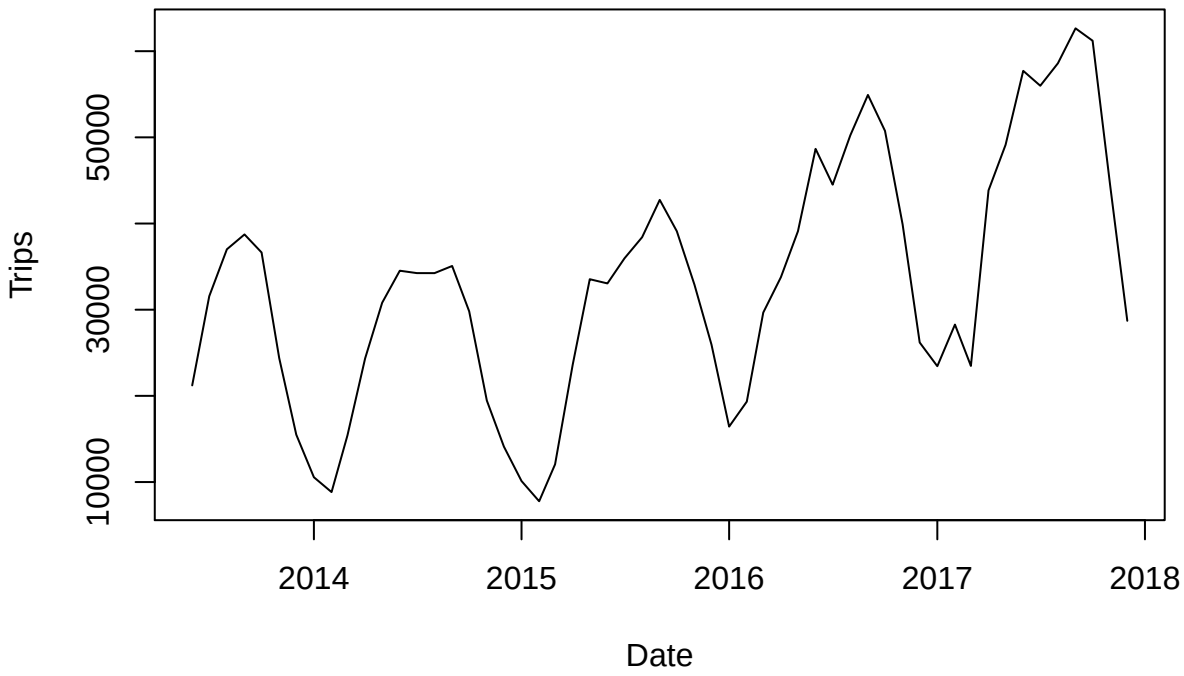
R Notebook

```
library("forecast")
```

```
data <- read.csv("/Users/chuamelia/Google Drive/Forecasting Time Series/citi-bike/aggTripsMonthly.csv")
date <- as.Date(data$MonthYear)
time <- 1:length(date)
trips <- data$woSeasonal
```

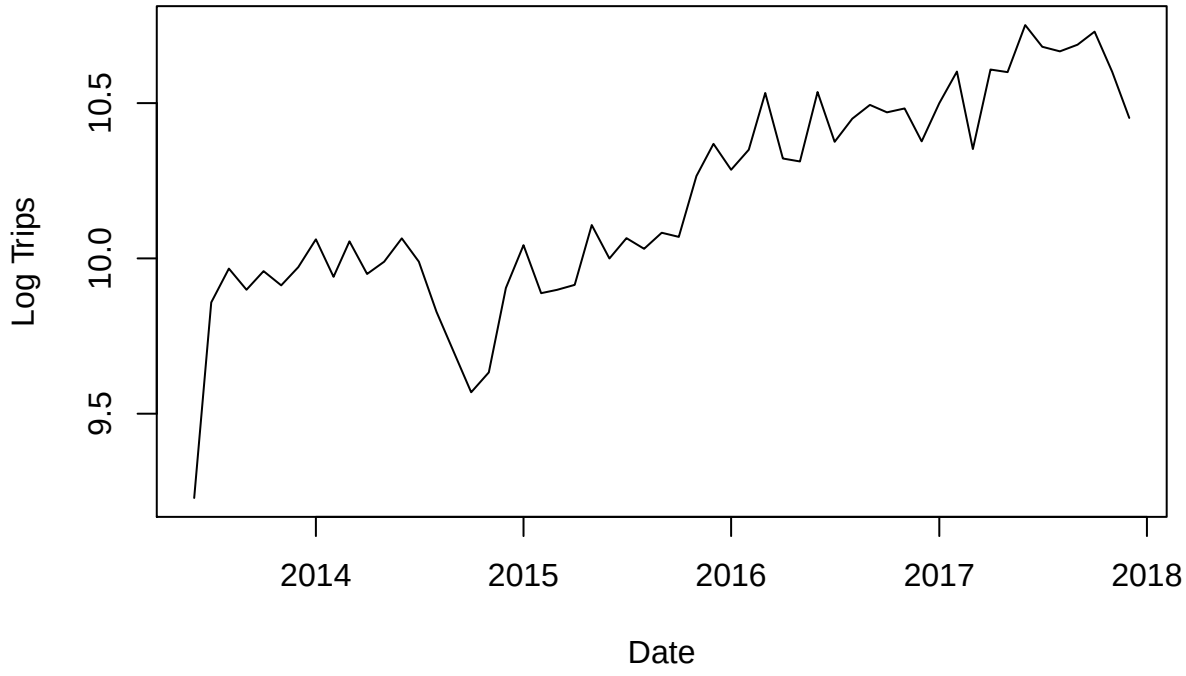
```
# Time series plot
plot(date, data$Trips, type="l",
     xlab="Date", ylab="Trips")
```



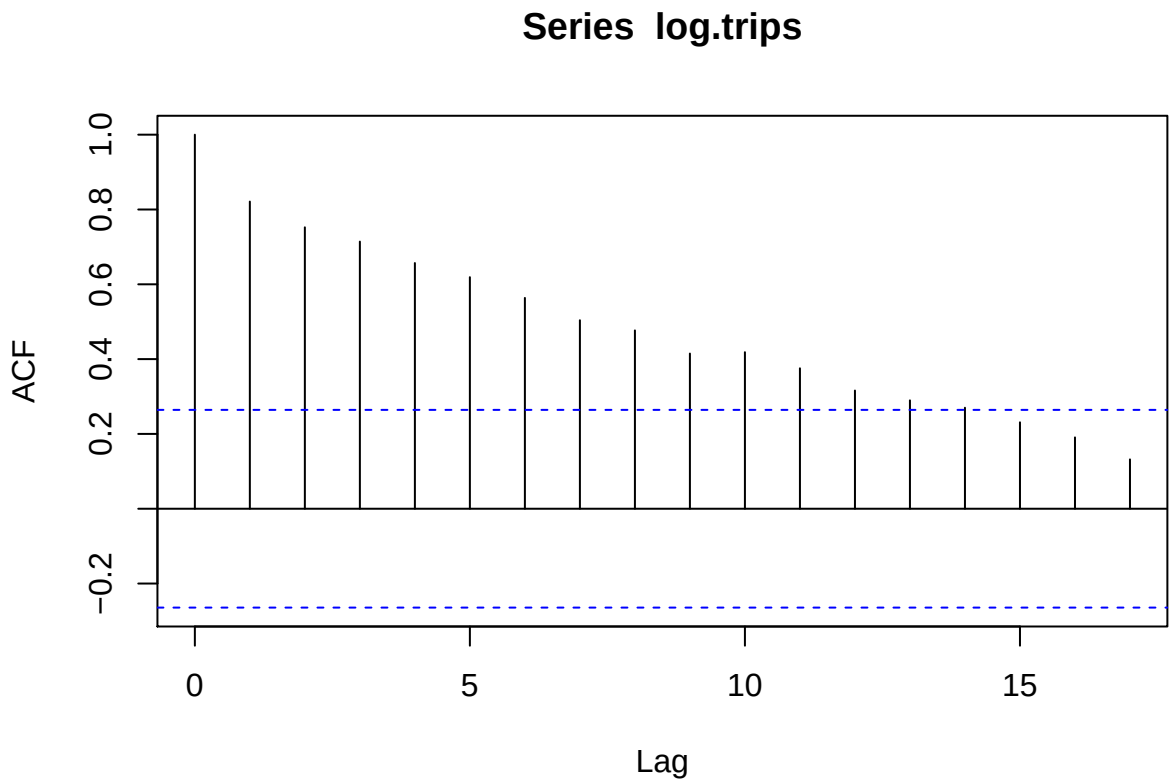

```
