pre-covid model

We will build a model for emergencies appointements for both Vila do Conde and Barcelos. We will then use this model to replace the covid period for our future analysis.

```
df <- read_csv("../data/transformed/tratadas_emergencias.csv")
```

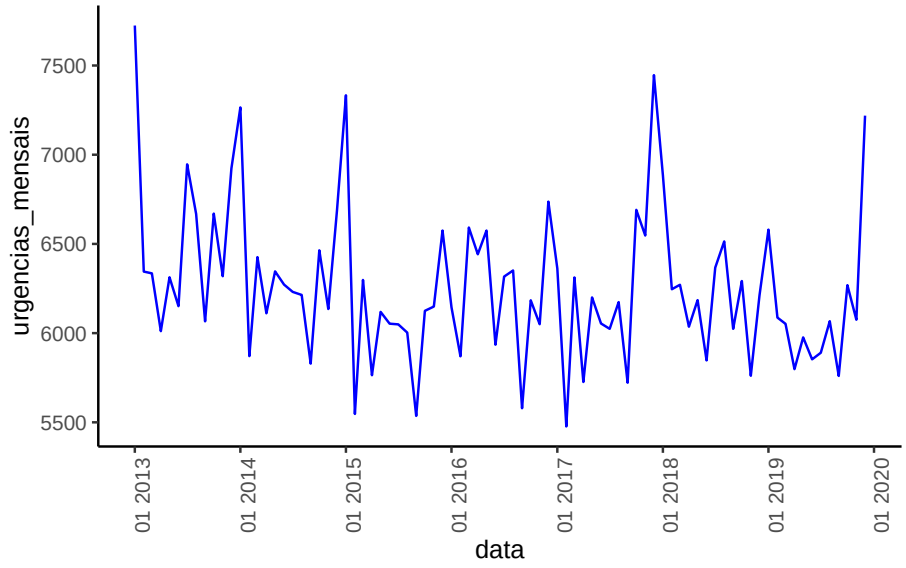
```
Rows: 282 Columns: 6
-- Column specification --------------------------------------------------------
Delimiter: ","
chr  (2): instituicao, unidade_saude
dbl  (3): urgencias_geral, year, urgencias_mensais
date (1): data

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
df <- df %>% filter(data < "2020-01-01")

df <- df %>%
    group_by(unidade_saude) %>%
    nest()
```

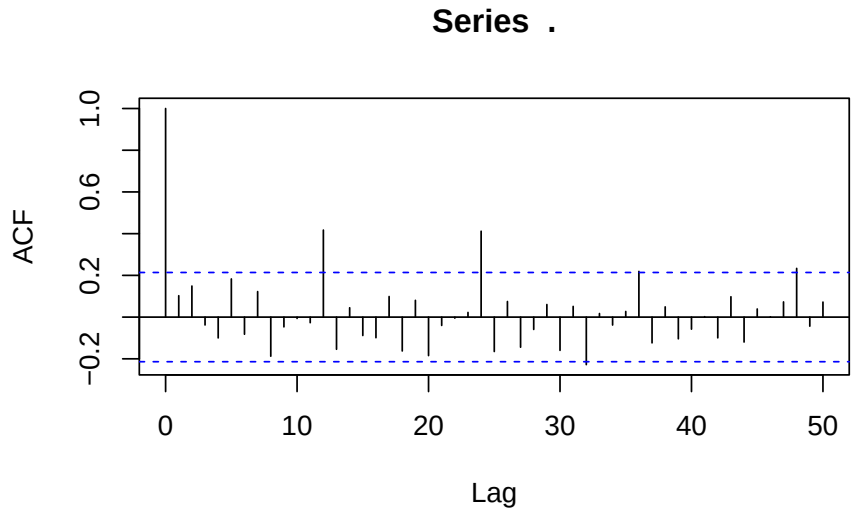
0.1. Vila do Conde

```
df %>%
    filter(unidade_saude == "Vila do Conde") %>%
    unnest(cols = c(data)) %>%
    ggplot(aes(x = data, y = urgencias_mensais )) +
    geom_line(color = "blue") +
    theme_classic() +
    scale_x_date(date_breaks = "1 year", date_labels = "%m %Y") +
    theme(axis.text.x = element_text(angle = 90, hjust = 1))
```

*Corresponding author




```
df %>%
    filter(unidade_saude == "Vila do Conde") %>%
    unnest(cols = c(data)) %>%
    pull(urgencias_mensais) %>%
    acf(lag.max = 50)
```



Initial visualizations show some signs of a seasonality with a regular high around te winter periods. It does show a sligth quadratic trend but is not very

pronounced from this visualization, it shows high levels of variability therefore we will consider the log of the series.

