

# Cement Production in Pakistan

*Zahid Asghar*

10 October 2022



# Cement Production

We have data from [Cement Manufacturers Association](#). Lets explore this data as follows:

```
1 library(tidyverse)
2 library(lubridate)
3 library(readr)
4 library(readxl)
5 library(collapse)
6 library(kableExtra)
7 library(dynlm)
8 library(forecast)
9 library(stargazer)
10 library(scales)
11 library(xts)
12 library(urca)
13 library(tsibble)
14 library(fpp2)
15 library(fpp3)
16 library(ggthemes)
17 library(DT)
18 library(fable)
19 library(fabletools)
```

# Read data

```

1 library(readr) # To load data from csv file, if one needs to load data from excel, then use excel file
2 cement<-read_csv("cement.csv")
3 glimpse(cement)
4 ## Rows: 1,116
5 ## Columns: 6
6 ## $ Date                <chr> "07/1991", "07/1991", "07/1991", "08...
7 ## $ `Series Name`       <chr> "Total Cement Sales", "Domestic Ceme...
8 ## $ Output              <dbl> 599, 599, NA, 632, 632, NA, 633, 633...
9 ## $ Unit                <chr> "Thousand Metric Ton", "Thousand Met...
10 ## $ `Observation Status` <chr> "Normal", "Normal", "Missing value",...
11 ## $ `Observation Status Comment` <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, ...

```

# Selecting and Renaming variables

```

1 cement<-cement %>% rename(Category=`Series Name` )
2
3 #cement<-ts(cement, frequency = 12, start=c(1991,7))
4 #cement
5 #cement<-cement %>% mutate(Month=yearmonth(Date)) %>% as_tsibble(index = Month)
6
7 cement<-cement %>% filter(Category=="Total Cement Sales") # Select only Total Cement Sales
8
9 cement$Date<-my(cement$Date) # Formatting required
10 cement$Date<-as_date(cement$Date,format="%Y-%m")
11
12 #cement<-tsibble(cement)
13 cement
14 ## # A tibble: 372 × 6
15 ##   Date          Category      Output Unit      Obser...1 Obser...2
16 ##   <date>        <chr>        <dbl> <chr>      <chr>    <lgl>
17 ## 1 1991-07-01 Total Cement Sales 599. Thousand Metric Ton Normal NA
18 ## 2 1991-08-01 Total Cement Sales 632. Thousand Metric Ton Normal NA
19 ## 3 1991-09-01 Total Cement Sales 622. Thousand Metric Ton Normal NA

