

Final report

Domas Junča, Tomas Arėška, Edoardo Da Besi

2023-05-15

Reads

<https://www.eia.gov/todayinenergy/detail.php?id=55920#:~:text=Europe%20became%20the%20primary%20destination,U.S.%20LNG%20exports%20to%20Europe.>

Analysing U.S. energy exports

1. Introduction

Energy commodities play a pivotal role in the global economy, as they are essential for economic growth and development. They are required for transportation, manufacturing, heating, cooling, and many other critical activities that drive economic activity.

The production, consumption, and trade of energy commodities have significant impacts on the global economy. The prices of energy commodities are influenced by supply and demand factors, geopolitical tensions, weather events, and technological advances.

The recent COVID-19 pandemic and the war in Ukraine have created unprecedented changes in the energy market, and understanding the impact of these events is crucial for policymakers, analysts, and investors. This study aims to examine how the Russian war on Ukraine affected energy exports of the United States to the OECD Europe countries.

OECD Europe was selected as the region, as because of geographical and historical relations it was most influenced region by the Russian war on Ukraine. The United States has been chosen as the focus because it is one of the leading producers and exporters of energy commodities, and its export trends have significant implications for the global energy market. Moreover, the United States serves as a significant alternative for OECD Europe countries to turn to after the imposition of sanctions on Russia.

In this study, we will analyze the impact of the war in Ukraine on US energy exports to OECD Europe. Specifically, we will focus on three main energy products: crude oil, petroleum products, and natural gas. By examining trends in these exports, we can better understand how the geopolitical events in Ukraine have affected the energy market in Europe. Additionally, we will investigate changes in OECD Europe's dependence on Russian gas to see how energy dynamics in the region have shifted.

2. Time Series Graphics

In this section of the report, we will be creating time series graphics using various techniques such as time plots, seasonal plots, seasonal subseries plots, and ACF(). These plots will provide insights into the different time series in the dataset and help us

understand their characteristics, such as seasonality, trend, and cyclicity. We will also look for unusual observations and changing patterns over time and explore the differences between the series in terms of these features. Additionally, for this analysis, we have decided to focus on three specific time series: US exports of Natural Gas to OECD, US Exports of Crude Oil to OECD, and US exports of Petroleum products to OECD. We will create the same types of graphs for all of these time series, including time plots, seasonal plots, seasonal subseries plots, and ACF plots. These time series span the last ten years from 2013 to 2023. Through these analyses, we aim to gain insights into the seasonality, cyclicity, trend, unusual observations, changing patterns, and differences between these time series, and determine if any of the series resemble a white noise process. By analyzing these time series graphics, we can gain a better understanding of the data and identify any potential patterns or anomalies that may exist within it.

2.1 Dataset

The data sets utilized in this study consist of monthly observations of Country's reliance on Russian oil products, USA's natural gas exports by country, USA's total crude oil exports by destination, and USA's petroleum products exports by destination.

Our primary source of data is the U.S. Energy Information Administration (EIA), an independent organization that collects, analyzes, and disseminates data and information related to energy production, consumption, and distribution in the United States. Also, for the purpose of this study we are using data from the International Energy Agency (IEA), which is an intergovernmental organization that was established in 1974 to promote global cooperation on energy issues.

In case of data on reliance on Russian Oil the dataset contains monthly total oil imports from Russia as a percentage of total oil imports for a selected group of European countries belonging to the Organization for Economic Co-operation and Development (OECD) from January 1990 to October 2022. The dataset includes the date, country name, and the percentage of total oil imports that came from Russia for each observation.

In this section, we will analyze three datasets that pertain to US exports of Crude oil, Petroleum Products, and Natural Gas, all measured in barrels. Prior to converting the data to tsibble format, we performed data cleaning and formatting. To simplify the data, we converted each dataset to have the same structure, with three columns: date, destination country, and the amount exported. After standardizing the format, we converted each dataset to tsibble format. We then filtered the data to only include countries that are part of the OECD and summarized the results based on the amount exported on each date, removing country names and leaving only the date and amount exported. Finally, we merged all three tsibble tables into one by using the left join function and pivoting the data, resulting in a final time series that represents US exports of the three different products in barrels to OECD countries. The tsibble contains three columns: Date, Export Type, and Amount.

```
library(fpp3)
library(tidyverse)
library(ggplot2)
```

```

library(dplyr)
library(tidyr)
library(lubridate)
library(scales)

#Total Crude Oil Exports by Destination

raw_crude_oil <- read_csv("https://raw.githubusercontent.com/edb-313/Energy-
trade-
forecasting/main/Data/csv/Total%20Crude%20Oil%20Exports%20by%20Destination_EI
A.csv", skip = 2,)

#Leaving only country names in column names

names(raw_crude_oil)[-2] <- str_replace_all(names(raw_crude_oil)[-2],
                                           c("U.S. Exports to " = "",
                                             " of Crude Oil \\\(Thousand
Barrels\\)" = ""))

#Pivoting (switching) columns starting from 2 column to rows

crude_oil <- raw_crude_oil %>%
  pivot_longer(cols = -Date,
               names_to = 'Destination country',
               values_to = 'Amount of Crude Oil (Thousand Barrels)')

#Filtering out values where Date is 'NA'

crude_oil <- crude_oil %>%
  drop_na(Date) %>%
  mutate(Date = as.Date(Date))

#Fixing date format, replacing 'na' values with 0

crude_oil$Date <- as.Date(crude_oil$Date)
crude_oil$`Amount of Crude Oil (Thousand Barrels)` <-
  replace_na(crude_oil$`Amount of Crude Oil (Thousand Barrels)`, 0)

#Setting up the tsibble

ts_crude_oil <- crude_oil %>%
  mutate(Date = yearmonth(Date)) %>%
  as_tsibble(
    index = Date,

```

```

    key = `Destination country`
  )

#OECD Europe Countries vector

oecd_europe <- c("Austria", "Belgium", "Czech Republic", "Denmark",
"Estonia",
               "Finland", "France", "Germany", "Greece", "Hungary",
"Iceland",
               "Ireland", "Italy", "Latvia", "Lithuania", "Luxembourg",
               "Netherlands", "Norway", "Poland", "Portugal", "Slovak
Republic",
               "Slovenia", "Spain", "Sweden", "Switzerland", "Turkey",
               "United Kingdom")

#filtering out only relevant values

oecd_crude_oil <- ts_crude_oil %>%
  filter(`Destination country` %in% oecd_europe) %>%
  filter(Date >= as.Date("2014-04-15"), Date <= as.Date("2023-01-15"))

oecd_crude_oil_agg <- oecd_crude_oil %>%
  index_by(Date) %>%
  summarize_at(vars(`Amount of Crude Oil (Thousand Barrels)`), sum) %>%
  ungroup()

#####
###

#Total Oil Products Exports by Destination

raw_oil_prod <- read_csv("https://raw.githubusercontent.com/edb-313/Energy-
trade-
forecasting/main/Data/csv/Total%20Oil%20Products%20Exports%20by%20Destination
_EIA.csv", skip = 2,)

names(raw_oil_prod)[-2] <- str_replace_all(names(raw_oil_prod)[-2],
                                           c("U.S. Exports to " = "",
                                             " of Total Petroleum Products
\\(Thousand Barrels\\)" = ""))
oil_prod <- raw_oil_prod %>%
  pivot_longer(cols = -Date,
               names_to = 'Destination country',
               values_to = 'Amount of total Petroleum Porducts (Thousand
Barrels)')

```

```

oil_prod$`Amount of total Petroleum Products (Thousand Barrels)` <-
replace_na(oil_prod$`Amount of total Petroleum Products (Thousand Barrels)`,
0)

oil_prod <- oil_prod %>%
  drop_na(Date) %>%
  mutate(Date = as.Date(Date))

ts_oil_prod <- oil_prod %>%
  mutate(Date = yearmonth(Date)) %>%
  as_tsibble(
    index = Date,
    key = `Destination country`
  )

oecd_oil_prod <- ts_oil_prod %>%
  filter(`Destination country` %in% oecd_europe) %>%
  filter(Date >= as.Date("2014-05-15"), Date <= as.Date("2023-01-15"))

oecd_oil_prod_agg <- oecd_oil_prod %>%
  index_by(Date) %>%
  summarize_at(vars(`Amount of total Petroleum Products (Thousand Barrels)`),
sum) %>%
  ungroup()

#####
###
#Reliance on Russian gas

raw_rel_rusoil <- read_csv("https://raw.githubusercontent.com/edb-313/Energy-
trade-forecasting/main/Data/csv/Monthly_Reliance_on_Russian_Oil.csv", skip =
1)

rel_rusoil <- raw_rel_rusoil %>%
  pivot_longer(cols = -...1,
    names_to = 'Date',
    values_to = 'Total Oil Imports from Russia / Total Oil Imports
(%)')

rel_rusoil <- rel_rusoil %>%
  rename(Country = ...1) %>%
  mutate(Date = lubridate::dmy(Date)) %>%
  select("Date", "Country", "Total Oil Imports from Russia / Total Oil Imports
(%)" )

```

```

ts_rel_rusoil <- rel_rusoil %>%
  as_tsibble(
    index = Date,
    key = Country)

oecd_relrus <- ts_rel_rusoil %>%
  filter(Country %in% oecd_europe)

#####
###
#Total Natural Gas Exports by Destination

raw_nat_gas <- read.csv('https://raw.githubusercontent.com/edb-313/Energy-
trade-
forecasting/main/Data/csv/Natural%20Gas%20Exports%20by%20Country_EIA.csv',hea
der=TRUE , check.names=FALSE, skip = 2)

# drop columns 2-3, 6-8, 54 and 57-72
raw_nat_gas <- subset(raw_nat_gas, select = -c(2:3, 6:8,54, 57:72))
raw_nat_gas[is.na(raw_nat_gas)] <- 0

#cimbining exports to the same countries

nat_gas <- raw_nat_gas %>%
  mutate(Mexico = `U.S. Natural Gas Pipeline Exports to Mexico (MMcf)` +
`Liquefied U.S. Natural Gas Exports by Truck to Mexico (Million Cubic Feet)`,
        Canada = `U.S. Natural Gas Pipeline Exports to Canada (MMcf)` +
`Liquefied U.S. Natural Gas Exports by Truck to Canada (Million Cubic Feet)`)

#removing duplicate columns

nat_gas <- nat_gas[, !(names(nat_gas) %in% c('U.S. Natural Gas Pipeline
Exports to Mexico (MMcf)',
                                           'Liquefied U.S. Natural Gas
Exports by Truck to Mexico (Million Cubic Feet)',
                                           'U.S. Natural Gas Pipeline
Exports to Canada (MMcf)',
                                           'Liquefied U.S. Natural Gas
Exports by Truck to Canada (Million Cubic Feet)'))]

#cleaning column names

colnames(nat_gas) <- gsub(pattern = ".*(to\\s)", "", colnames(nat_gas))
colnames(nat_gas) <- gsub("\\(.*?\\)", "", colnames(nat_gas))

#pivoting data
gastemp <- nat_gas %>%
  pivot_longer(cols = -Date,
               names_to = 'Destination country',

```

```

        values_to = 'Amount of Natural gas (MMcf)')

#removing n/as in date column, fixing date formatting, fixing values
(romoving space in the end)
library(stringr) #for trimming the strings

gastemp <- gastemp %>%
  drop_na(Date) %>%
  mutate(Date = as.Date(Date)) %>%
  mutate(`Destination country` = str_trim(`Destination country`))

#summing duplicates

gastemp <- gastemp %>%
  group_by(Date, `Destination country`) %>%
  summarize_at(vars(`Amount of Natural gas (MMcf)`), sum) %>%
  ungroup() %>%
  drop_na(Date)

#creating a tsibble

ts_nat_gas <- gastemp %>%
  mutate(Date = yearmonth(Date)) %>%
  as_tsibble(
    index = Date,
    key = `Destination country`
  )

#filtering only relevant data

oecd_nat_gas <- ts_nat_gas %>%
  filter(`Destination country` %in% oecd_europe) %>%
  filter(Date >= as.Date("2014-05-15"), Date <= as.Date("2023-01-15"))

oecd_nat_gas_agg <- oecd_nat_gas %>%
  index_by(Date) %>%
  summarize_at(vars(`Amount of Natural gas (MMcf)`), sum) %>%
  ungroup()

oecd_nat_gas_agg <- oecd_nat_gas_agg %>%
  mutate(`Amount of Natural gas (BOE)` = `Amount of Natural gas (MMcf)` *
1000 / 6000 )

oecd_nat_gas_agg <- oecd_nat_gas_agg %>%
  select(Date, `Amount of Natural gas (BOE)`)

#####
###

