ACF

Saurabh

2025-05-08

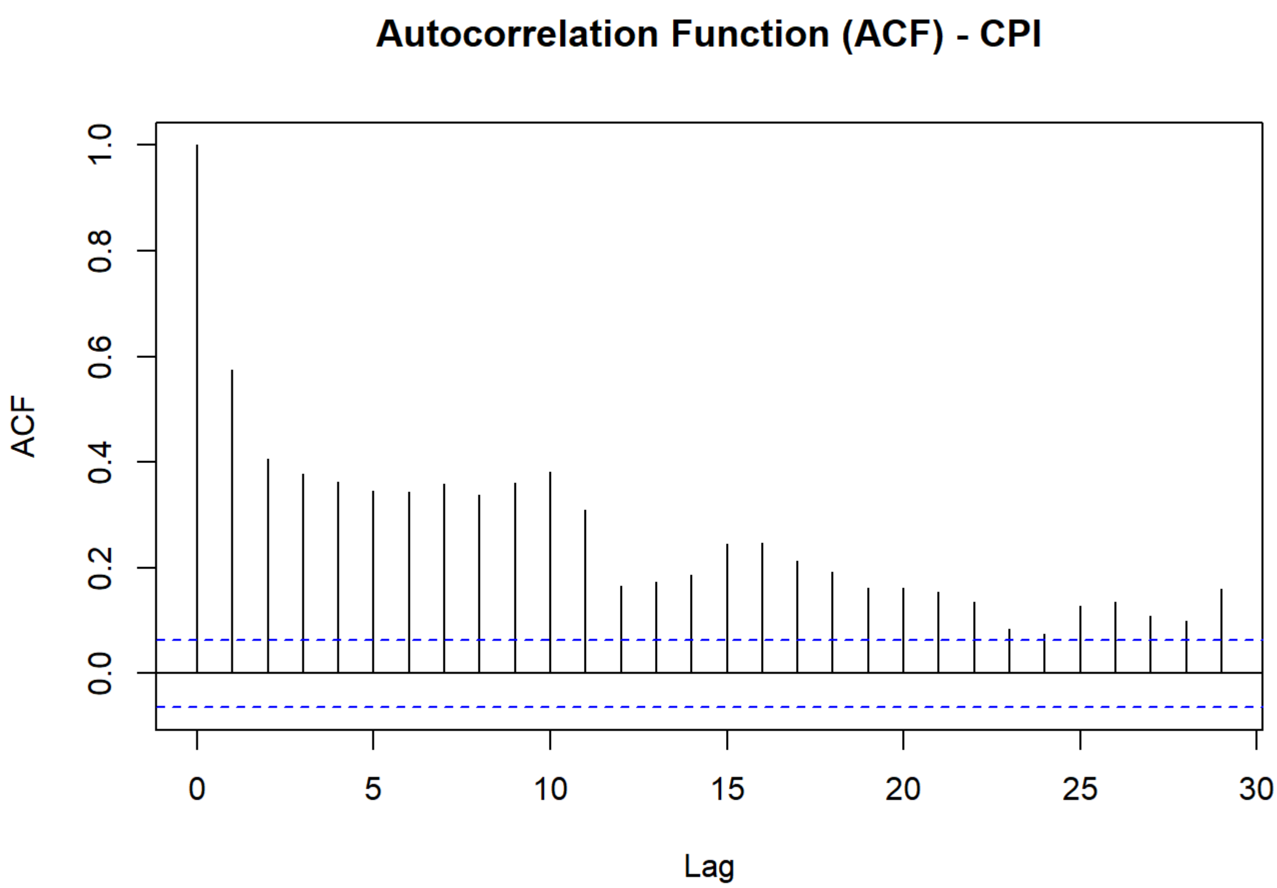
```
library(readr)
library(dplyr)
library(tidyr)
library(ggplot2)

cpi_data <- read_csv("No Header_ConsumerPriceIndex.csv")


cpi_long <- cpi_data |>
  pivot_longer(cols = -Year, names_to = "month", values_to = "cpi") |>
  rename(year = Year) |>
  mutate(date = as.Date(paste(year, month, "01", sep = "-"), format = "%Y-%B-%d")) |>
  arrange(date)


cpi_ts <- ts(na.omit(cpi_long$cpi))


acf(cpi_ts, main = "Autocorrelation Function (ACF) - CPI")
```



Ans:

📊 ACF Plot Explanation: “Autocorrelation of CPI” 🔹 Y-axis (ACF Values): This axis shows the autocorrelation coefficient, which ranges from -1 to 1. At lag 0, it’s always 1.0 — because any time series is perfectly correlated with itself at lag 0.

🔸 X-axis (Lag): Represents the number of time periods by which the CPI data is shifted (lagged). For example, lag 1 compares CPI values with those from the previous month, lag 2 with two months ago, and so on.

📈 Bars in the Plot: Each vertical line shows the strength of autocorrelation at a particular lag.

At lag 0, ACF = 1.0 (perfect correlation).

As the lags increase (lag 1, 2, …, 10), the autocorrelation gradually decreases.

Around lag 4, it drops below 0.4.

By lag 10, it’s still around 0.4, showing a slowly decreasing trend.

🔵 Blue Shaded Bands (Confidence Interval):

If an ACF spike goes beyond the blue bands, it’s statistically significant (not just random noise).

Since many spikes are outside the blue bands, this indicates strong autocorrelation — past CPI values significantly influence future ones.