log.trips <- log(trips)
diff.log.trips <- c(NA, diff(log.trips))
diff2.log.trips <- c(NA, diff(diff.log.trips))
```

```
# Time series plot
plot(date, trips, type="l",
     xlab="Date", ylab="Trips w/o Seasonal")
```

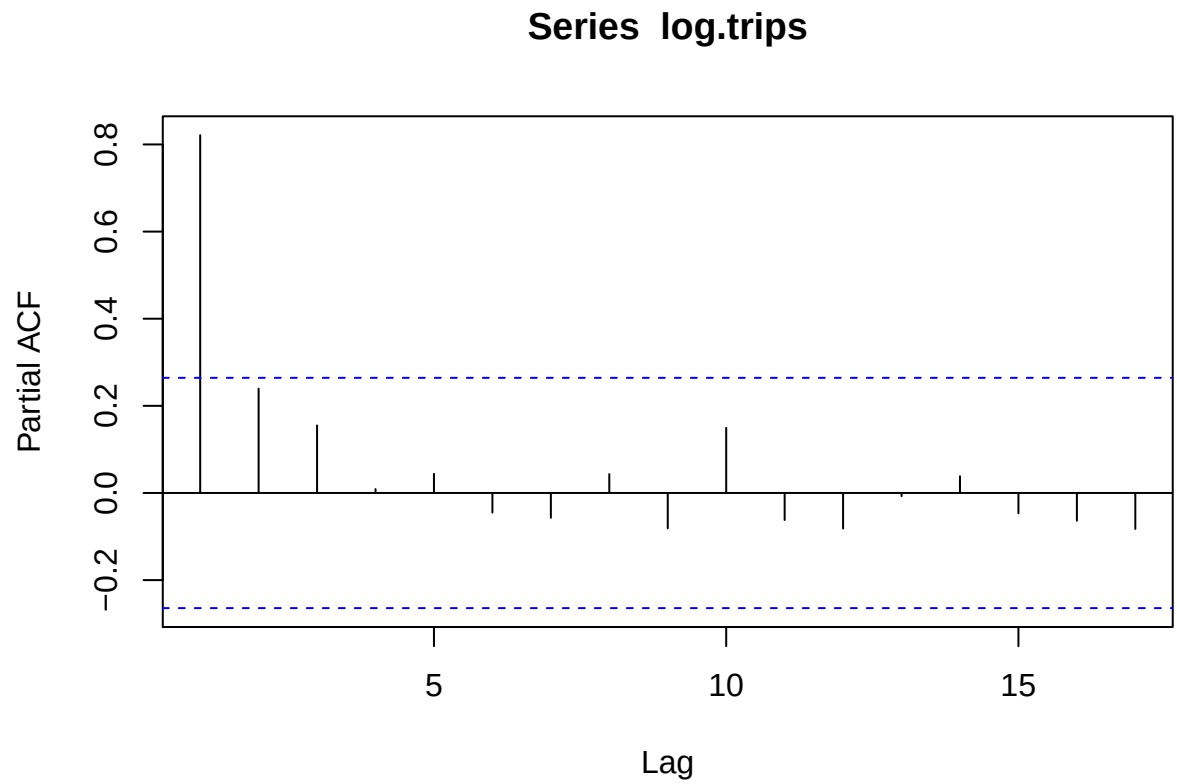

```
# Time series plot
plot(as.Date(data$MonthYear), log.trips, type="l",
     xlab="Date", ylab="Log Trips")
```



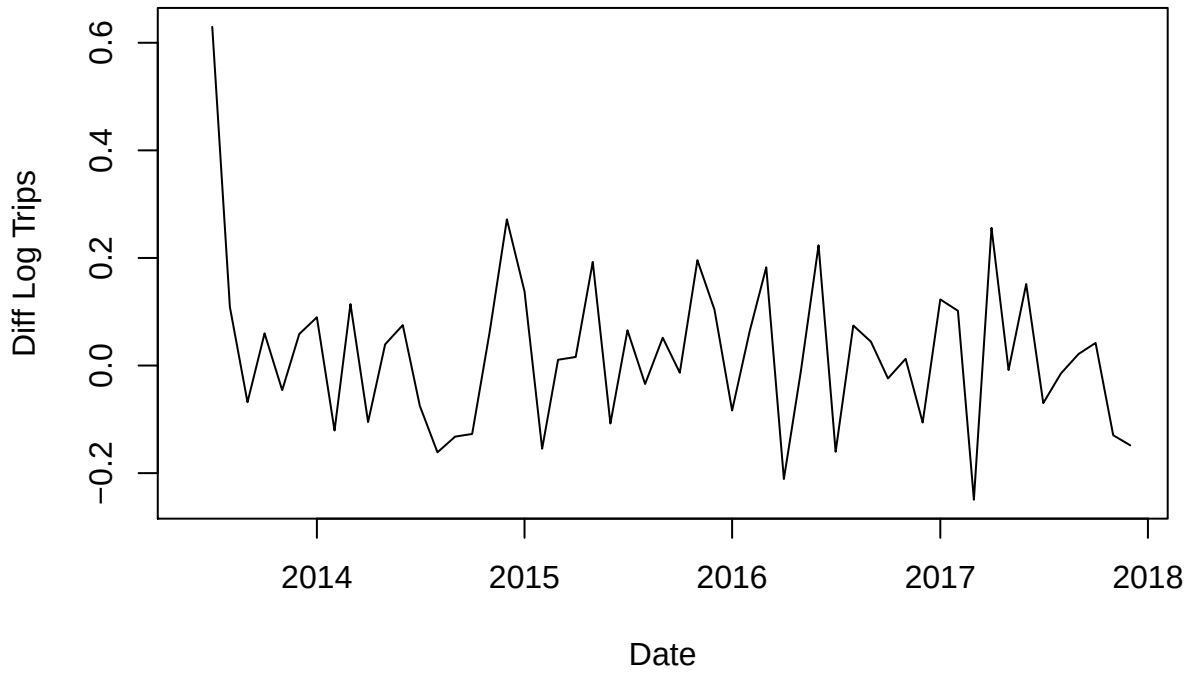

```
# ACF and PACF
acf(log.trips, na.action = na.pass)
```