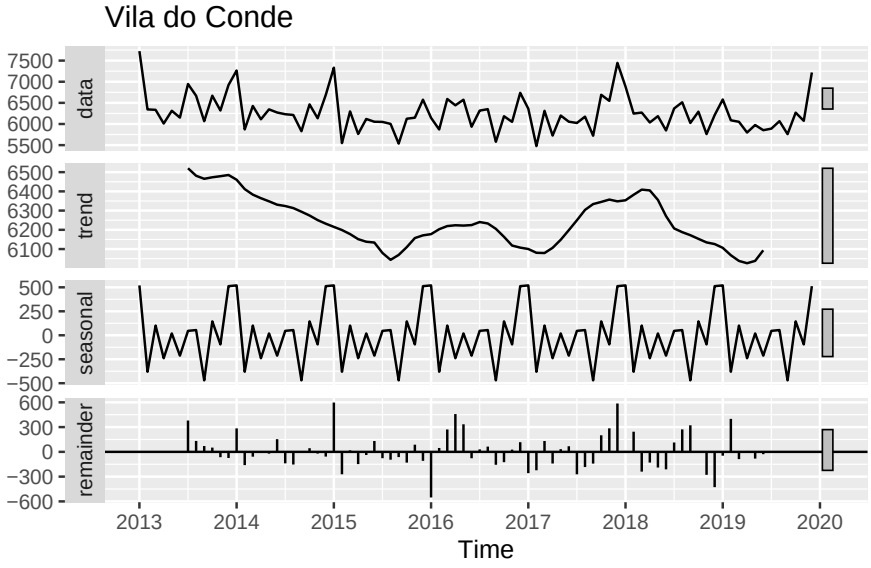
```
# Create time series for each serires

df <- df %>%
    mutate(
        data.ts = map(
            .x = data,
            .f = tk_ts,
            select   = c(-data, -year, -urgencias_geral),
            start = 2013,
            freq = 12
        )
    )
```

```
Warning: There were 2 warnings in `mutate()`.
The first warning was:
i In argument: `data.ts = map(...)`.
i In group 1: `unidade_saude = "Barcelos"`.
Caused by warning:
! Non-numeric columns being dropped: instituicao
i Run `dplyr::last_dplyr_warnings()` to see the 1 remaining warning.
```

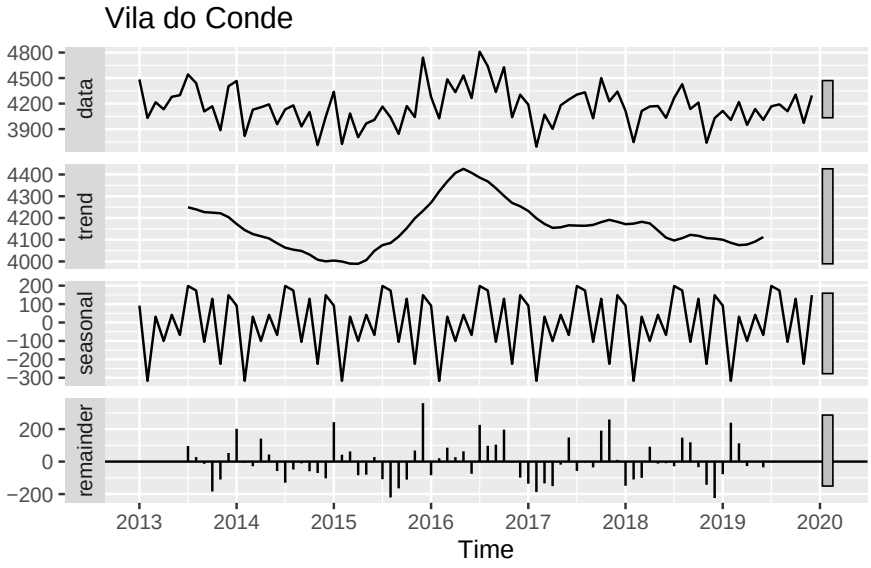
```
df
```

```
# A tibble: 2 x 3
# Groups:   unidade_saude [2]
  unidade_saude data              data.ts
  <chr>         <list>            <list>
1 Barcelos      <tibble [84 x 5]> <ts [84 x 1]>
2 Vila do Conde <tibble [84 x 5]> <ts [84 x 1]>
```

```
df %>%
    filter(unidade_saude == "Vila do Conde") %>%
    pull(data.ts) %>%
    .[[1]] %>%
    decompose(type = "additive") %>%
    autoplot() +
    ggtitle("Vila do Conde")
```

