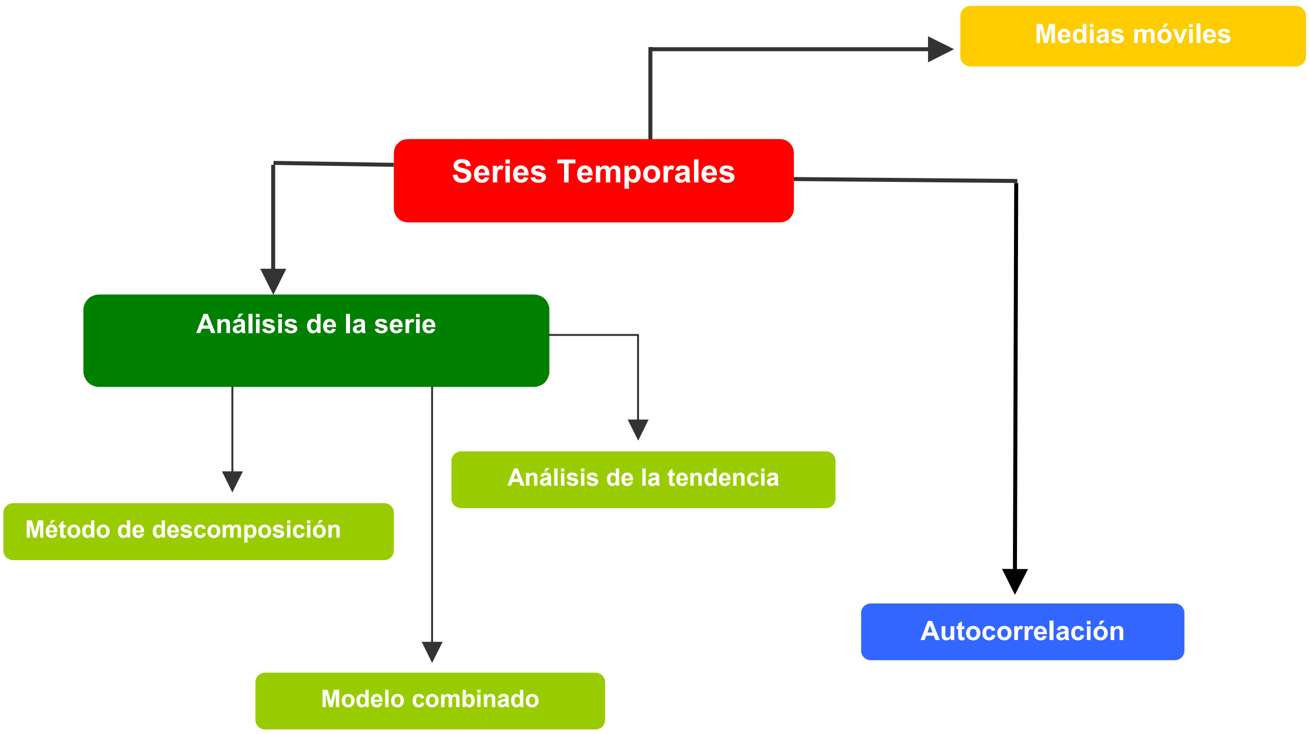
SERIES TEMPORALES

Autores: Manuel Terrádez (mterradez@uoc.edu), Ángel A. Juan (ajuanp@uoc.edu)

ESQUEMA DE CONTENIDOS



INTRODUCCIÓN

Una serie temporal es un conjunto de observaciones ordenadas en el tiempo, que pueden representar la evolución de una variable (económica, física, etc.) a lo largo de él.

El objetivo del análisis de una serie temporal es el conocimiento de su patrón de comportamiento, para así prever su evolución futura, suponiendo que las condiciones no variarán.

Dado que no se trata de fenómenos deterministas, sino sujetos a una aleatoriedad, el estudio del comportamiento pasado ayuda a inferir la estructura que permita predecir su comportamiento futuro, pero es necesaria una gran cautela en la previsión debido a la inestabilidad del modelo.

La particular forma de la información disponible de una serie cronológica (se dispone de datos en periodos regulares de tiempo) hace que las técnicas habituales de inferencia estadística no sean válidas para estos casos, ya que nos encontramos ante n muestras de tamaño 1 procedentes de otras tantas poblaciones de características y distribución desconocidas.

