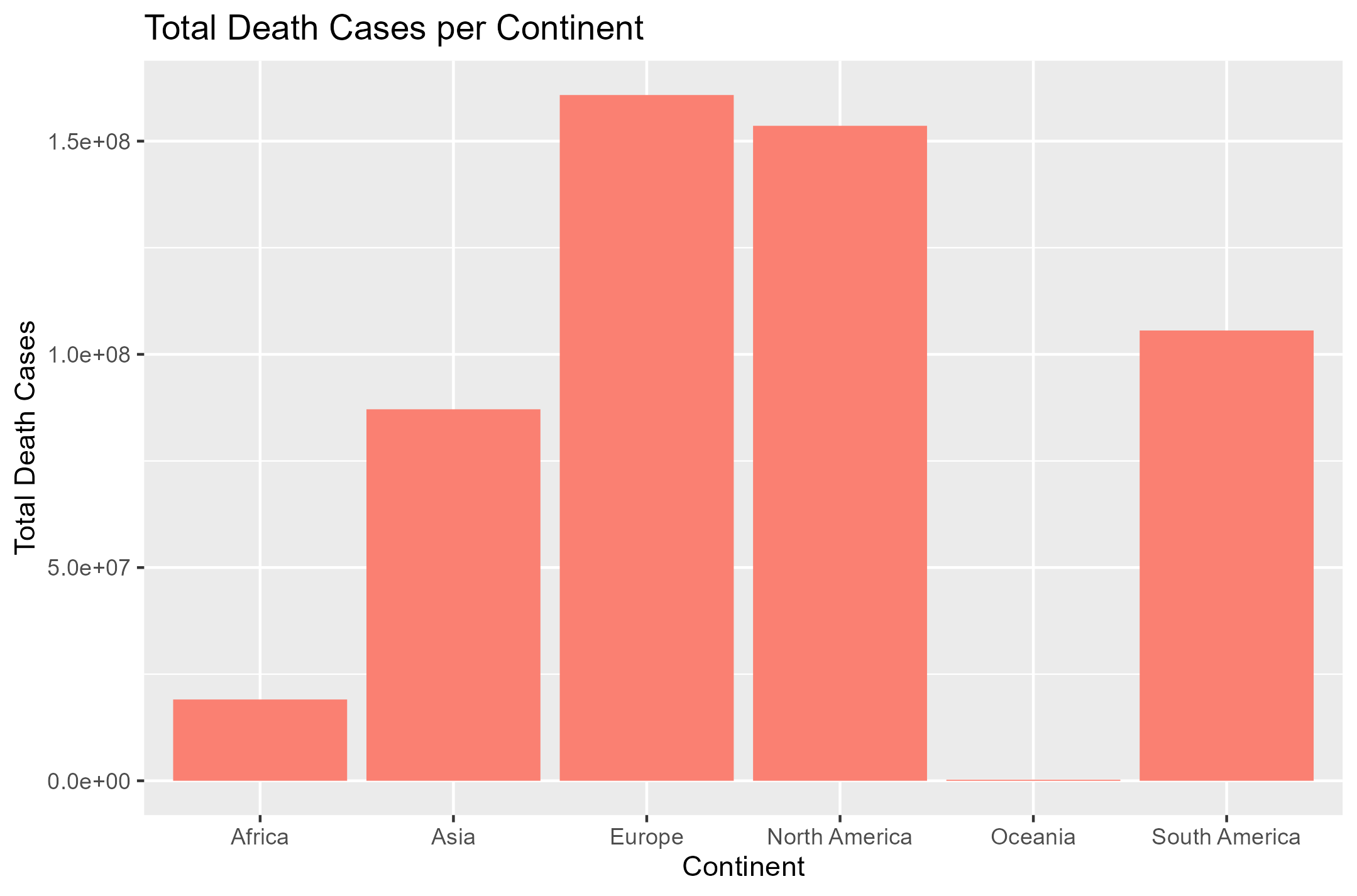
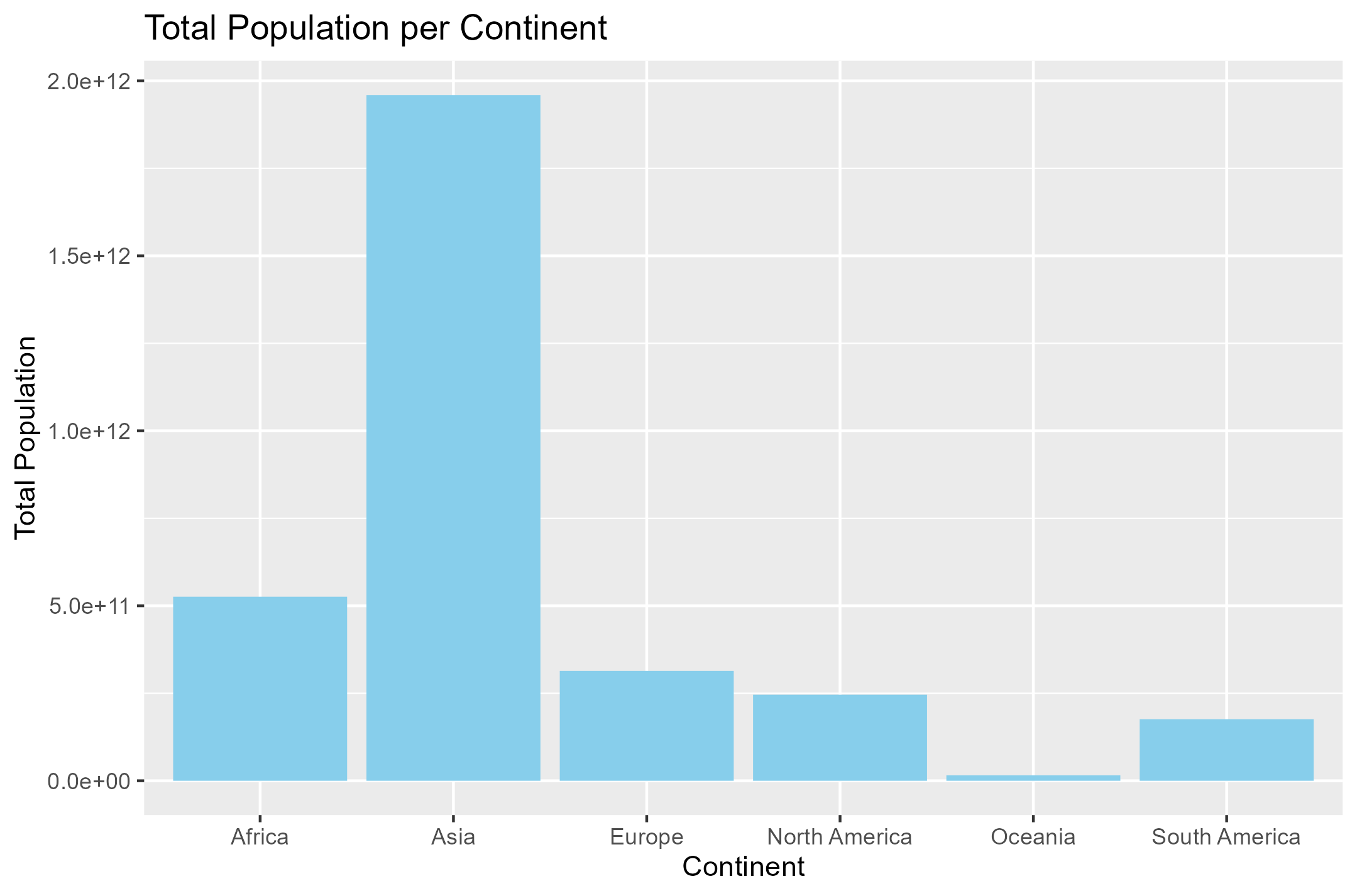
print(deaths\_table)