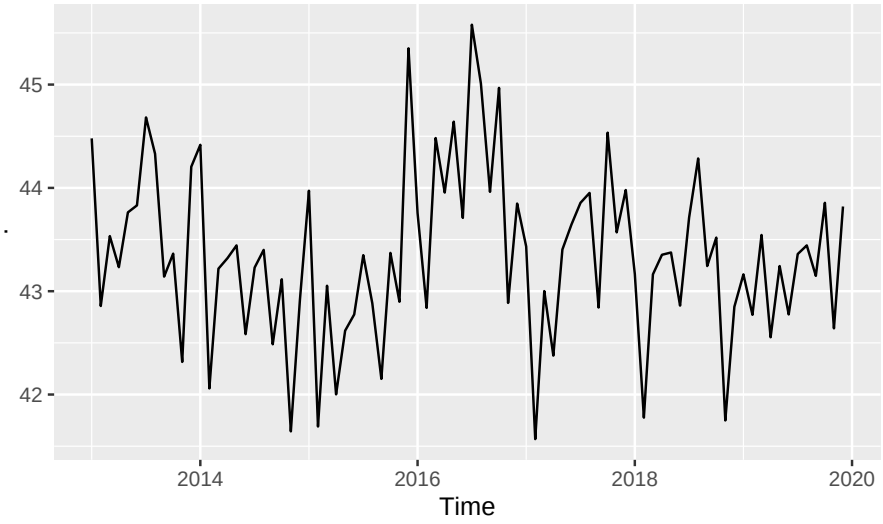



```
df %>%
    filter(unidade_saude == "Barcelos") %>%
    pull(data.ts) %>%
    .[[1]] %>%
    decompose(type = "additive") %>%
    autoplot() +
    ggtitle("Vila do Conde")
```



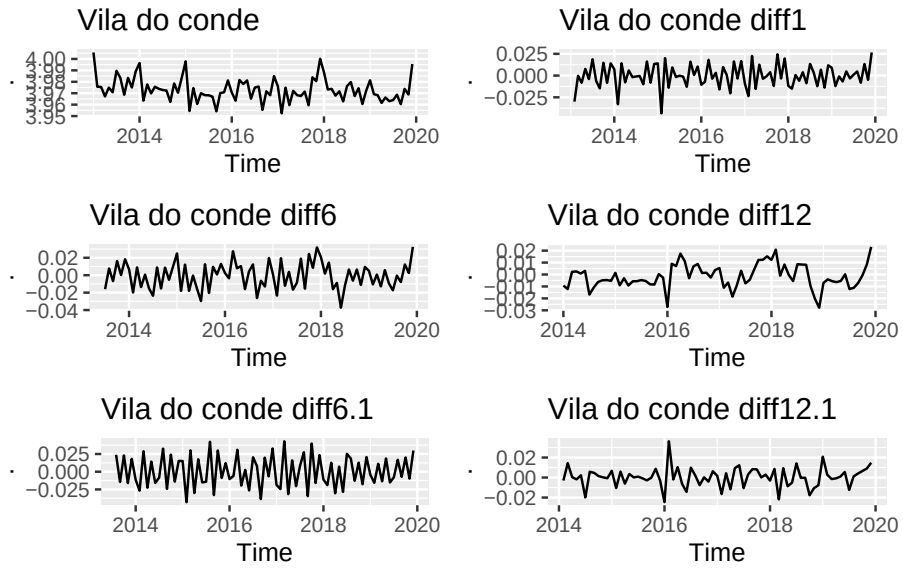
```
df_transform <- df %>%
    mutate(
        boxcox.lambda = map(data.ts, BoxCox.lambda),
        data.ts.box = map2(
            .x = data.ts,
            .y = boxcox.lambda,
            .f = ~BoxCox(.x, .y))
        )
```

```
df_transform  %>%
    filter(unidade_saude == "Barcelos") %>%
    pull(data.ts.box) %>%
    .[[1]] %>%
    autoplot()
```