```

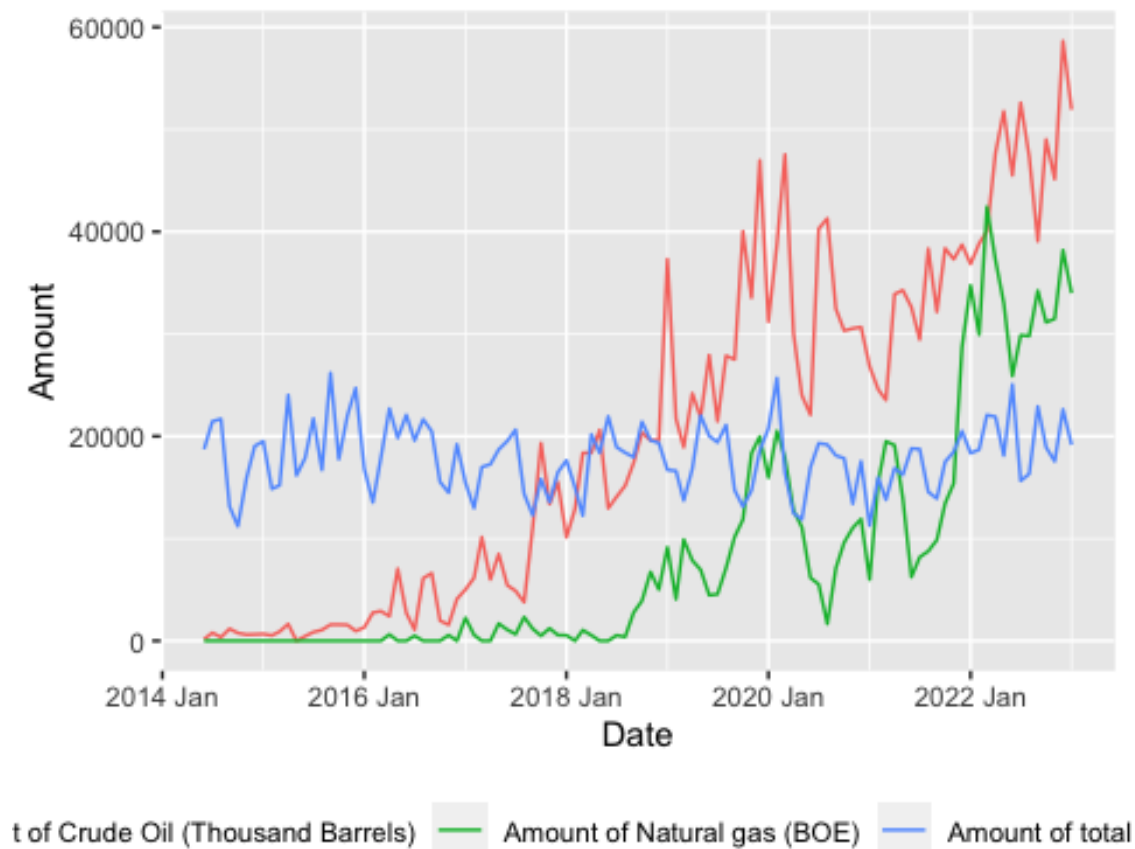
```

oecd_merged <- left_join(oecd_nat_gas_agg, oecd_oil_prod_agg, by = "Date")
%>%
  left_join(oecd_crude_oil_agg, by = "Date")

total_exp_oecd <- oecd_merged %>%
  pivot_longer(cols = -Date, names_to = "Export Type", values_to = "Amount")

total_exp_oecd %>% ggplot(
  aes(Date, Amount, color = `Export Type`) +
  geom_line() +
  labs(x = "Date", y = "Amount") +
  theme(legend.position = "bottom")

```



Comment on Nat Gas export zeros

2.2 Time Series Plots

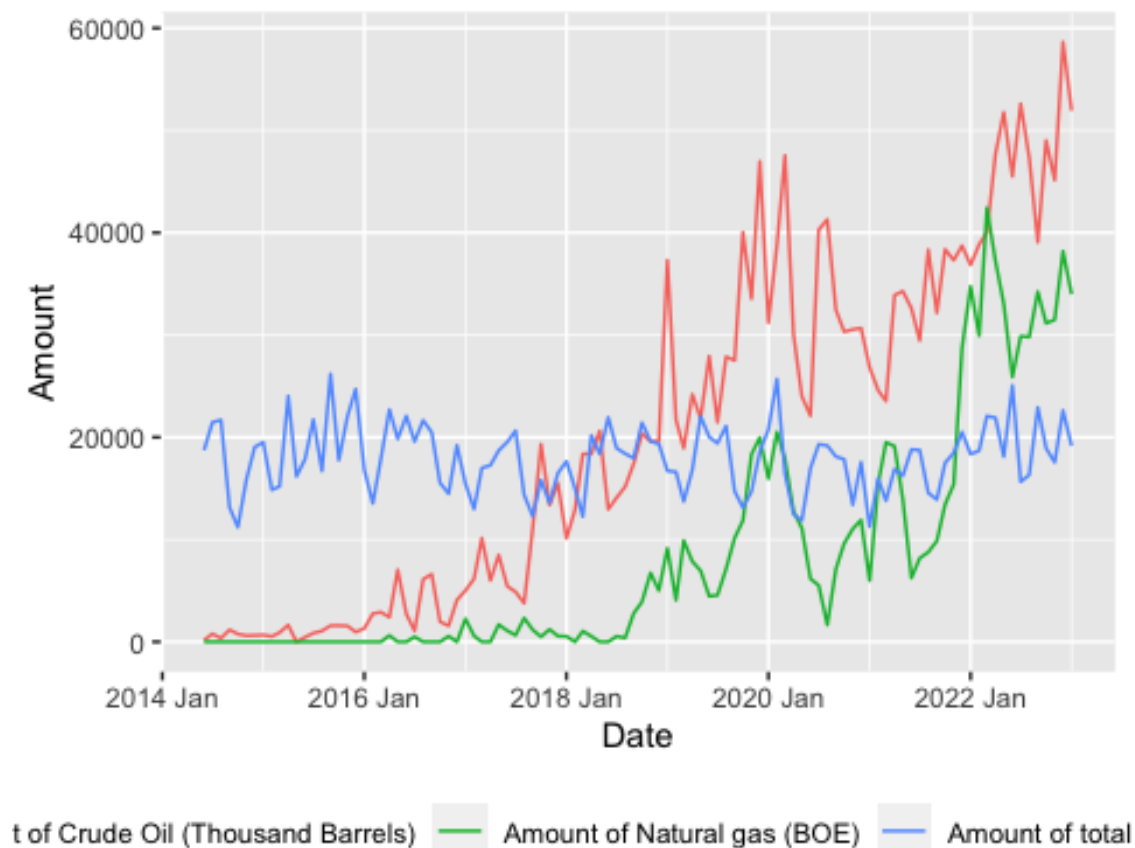
Firstly, we will examine a general overview of the three products: US exports of Crude Oil, Petroleum Products, and Natural Gas to OECD countries in barrels over the last ten years. This analysis will provide us with an understanding of the overall trend and magnitude of the exports. By examining the trend, we can identify any significant changes in the exports

over the years and evaluate the magnitude to determine which product has the highest export value. This overview will be a crucial step before we dive deeper into the individual time series of each product.

Including Plots

The graph below reveals significant growth in Crude oil and Natural gas exports since around 2015, with a clear year-on-year increase in both goods, indicating a noticeable trend. Moreover, it appears that the time series for these two products are non-stationary, possibly suggesting the presence of a trend, seasonality, or both. In 2022/2023, Crude oil and Natural gas exports reached their peak, with crude oil almost 60,000 and natural gas around 40,000. Conversely, the time series for Petroleum products appears to be stationary, with a slight increasing trend that may require further analysis.

```
total_exp_oecd %>% ggplot(  
  aes(Date, Amount, color = `Export Type`) ) +  
  geom_line() +  
  labs(x = "Date", y = "Amount") +  
  theme(legend.position = "bottom")
```



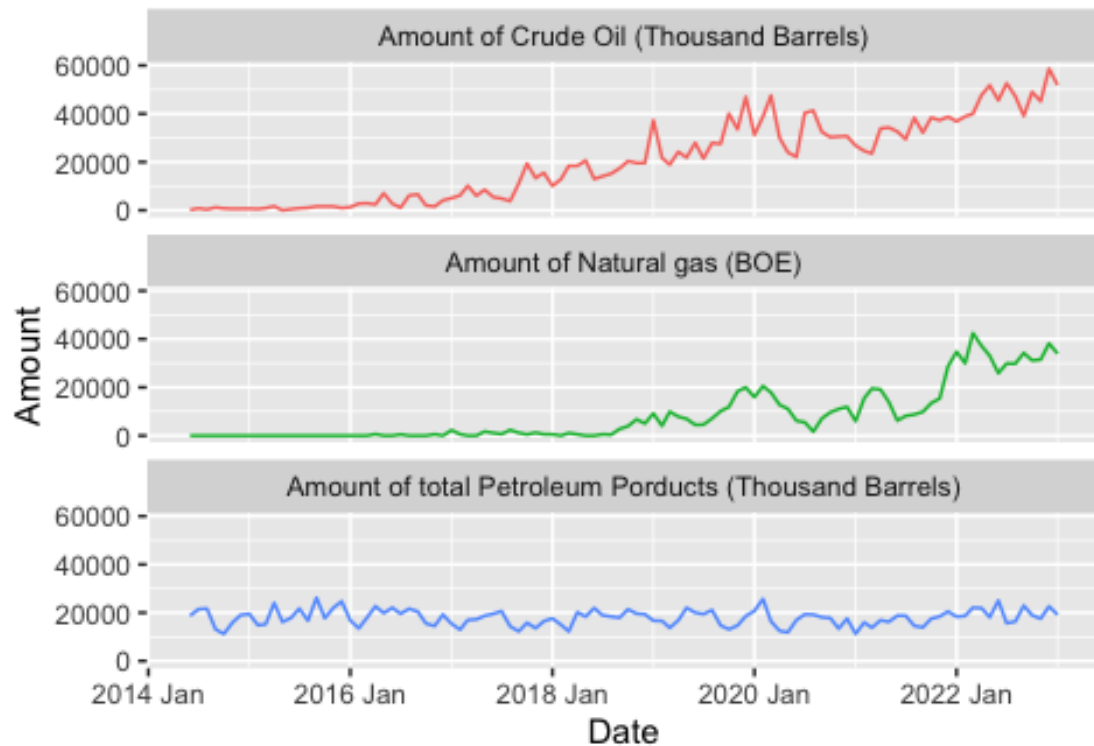
In the graph below, we have plotted each exported product separately to provide a better picture of each time series. The data shows that crude oil exports began in 2016 with lower amounts initially, but within a few years, exports boomed. Since 2016, there has been a

significant positive trend, possibly with seasonality. In 2020, crude oil exports reached a peak of around 50,000 barrels. Then, due to the COVID-19 pandemic, exports decreased for two years. However, in the beginning of 2022, exports increased significantly, most likely due to the Russian invasion in Ukraine. Currently, exports of crude oil are at a peak, reaching almost 60,000 barrels.

Looking at natural gas exports, we can see that the US did not export to OECD countries until after 2016, when low amounts of natural gas were first exported. In 2018, exports increased significantly, reaching their first peak in 2020. We can see a similar positive trend and possible seasonality as in the crude oil graph. During the COVID-19 pandemic, exports decreased significantly, but again, after the Russian invasion, exports increased significantly, reaching 40,000 barrels. This suggests that OECD countries stopped buying natural gas from Russia.

