Lab 2: Trend Seasonality and Sample Autocorrelation

Questions

Dataset "dataTempPG.csv"

1. **Features of the data**

Read the data into R using `read.csv()`.

```
df1 = read.csv("dataTempPG.csv")
```

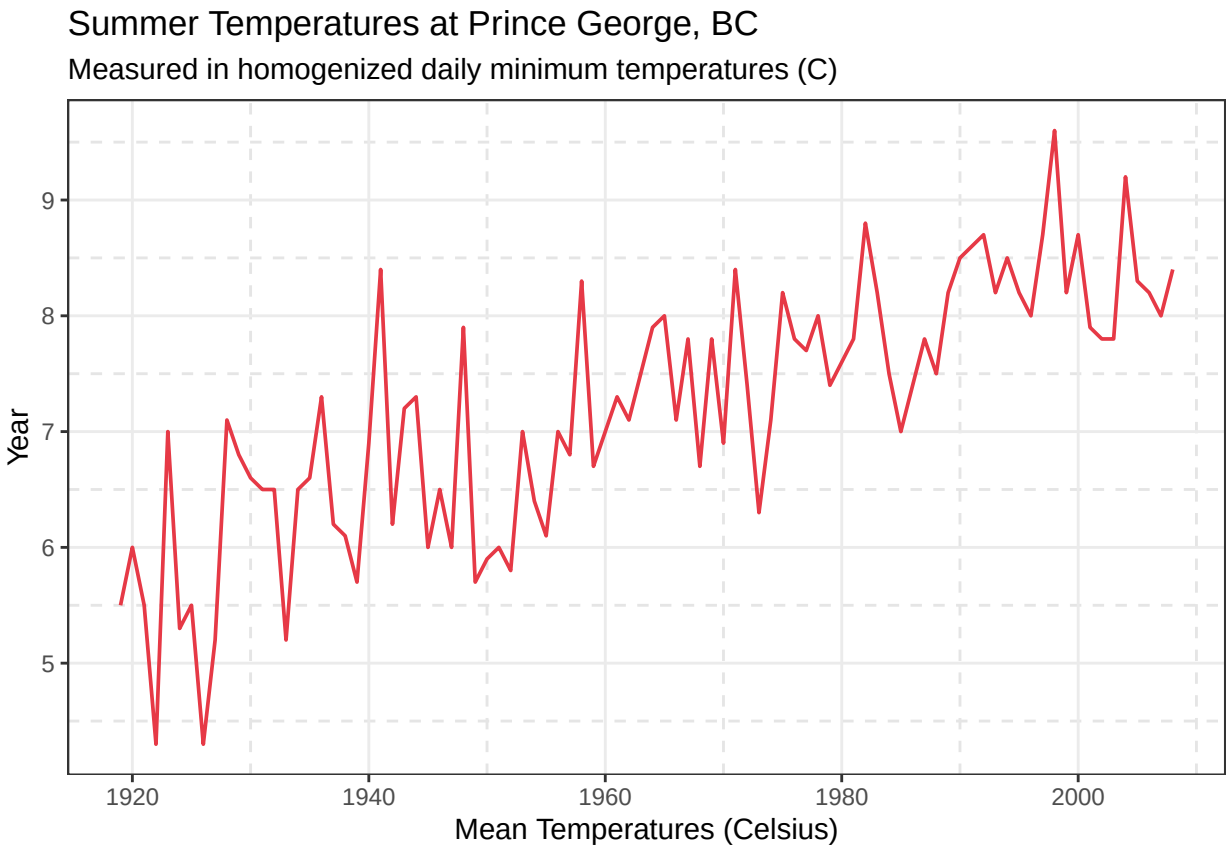
Extract the "Summer" column and convert it into a time series in R.

```
summer_data = df1$Summer

# the data are yearly starting in 1919
summer_series = ts(data = summer_data,
                   start = c(1919),
                   frequency = 1)
```

Plot the time series with appropriate labels.

```
p1data = fortify.zoo(summer_series)
p1 <- ggplot(p1data, aes(x = Index, y = summer_series)) +
  geom_line(color = "#E63946", linewidth = 0.65) +
  labs(
    title = "Summer Temperatures at Prince George, BC",
    subtitle = "Measured in homogenized daily minimum temperatures (C)",
    x = "Mean Temperatures (Celsius)",
    y = "Year"
  ) + theme_bw() +
  theme(panel.grid.minor = element_line(
    color = "grey90",
    linetype = "dashed",
    linewidth = 0.5
  ))
print(p1)
```



Comment on the main features of this series (e.g., seasonality, trend).

There doesn't appear to be much seasonality in the data. This makes sense, considering we are already isolating for a season (summer.) However, there does seem to be an increasing (positive) trend to the data.