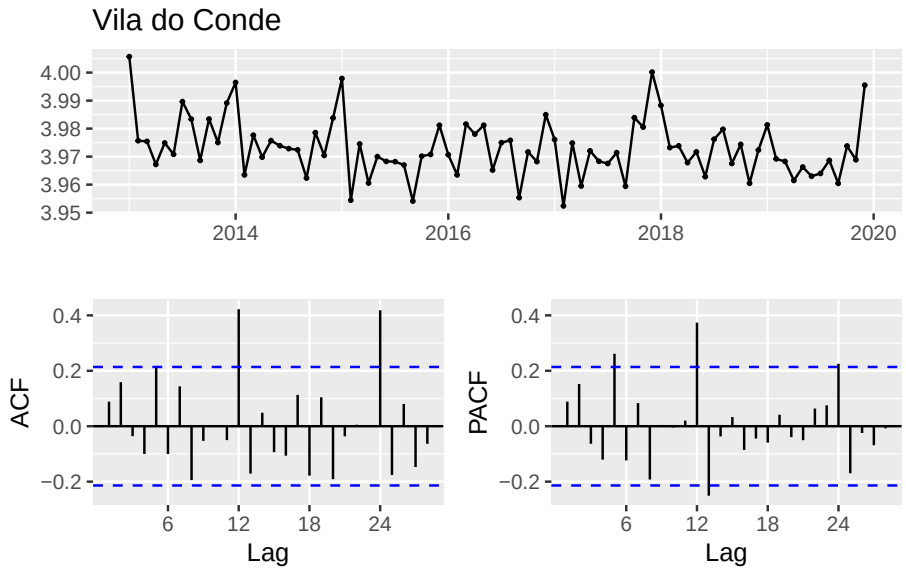



```
df_transform <- df_transform %>%
    mutate(
        stat.diffseason = map(.x = data.ts.box, nsdiffs),
        stat.diff6 = map(.x = data.ts.box,.f = ~ndiffs(diff(.x, 6))),
        stat.diff12 = map(.x = data.ts.box,.f = ~ndiffs(diff(.x, 12))),
        diff1 = map(.x = data.ts.box, .f = ~diff(.x, 1)),
        diff6 = map(.x = data.ts.box, .f = ~diff(.x, 6)),
        diff12 = map(.x = data.ts.box, .f = ~diff(.x, 12)),
        diff6.1 = map(.x = data.ts.box, .f = ~diff(diff(.x, 6),1)),
        diff12.1 = map(.x = data.ts.box, .f = ~diff(diff(.x, 12),1))
    )
```

```r
pl1 <- df_transform  %>%
    filter(unidade_saude == "Vila do Conde") %>%
    pull(data.ts.box) %>%
    .[[1]] %>%
    autoplot() +
    ggtitle("Vila do conde")


pl2 <- df_transform  %>%
    filter(unidade_saude == "Vila do Conde") %>%
    pull(diff1) %>%
    .[[1]] %>%
    autoplot() +
    ggtitle("Vila do conde diff1")

pl3 <- df_transform  %>%
    filter(unidade_saude == "Vila do Conde") %>%
    pull(diff6) %>%
    .[[1]] %>%
    autoplot() +
    ggtitle("Vila do conde diff6")

pl4 <- df_transform  %>%
    filter(unidade_saude == "Vila do Conde") %>%
    pull(diff12) %>%
    .[[1]] %>%
    autoplot() +
    ggtitle("Vila do conde diff12")

pl5 <- df_transform  %>%
    filter(unidade_saude == "Vila do Conde") %>%
    pull(diff6.1) %>%
    .[[1]] %>%
    autoplot() +
    ggtitle("Vila do conde diff6.1")

pl6 <- df_transform  %>%
    filter(unidade_saude == "Vila do Conde") %>%
    pull(diff12.1) %>%
    .[[1]] %>%
    autoplot() +
    ggtitle("Vila do conde diff12.1")

grid.arrange(pl1,pl2,pl3,pl4,pl5,pl6, ncol = 2)
```