Finally, looking at petroleum product exports, we can see some fluctuations since 2010, but no clear trend. Seasonality may exist. The changes during COVID-19 and after the Russian invasion were not as significant as in the other two graphs.

```
total_exp_oecd %>%  
  ggplot(aes(Date, Amount, color = `Export Type`)) +  
  geom_line()+  
  labs(x = "Date", y = "Amount")+  
  theme(legend.position = "bottom") +  
  facet_wrap(~`Export Type`, ncol = 1)
```

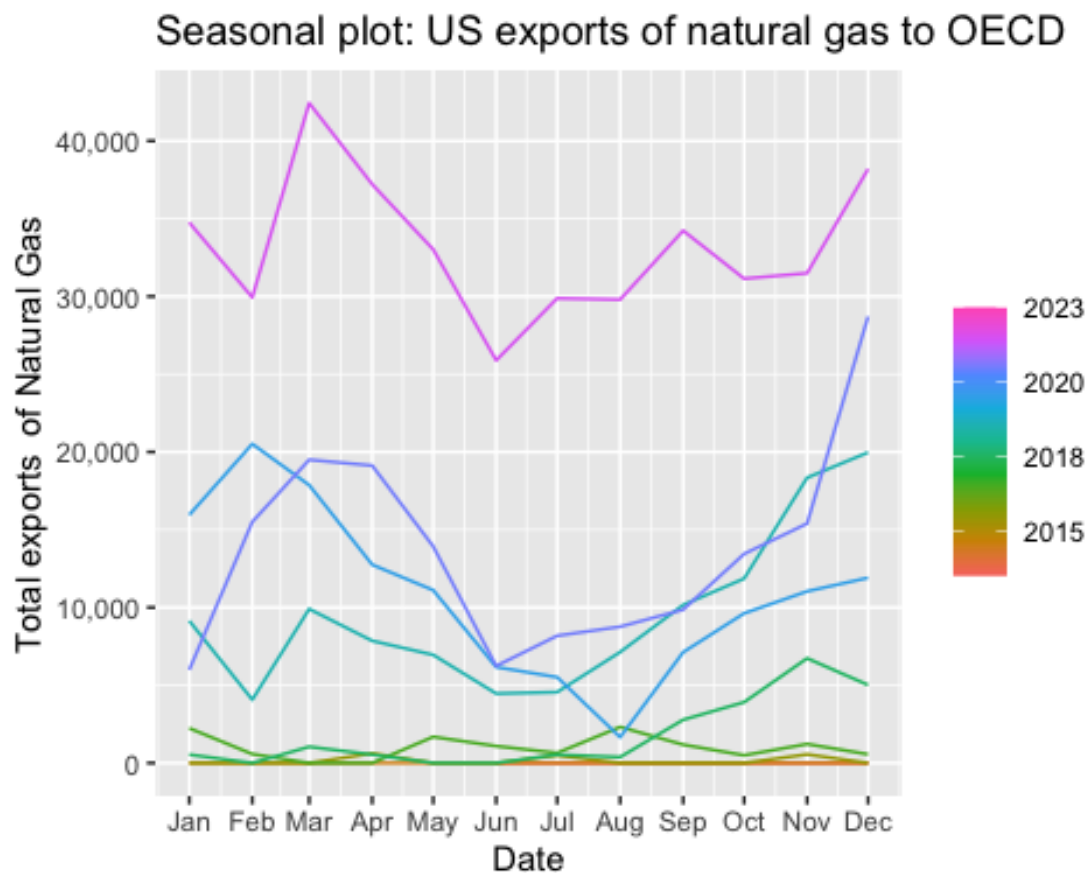


t of Crude Oil (Thousand Barrels) Amount of Natural gas (BOE) Amount of total

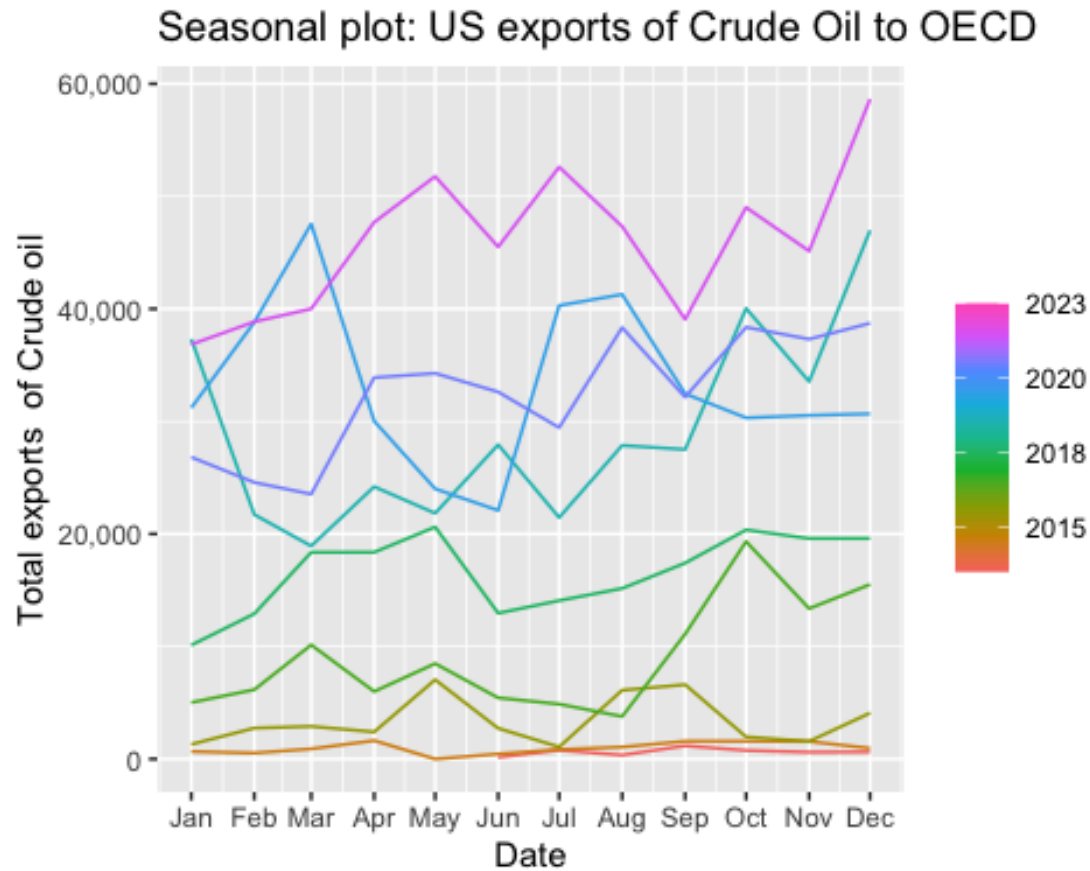
####

2.3 Time Series Patterns

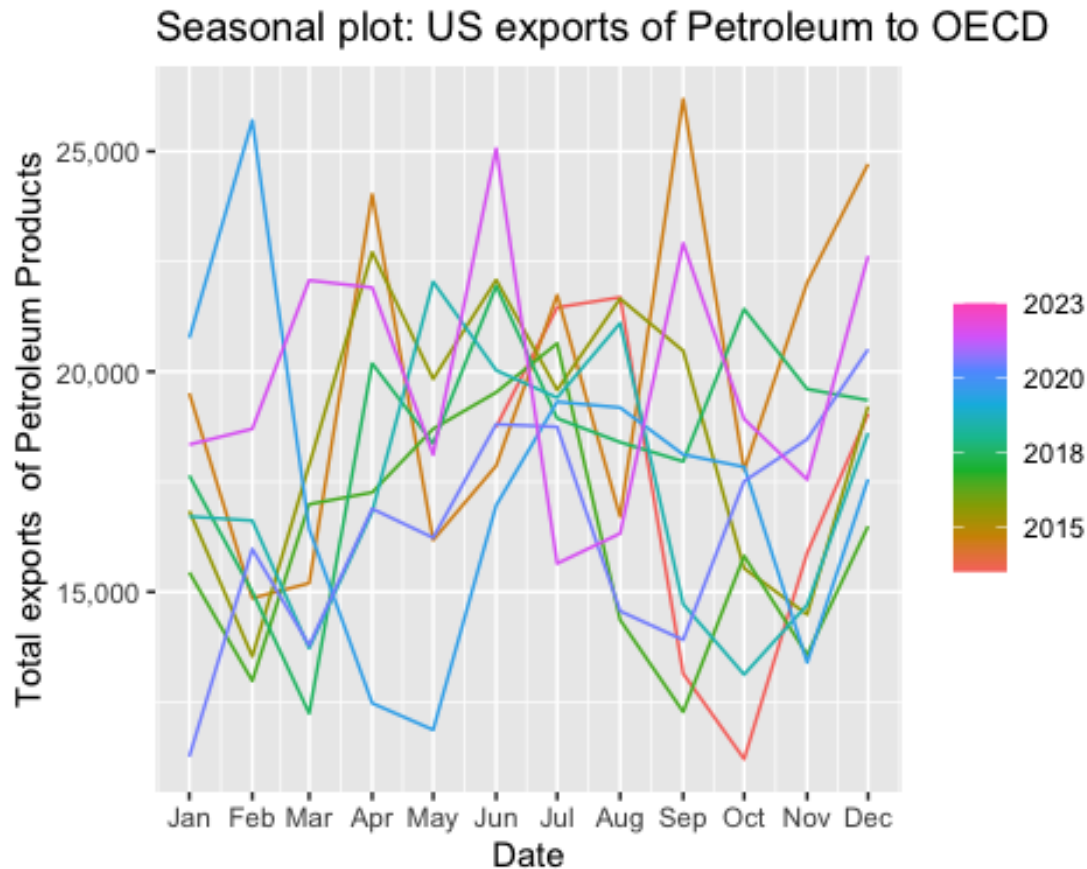
```
# create seasonal plot of total exports monthly of Natural gas
total_exp_oecd %>% filter(`Export Type` == "Amount of Natural gas (BOE)") %>%
  index_by(Date) %>%
  summarise(Total_Exports = sum(`Amount`)) %>%
  gg_season(Total_Exports) +
  ylab("Total exports of Natural Gas") +
  ggtitle("Seasonal plot: US exports of natural gas to OECD") +
  scale_y_continuous(labels = comma_format())
```



```
# create seasonal plot of total exports monthly of Crude oil
total_exp_oecd %>% filter(`Export Type` == "Amount of Crude Oil (Thousand
Barrels)") %>%
  index_by(Date) %>%
  summarise(Total_Exports = sum(`Amount`)) %>%
  gg_season(Total_Exports) +
  ylab("Total exports of Crude oil") +
  ggtitle("Seasonal plot: US exports of Crude Oil to OECD") +
  scale_y_continuous(labels = comma_format())
```



```
# create seasonal plot of total exports monthly of Petroleum products
total_exp_oecd %>% filter(`Export Type` == "Amount of total Petroleum
Products (Thousand Barrels)") %>%
  index_by(Date) %>%
  summarise(Total_Exports = sum(`Amount`)) %>%
  gg_season(Total_Exports) +
  ylab("Total exports of Petroleum Products") +
  ggtitle("Seasonal plot: US exports of Petroleum to OECD") +
  scale_y_continuous(labels = comma_format())
```



To further elaborate, the data analysis of US exports of oil products to OECD shows that seasonal patterns can differ depending on the type of product being exported. While natural gas and crude oil both display some level of seasonality, petroleum products do not show any significant trends.

In the case of natural gas, we can see a clear increase in exports during the colder months, which could be due to increased demand for heating purposes. Conversely, during the warmer months, exports tend to decrease. This pattern is consistent throughout the years, with the exception of some yearly fluctuations.

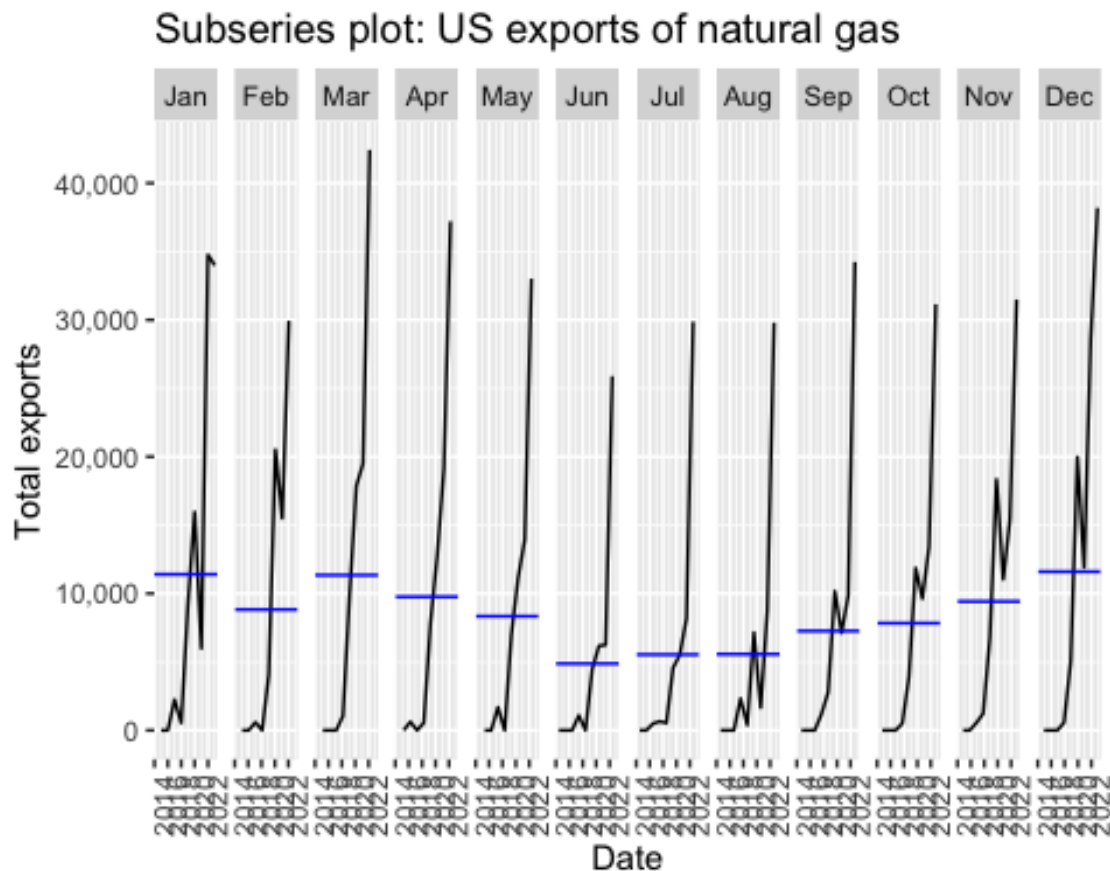
Regarding crude oil, the seasonal pattern is not as clear as with natural gas. However, we still observe a similar trend where exports tend to increase during the colder months and decrease during the warmer months. Interestingly, we see a deviation from this pattern in 2022, with exports remaining high even during spring and summer months. This could be attributed to the political tensions caused by the Russian invasion in Ukraine.

