

Reporte Series de Tiempo

Nicolas Cardenas A01114959

2022-11-17

Módulo 5: Estadística Avanzada para ciencia de datos y nombre de la concentración

Grupo 502

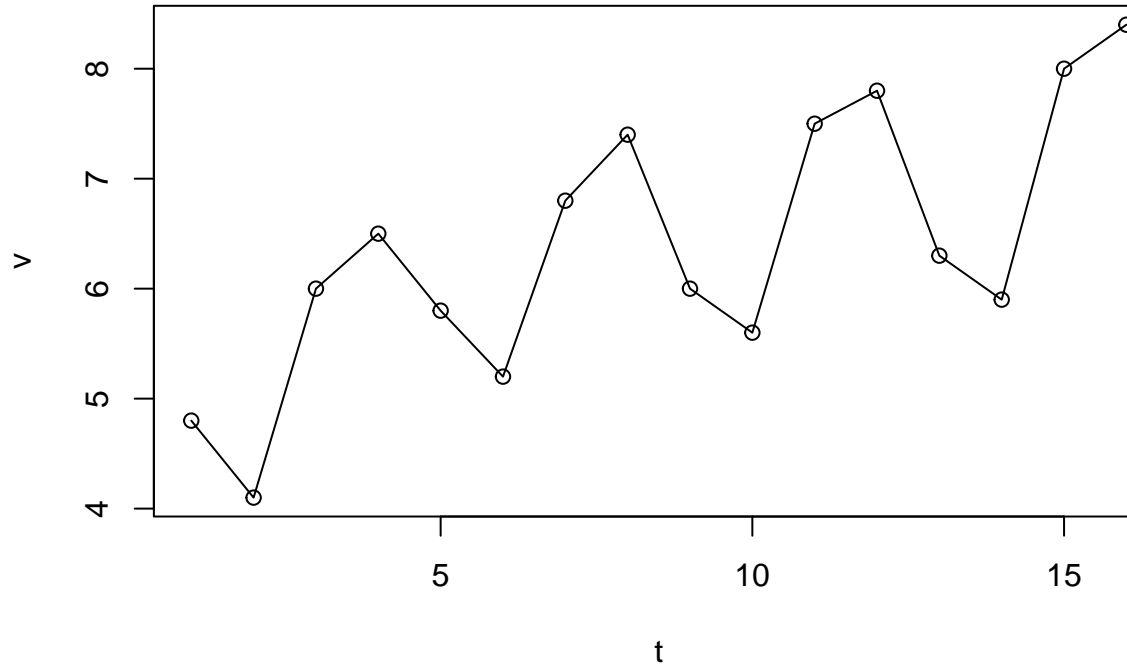
RESUMEN

En este reporte analizamos las ventas de televisores a lo largo del tiempo, utilizamos herramientas vistas en clase para descomponer esta serie de tiempo para poder observar comportamientos y convertir esta informacion a acciones que podrian ser buenos para el negocio. Utilizamos promedios moviles, los datos descompuestos y sin estacionalizacion para poder realizar un modelo que pueda hacer predicciones tomando en cuenta todos estos factores.

Contamos con dos situaciones: una que es la venta de televisores y otra que es la ventas de libros para la escuela. Para ambas tenemos informacion de ventas por periodo.

INTRODUCCION

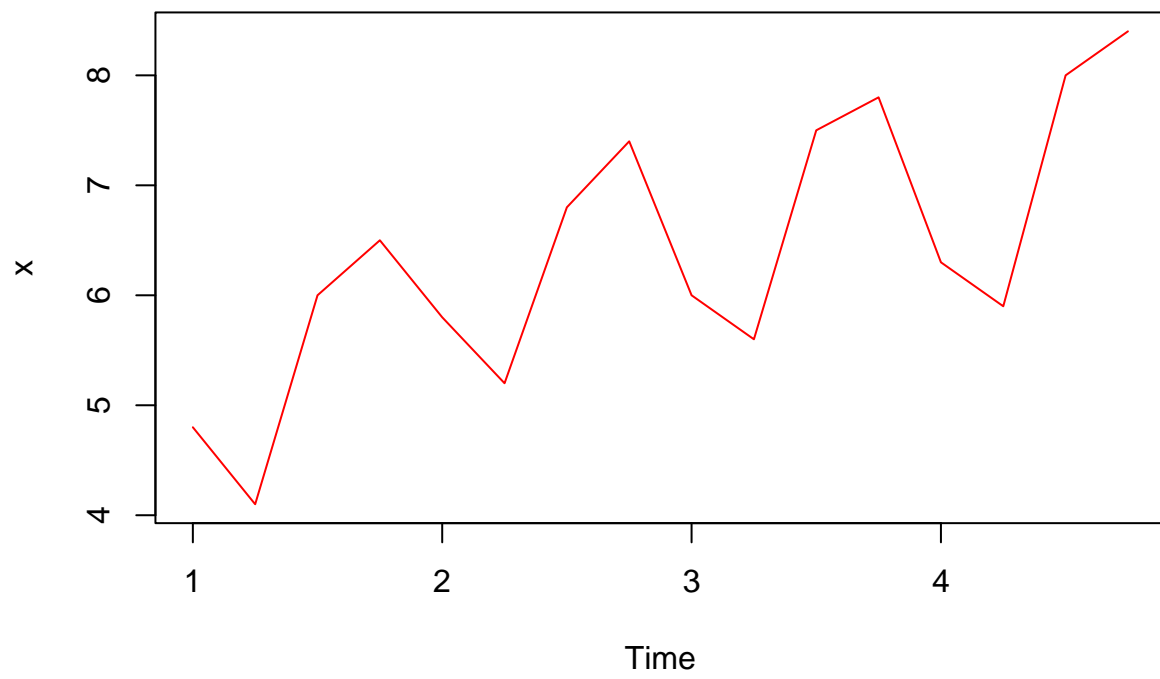
Primero que nada tenemos que entender los datos. Tenemos las ventas en miles en cada de 4 semestres por 1