```

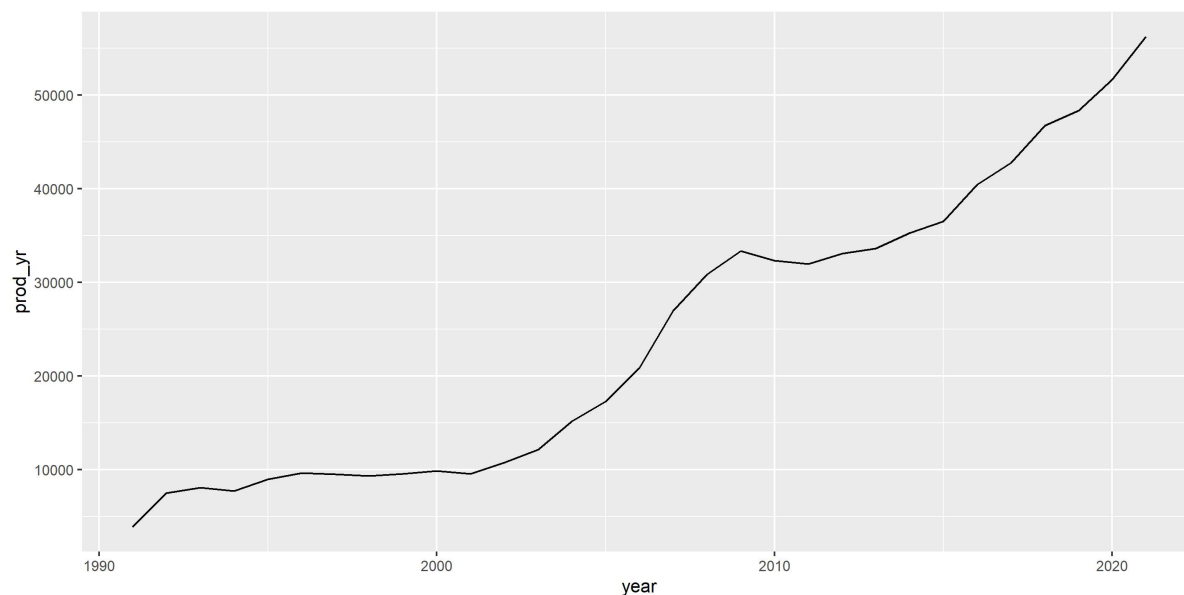
# Annual cement sale

What happens if year 2022 is included?

Why annual is less informative?

```
1 p1<-ggplot(cement_yrly)+aes(x=year,y=prod_yr)+geom_line()  
2 p1
```

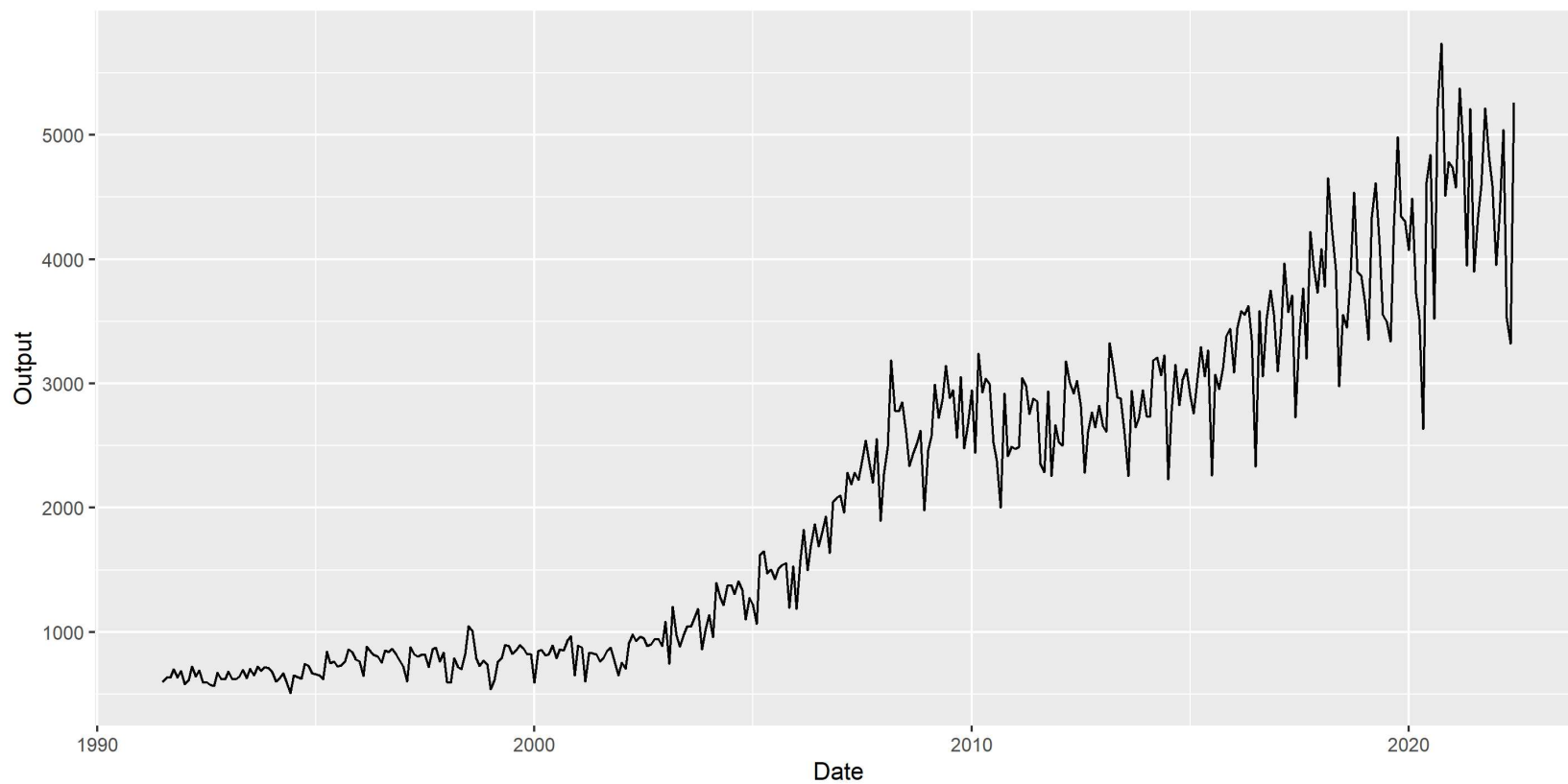
```
1 #forecast::naive(cement_yrly$prod_yr, h=4) %>% autoplot()
```