On the other hand, the export of petroleum products does not appear to have any notable seasonal patterns. While there may be a slight increase towards the end of each year, there is no clear trend observed year-over-year.

Overall, analyzing seasonal patterns in the US export of oil products can provide insights into demand patterns and other underlying factors that affect the global oil market.

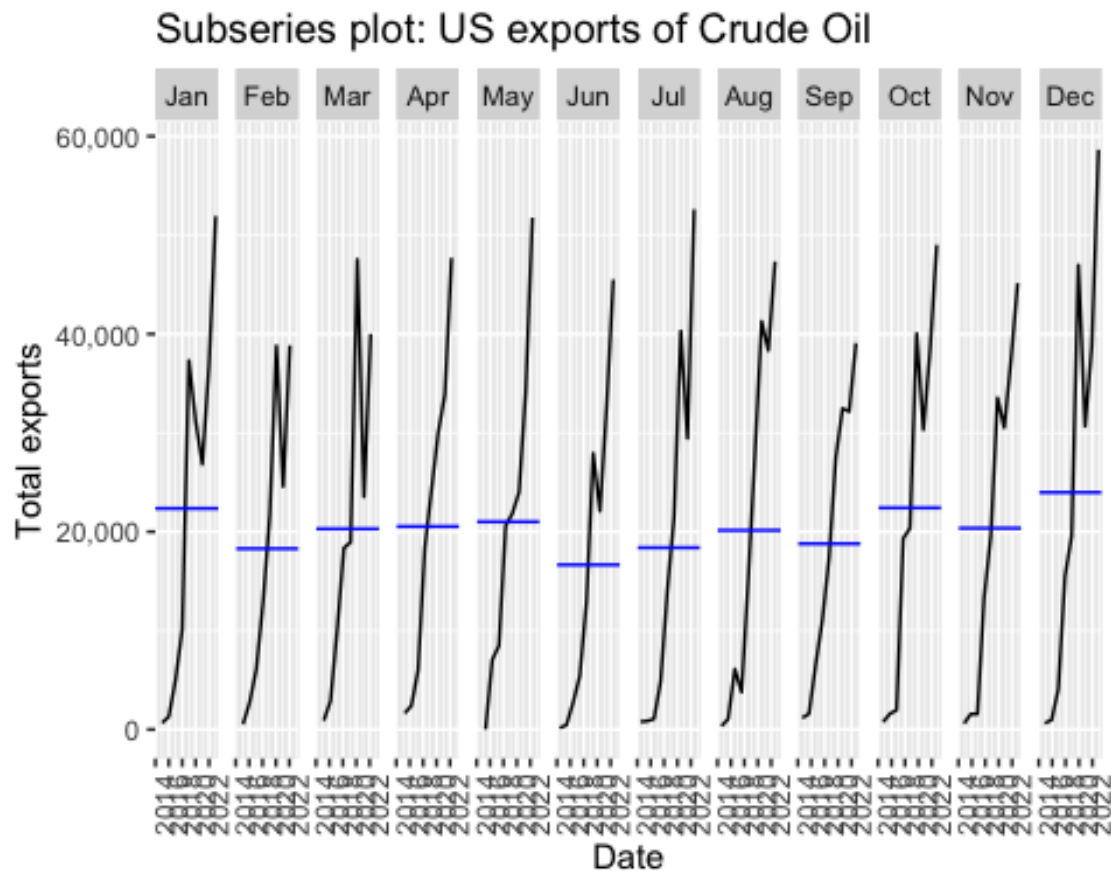
#Subseries plots Natural gas

```
total_exp_oecd %>% filter(`Export Type` == "Amount of Natural gas (BOE)") %>%
  index_by(Date) %>%
  summarise(Total_Exports = sum(`Amount`)) %>%
  gg_subseries(Total_Exports) +
  ylab("Total exports") +
  ggtitle("Subseries plot: US exports of natural gas")+
  scale_y_continuous(labels = comma_format())
```

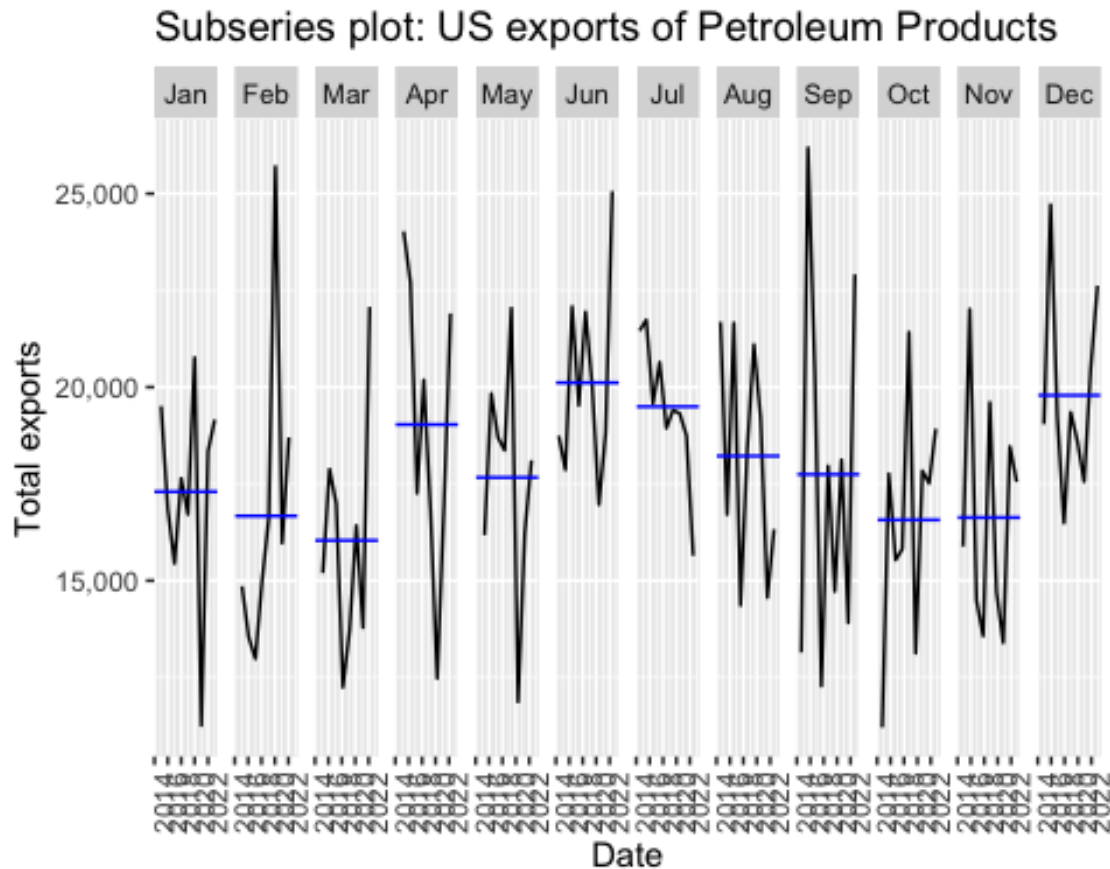


#Subseries plots Crude oil

```
total_exp_oecd %>% filter(`Export Type` == "Amount of Crude Oil (Thousand Barrels)") %>%
  index_by(Date) %>%
  summarise(Total_Exports = sum(`Amount`)) %>%
  gg_subseries(Total_Exports) +
  ylab("Total exports") +
  ggtitle("Subseries plot: US exports of Crude Oil")+
  scale_y_continuous(labels = comma_format())
```



```
#Subseries plots Petroleum products
total_exp_oecd %>% filter(`Export Type` == "Amount of total Petroleum
Products (Thousand Barrels)") %>%
  index_by(Date) %>%
  summarise(Total_Exports = sum(`Amount`)) %>%
  gg_subseries(Total_Exports) +
  ylab("Total exports") +
  ggtitle("Subseries plot: US exports of Petroleum Products")+
  scale_y_continuous(labels = comma_format())
```

To gain a better understanding of seasonality, we created additional plots known as subseries seasonal plots. These graphs provide a clearer view of monthly exports and also display the means over each month for the analyzed time series period.

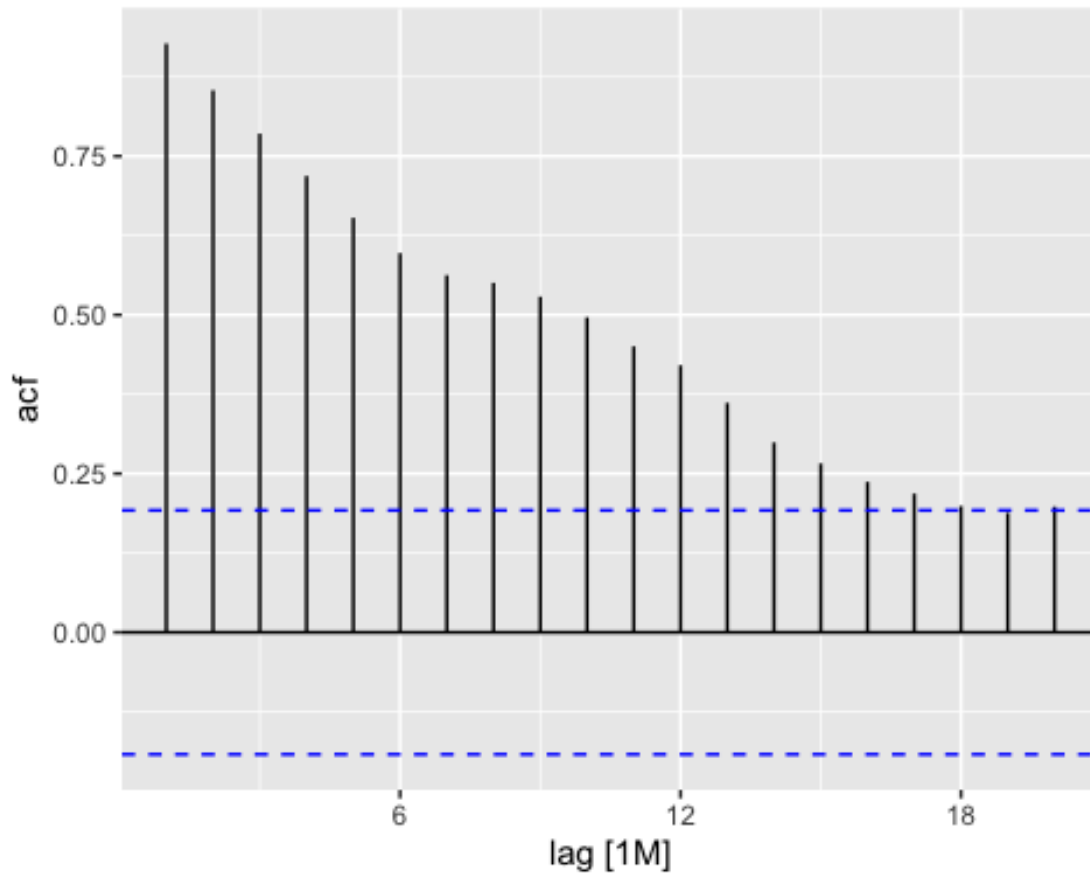
When we look at the subseries seasonal plot for natural gas exports, we can see a clear seasonality of increased natural gas exports during the colder seasons, starting from September and decreasing in spring. From the crude oil subseries seasonal plot, we can notice a huge increase in December and January as well, indicating that these months have a strong seasonal effect on crude oil exports. Finally, from the subseries seasonal plot of Petroleum products, we can notice the same increase in December, but there is a huge decrease in February and March. All other months are fluctuating more randomly without displaying any clear seasonality.

In general, the seasonal subseries plots for all three types of oil exports to OECD reveal similar patterns to those seen in the seasonal plots. There is a clear increase in exports during the colder seasons, particularly in December and January, and a decrease during the warmer seasons. However, there are some fluctuations and irregularities throughout the year that do not follow a clear seasonal pattern.