2. **Sample Autocorrelation**

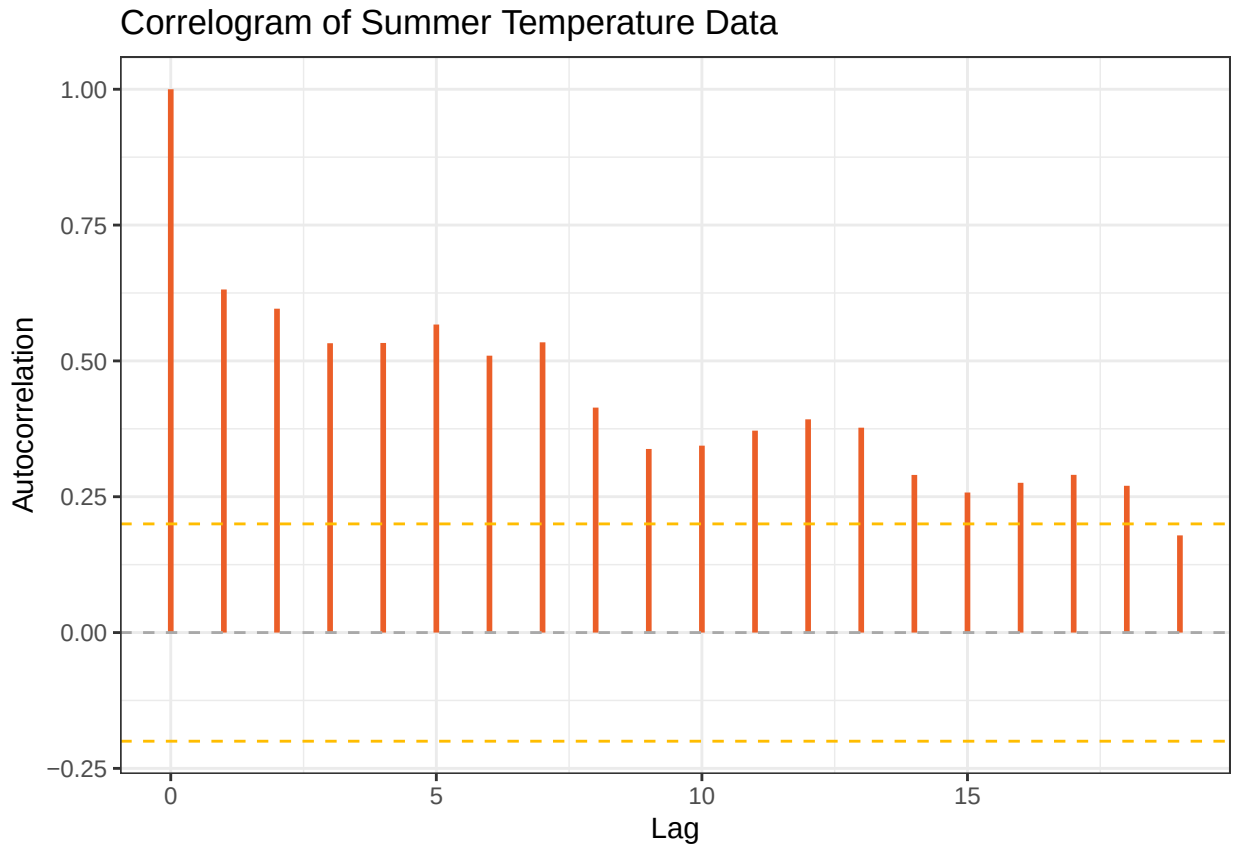
   - Use `acf()` to create the autocorrelation function for the mean summer temperature data.

here, we calculate and plot the ACF.

```
p2data = data.frame(
  h = 0:19,
  rh = acf(summer_series, plot = FALSE)$acf
)

p2 <- ggplot(p2data, aes(x = h, y = rh)) +
  geom_segment(aes(xend = h, yend = 0),
               color = "#eb5e28",
               linewidth = 1) +
  # geom_point(
  #   col = "#f6bd60",
  #   fill = "#f7ede2",
  #   stroke = 1,
  #   size = 3,
  #   shape = 21
  # ) +
```

```
  geom_hline(yintercept = 0.2, linetype = "dashed", col = "#ffbd00")+
  geom_hline(yintercept = -0.2, linetype = "dashed", col = "#ffbd00")+
  ylim(-0.2, 1)+
  geom_hline(yintercept = 0,
             linetype = "dashed",
             color = "darkgray") +
  labs(x = "Lag", y = "Autocorrelation",
       title = "Correlogram of Summer Temperature Data") +
  theme_bw()
print(p2)
```



Comment on the behavior of the sample autocorrelation function.

3. **Smoothing**

Use `window()` to extract the portion of the time series between 1968 and 2008.

Plot this recent record of the data.

Use `rollmean()` with a moving average parameter of k=5 to add the smoothed series to the plot.

Add a legend to your plot.

Dataset "LakeLevels.csv"

1. **Sample Autocorrelation Function**

- Read the data into R and convert it into a time series.
- Use `acf()` to create the sample autocorrelation function for the lake levels.
- Comment on the behavior of this autocorrelation function.

2. **Decomposition (Additive Model)**

   - Use `decompose()` to decompose the time series into trend, seasonal component, and error (additive model).
   - Plot the trend, seasonal component, and error.

3. **Decomposition (Loess Method)**

   - Use `stl()` with `s.window` set to "periodic" for decomposition.
   - Plot the trend, seasonal component, and error using the loess method.