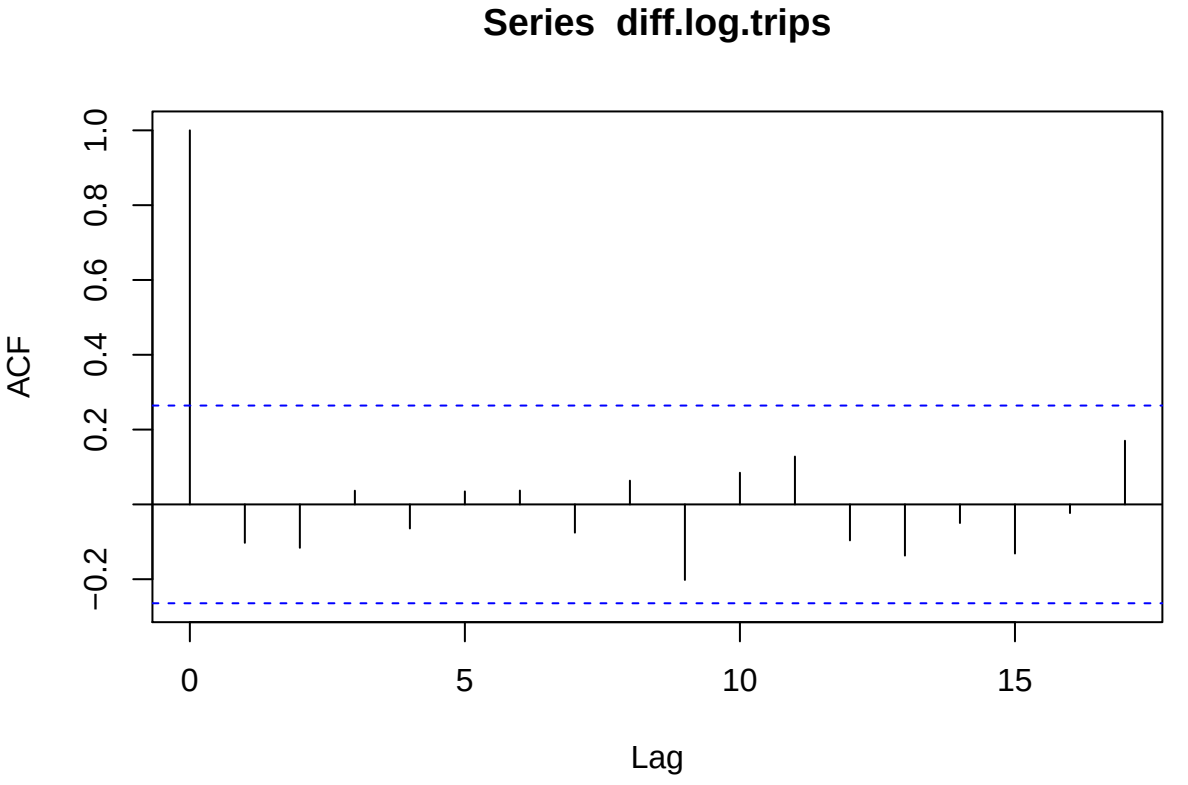

```
pacf(log.trips, na.action = na.pass)
```



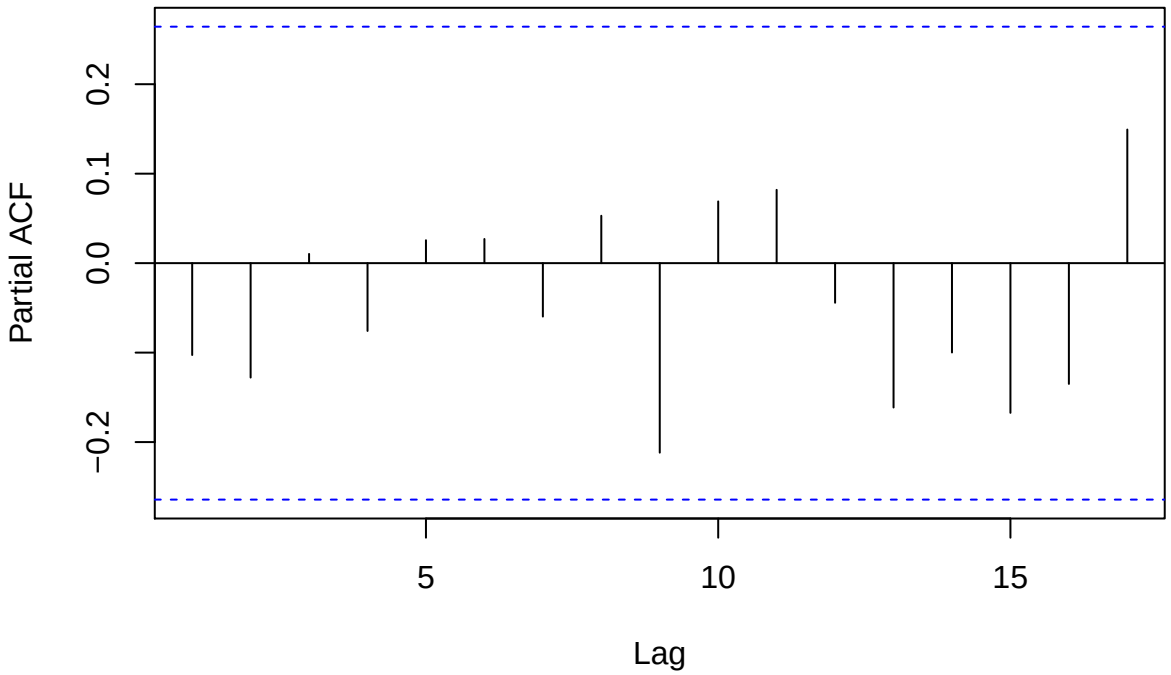

```
# Time series plot
plot(as.Date(data$MonthYear), diff.log.trips, type="l",
     xlab="Date", ylab="Diff Log Trips")
```




```
# ACF and PACF
acf(diff.log.trips, na.action = na.pass)
```




```
pacf(diff.log.trips, na.action = na.pass)
```




```
Arima(log.trips, c(0, 0, 0), include.constant=FALSE)$aicc
```

```
## [1] 413.5681
```