####2.4 ACF and White Noise

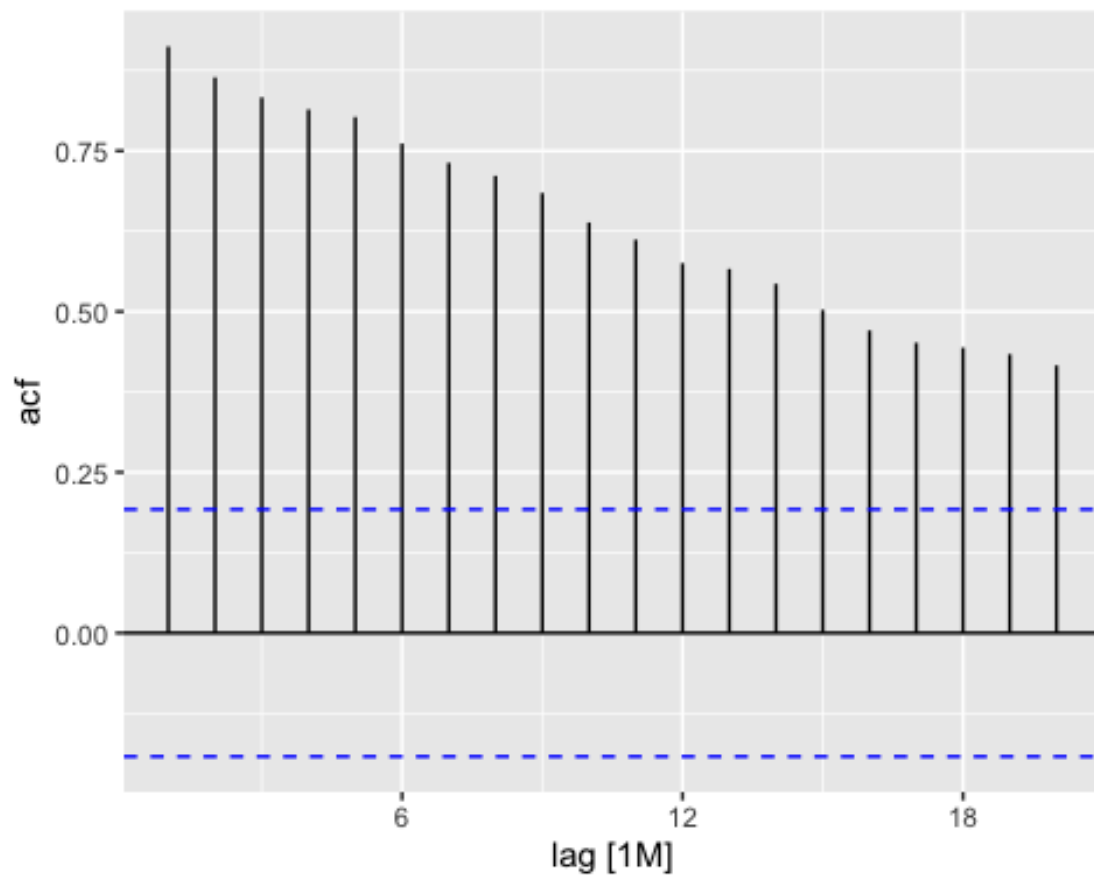
#ACF plot without Lag Natural Gas

```
total_exp_oecd %>% filter(`Export Type` == "Amount of Natural gas (BOE)") %>%
  index_by(Date) %>%
  summarise(Total_Exports = sum(`Amount`)) %>%
  ACF() %>% autoplot()
```



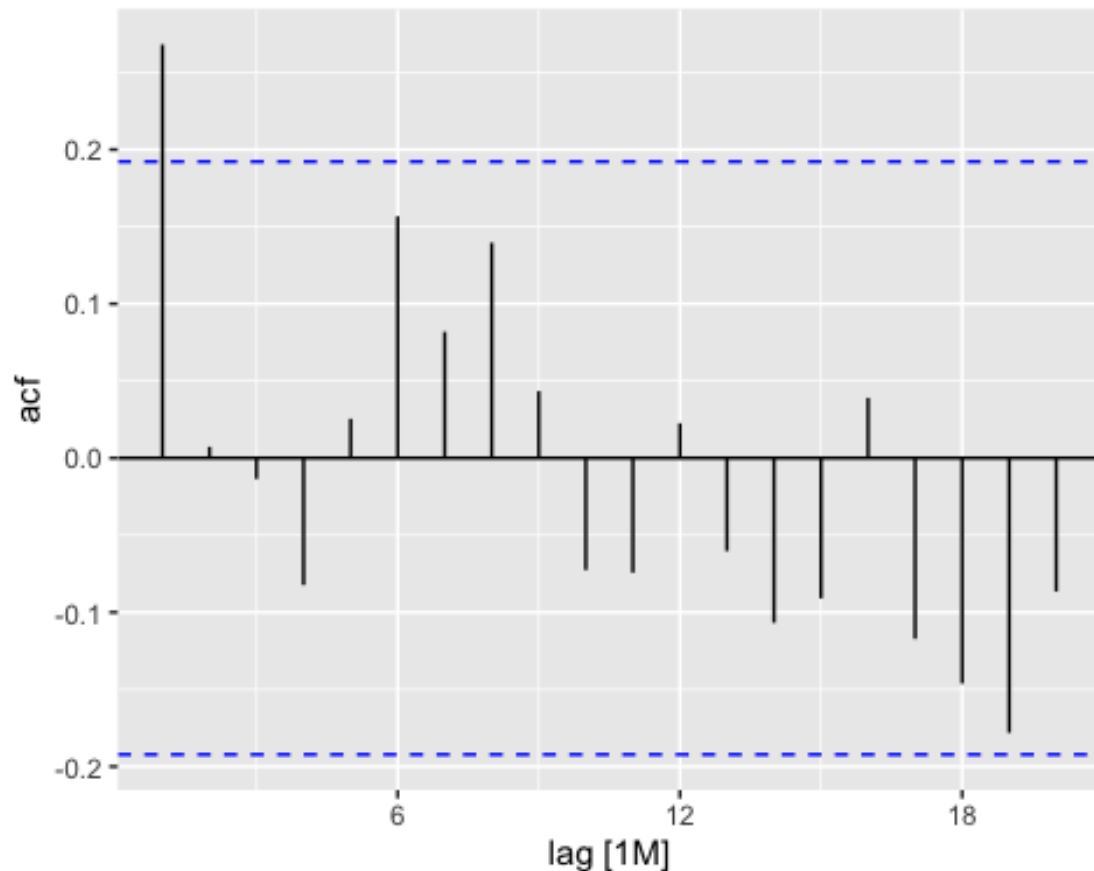
#ACF plot without Lag Crude Oil

```
total_exp_oecd %>% filter(`Export Type` == "Amount of Crude Oil (Thousand
Barrels)") %>%
  index_by(Date) %>%
  summarise(Total_Exports = sum(`Amount`)) %>%
  ACF() %>% autoplot()
```



#ACF plot without Lag Petroleum

```
total_exp_oecd %>% filter(`Export Type` == "Amount of total Petroleum
Products (Thousand Barrels)") %>%
  index_by(Date) %>%
  summarise(Total_Exports = sum(`Amount`)) %>%
  ACF() %>% autoplot()
```



The autocorrelation function (ACF) plot displays the correlation between a time series and its own lagged values. A high ACF value for a particular lag indicates that the current values of the time series are significantly correlated with the past values at that lag.

The ACF plots for US exports of Natural Gas, Crude Oil to OECD reveal interesting insights regarding the seasonality and trends in the time series. In the case of Natural Gas and Crude Oil, the ACF plots demonstrate a significant positive correlation, indicating the presence of high trend component with slight seasonality patterns observed as. However, these plots reflect a dominant upward trend, as evidenced by the substantial increase in exports. The trend component seems to be overtaking the seasonality, suggesting that the overall upward trend in exports is more prominent than the periodic variation in the data.

The ACF plot for US exports of Petroleum Products to OECD indicates that there is no noticeable seasonality in the time series. The first three lags show significant positive correlations, suggesting some short-term dependence between current exports and those from 1 to 3 months ago. However, beyond the third lag, the ACF values fall within the confidence interval, indicating a weakening of the correlation. This suggests that there is no clear and consistent pattern of seasonality in the exports of petroleum products.

Finally, based on the ACF plots of the US exports of natural gas, crude oil, and petroleum products to OECD, we can conclude that the time series do not appear to be white noise.

3. Time series decomposition

3.1. Adjustments

After a thorough review of the data, we determined that no adjustments were necessary. We did not observe any calendar or population effects on the monthly exports data, and the exports data is already presented in barrels or barrels of oil equivalent (BOE), so no inflation adjustments were required.

3.2. Transformations

Given the nature of the three data sets that are being used: Crude Oil Exports, Petroleum exports and Natural gas exports we decided that certain transformations might be necessary for further analysis. Not normal distribution of the data and differing variance were seen as potential threat for correct conclusions. We identified Box-Cox transformation as necessary to continue since it stabilizes the variance of the data and makes the distribution more normal. This will improve the reliability of the statistical tests and provide more accurate results.

For Box-Cox transformation to be performed we found lambda for each time series and then use it to transform the time series.

```
lambda1 <- total_exp_oecd %>%
  filter(`Export Type` == "Amount of Crude Oil (Thousand Barrels)") %>%
  index_by(Date) %>%
  summarise(Total_Exports = sum(Amount)) %>%
  features(Total_Exports, features = guerrero) %>%
  pull(lambda_guerrero)

crude_oil_bx <- total_exp_oecd %>%
  filter(`Export Type` == "Amount of Crude Oil (Thousand Barrels)") %>%
  index_by(Date) %>%
  summarise(Total_Exports = sum(Amount)) %>%
  mutate(Total_Exports = box_cox(Total_Exports + 1, lambda1))

crude_oil_bx

## # A tsibble: 104 x 2 [1M]
##       Date Total_Exports
##       <mth>         <dbl>
## 1 2014 Jun           13.9
## 2 2014 Jul           26.7
## 3 2014 Aug           19.3
## 4 2014 Sep           31.2
## 5 2014 Oct           26.4
## 6 2014 Nov           24.1
## 7 2014 Dec           24.5
## 8 2015 Jan           25.0
## 9 2015 Feb           22.9
```

```
## 10 2015 Mar          28.2
## # ... with 94 more rows

lambda2 <- total_exp_oecd %>%
  filter(`Export Type` == "Amount of total Petroleum Products (Thousand
Barrels)") %>%
  index_by(Date) %>%
  summarise(Total_Exports = sum(Amount)) %>%
  features(Total_Exports, features = guerrero) %>%
  pull(lambda_guerrero)

pet_prod_bx <-total_exp_oecd %>%
  filter(`Export Type` == "Amount of total Petroleum Products (Thousand
Barrels)") %>%
  index_by(Date) %>%
  summarise(Total_Exports = sum(Amount)) %>%
  mutate(Total_Exports = box_cox(Total_Exports + 1, lambda2))

lambda3 <- total_exp_oecd %>%
  filter(`Export Type` == "Amount of Natural gas (BOE)") %>%
  index_by(Date) %>%
  summarise(Total_Exports = sum(Amount)) %>%
  features(Total_Exports, features = guerrero) %>%
  pull(lambda_guerrero)

nat_gas_bx <-total_exp_oecd %>%
  filter(`Export Type` == "Amount of Natural gas (BOE)") %>%
  index_by(Date) %>%
  summarise(Total_Exports = sum(Amount)) %>%
  mutate(Total_Exports = box_cox(Total_Exports + 1, lambda3))

nat_gas_bx <- filter(nat_gas_bx, Date >= as.Date("2016-01-01"))
```