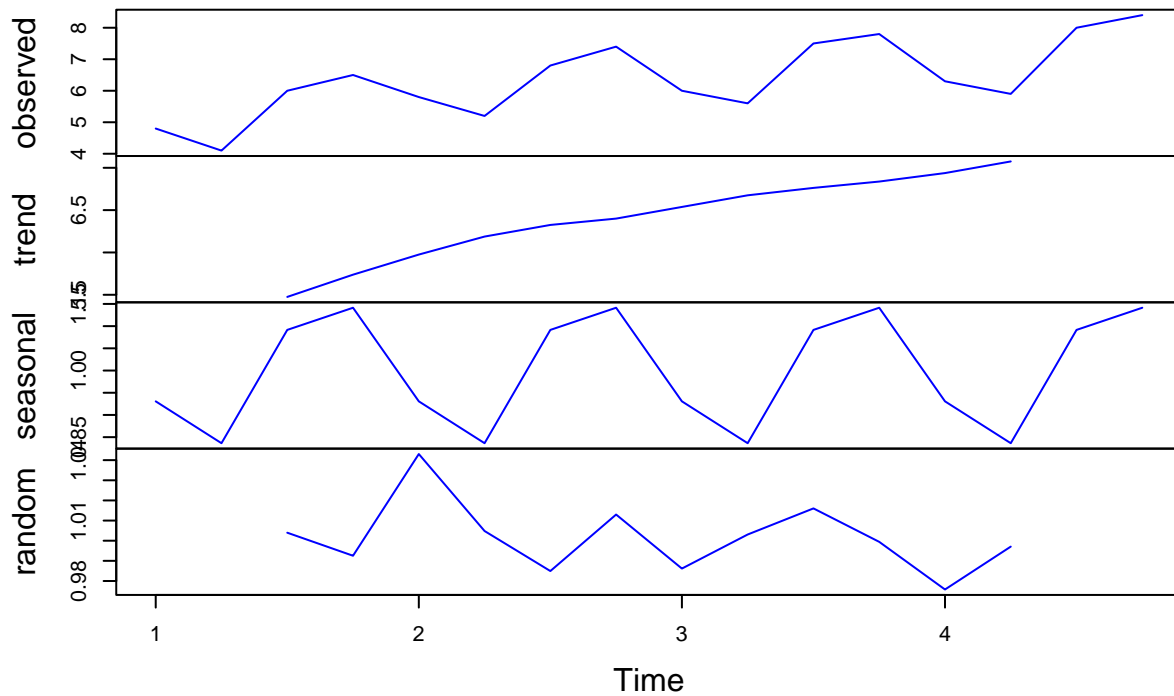
anio por 4 anios (total 16):

ANALISIS

Lo mas importante de las ventas es descomponerlo en sus diferentes componentes (trend, seasonality, irregularity). Usaremos el tipo multiplicativo porque vemos que esta aumentando cada vez mas nuestra grafica.



Decomposition of multiplicative time series

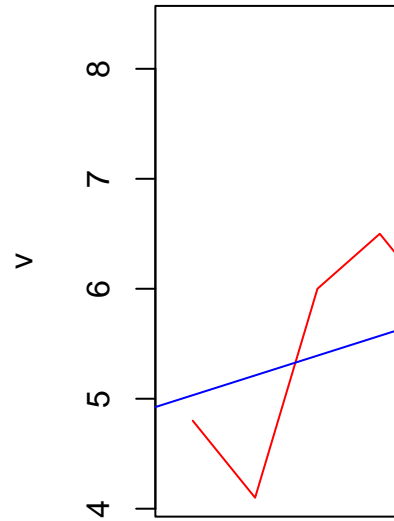


Como podemos ver, los diferentes componentes nos cuentan mucho de como se comporta la compra y venta de televisores: la tendencia es que esta incrementando, y con nuestro ciclo vemos que hay estacionalidad de cuatro, observamos diferentes comportamientos para cada trimestre de ventas.

```
##          Qtr1      Qtr2      Qtr3      Qtr4
## 1 0.9306617 0.8363763 1.0915441 1.1414179
## 2 0.9306617 0.8363763 1.0915441 1.1414179
## 3 0.9306617 0.8363763 1.0915441 1.1414179
## 4 0.9306617 0.8363763 1.0915441 1.1414179
```

Podemos ver que son similares pero la mejor temporada es el Q4 mientras que la peor es el Q2