## # A tibble: 6 × 2  
## continent total\_deaths  
## <chr> <dbl>  
## 1 Africa 19064923  
## 2 Asia 87116891  
## 3 Europe 160813747  
## 4 North America 153596231  
## 5 Oceania 256504  
## 6 South America 105595385

print(cases\_table)

## # A tibble: 6 × 2  
## continent total\_cases  
## <chr> <dbl>  
## 1 Africa 755404142  
## 2 Asia 5244873657  
## 3 Europe 5919901332  
## 4 North America 5899927388  
## 5 Oceania 9844614  
## 6 South America 3726062857

## Visualizations

Total Cases and Death Cases by Continent 

## Conclusion

This analysis provides insights into the distribution of COVID-19 cases and death cases across different continents, highlighting variations in population, disease burden and death cases. From the statistical analysis, Asia has the largest population and Ocenia with the lowest population among the continents. Europe has the highest record of COVID\_19 cases and total death cases. It is the concluded from the analysis that Europe encountered the highest loss of lives and has the highest disease burden of COVID\_19.

In this R Markdown document:

* The setup code chunk loads necessary libraries and reads the COVID-19 dataset.
* Summary statistics for the dataset, total population per continent, and total cases and death cases per continent are provided.
* Visualizations are included to depict the total cases and death cases by continent.
* The conclusion summarizes the findings of the analysis.