zahid asghar

# Monthly cement production

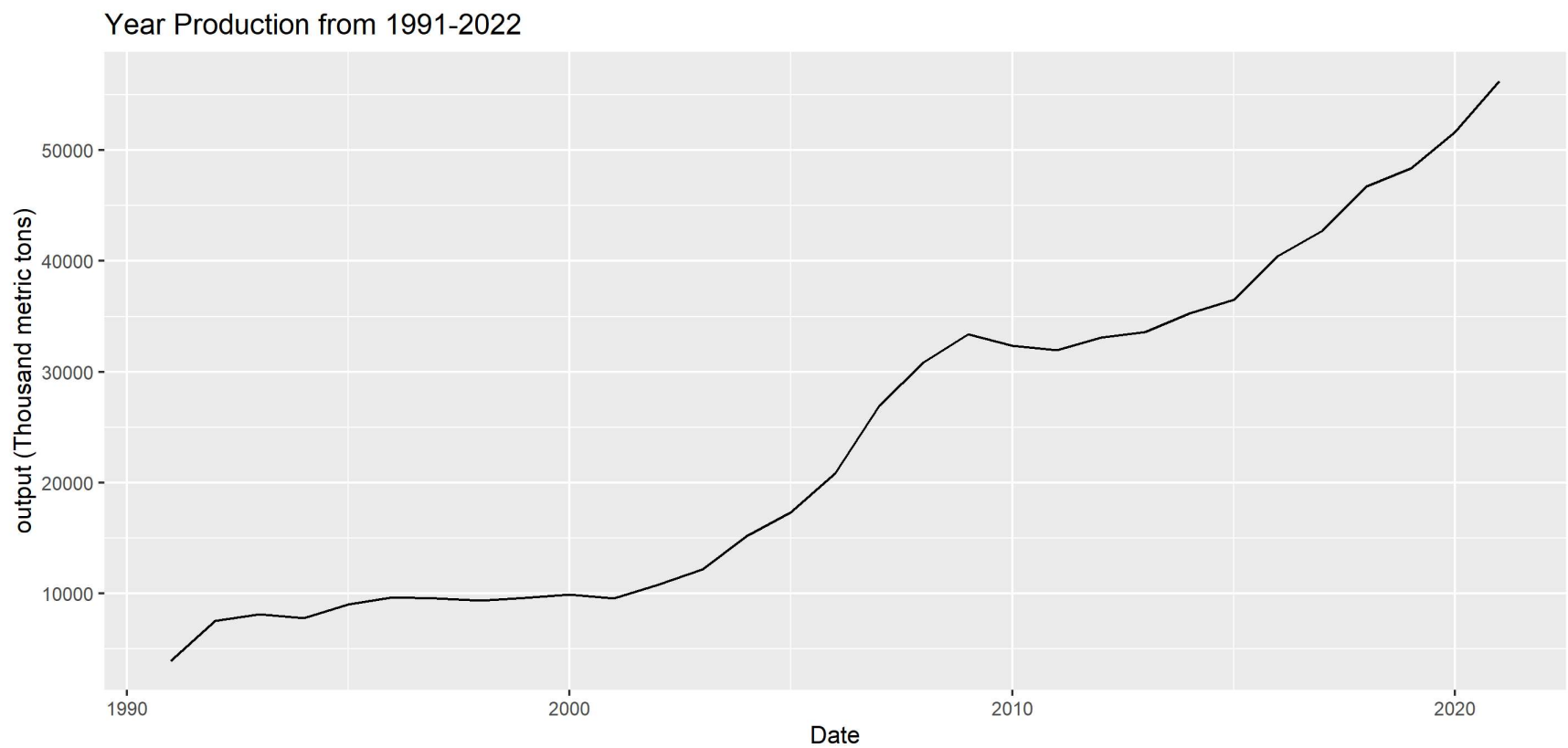
```
1 pm<-ggplot(cement)+aes(x=Date,y=Output)+geom_line()  
2 pm
```