PREDICCIONES

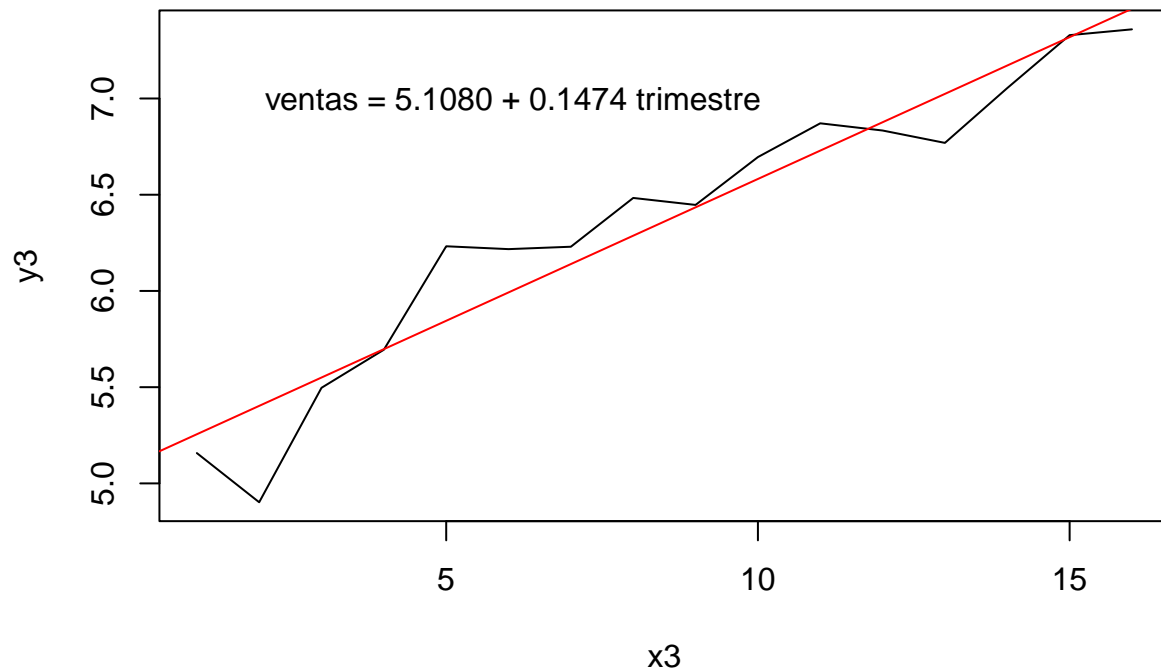


Primero veremos como se ve nuestra formula de regresion para observar el comportamiento:

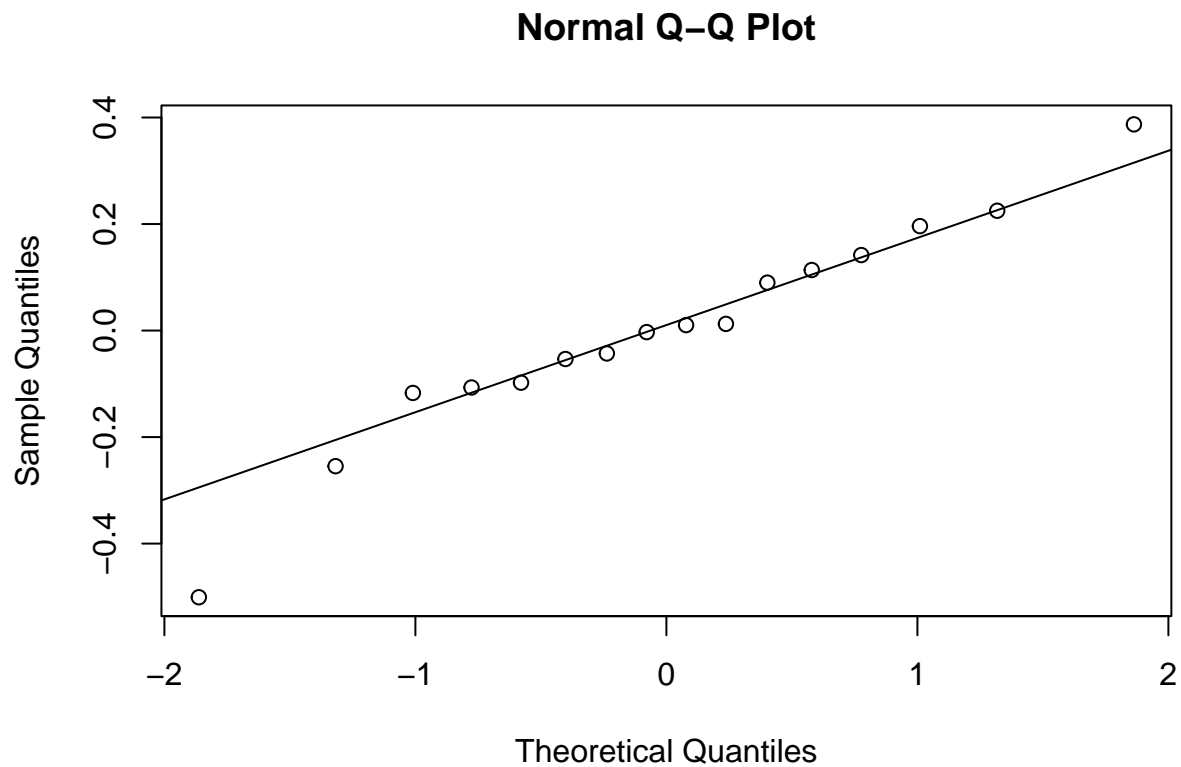
```
##  
## Call:  
## lm(formula = v ~ t)  
##  
## Coefficients:  
## (Intercept)          t  
##      4.8525      0.1799
```

Veremos ahora las ventas desestacionalizadas:

```
##  
## Call:  
## lm(formula = y3 ~ x3)  
##  
## Coefficients:  
## (Intercept)          x3  
##      5.1080      0.1474
```



```
##
## Call:
## lm(formula = y3 ~ x3)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.5007 -0.1001  0.0037  0.1207  0.3872
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  5.10804    0.11171   45.73  < 2e-16 ***
## x3           0.14738    0.01155   12.76 4.25e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.213 on 14 degrees of freedom
## Multiple R-squared:  0.9208, Adjusted R-squared:  0.9151
## F-statistic: 162.7 on 1 and 14 DF, p-value: 4.248e-09
```