```
#Arima(log.trips, c(1, 0, 0), include.constant=FALSE)$aicc
Arima(log.trips, c(1, 0, 0), include.constant=TRUE)$aicc
```

```
## [1] -46.71975
```

```
Arima(log.trips, c(1, 1, 0), include.constant=TRUE)$aicc
```

```
## [1] -49.11246
```

```
Arima(log.trips, c(2, 1, 0), include.constant=TRUE)$aicc
```

```
## [1] -48.63872
```

```
Arima(log.trips, c(1, 2, 0), include.constant=TRUE)$aicc
```

```
## [1] -28.3975
```

```
Arima(log.trips, c(1, 1, 1), include.constant=TRUE)$aicc
```

```
## [1] -52.05159
```

```
Arima(log.trips, c(1, 1, 2), include.constant=TRUE)$aicc
```

```
## [1] -49.63884
```

```
fit <- Arima(log.trips, c(1, 1, 1), include.constant=TRUE)
print(fit)
```

```
## Series: log.trips
## ARIMA(1,1,1) with drift
```

```
##
## Coefficients:
##          ar1      ma1   drift
##       0.6491  -1.0000  0.0191
## s.e.  0.1290   0.0513  0.0030
##
## sigma^2 estimated as 0.01917:  log likelihood=30.43
## AIC=-52.87   AICc=-52.05   BIC=-44.91
```