zahid asghar

# Labels, data source, theme

```
1 p1+labs(x="Date",y="output (Thousand metric tons)", title = "Year Production from 1991-2022", caption = "By Zahid Asghar")
```

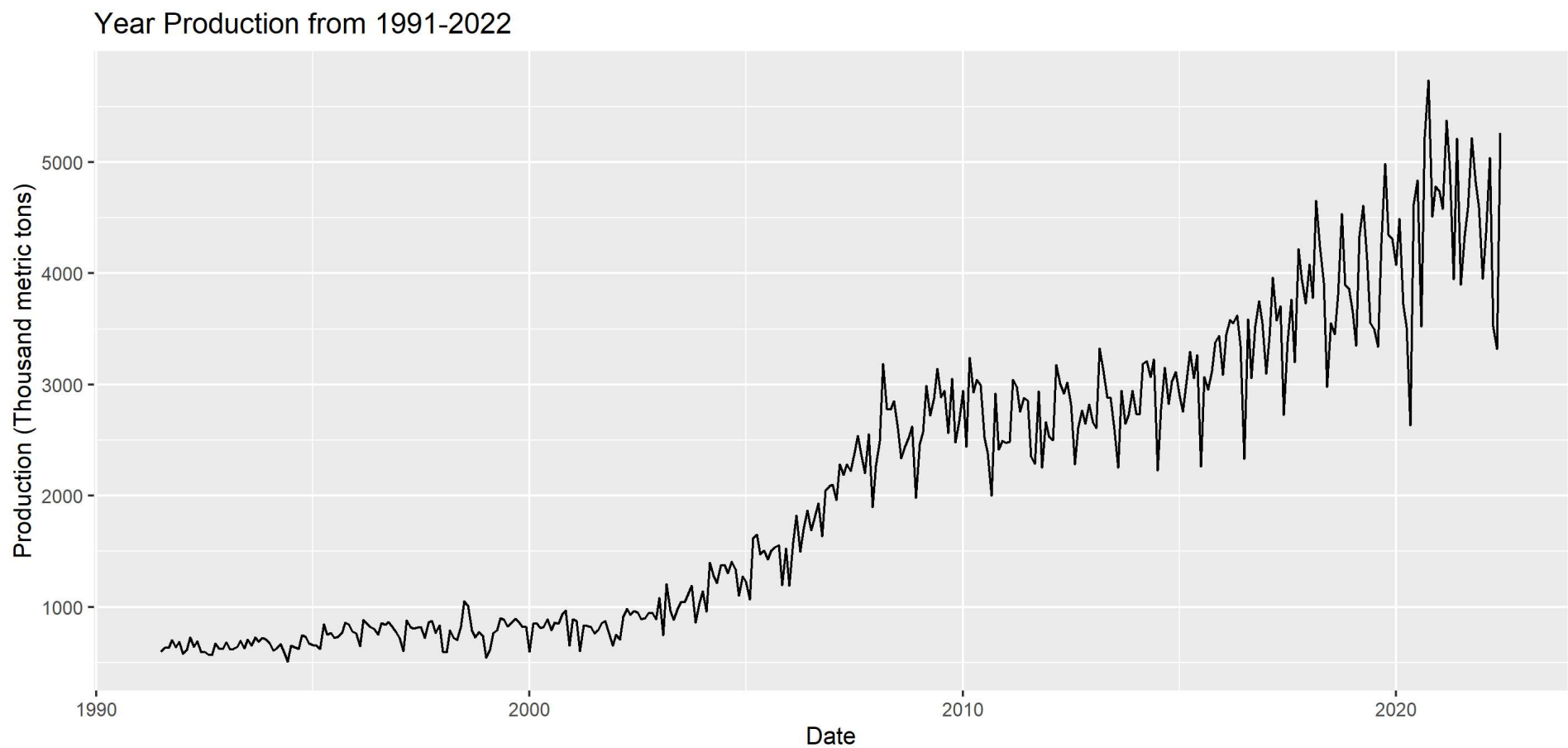


By Zahid Asghar, Source:APCMA, Pakistan

zahid asghar



```
1 pm+labs(x="Date",y="Production (Thousand metric tons)", title = "Year Production from 1991-2022", caption = "By Zahid
```



By Zahid Asghar, Source:APCMA, Pakistan

zahid asghar

# Last 10 years data

```

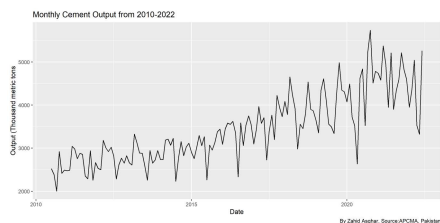
1 cement_2010<-cement %>% filter(Date>="2010-6-30")
2 cement_2010
3 ## # A tibble: 144 × 7
4 ##   Date          Category      Output Unit      Obser...1 Obser...2 year
5 ##   <date>        <chr>          <dbl> <chr>      <chr>    <lgl>    <dbl>
6 ## 1 2010-07-01 Total Cement Sales 2524. Thousand Met... Normal  NA      2010
7 ## 2 2010-08-01 Total Cement Sales 2383. Thousand Met... Normal  NA      2010
8 ## 3 2010-09-01 Total Cement Sales 2002. Thousand Met... Normal  NA      2010
9 ## 4 2010-10-01 Total Cement Sales 2920. Thousand Met... Normal  NA      2010
10 ## 5 2010-11-01 Total Cement Sales 2416. Thousand Met... Normal  NA      2010
11 ## 6 2010-12-01 Total Cement Sales 2493. Thousand Met... Normal  NA      2010
12 ## 7 2011-01-01 Total Cement Sales 2473. Thousand Met... Normal  NA      2011
13 ## 8 2011-02-01 Total Cement Sales 2488. Thousand Met... Normal  NA      2011
14 ## 9 2011-03-01 Total Cement Sales 3043. Thousand Met... Normal  NA      2011
15 ## 10 2011-04-01 Total Cement Sales 2980. Thousand Met... Normal  NA      2011
16 ## # ... with 134 more rows, and abbreviated variable names
17 ## #   1`Observation Status`, 2`Observation Status Comment`
18 ## # i Use `print(n = ...)` to see more rows