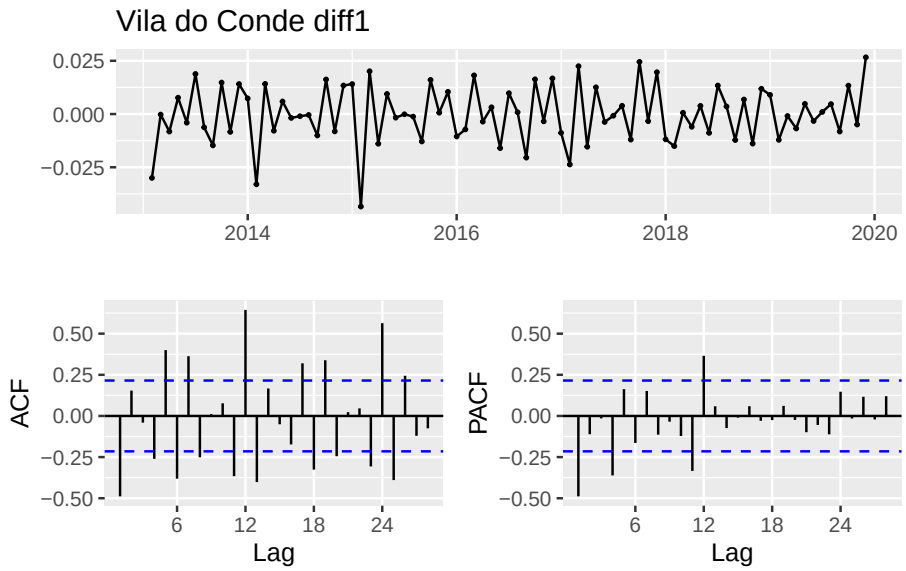



```
acf1 <- df_transform  %>%
    filter(unidade_saude == "Vila do Conde") %>%
    pull(data.ts.box) %>%
    .[[1]] %>%
    ggtsdisplay(main = "Vila do Conde")
```

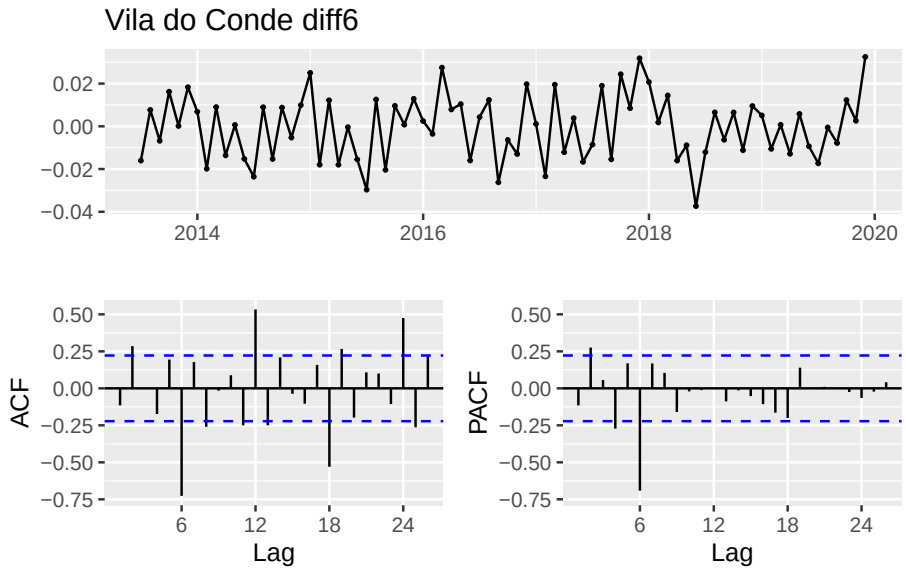



```
acf2 <- df_transform  %>%
    filter(unidade_saude == "Vila do Conde") %>%
```

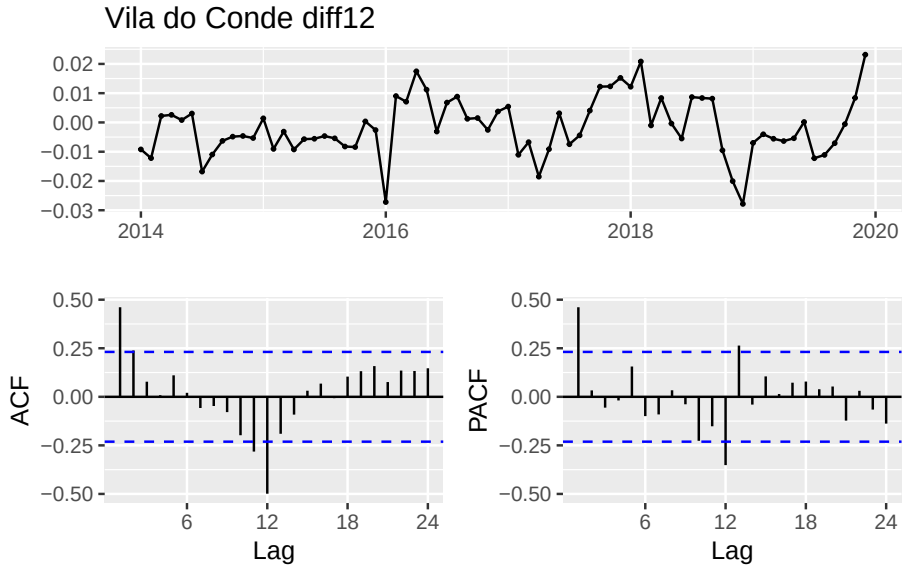
```
    pull(diff1) %>%
    .[[1]] %>%
    ggtsdisplay(main = "Vila do Conde diff1")
```




```
acf3 <- df_transform  %>%
    filter(unidade_saude == "Vila do Conde") %>%
    pull(diff6) %>%
    .[[1]] %>%
    ggtsdisplay(main = "Vila do Conde diff6")
```