```
# Hint: use Box.test function.  You must call this function four times,
# for lags 12, 24, 36, and 48.
Box.test(log.trips, lag = 12, type = c("Box-Pierce", "Ljung-Box"), fitdf = 0)
```

```
##
##  Box-Pierce test
##
## data:  log.trips
## X-squared = 217.4, df = 12, p-value < 2.2e-16
```

```
Box.test(log.trips, lag = 24, type = c("Box-Pierce", "Ljung-Box"), fitdf = 0)
```

```
##
##  Box-Pierce test
##
## data:  log.trips
## X-squared = 237.7, df = 24, p-value < 2.2e-16
```

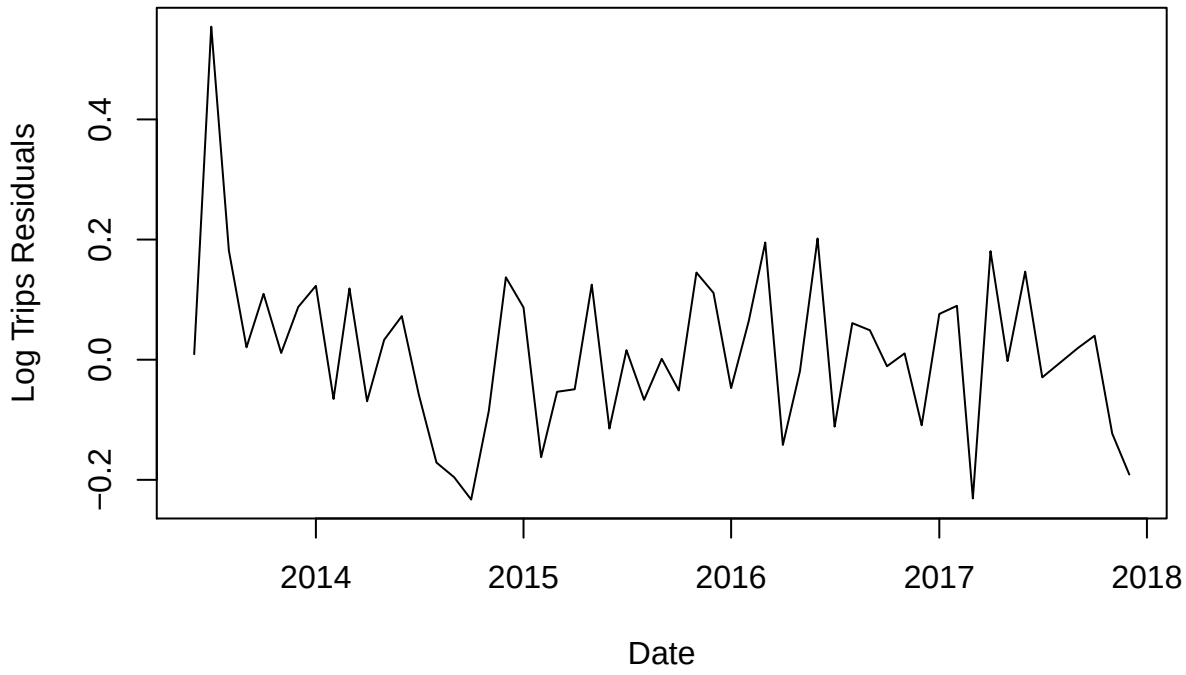
```
Box.test(log.trips, lag = 36, type = c("Box-Pierce", "Ljung-Box"), fitdf = 0)
```

```
##
##  Box-Pierce test
##
## data:  log.trips
## X-squared = 317.88, df = 36, p-value < 2.2e-16
```

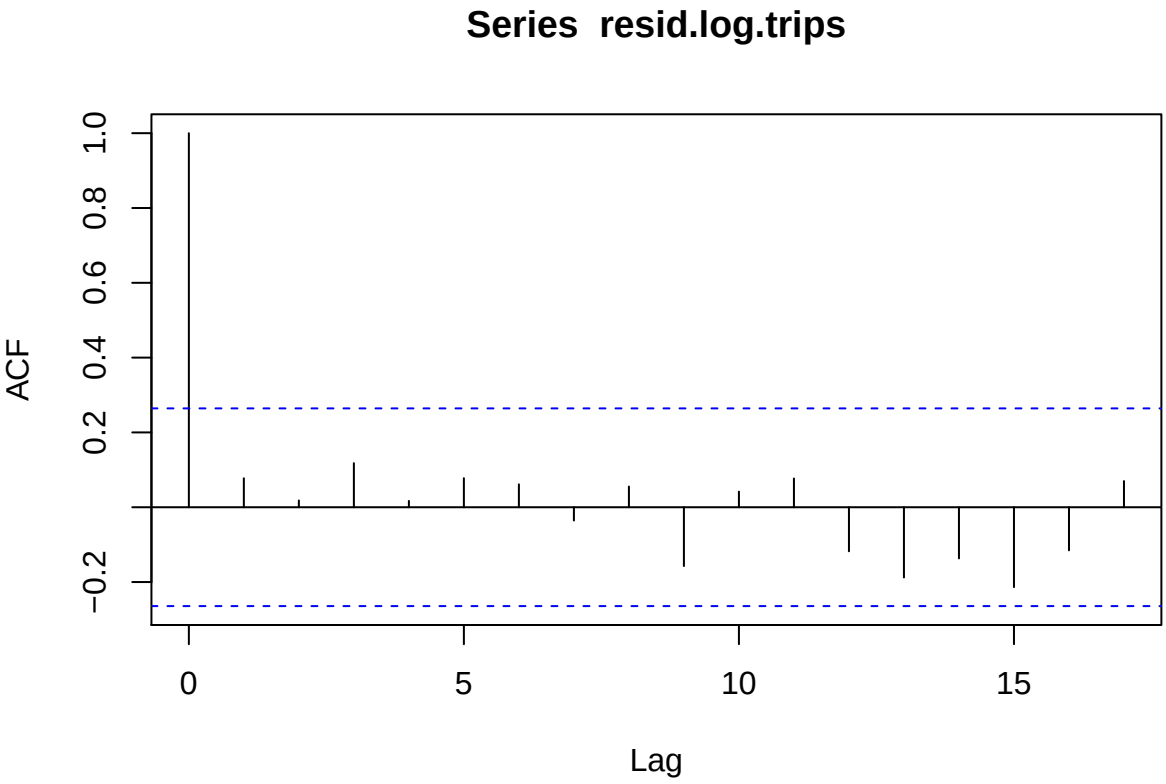
```
Box.test(log.trips, lag = 48, type = c("Box-Pierce", "Ljung-Box"), fitdf = 0)
```

```
