rail_transport.R

Portatil

2024-02-14

Data Loading The data has been obtained through the INE, total passengers transported per quarter. It is worth noting that the data has been cleaned through Excel. The information was obtained from a survey of freight and passenger transport companies. The database covers from the year 2010 to the year 2023, but given the presence of COVID, I have decided not to include it in the study due to its atypical component that would affect this investigation.

```
library(forecast)
```

```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
```

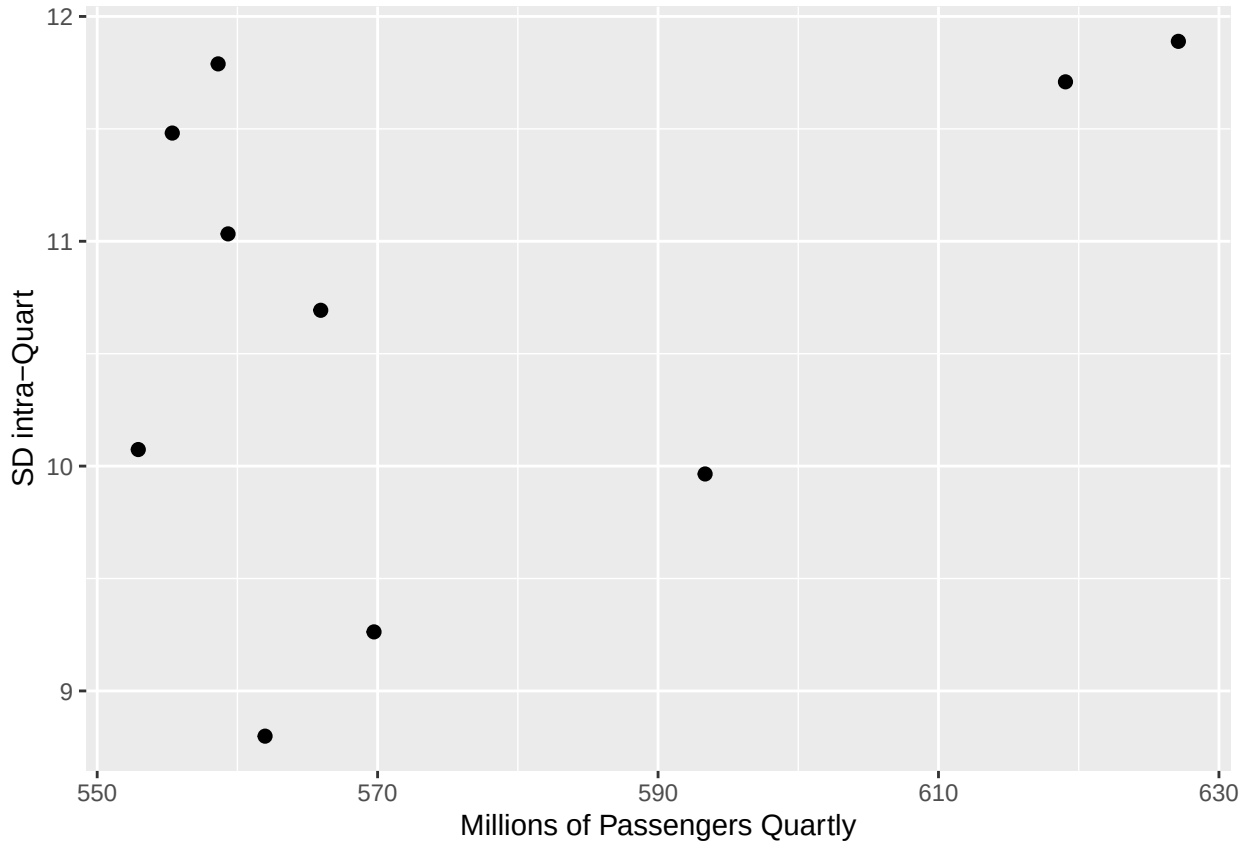
```
library(readxl)
library(ggplot2)

data <- read_excel("ferrocarril_trimestre.xlsx")
```

```
data$Trimestre <- as.integer(data$Trimestre)

ts_data <- ts(data$Total/1000000, start = c(2010, 1), end = c(2019, 4), frequency = 4)  # 4 Quarts per

CasosTrimestre = aggregate(ts_data, FUN = sum)
DesviacionTrimestre = aggregate(ts_data, FUN = sd)
```