OBJETIVOS

- Entender la estructura especial de la información en una serie temporal.
- Comprender qué está sucediendo con los datos (patrón de comportamiento).
- Predecir valores futuros.

CONOCIMIENTOS PREVIOS

Aparte de estar iniciado en el uso del paquete estadístico Minitab, resulta muy conveniente haber leído con profundidad los siguientes *math-blocks*:

- Estadística descriptiva.
- Análisis de regresión y correlación lineal.

CONCEPTOS FUNDAMENTALES

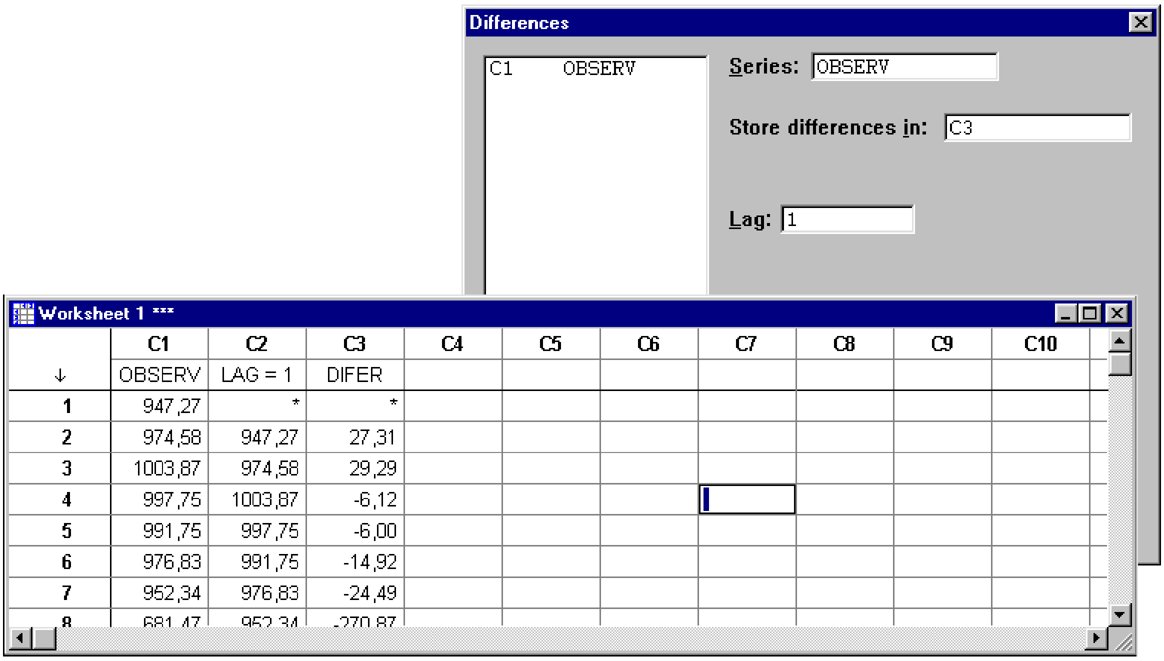
Consideraciones previas al análisis

Normalmente, la mejor forma de comenzar a analizar los datos de una serie temporal es representar las observaciones vs. el tiempo a fin de detectar tendencias, patrones estacionarios, y outliers.