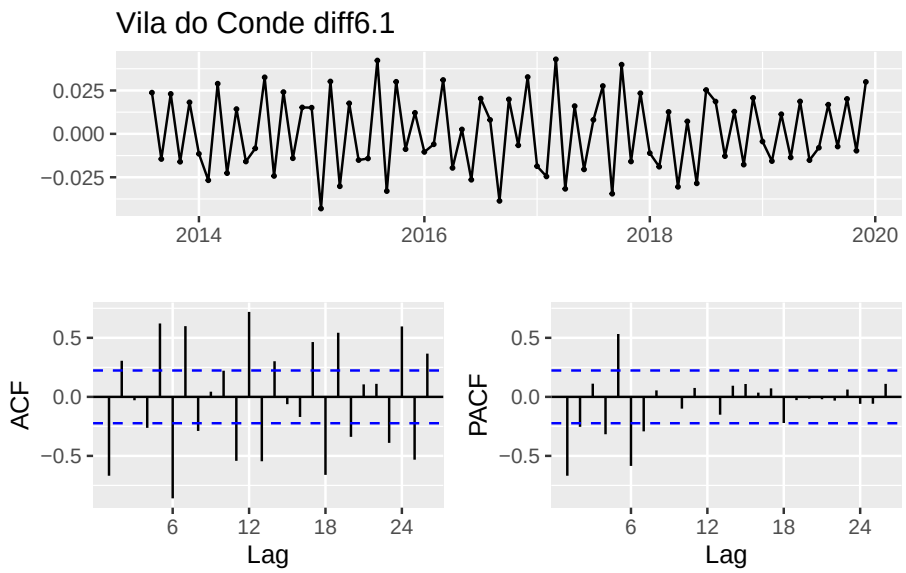



```
acf4 <- df_transform  %>%
    filter(unidade_saude == "Vila do Conde") %>%
    pull(diff12) %>%
    .[[1]] %>%
    ggtsdisplay(main = "Vila do Conde diff12")
```

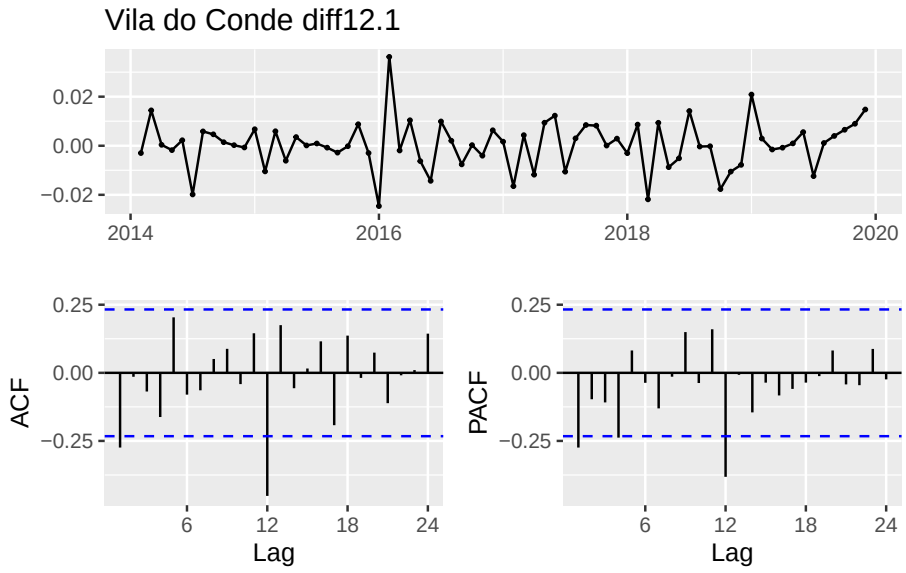



```
acf5 <- df_transform  %>%
    filter(unidade_saude == "Vila do Conde") %>%
```

```
    pull(diff6.1) %>%
    .[[1]] %>%
    ggtsdisplay(main = "Vila do Conde diff6.1")
```




```
acf6 <- df_transform  %>%
    filter(unidade_saude == "Vila do Conde") %>%
    pull(diff12.1) %>%
    .[[1]] %>%
    ggtsdisplay(main = "Vila do Conde diff12.1")
```