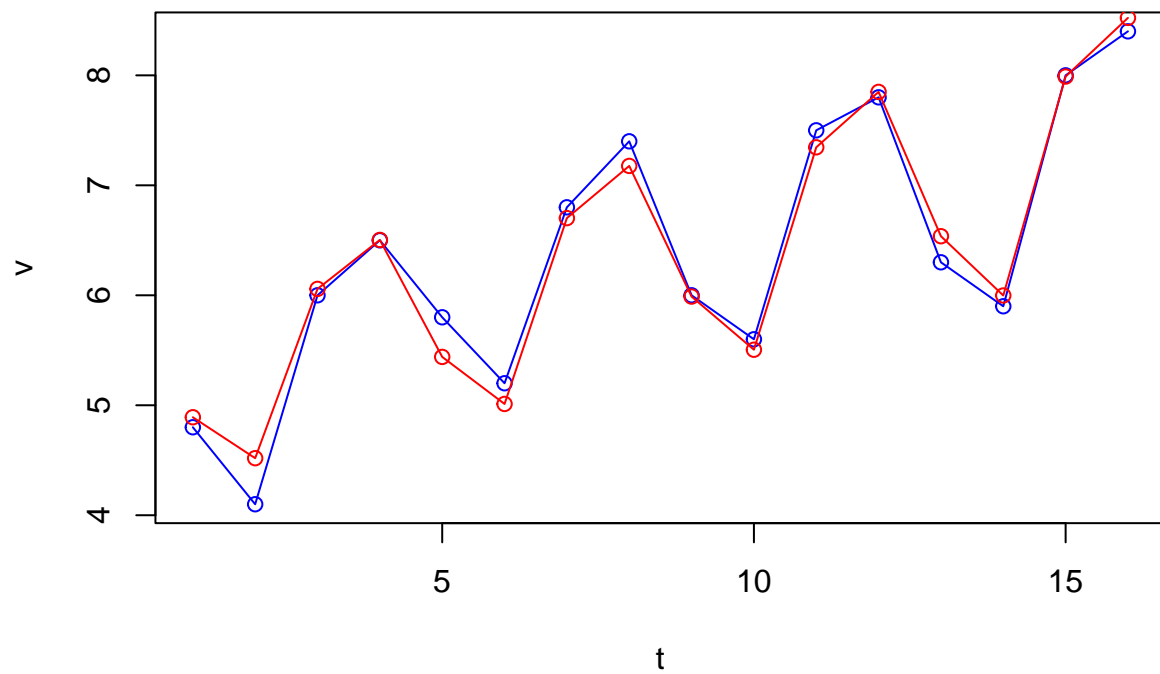
Nuestras “predicciones” siguen siendo lineales.

Con solo estos datos, incluso antes de volver la estacionalidad a nuestra predicciones, podemos ver que es un buen modelo. Tenemos que ambos coeficientes son significantes. Observamos normalidad en los residuos, y explica 92% de variabilidad el modelo. Es un buen modelo que servira.

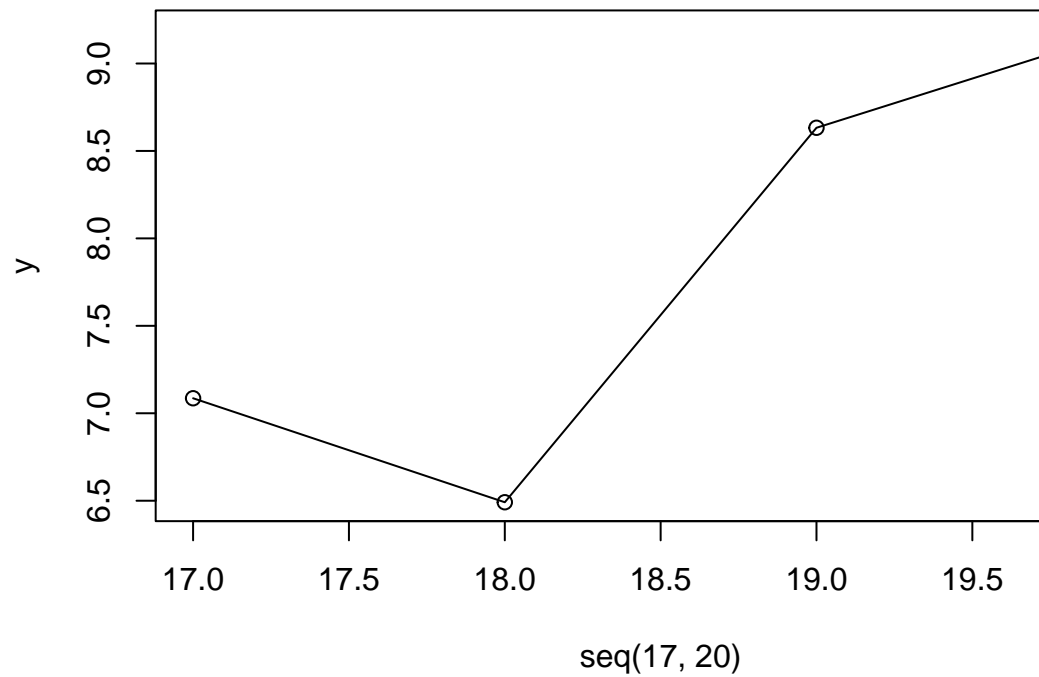
Hacemos predicciones y volvemos la estacionalidad para ver nuestros errores.

```
## [1] 0.03302078
```

```
## [1] 0.02439396
```



Desde esta grafica vemos que es un extremadamente buen modelo. Vemos que tanto el CME y el EPAM son extremadamente bajos, tenemos un excelente modelo.