```

```

1 p11<-ggplot(cement_2010)+aes(x=Date,y=Output)+geom_line()
2 p11+labs(x="Date",y="Output (Thousand metric tons)", title = "Monthly Cement Output from 2010-2022", caption = "By Zahid Asghar")

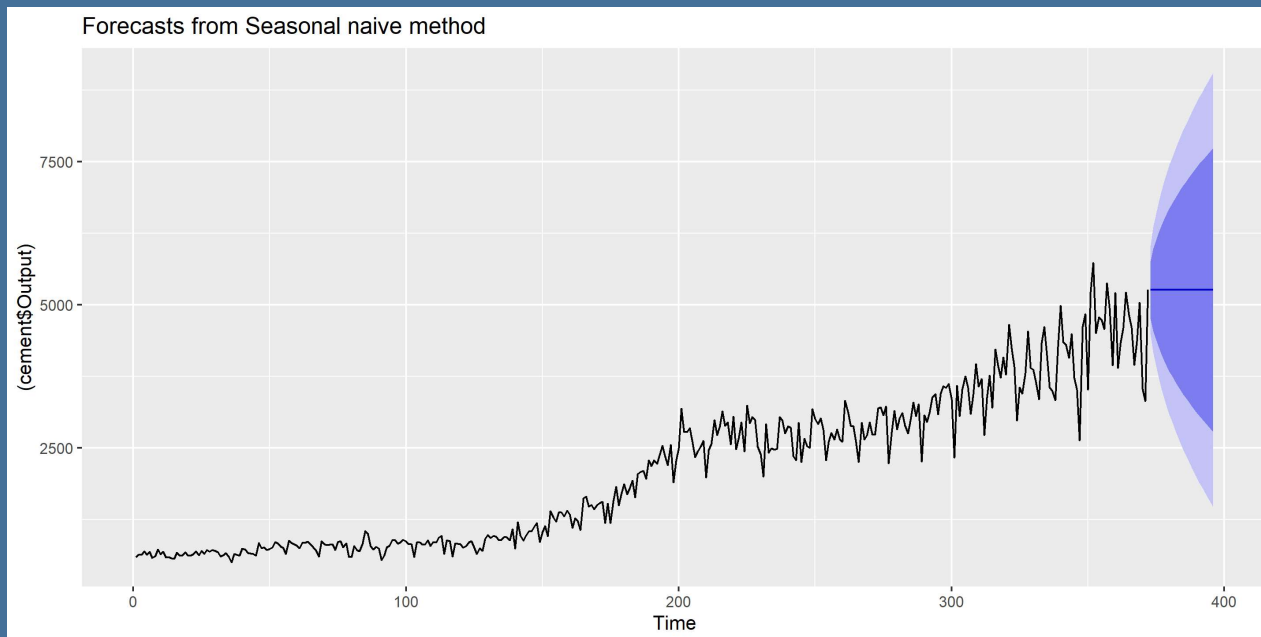
```



# Forecasting

Why series seems reversed? This is because date is placed in reverse order. So lets use a verb for sorting date in ascending order by **sort**

```
1  
2 forecast::snaive((cement$Output), h = 24) %>% autoplot()
```



zahid asghar

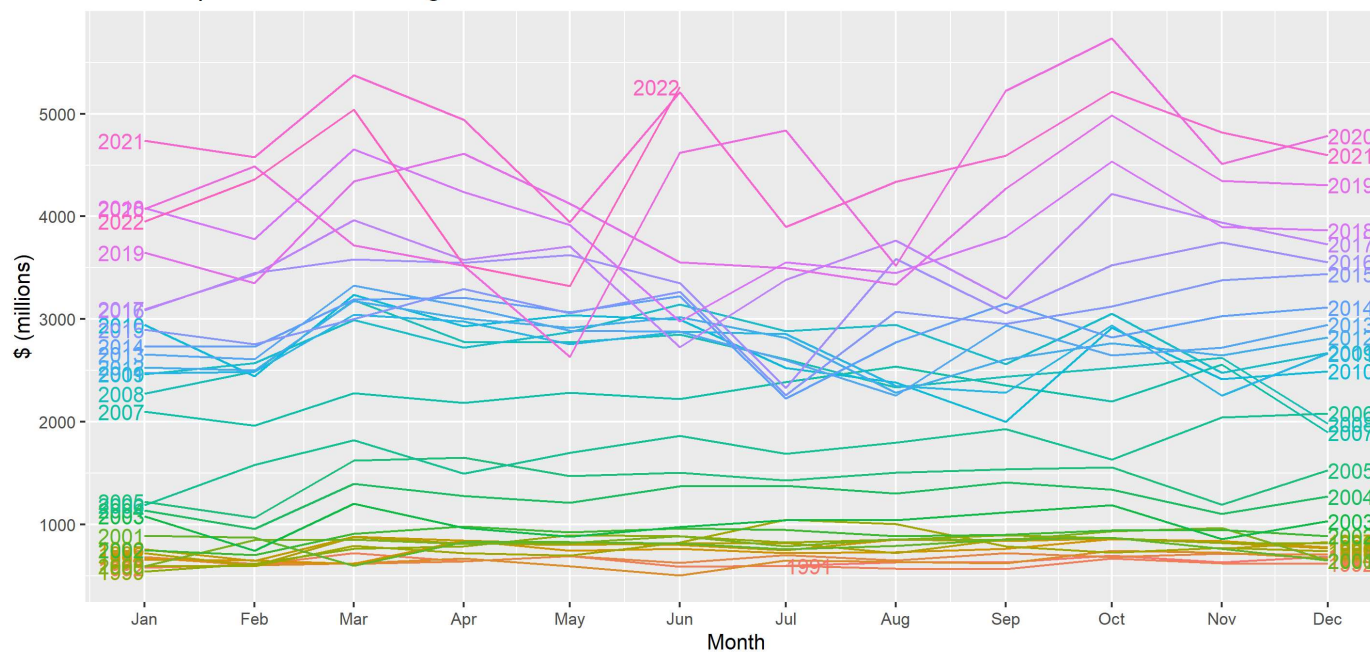
# ARIMA models

```
## Series: cement$Output
## ARIMA(4,1,1) with drift
##
## Coefficients:
##          ar1      ar2      ar3      ar4      ma1      drift
##          0.137  -0.106  0.082  -0.169  -0.83    10.54
## s.e.    0.064   0.059  0.059   0.057   0.04     2.67
##
## sigma^2 = 100902:  log likelihood = -2661
## AIC=5337   AICc=5337   BIC=5364
##
## Training set error measures:
##              ME RMSE MAE      MPE MAPE  MASE      ACF1
## Training set -0.403  315 202  -2.99 10.1  0.811 -0.00574
```

