Take Home Quiz 1

Solutions

Due Tuesday Feb. 18 at 11:00am

This quiz should take you approximately 35 minutes. Place your answers into this markdown document, knit it, and hand in the result as a PDF or Word document. You may use R, any reference material, and information already available on the internet. Do not work together and do not get help from other people or from AI. If you have questions, ask Dr. Clair.

```
knitr::opts_chunk$set(fig.width=5, fig.height=3)
library(fpp3)
```

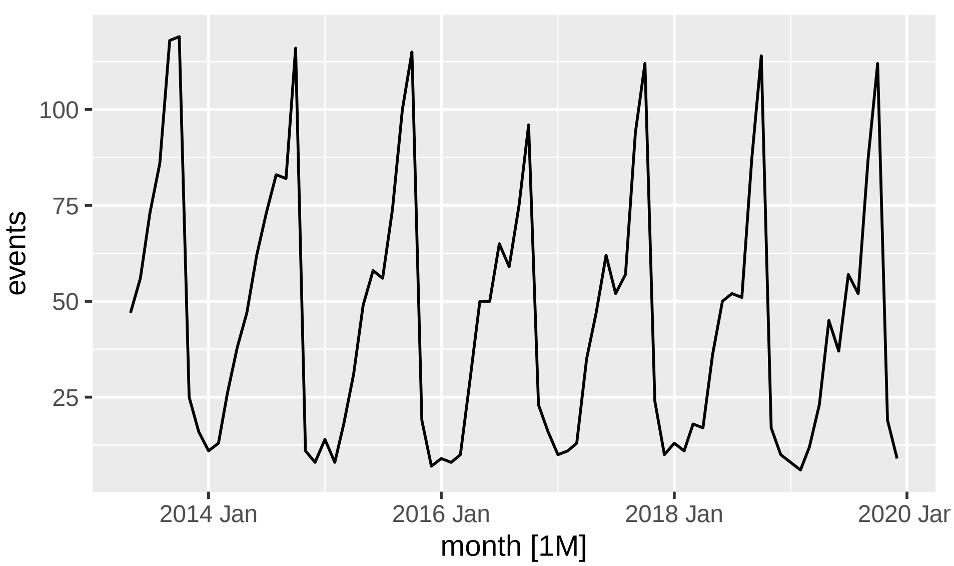
The first three problems use the data `stl-event-permits.csv` which contains monthly counts of event or party permits issued by the St. Louis Streets Department from 2013-2019. It is available on our website at https://turtlegraphics.org/timeseries/data/stl-event-permits.csv

Problem 1 (10 points)

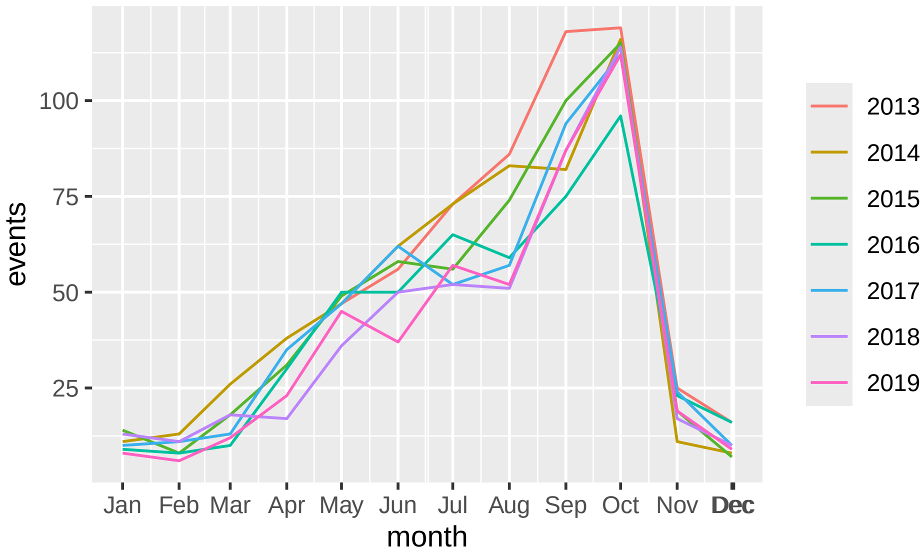
a. Load the `stl-event-permits.csv` data, convert to a tsibble, and produce a time plot of the number of events.
b. Produce a seasonal plot of the number of events. What is the peak month for street parties in St. Louis?

Solution

```
se_raw <- read.csv("https://turtlegraphics.org/timeseries/data/stl-event-permits.csv")
stlevents <- se_raw |> mutate(month = make_yearmonth(year,month)) |> select(-year) |>
  as_tsibble(index = month)
stlevents |> autoplot(events)
```



```r
stlevents |> gg_season(events)
```



October is the peak month for street events in St. Louis.