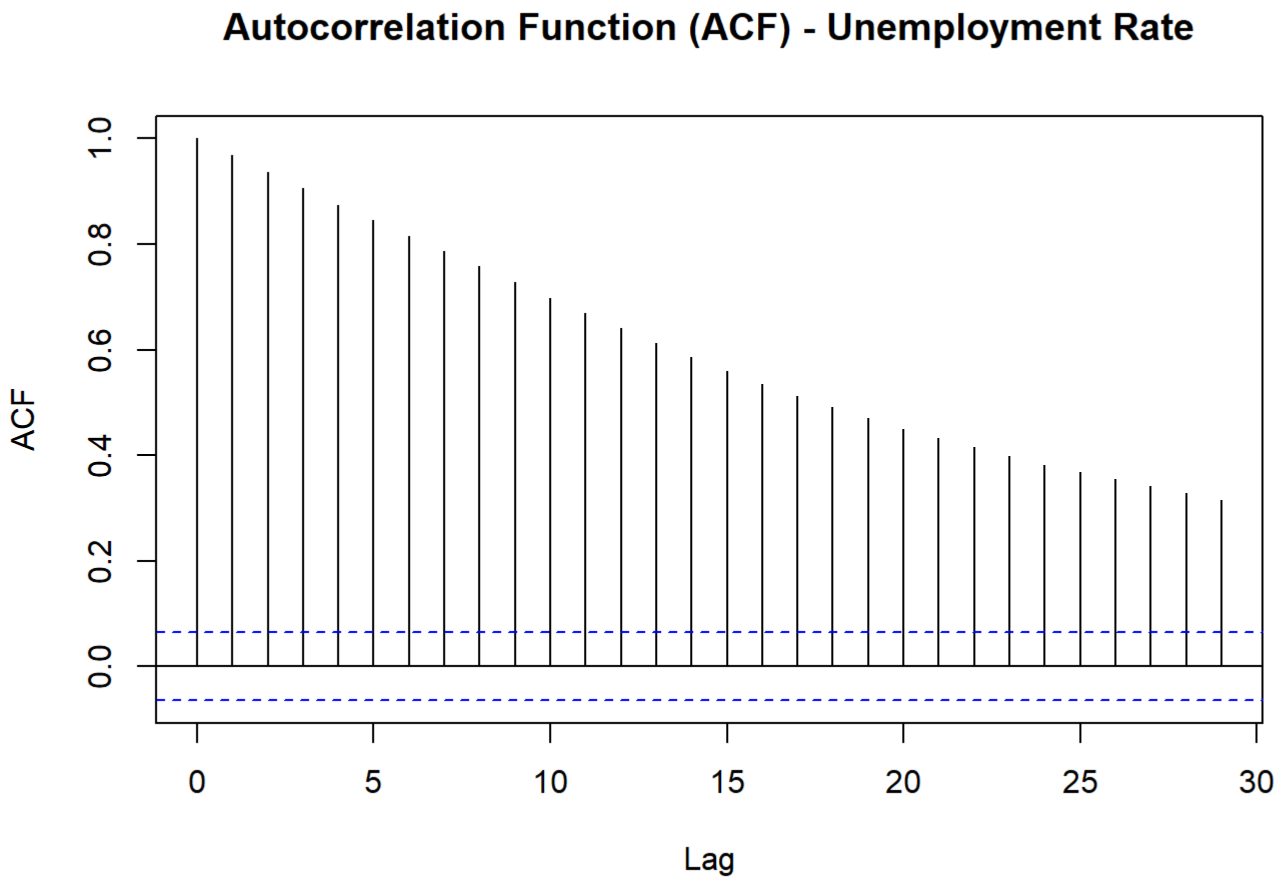
💡 Insight: The CPI time series is highly autocorrelated, especially at lower lags. This suggests that past inflation values (CPI) can help predict future values — a key insight for building ARIMA or other time series models.

```
unemp_data <- read_csv("NoHeader_UnemploymentRate.csv")


unemp_long <- unemp_data |>
  pivot_longer(cols = -Year, names_to = "month", values_to = "unemployment_rate") |>
  rename(year = Year) |>
  mutate(date = as.Date(paste(year, month, "01", sep = "-"), format = "%Y-%B-%d")) |>
  arrange(date)


unemp_ts <- ts(na.omit(unemp_long$unemployment_rate))

acf(unemp_ts, main = "Autocorrelation Function (ACF) - Unemployment Rate")
```



📉 ACF Plot Explanation: "Autocorrelation of Unemployment Rate" 🔹 Y-axis (ACF Values): Represents the autocorrelation coefficient, ranging from -1.0 to 1.0.

At lag 0, ACF = 1.0, meaning the series is perfectly correlated with itself.

🔸 X-axis (Lag): Shows the time lag from 0 to 30. Each point compares unemployment values with values from earlier months.

📈 Vertical Lines (ACF spikes):

All lines are outside the horizontal blue confidence bands, indicating statistically significant autocorrelation at every lag.

The ACF values start from 1.0 at lag 0 and slowly decrease with each increasing lag.

By lag 30, the ACF drops below 0.4, suggesting the correlation weakens but remains present even at longer lags.

🔵 Blue Horizontal Lines: These represent the 95% confidence interval. Spikes outside these lines show real autocorrelation, not random noise.

💡 Insight: The Unemployment Rate is highly autocorrelated, especially across short and medium lags. This means past unemployment levels have a strong influence on future values — useful for forecasting models like ARIMA or exponential smoothing.