##
##  Box-Pierce test
##
## data:  log.trips
## X-squared = 355.33, df = 48, p-value < 2.2e-16
```

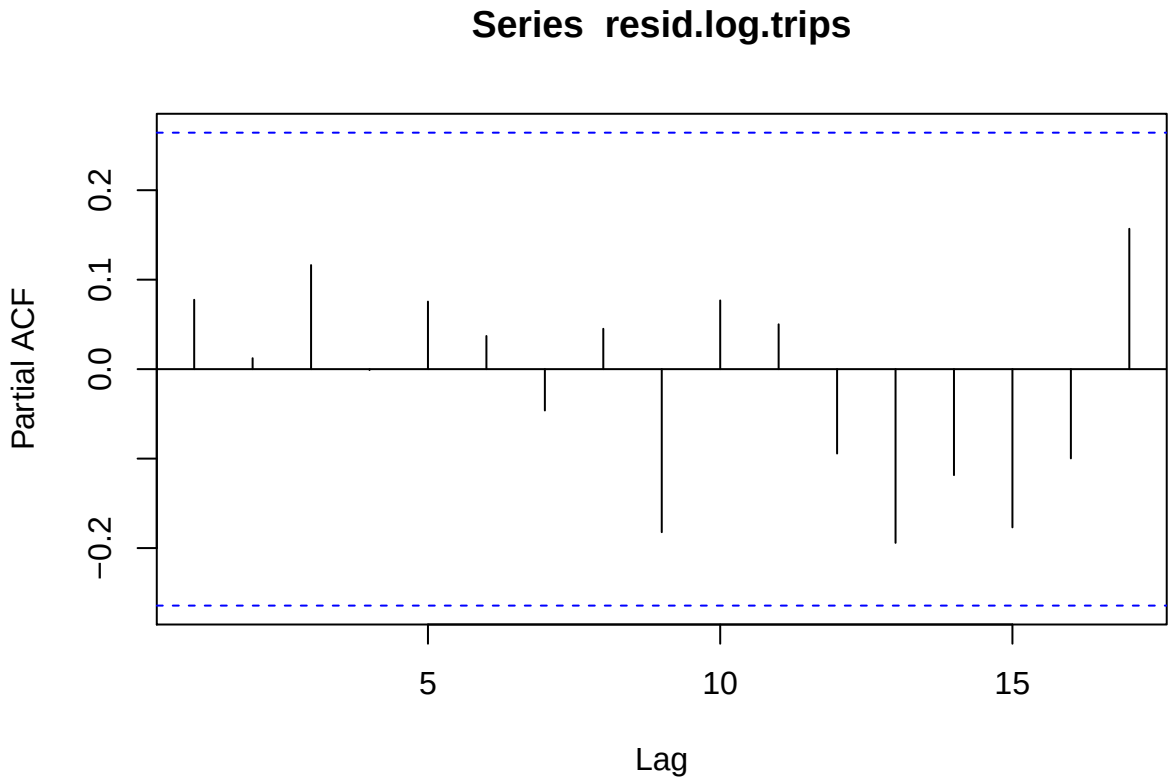
```
# Add TS plot, ACF, and PACF
resid.log.trips <- residuals(fit)
# Time series plot
plot(as.Date(data$MonthYear), resid.log.trips, type="l",
     xlab="Date", ylab="Log Trips Residuals")
```




```
# ACF and PACF
acf(resid.log.trips, na.action = na.pass)
```




```
pacf(resid.log.trips, na.action = na.pass)
```




```
plot(forecast(fit, h=30, level=95), col=2)
```