# Seasonal Plot

```
1 cement<-cement %>% mutate(Month=yearmonth(Date)) %>% as_tsibble(index=Month)
2
3 cement %>%
4   gg_season(Output, labels = "both")+
5   labs(y = "$ (millions)",
6        title = "Seasonal plot: Antidiabetic drug sales")
```

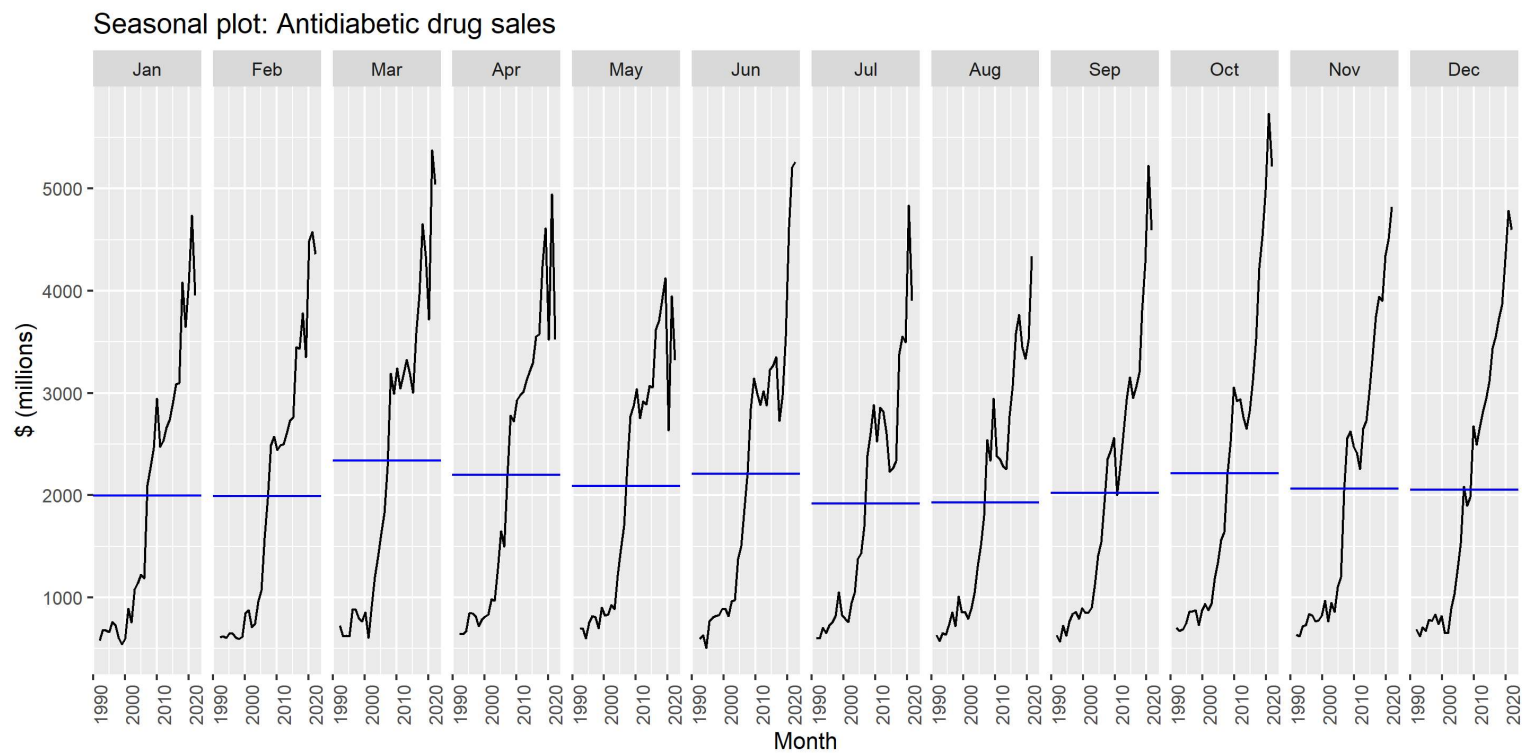
Seasonal plot: Antidiabetic drug sales



zahid asghar

# Monthly Analysis

```
1 cement %>%
2   gg_subseries(Output, labels = "both")+
3   labs(y = "$ (millions)",
