Practical 1: US Population Time Series Analysis

Praveer Raj

2026-02-07

Practical 1: US Population Time Series Analysis

Name: Praveer Raj
Roll: 1
Reg. No.: 230957002

---

1. Title

For a given dataset of the US Population (in millions), perform the following tasks in the R environment:

(a) Enter the data into an Excel file.

(b) Import the data from the Excel file into R.

(c) Call a suitable package for time series analysis.

(d) Convert the dataset into a time series object and explain its sampling frequency.

(e) Plot the data and identify the dominating component.

Data:

3929214, 5308483, 7239881, 9638453, 12860702, 17063353,
23191876, 31443321, 38558371, 50189209, 62979666, 76212168,
92228496, 106021537, 123202624, 132164569, 151325798,
179323175, 203302031, 226542203, 248709873

---

2. Objective

To create an Excel dataset of US population, import it into R, convert it into a time series object, and analyze its trend using graphical representation.

---

3. R Code

```r
# Creating Excel file

Year <- seq(1790, 1990, by = 10)

US_Population <- c(
  3929214, 5308483, 7239881, 9638453, 12860702,
  17063353, 23191876, 31443321, 38558371, 50189209,
  62979666, 76212168, 92228496, 106021537, 123202624,
  132164569, 151325798, 179323175, 203302031,
  226542203, 248709873
)

population_df <- data.frame(
  Year = Year,
  Population = US_Population
)

write.xlsx(population_df, "US_Population.xlsx")

# Importing Excel file
dataset <- read_excel("US_Population.xlsx")

# Converting to time series
us_pop_ts <- ts(
  dataset$Population / 1e6,
  start = 1790,
  frequency = 0.1
)

# Formatting Y-axis labels
vals <- pretty(us_pop_ts)

# Plotting time series
plot(
  us_pop_ts,
  main = "US Population Growth (1790-1990)",
  xlab = "Year",
  ylab = "Population(in millions)",
  yaxt = "n",
  col = "blue",
  lwd = 2
)

axis(
  2,
  at = vals,
  labels = format(vals, big.mark = ",", scientific = FALSE)
)
```
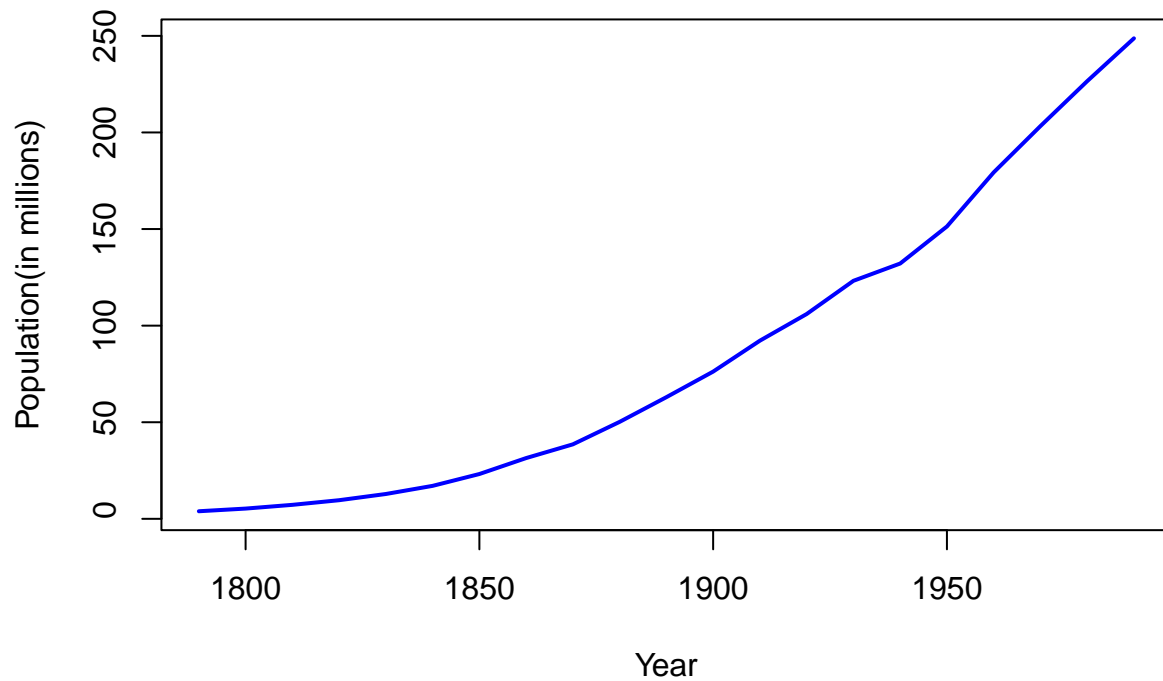
US Population Growth (1790–1990)

4. Output

The time series plot of US population is shown above.

5. Explanation

The dataset represents **decennial population data**, meaning the sampling frequency is once every 10 years. After converting the dataset into a time series object, the plot shows a steady upward movement.

6. Conclusion

The time series graph shows a continuous upward trend in US population from 1790 to 1990. The dominant component observed is the **trend component**, indicating steady long-term population growth.