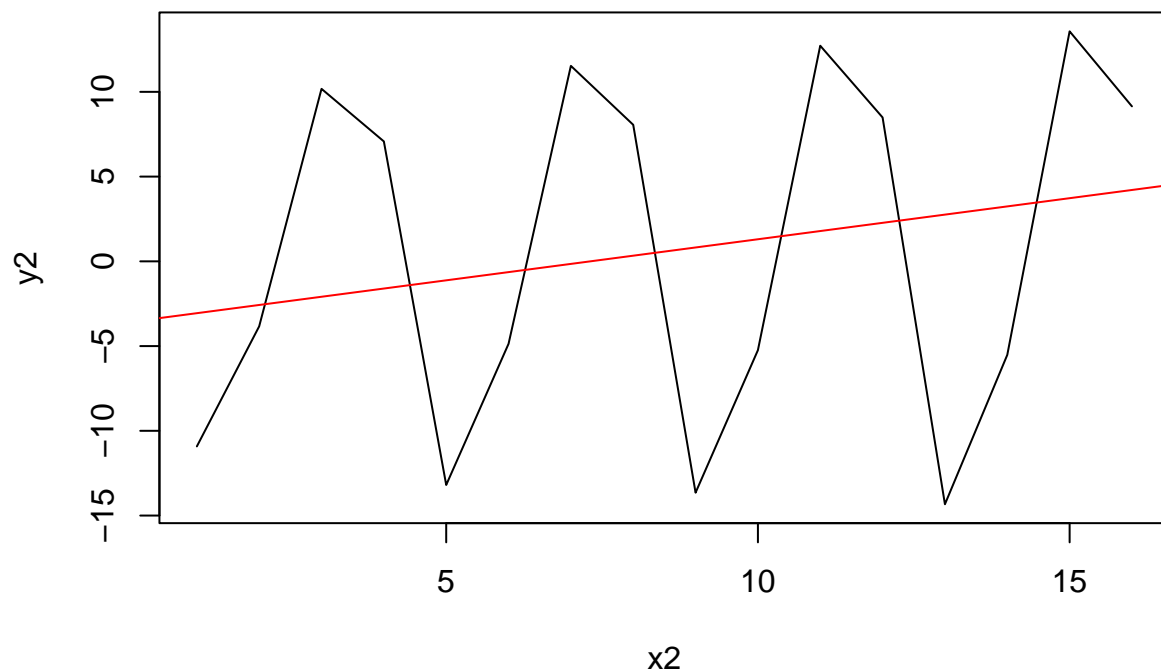
Predicciones para el proximo anio:

Aqui podemos notar el modelo y la estacionalizacion es su maximo esplendor, se nota el efecto en cada trimestre de nuestras predicciones.

COMPARACION

A continuacion: haremos lo mismo que acabamos de hacer pero ahora utilizaremos el modelo de manera aditiva con el proposito de comparar modelos.

```
##
## Call:
## lm(formula = y2 ~ x2)
##
## Coefficients:
## (Intercept)          x2
##      -3.5443         0.4847
```



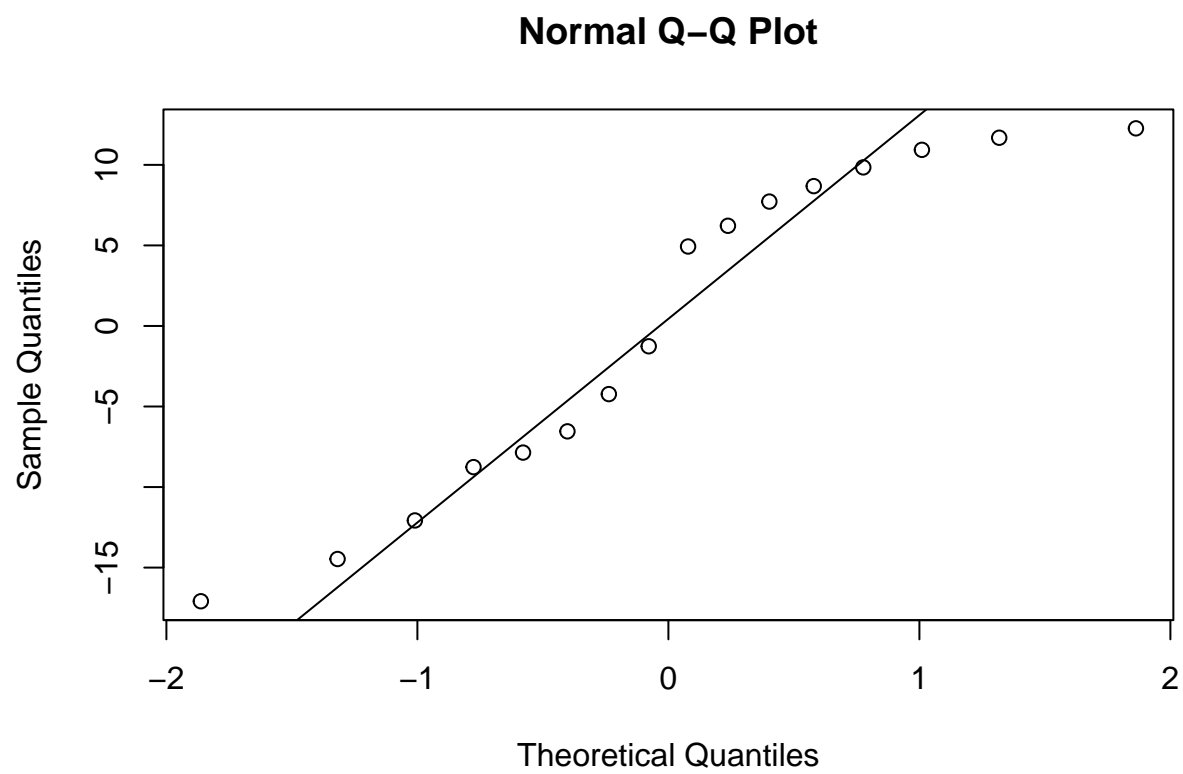
```
##
## Call:
## lm(formula = y2 ~ x2)
##
## Residuals:
```

	Min	1Q	Median	3Q	Max
	-17.088	-8.085	1.836	8.971	12.267

```
##
## Coefficients:
```