3.3. Decomposition

###STL Decomposition

```
dcmp_crude_oil <- crude_oil_bx %>%
  model(stl = STL(Total_Exports))

components(dcmp_crude_oil)

## # A dable: 104 x 7 [1M]
## # Key:      .model [1]
## # :      Total_Exports = trend + season_year + remainder
##   .model      Date Total_Exports trend season_year remainder season_adjust
##   <chr>      <mth>      <dbl> <dbl>      <dbl>      <dbl>      <dbl>
## 1 stl      2014 Jun          13.9  24.0        -5.85        -4.28        19.7
## 2 stl      2014 Jul          26.7  24.1        -5.38         8.00        32.1
## 3 stl      2014 Aug          19.3  24.1        -1.83        -2.96        21.2
```

```
## 4 stl      2014 Sep      31.2  24.2      3.74      3.24      27.4
## 5 stl      2014 Oct      26.4  24.3      2.83     -0.785     23.5
## 6 stl      2014 Nov      24.1  24.4     -1.81      1.52     26.0
## 7 stl      2014 Dec      24.5  24.6      0.841    -0.900     23.7
## 8 stl      2015 Jan      25.0  24.8      0.374    -0.118     24.7
## 9 stl      2015 Feb      22.9  25.0      0.775    -2.87      22.1
## 10 stl     2015 Mar      28.2  25.2      4.84     -1.86      23.3
## # ... with 94 more rows
```

```
dcmp_oil_prod <- pet_prod_bx %>%
  model(stl = STL(Total_Exports))
```

```
components(dcmp_oil_prod)
```

```
## # A dable: 104 x 7 [1M]
## # Key:      .model [1]
## # :      Total_Exports = trend + season_year + remainder
##   .model      Date Total_Exports trend season_year remainder season_adjust
##   <chr>      <nth>      <dbl> <dbl>      <dbl>      <dbl>      <dbl>
## 1 stl      2014 Jun      548.  508.      35.4       4.41      513.
## 2 stl      2014 Jul      594.  511.      44.4       38.1      549.
## 3 stl      2014 Aug      597.  514.      17.6       65.9      580.
## 4 stl      2014 Sep      445.  517.      -5.36      -66.5      450.
## 5 stl      2014 Oct      405.  520.      -34.6      -80.1      439.
## 6 stl      2014 Nov      497.  522.      -22.0      -3.25      519.
## 7 stl      2014 Dec      553.  525.      30.6       -2.55      523.
## 8 stl      2015 Jan      561.  529.      -10.2      43.0       572.
## 9 stl      2015 Feb      478.  532.      -44.7      -8.88      523.
## 10 stl     2015 Mar      485.  535.      -44.0      -6.38      528.
## # ... with 94 more rows
```

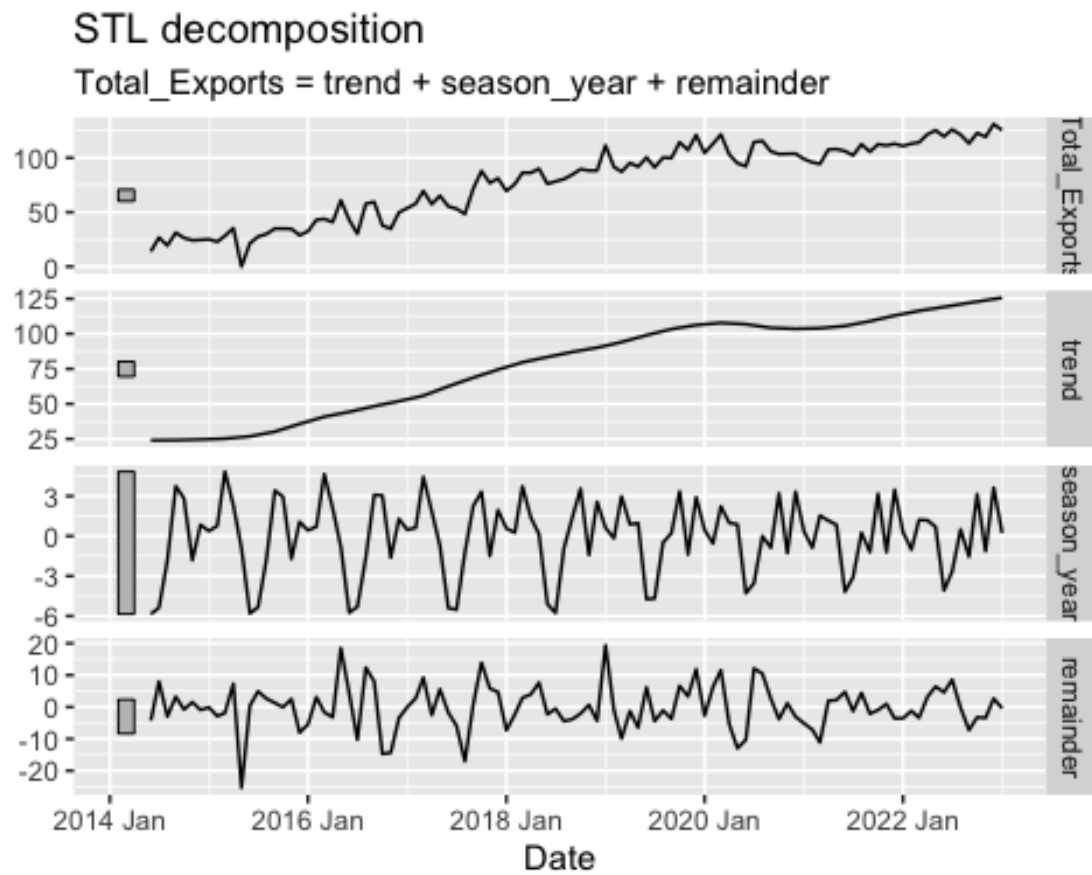
```
dcmp_nat_gas <- nat_gas_bx %>%
  model(stl = STL(Total_Exports))
```

```
components(dcmp_nat_gas)
```

```
## # A dable: 85 x 7 [1M]
## # Key:      .model [1]
## # :      Total_Exports = trend + season_year + remainder
##   .model      Date Total_Exports trend season_year remainder season_adjust
##   <chr>      <nth>      <dbl> <dbl>      <dbl>      <dbl>      <dbl>
## 1 stl      2016 Jan      0      1.43      10.3      -11.7      -10.3
## 2 stl      2016 Feb      0      2.72     -0.425     -2.30       0.425
## 3 stl      2016 Mar      0      4.02      4.69      -8.71      -4.69
## 4 stl      2016 Apr      28.9  5.31      3.60      20.0       25.3
## 5 stl      2016 May      0      6.60     -3.05     -3.56       3.05
## 6 stl      2016 Jun      0      7.90     -13.8      5.88       13.8
## 7 stl      2016 Jul      26.2  9.20     -5.88      22.8       32.0
## 8 stl      2016 Aug      0     10.6     -10.7      0.105      10.7
## 9 stl      2016 Sep      0     11.9     -1.78     -10.2       1.78
```

```
## 10 stl      2016 Oct           0  13.3      -1.16    -12.2           1.16
## # ... with 75 more rows

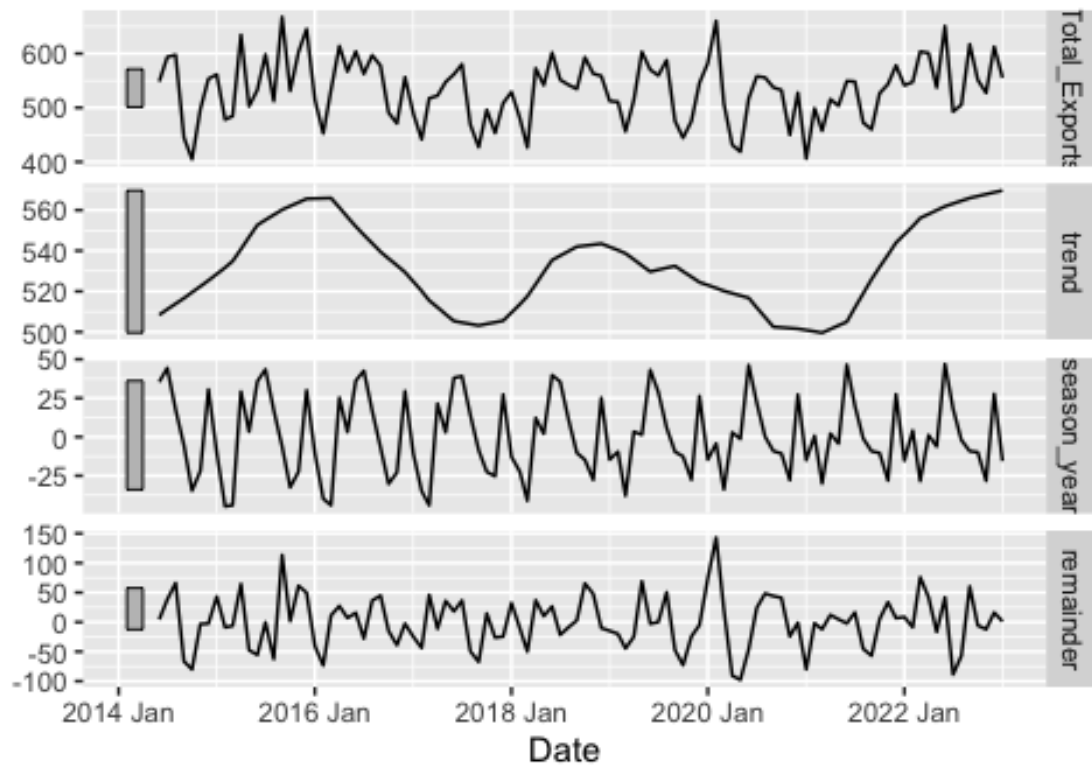
components(dcmp_crude_oil) %>% autoplot()
```



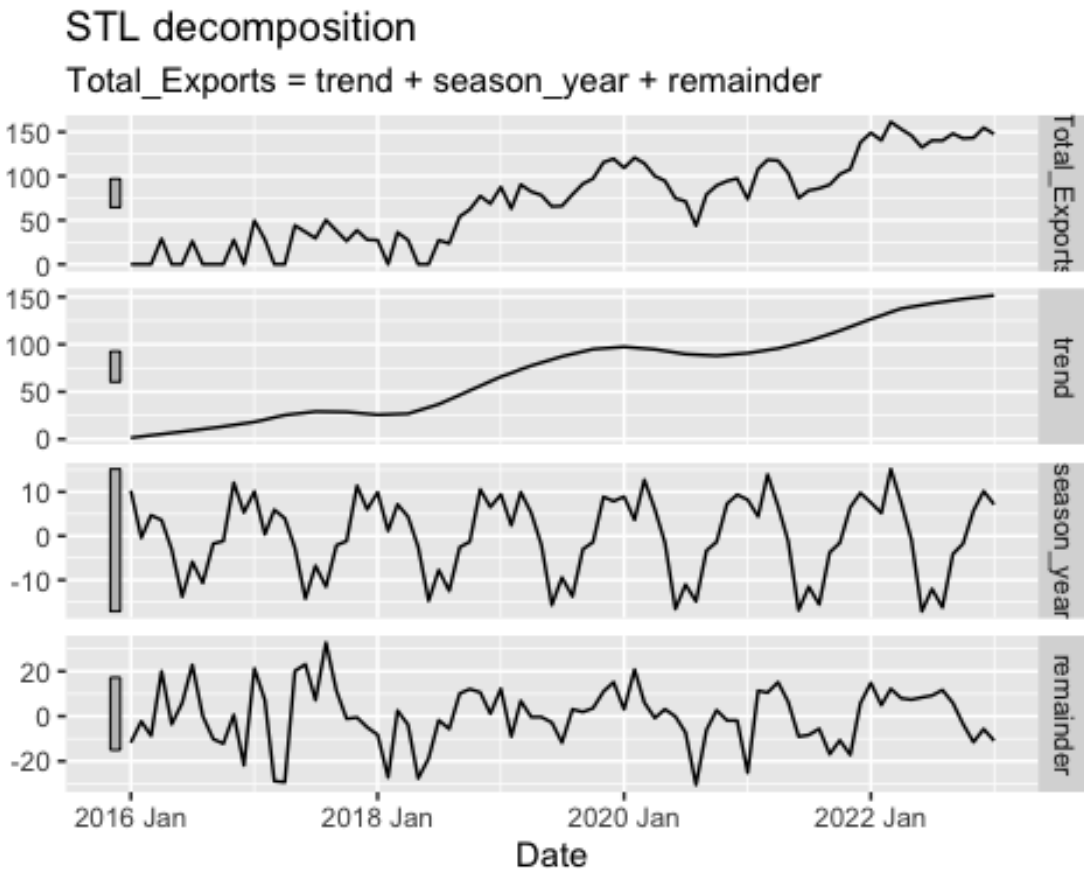
```
components(dcmp_oil_prod) %>% autoplot()
```


STL decomposition

Total_Exports = trend + season_year + remainder



```
components(dcmp_nat_gas) %>% autoplot()
```



We used STL decomposition to decompose the Box-Cox transformed time series into three components: trend, seasonal, and residual. Four graphs were plotted, three of which display the trend, seasonal, and residual components, respectively while fourth one, the one at the top, represent the export amounts from 2013 to 2023.

Looking at the graphs we see that in case of natural gas exports actual patterns can only be seen appearing after approximately 2016 when the U.S.

There are several conclusions to be drawn looking at the STL decomposition. First of all, we can see clear trend in exports of crude oil and natural gas while petroleum products does not have clear trend. In both cases of exports of crude oil and natural gas we can clearly see slight drops in exports after 2020 which can be explained by COVID 19 pandemic and continued growths after situation normalized. Another observation worth mentioning is the fact that exports of two mentioned products hits all time high around the year 2022. This can be explained by huge declines of both products from Russia after sanction were imposed.