Problem 2 (10 points)

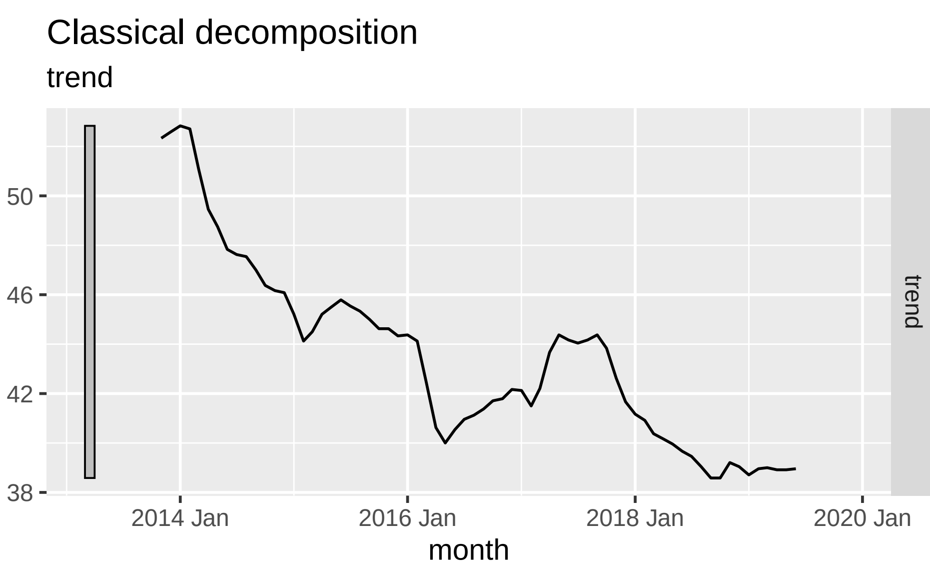
a. Perform a classical seasonal decomposition of the number of events.

b. Plot the trend component. Does there appear to be a trend over time?

Solution

```r
stl_cd <- stlevents |> model(classical_decomposition(events))
stl_cd |> components() |> autoplot(trend)
```

```
## Warning: Removed 12 rows containing missing values or values outside the scale range
## (`geom_line()`).
```



Definitely a slight downward trend over the seven years of data.

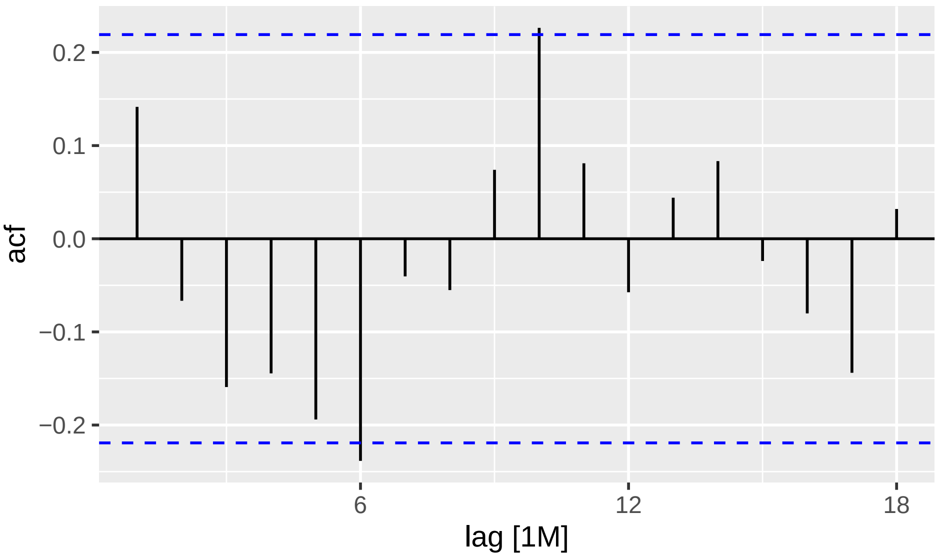
Problem 3 (10 points)

a. Create an autocorrelogram of the random components of the seasonal decomposition from problem 2.

b. Does the random component appear to be white noise? Check with a Ljung-Box test using 18 lags.

Solution

```
stl_cd |> components() |> ACF(random) |> autoplot()
```




```
stl_cd |> components() |> features(random, ljung_box, lag=18)
```

```
## # A tibble: 1 x 3
##   .model                          lb_stat lb_pvalue
##   <chr>                             <dbl>     <dbl>
## 1 classical_decomposition(events)    21.7     0.248
```

The random component is not significantly different from white noise.

Problem 4 (10 points)

Consider the time series 1, 2, 3, 1, 2, 3, 1, 2, 3, 1, 2, 3...

a. Calculate the 2-moving average.

b. Calculate the 3-moving average.

Solution

a. 1.5, 2.5, 2, 1.5, 2.5, 2, 1.5, 2.5, 2, ...
b. 2, 2, 2, 2, 2, 2, 2, ...

Problem 5 (10 ponts)

The series `souvenirs` from `fpp3` is heteroscedastic. Explain what this means, and how you might stabilize the variance of the series.

Solution

The variance of the series is not constant. For this series, the variance appears to grow over time, along with the size of the series itself. You could stabilize it with a log transform or a Box-Cox transform.