4        title = "Seasonal plot: Antidiabetic drug sales")
```



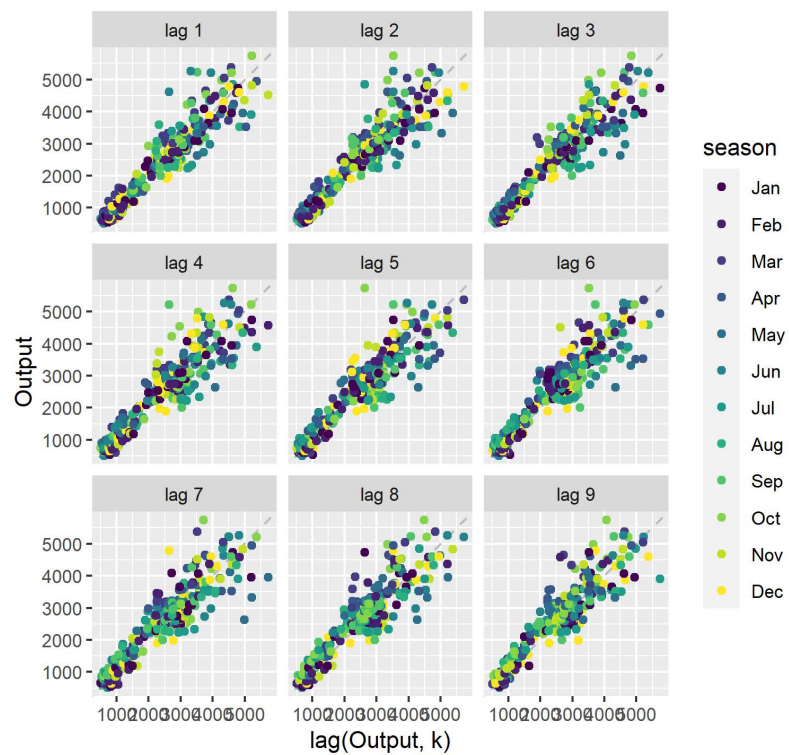
zahid asghar

# Lag Plots

```

1
2 cement %>%
3   gg_lag(Output, geom = "point") +
4   labs(x = "lag(Output, k)")

```



zahid asghar

# Shiny Interactive View

```
1 library(seasonalview)
2 library(shiny)
3 #cement_prod<-ts(data = cement$Output,frequency = 12, start=c(1991,7))
4 #view(seas(cement_prod))
```

```
1 quarto render cement.qmd --to pdf
```