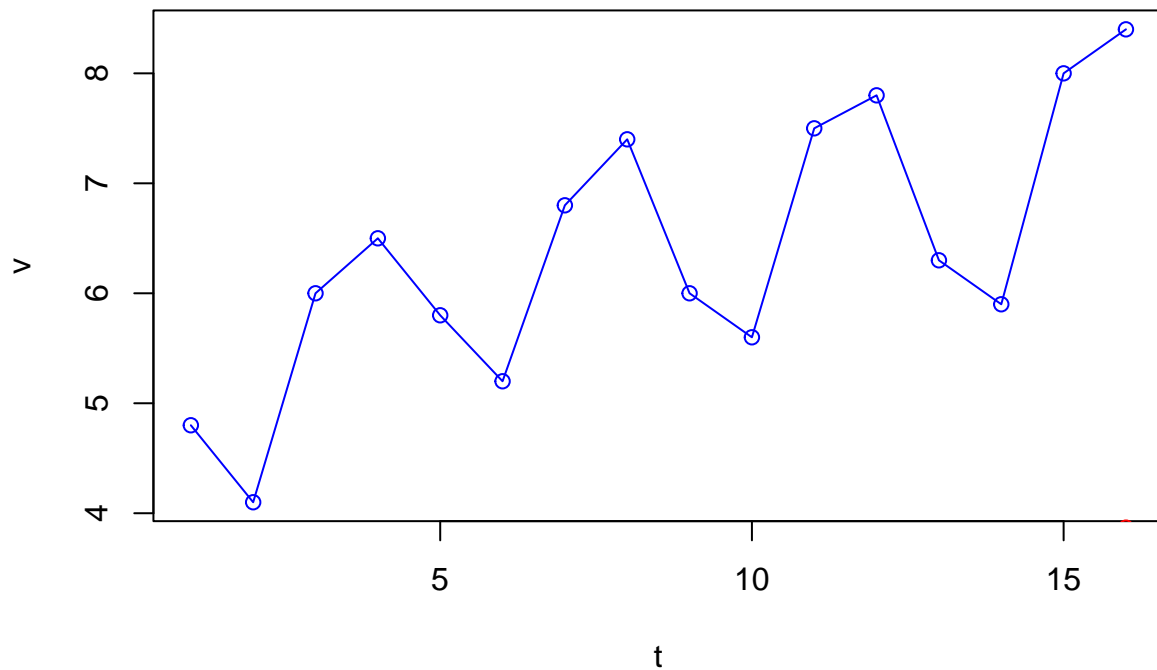
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-3.5443	5.5166	-0.642	0.531
x2	0.4847	0.5705	0.850	0.410

```
##
## Residual standard error: 10.52 on 14 degrees of freedom
## Multiple R-squared:  0.04902,    Adjusted R-squared:  -0.0189
## F-statistic: 0.7217 on 1 and 14 DF,  p-value: 0.4099
```



```
## [1] 39.86844
```

```
## [1] 0.9488595
```

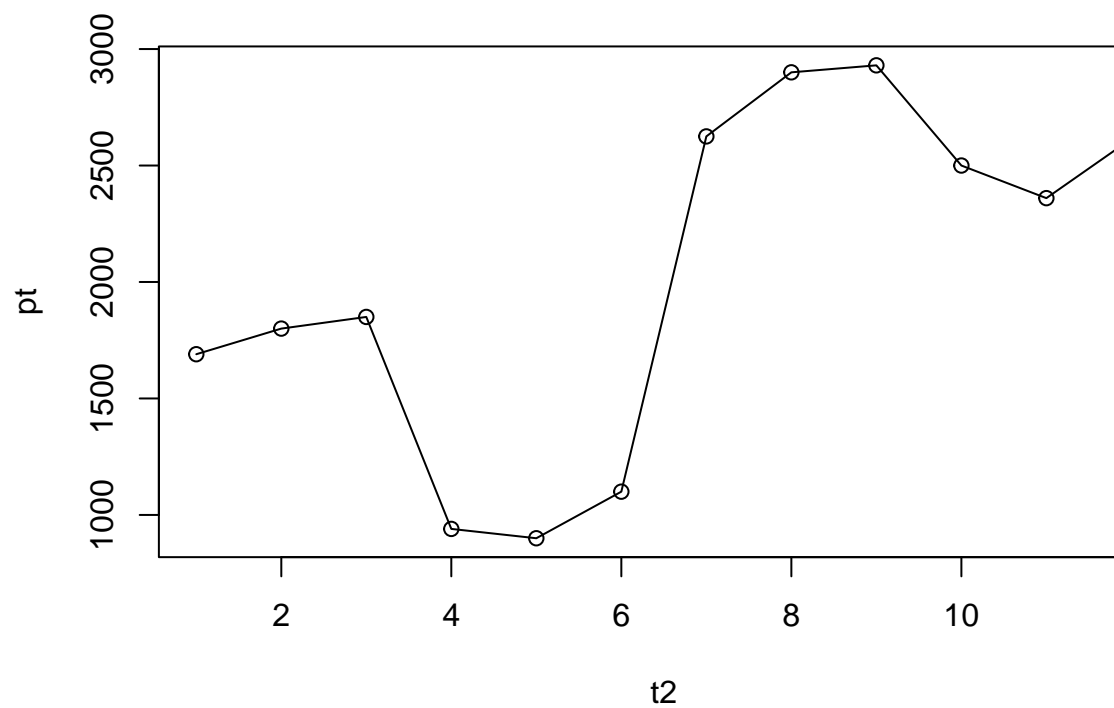


Este nuevo modelo (aditivo) podemos ver tanto en las metricas (normalidad de residuos, R squared, etc.) Como en la grafica (que ni si quiera se observan las predicciones de lo malo que es el modelo) que es un modelo horrible.

VENTA DE LIBROS:

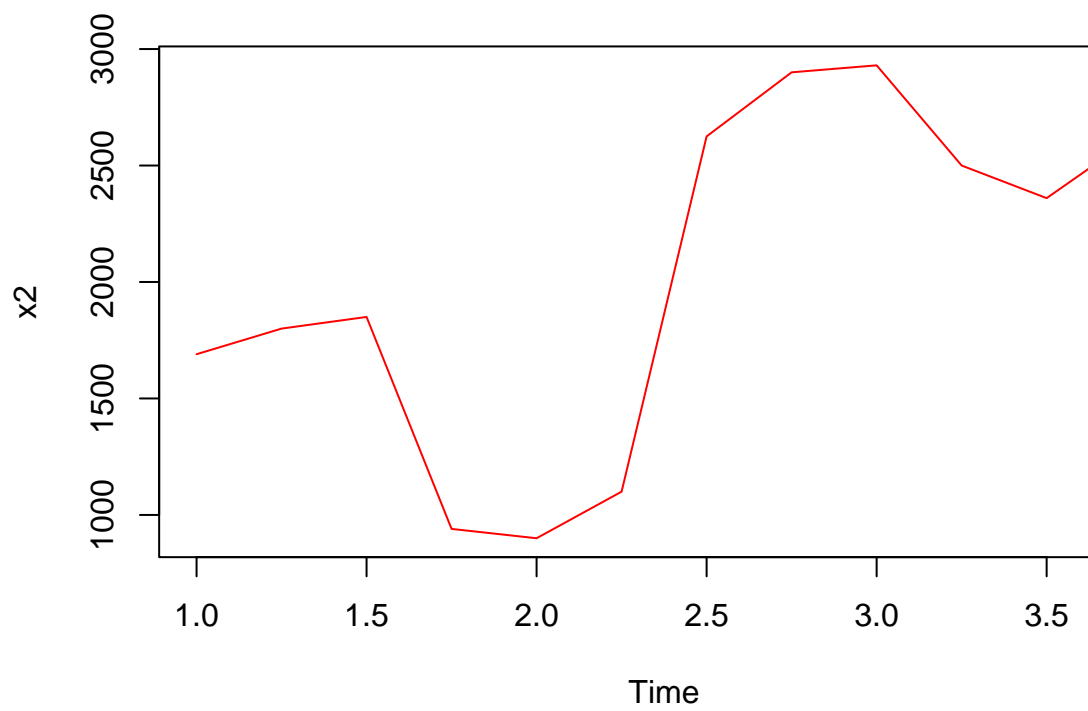
Ahora tenemos un set extra de datos de nuestro problema secundario: analizar los comportamientos de la venta de libros escolares.

Esta es la informacion con la que contamos, igualmente esta dividido por anio pero ahora son solo 3 anios (4 trimestres por anio).



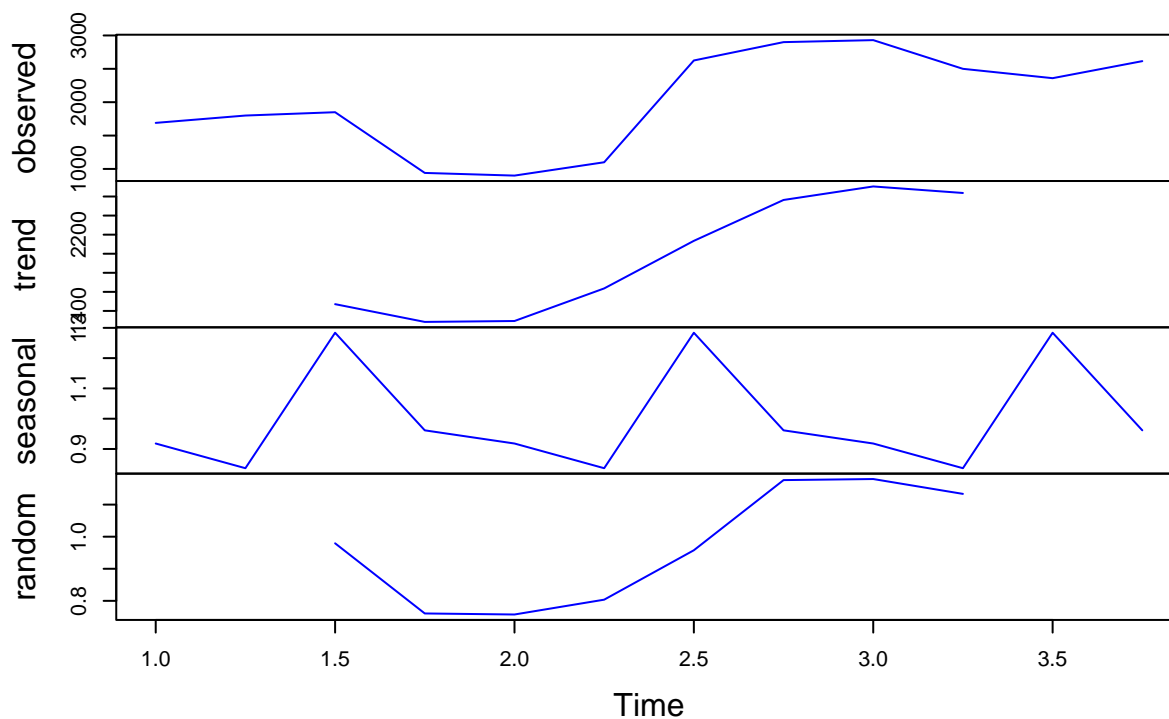
Aqui podemos ver los datos:

Haremos igualmente la decomposicion del modelo, usaremos modelo multiplicativo debido a que igualmente



esta incrementando cada anio:

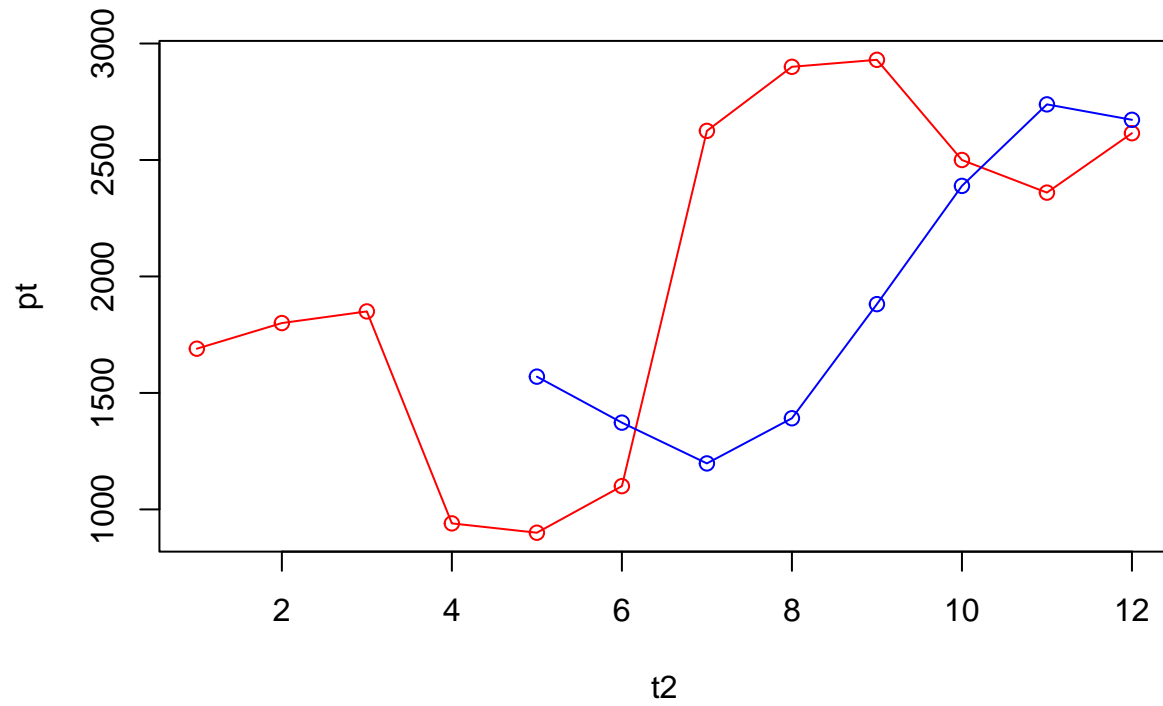
Decomposition of multiplicative time series



```
##          Qtr1      Qtr2      Qtr3      Qtr4
## 1 0.9180147 0.8365072 1.2838778 0.9616003
## 2 0.9180147 0.8365072 1.2838778 0.9616003
## 3 0.9180147 0.8365072 1.2838778 0.9616003
```

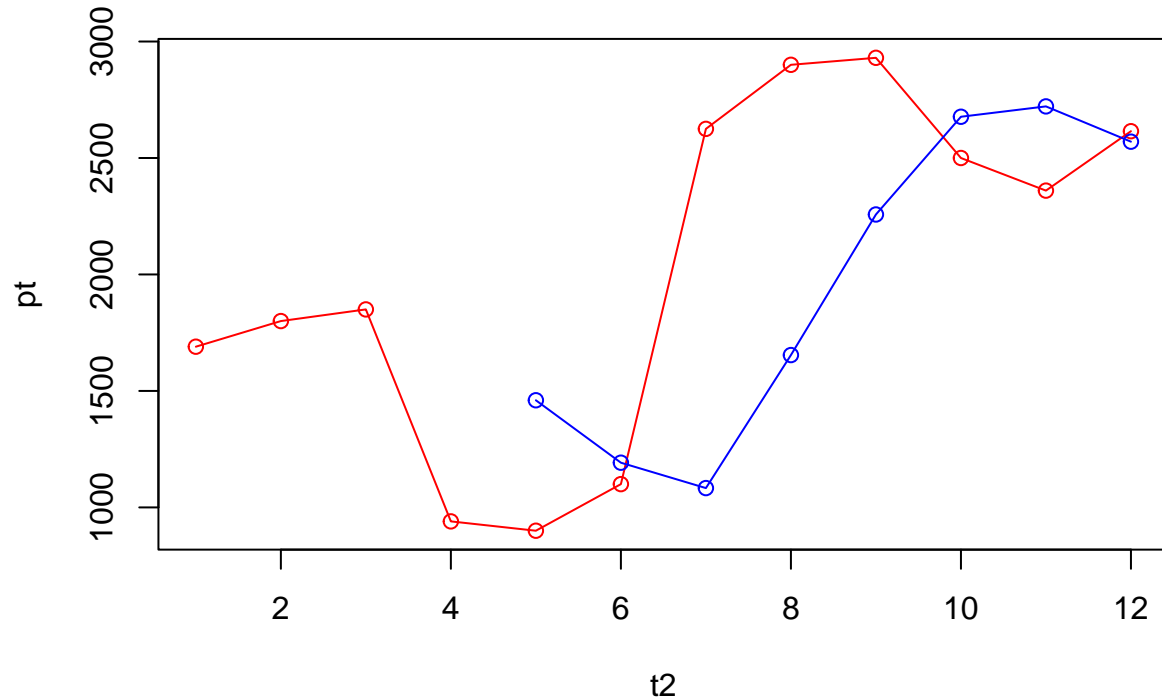
Aqui observamos estacionalizacion mucho mas drastica y podemos hacer mas deducciones: Q3 tiene el mayor indice estacional, esto puede ser porque en verano cuando todos vuelven a clases y necesitan libros nuevos de nuevas materias. El mas bajo es justo despues de verano donde ya todos tienen sus libros, han empezado clases y ya no necesitan libros. Tiene mucho sentido lo que nos mencionan las graficas.

PROMEDIO MOVIL



Utilizando el promedio movil nos podemos dar una idea de como se esta comportando la serie de tiempo.

PROMEDIO MOVIL CENTRADO



El promedio centrado igualmente nos da una idea de como se esta comportando la serie de tiempo.

CONCLUSION

Descomponer las series de tiempo en sus diferentes componentes resulta ser beneficiosos no solo en temas de negocios o ventas, si no tambien que se pueden aplicar en temperaturas, consumo de energia, entre otras cosas. Puede ser extremadamente util para estimar y podemos sacarle el maximo provecho a las series de tiempo. Con los analisis que hicimos hoy podemos tomar acciones para aumentar las ventas y poder maximizarlas.

BIBLIOGRAFIA

N / A

ANEXOS

Drive: https://drive.google.com/drive/folders/16Y6_cbbXaWo_AuxQGE0QRvrL4VdnlmkS?usp=share_link