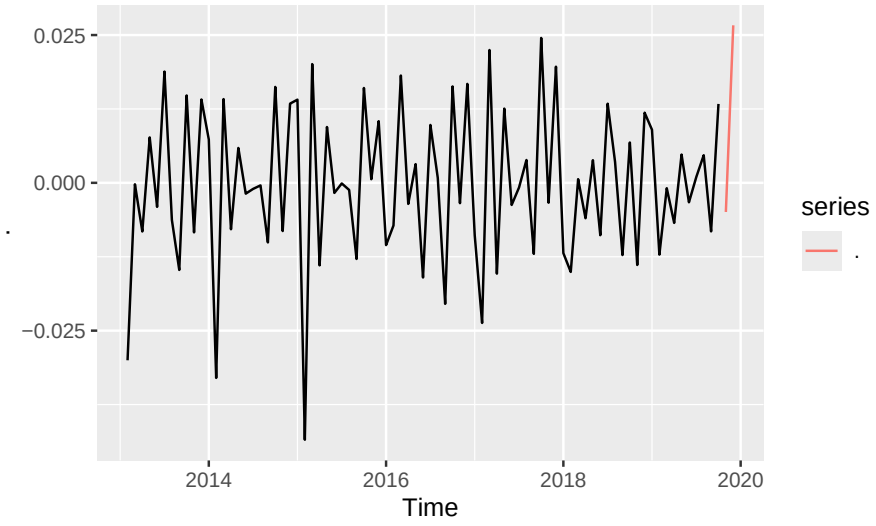


0.3. Split data

```
df_train_test <- df_transform %>%
    mutate(
        train = map(
            .x = diff1,
            .f = ~window(.x, end = c(2019, 10))
        ),
        test = map(
            .x = diff1,
            .f = ~window(.x, start = c(2019, 11))
        )
    )

df_train_test
```

```
# A tibble: 2 x 15
# Groups:   unidade_saude [2]
  unidade_saude data     data.ts       boxcox.lambda data.ts.box stat.diffseason
  <chr>         <list>   <list>        <list>        <list>      <list>
1 Barcelos      <tibble> <ts [84 x 1]> <dbl [1]>     <ts[...]>   <dbl [1]>
2 Vila do Conde <tibble> <ts [84 x 1]> <dbl [1]>     <ts[...]>   <dbl [1]>
# i 9 more variables: stat.diff6 <list>, stat.diff12 <list>, diff1 <list>,
#   diff6 <list>, diff12 <list>, diff6.1 <list>, diff12.1 <list>, train <list>,
#   test <list>
```

```
tr_pl <- df_train_test %>%
    filter(unidade_saude == "Vila do Conde") %>%
    pull(train) %>%
    .[[1]] %>%
    autoplot()

layer <- df_train_test %>%
    filter(unidade_saude == "Vila do Conde") %>%
    pull(test) %>%
    .[[1]] %>%
    autolayer()

tr_pl + layer
```




```
kpss.test(df_transform$diff1[[1]])
```

```
Warning in kpss.test(df_transform$diff1[[1]]): p-value greater than printed
p-value
```

```

    KPSS Test for Level Stationarity

data:  df_transform$diff1[[1]]
KPSS Level = 0.048206, Truncation lag parameter = 3, p-value = 0.1
```

```
adf.test(df_transform$diff1[[1]])
```

```
Warning in adf.test(df_transform$diff1[[1]]): p-value smaller than printed
```