Creating the Time Series I have calculated the sum and the standard deviation of the quarterly data and plotted this relationship to analyze its scheme type.

```
ggplot() +
  geom_point(aes(x = CasosTrimestre, y = DesviacionTrimestre), size = 2) +
  xlab("Millions of Passengers Quartly") +
  ylab("SD intra-Quart")
```

```
## Don't know how to automatically pick scale for object of type <ts>. Defaulting
## to continuous.
## Don't know how to automatically pick scale for object of type <ts>. Defaulting
## to continuous.
```

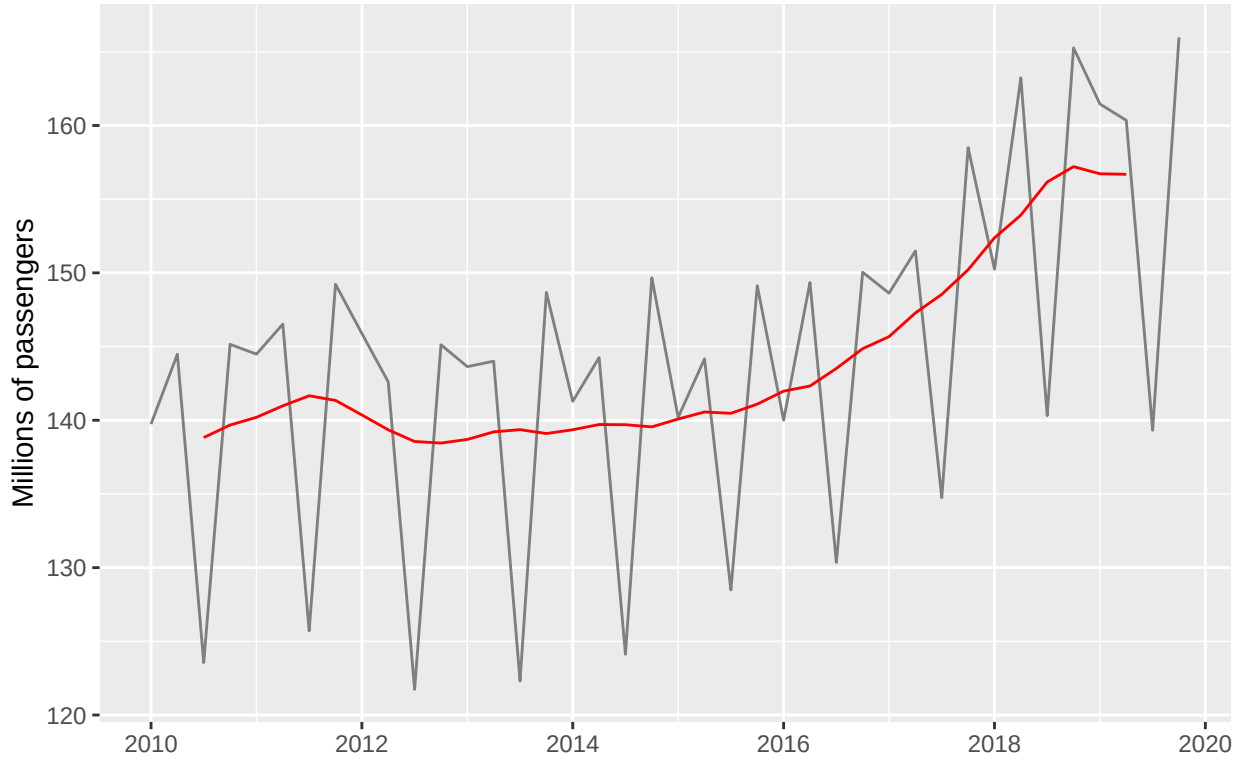



```
# It seems like an additive model because there is no additional standard
# deviation as passengers increase
```

```
pasDesAdi <- decompose(ts_data,
                       type = "addi")


autoplot(ts_data, series="Passengers",
         xlab = "",
         ylab = "Millions of passengers",
         main = "") +
  autolayer(trendcycle(pasDesAdi), series="Trend") +
  scale_colour_manual(values=c("Millions of passengers"="black","Trend"="red"),
                      breaks=c("Millions of passengers","Tendencia"))
```

Analysis of trend and seasonality.

```
## Warning: Removed 4 rows containing missing values (‘geom_line()‘).
```



The growing demand for rail transportation since 2016 may be attributed to the fact that in this year, the company Renfe (the main passenger and freight railway transport company in Spain) set the objective of "selling every seat on every train" through strategies of low prices and improvement of stations and service quality. They put this into practice with strategies such as the low-cost AVE service (known as EVA) in 2019, a five-year plan, and even with the support of the State, through the Rail Freight Transport Boost Plan 2017-2023.