Moreover, we can see seasonality in exports of all three products. That is no surprise, given that all energy products are highly reactive to different seasons.

After decomposing the data of all three time series and looking at each components separately we decided that exports of petroleum products will not be analysed further.

Exports of petroleum products to OECD Europe markets does not have clear trend therefore indicating that Russian invasion in Ukraine did not have positive effect on growing imports from US. This can be explained by the fact that embargo on refined oil products from EU 27 was introduced as of 5 February 2023. Therefore, Russian war on Ukraine and sanctions on Russia that followed did not have any visible effect on US exports of petroleum products to OECD Europe markets.

4. The forecaster's toolbox

4.1. Finding the fit

When choosing from four possible simple forecasting methods the intuition tells us that the most appropriate method in our case is Drift method. Clear trend is seen in both crude oil and natural gas exports data and drift method allows the forecasts to increase or decrease over time based on the change in the historical data.

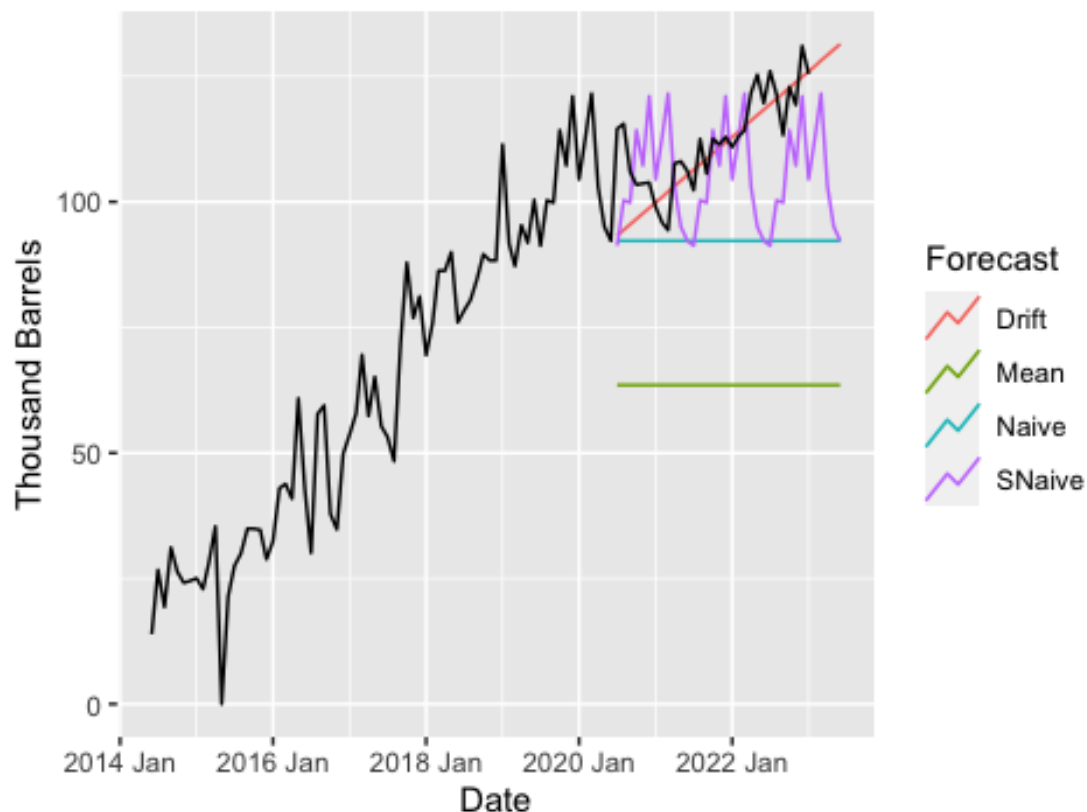
```
crude_oil_tr <- crude_oil_bx %>%
  filter(Date >= as.Date("2010-06-01"), Date <= as.Date("2020-06-01"))

crude_oil_fit <- crude_oil_tr %>%
  model(
    Mean = MEAN(Total_Exports),
    Naive = NAIVE(Total_Exports),
    SNaive = SNAIVE(Total_Exports),
    Drift = RW(Total_Exports ~ drift())
  )

crude_oil_fc <- crude_oil_fit %>%
  forecast(h = 36)

crude_oil_fc %>%
  autoplot(
    crude_oil_bx,
    level = NULL
  ) +
  labs(
    y = "Thousand Barrels",
    title = "Forecasts for U.S. Crude Oil exports to OECD Europe"
  ) +
  guides(colour = guide_legend(title = "Forecast"))
```

Forecasts for U.S. Crude Oil exports to OECD Europe



```

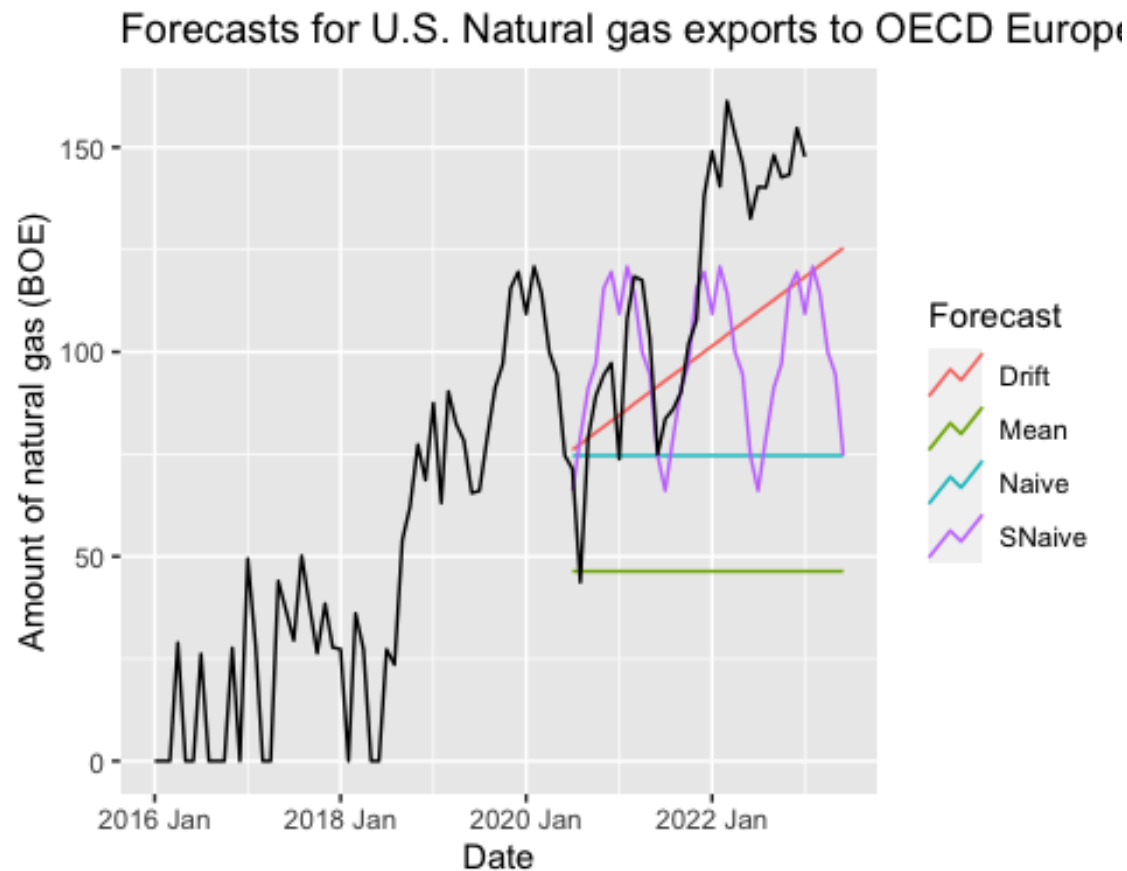
nat_gas_tr <- nat_gas_bx %>%
  filter(Date <= as.Date("2020-06-01"))

nat_gas_fit <- nat_gas_tr %>%
  model(
    Mean = MEAN(Total_Exports),
    Naive = NAIVE(Total_Exports),
    SNaive = SNAIVE(Total_Exports),
    Drift = RW(Total_Exports ~ drift())
  )

nat_gas_fc <- nat_gas_fit %>%
  forecast(h = 36)

nat_gas_fc %>%
  autoplot(
    nat_gas_bx,
    level = NULL) +
  labs(
    y = "Amount of natural gas (BOE)",
    title = "Forecasts for U.S. Natural gas exports to OECD Europe") +
  guides(colour = guide_legend(title = "Forecast"))

```



Only looking at plots in Figure [] we can see that our intuition was right and drift method still seems as most appropriate after conducting visual inspection. Though, it is important to highlight that if in case of crude oil exports Drift method was actually quite good at forecasting export amounts, in the case of natural gas exports Drift method looked as best of all but still quite far away from actual data. The important takeaway at this point is the fact that the average growth rate calculated from historical data was strongly effected by dropped exports during the COVID pandemic.

After inspecting data visually we will conduct cross validation to see if our initial predictions that Drift method is most appropriate to forecast both time series was right.

#Traditional accuracy

```
accuracy1 <- accuracy(crude_oil_fit) %>%
  arrange(.model) %>%
  select(.model, .type, RMSE, MAE, MAPE, MASE, RMSSE)
```

```
accuracy1
```

```
## # A tibble: 4 × 7
##   .model .type      RMSE    MAE  MAPE  MASE RMSSE
##   <chr>  <chr>    <dbl> <dbl> <dbl> <dbl> <dbl>
```

```
## 1 Drift Training 11.3 8.63 Inf 0.509 0.538
## 2 Mean Training 30.9 27.4 Inf 1.62 1.47
## 3 Naive Training 11.4 8.68 Inf 0.512 0.540
## 4 SNaive Training 21.0 17.0 24.7 1 1
```

```
accuracy2 <- accuracy(nat_gas_fit) %>%
  arrange(.model) %>%
  select(.model, .type, RMSE, MAE, MAPE, MASE, RMSSE)
```

```
accuracy2
```

```
## # A tibble: 4 × 7
##   .model .type      RMSE    MAE  MAPE  MASE RMSSE
##   <chr>  <chr>    <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 Drift Training 19.3 15.4 Inf 0.439 0.489
## 2 Mean Training 38.3 33.3 Inf 0.946 0.972
## 3 Naive Training 19.3 15.1 Inf 0.430 0.490
## 4 SNaive Training 39.4 35.2 Inf 1 1
```

#Time series cross validation

```
crude_oil_tr2 <- crude_oil_bx %>%
  stretch_tsibble(.init = 3, .step = 1) %>%
  relocate(Date, Total_Exports, .id)
```

```
crude_oil_tr2 %>%
  model(Drift = RW(Total_Exports ~ drift())) %>%
  forecast(h = 12) %>%
  accuracy(crude_oil_bx)
```

```
## # A tibble: 1 × 10
##   .model .type    ME  RMSE  MAE  MPE  MAPE  MASE  RMSSE  ACF1
##   <chr>  <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 Drift Test -2.03 14.9 10.9 -Inf Inf 0.708 0.786 0.716
```

```
crude_oil_bx |>
  model(RW(Total_Exports ~ drift())) |>
  accuracy()
```

```
## # A tibble: 1 × 10
##   .model .type    ME  RMSE  MAE  MPE  MAPE  MASE  RMSSE
##   <chr>  <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 RW(Total_Exports ~ Trai... -3.97e-15 10.2 7.65 -Inf Inf 0.496 0.539
## -0.319
```

4.3. Residual diagnostics

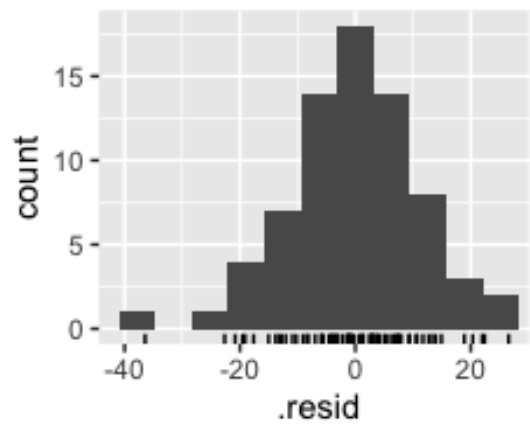
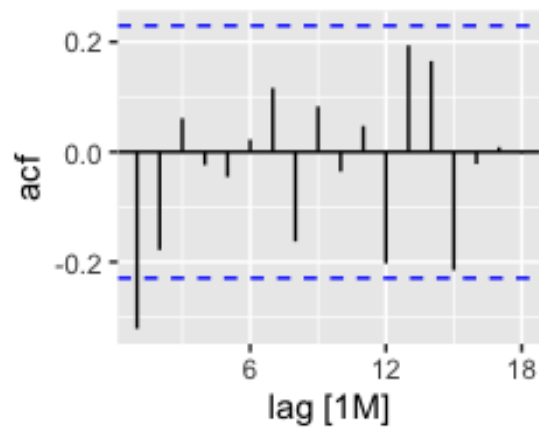
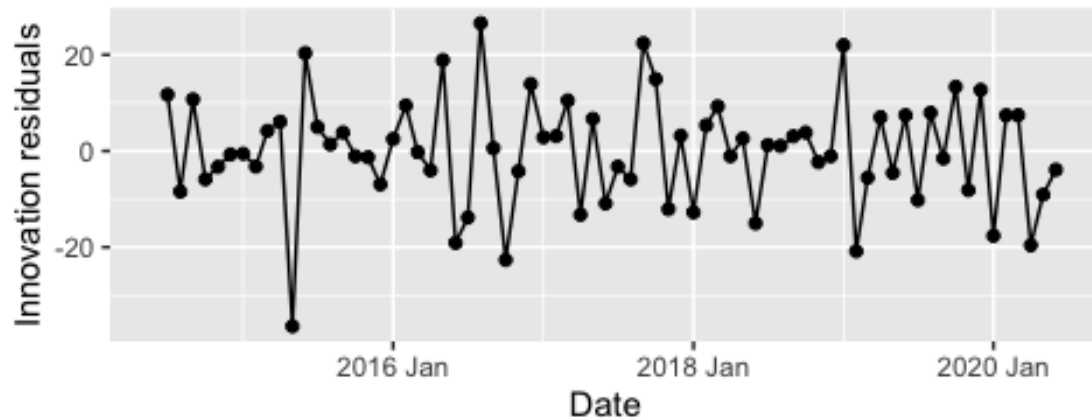
First we inspect residuals of drift and naive methods for crude oil exports. Both, according to previous analysis were most accurate.