p-value

```
    Augmented Dickey-Fuller Test

data:  df_transform$diff1[[1]]
Dickey-Fuller = -4.7912, Lag order = 4, p-value = 0.01
alternative hypothesis: stationary
```

```
df_train_test <- df_train_test %>%
    mutate(
        model.fit = map(.x = train, ~auto.arima(.x)),
        test.forecast = map2(.x = model.fit ,.y = test , .f = ~forecast(.x, h = length(.y)))
        test.mse = map2(.x = test.forecast , .y = test, .f = ~mean((.x$mean - .y)^2))
    )
```

```
t <- df_train_test %>%
    mutate(
        forecast.covid =  map(.x = model.fit, .f = ~forecast(.x, h = 22))
    )

t$forecast.covid
```

```
[[1]]
         Point Forecast        Lo 80      Hi 80       Lo 95      Hi 95
Nov 2019   -1.324106638 -2.058526250 -0.5896870 -2.44730451 -0.2009088
Dec 2019    1.247806518  0.482168639  2.0134444  0.07686443  2.4187486
Jan 2020   -0.342447748 -1.117200488  0.4323050 -1.52732981  0.8424343
Feb 2020   -1.603740541 -2.377824681 -0.8296564 -2.78760007 -0.4198810
Mar 2020    1.282409259  0.507448847  2.0573697  0.09720959  2.4676089
Apr 2020   -0.391330883 -1.166561802  0.3839000 -1.57694426  0.7942825
May 2020    0.503631399 -0.271683056  1.2789459 -0.68210974  1.6893725
Jun 2020   -0.329212447 -1.104552664  0.4461278 -1.51499298  0.8565681
Jul 2020    0.808822616  0.033474581  1.5841707 -0.37696988  1.9946151
Aug 2020   -0.012292850 -0.787642846  0.7630571 -1.19808834  1.1735026
Sep 2020   -0.956298128 -1.731647261 -0.1809490 -2.14209230  0.2294960
Oct 2020    0.852463180  0.077119072  1.6278073 -0.33332331  2.0382497
Nov 2020   -1.348361464 -2.125542811 -0.5711801 -2.53695776 -0.1597652
Dec 2020    1.250530885  0.473248494  2.0278133  0.06178005  2.4392817
Jan 2021   -0.222354473 -0.999474718  0.5547658 -1.41085733  0.9661484
Feb 2021   -1.400788966 -2.175467319 -0.6261106 -2.58555727 -0.2160207
Mar 2021    1.205071633  0.430389396  1.9797539  0.02029739  2.3898459
Apr 2021   -0.485084150 -1.259767585  0.2895993 -1.66986022  0.6996919
May 2021    0.534890241 -0.239793555  1.3095740 -0.64988638  1.7196669
Jun 2021   -0.350805745 -1.125489621  0.4238781 -1.53558249  0.8339710
Jul 2021    0.773264686 -0.001419116  1.5479485 -0.41151195  1.9580413
Aug 2021    0.003320155 -0.771363304  0.7780036 -1.18145596  1.1880963
```

```
[[2]]
         Point Forecast        Lo 80        Hi 80        Lo 95        Hi 95
Nov 2019  -0.0072963594 -0.016366429  0.001773710 -0.021167834  0.006575115
Dec 2019   0.0144575476  0.004239139  0.024675956 -0.001170159  0.030085255
Jan 2020  -0.0000291316 -0.010396775  0.010338512 -0.015885074  0.015826811
Feb 2020  -0.0205561770 -0.030917319 -0.010195035 -0.036402176 -0.004710178
Mar 2020   0.0089258534 -0.001438550  0.019290257 -0.006925134  0.024776841
Apr 2020  -0.0085128345 -0.018877713  0.001852044 -0.024364549  0.007338880
May 2020   0.0064199755 -0.003944973  0.016784924 -0.009431845  0.022271796
Jun 2020  -0.0057872029 -0.016152161  0.004577755 -0.021639039  0.010064633
Jul 2020   0.0054321340 -0.004932826  0.015797094 -0.010419704  0.021283972
Aug 2020   0.0018613013 -0.008503658  0.012226261 -0.013990536  0.017713139
Sep 2020  -0.0122889011 -0.022653859 -0.001923943 -0.028140736  0.003562934
Oct 2020   0.0148457396  0.004480795  0.025210684 -0.001006075  0.030697554
Nov 2020  -0.0067833226 -0.017392174  0.003825529 -0.023008161  0.009441516
Dec 2020   0.0146533753  0.003980810  0.025325941 -0.001668905  0.030975655
Jan 2021   0.0000456164 -0.010628943  0.010720176 -0.016279713  0.016370946
Feb 2021  -0.0205276455 -0.031175819 -0.009879472 -0.036812622 -0.004242669
Mar 2021   0.0089367440 -0.001711618  0.019585106 -0.007348521  0.025222009
Apr 2021  -0.0085086775 -0.019157068  0.002139712 -0.024793985  0.007776630
May 2021   0.0064215623 -0.004226832  0.017069956 -0.009863751  0.022706876
Jun 2021  -0.0057865972 -0.016434992  0.004861797 -0.022071911  0.010498717
Jul 2021   0.0054323651 -0.005216029  0.016080760 -0.010852949  0.021717679
Aug 2021   0.0018613896 -0.008787005  0.012509784 -0.014423924  0.018146703
```