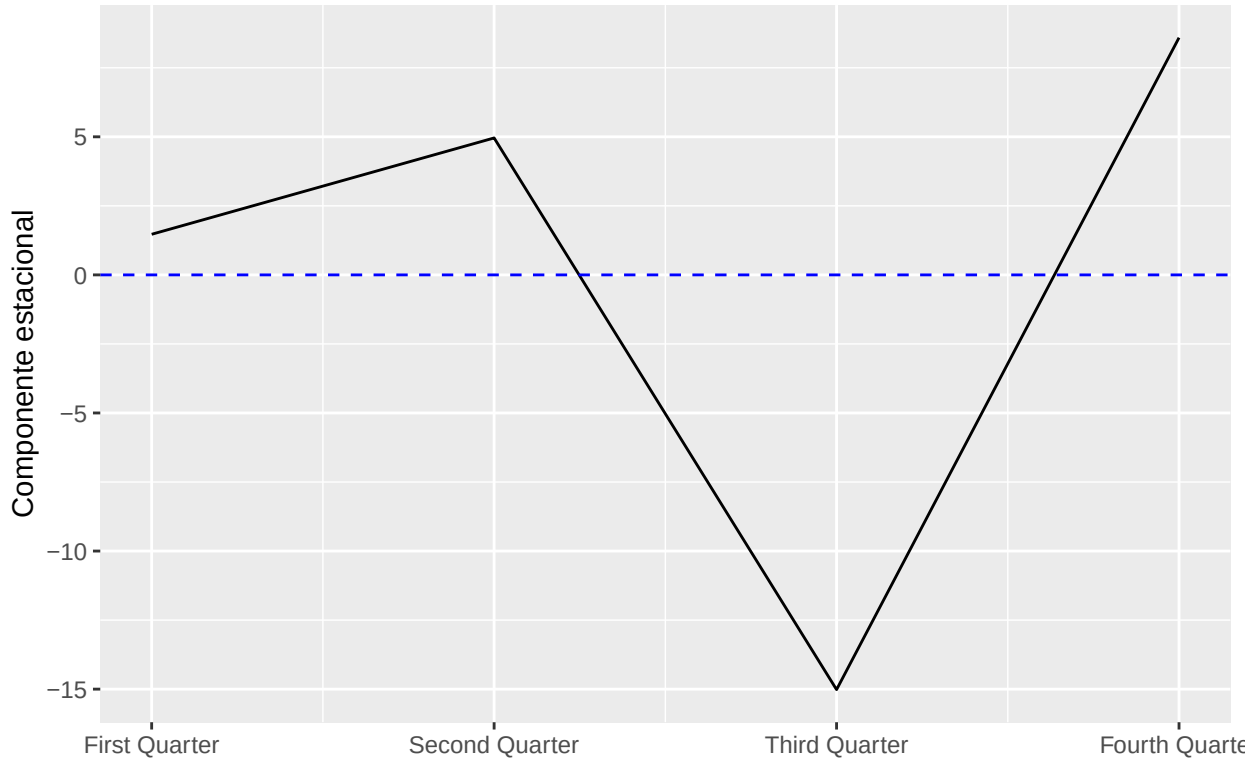
```
componenteEstacional <- tapply(ts_data - mean(ts_data), #respecto a la media
                               cycle(ts_data),
                               FUN = mean)

round(pasDesAdi$figure,0)
```

```
## [1]   2   5 -15   8
```

In the first quarter, there are 2M passengers above the average, in the second quarter 5M above, in the third quarter 15M below, and in the fourth quarter 8M above. The noticeable drop in the number of passengers during the summer may indicate that railway usage is mainly for work and school-related reasons, so demand drops during vacations.

```
ggplot() +
  geom_line(aes(x = 1:4, y = componenteEstacional)) +
  geom_hline(yintercept = 0, colour = "blue", lty = 2) +
  ggtitle("") +
  xlab("") +
```

```
  ylab("Componente estacional") +
  scale_x_continuous(breaks= 1:4,
                     labels = c("First Quarter", "Second Quarter", "Third Quarter", "Fourth Quarter"))
```



Conclusion In conclusion, the data shows a clear relationship between seasonality and demand for railway transportation, with peaks and valleys reflecting seasonal patterns and consumption trends. These findings are essential to understand and plan business strategies, adjusting services and fares according to seasonal demand, and ensuring efficient management of railway transportation resources and capacities.