#Inspecting residuals crude oil

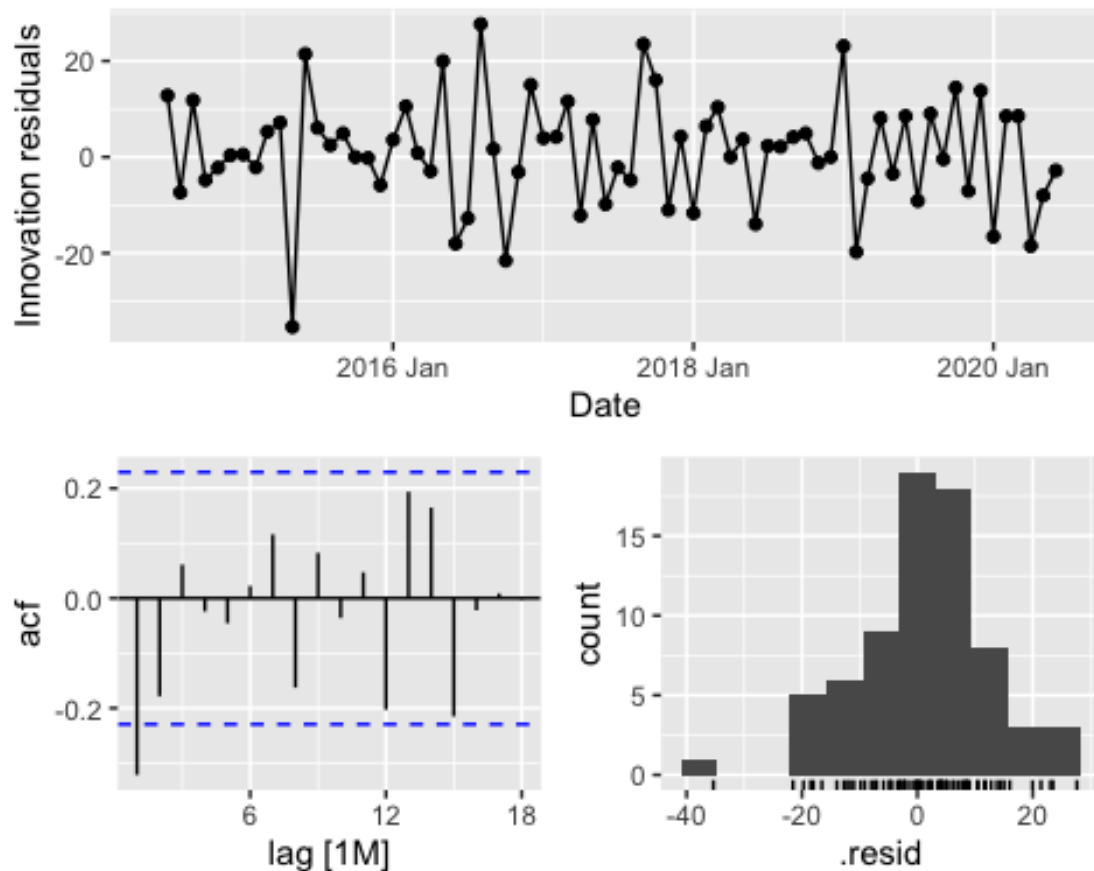
```
drift_crude_oil <- crude_oil_tr %>%  
  model(RW(Total_Exports ~ drift()))
```

```
naive_crude_oil <- crude_oil_tr %>%  
  model(NAIVE(Total_Exports))
```

```
gg_tsresiduals(drift_crude_oil)
```



```
gg_tsresiduals(naive_crude_oil)
```



The graphs demonstrate that the naive and drift forecasting methods provide crude oil export predictions that seem to capture all the relevant information available. The residuals, or the differences between the predicted values and actual values, have a mean close to zero and do not show any significant correlation, as evidenced by the lack of discernible patterns in the residual series. The time plot of the residuals indicates that their variance remains relatively constant across the historical data, except for one outlier at the beginning of the time series (which corresponds to the zeros observed during that time period). This constant residual variance is further supported by the histogram of the residuals.

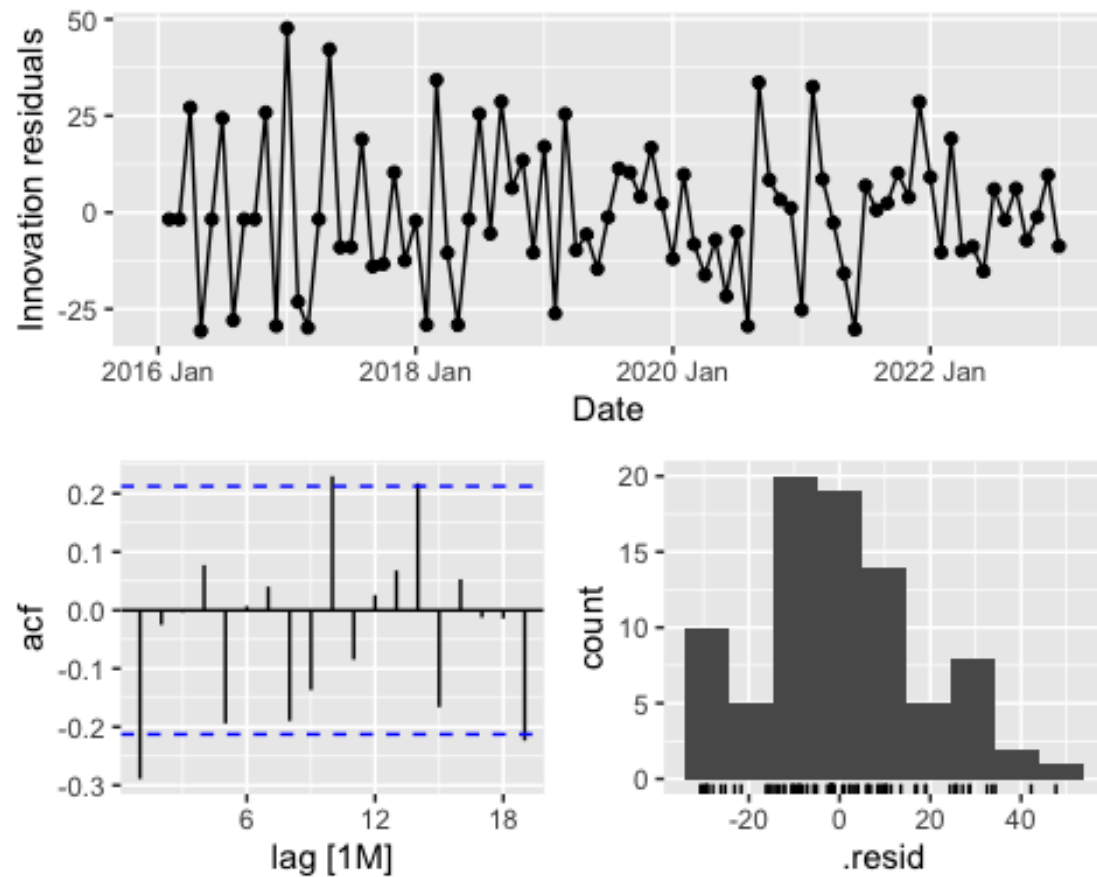
We conduct the same analysis with drift and naive methods on natural gas exports.

#Inspecting residuals nat gas

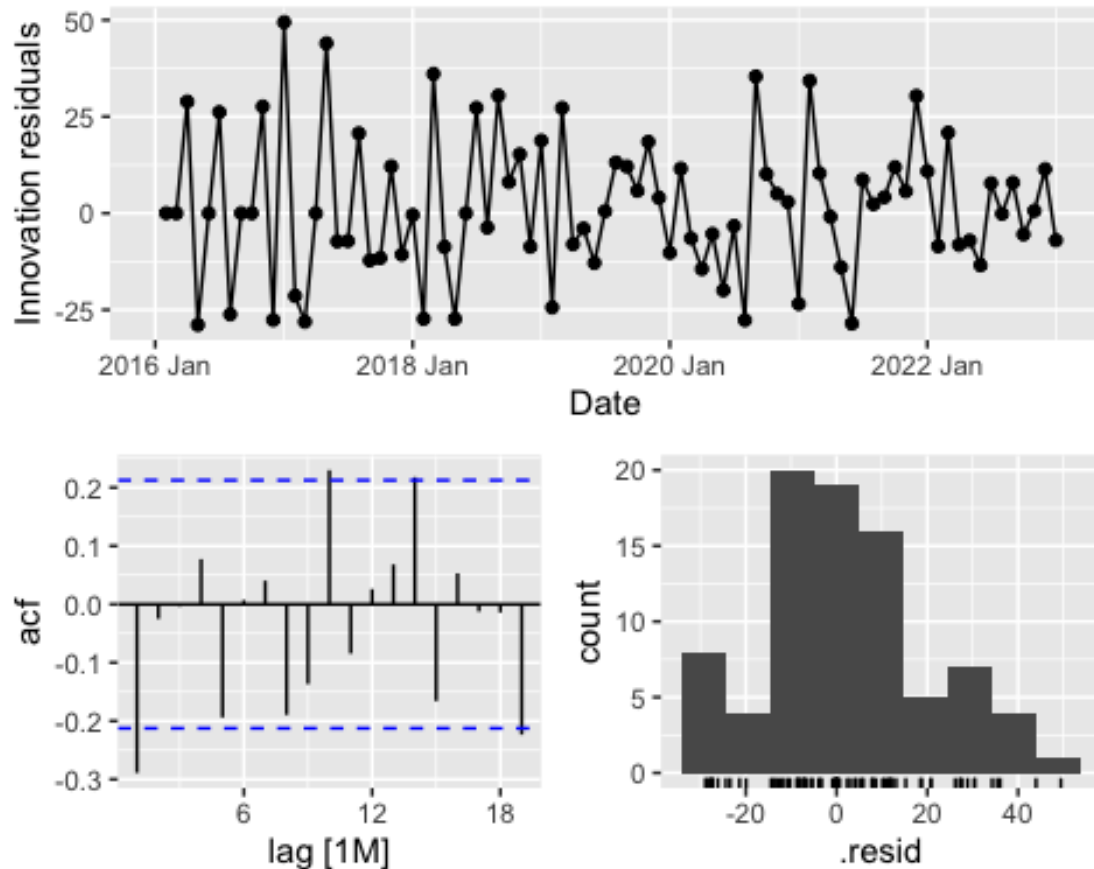
```
drift_nat_gas <- nat_gas_bx %>%
  model(RW(Total_Exports ~ drift()))

naive_nat_gas <- nat_gas_bx %>%
  model(NAIVE(Total_Exports))

gg_tsresiduals(drift_nat_gas)
```

```
gg_tsresiduals(naive_nat_gas)
```



Looking at the graphs we see that in cases of both methods the residuals seem to stay close to zero therefore implying that there is no correlation between the residuals and thus suggesting that the models are accurate in the predictions of the data.

5. Exponential smoothing

Yet to be started

6. ARIMA

Yet to be started

7. Other forecasting methods

Yet to be started