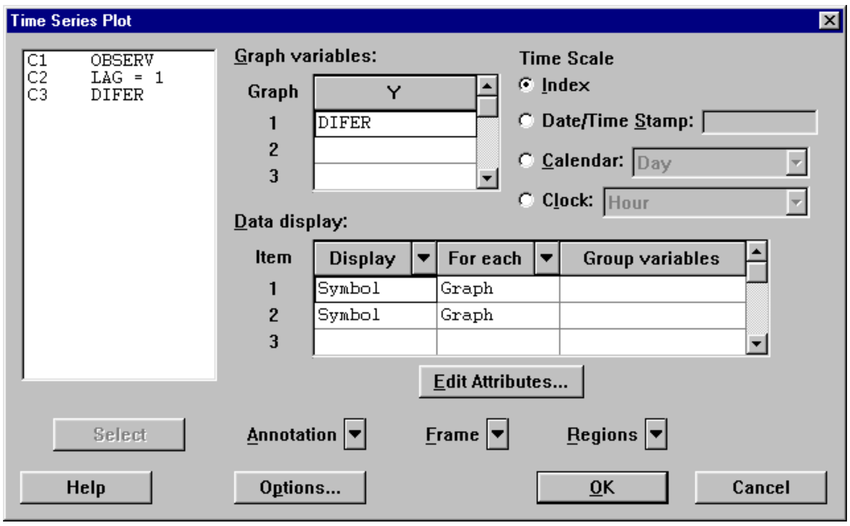
Si la variabilidad de la serie cambia con el tiempo, es conveniente aplicar una transformación a los datos que **estabilice la varianza**. Se suele utilizar una transformación logarítmica o, en ocasiones, considerar el cambio porcentual de cada observación a la siguiente (en lugar de las propias observaciones). Para estudiar dicha variabilidad, podemos hacer (con ayuda de un programa estadístico, por ejemplo `Minitab`) lo siguiente:

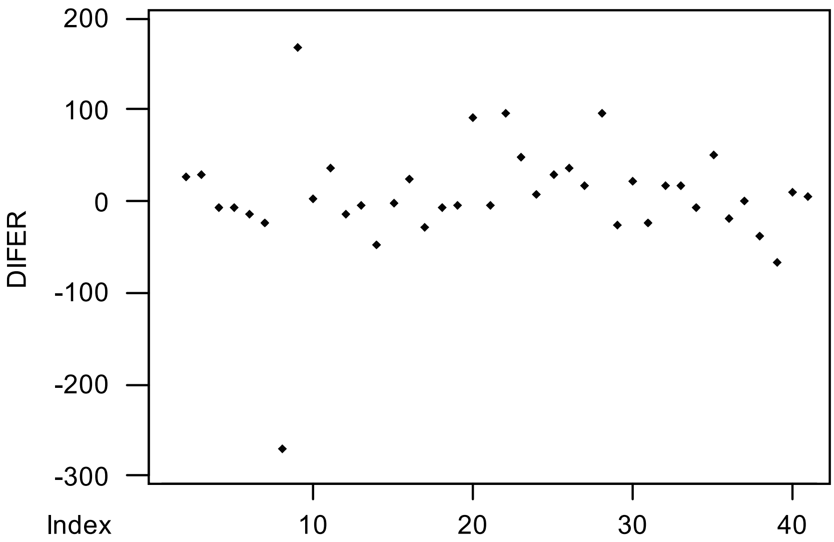
1. Colocar las observaciones en la columna 1
2. Colocar en la columna 2 las observaciones desplazadas en un lugar (lag = 1)
3. Calcular en la columna 3 la diferencia entre los datos de ambas columnas

Los pasos 2 y 3 se pueden hacer con `Minitab` mediante las opciones ***Stat > Time Series > Lag***, y ***Calc > Calculator***, o bien directamente con la opción ***Stat > Time Series > Differences,*** como se muestra a continuación:



Representando ahora la columna de las diferencias podemos saber si la varianza permanece aproximadamente constante.





En el ejemplo anterior, se observa que la varianza permanece aproximadamente constante (con excepción de los dos *outliers*), por lo que no parece necesario aplicar ninguna transformación a los datos.

Es frecuente que nos interese comparar el valor observado en un instante temporal determinado con el valor observado en algún instante anterior. Así, podríamos estar interesados en comparar los datos de la columna OBSERV con los de la columna LAG = 1 (los mismos datos pero desplazados en una unidad temporal). Esta comparación nos puede permitir determinar el coeficiente de correlación entre ambos conjuntos de datos, lo cual será útil a la hora de realizar predicciones. Para determinar dichas correlaciones haremos uso de la **función de autocorrelación** y de la **función de autocorrelación parcial**.

En el análisis de las series temporales se considera que las observaciones contienen: (a) un patrón sistemático, y (b) un componente de error aleatorio al que llamaremos **ruido**. La mayoría de las técnicas que veremos tendrán como objetivo "filtrar" dicho ruido.

Análisis de la tendencia

El **análisis de la tendencia** es un método que consiste en ajustar un modelo de tendencia general a una serie temporal con el fin de realizar predicciones. Se suele utilizar cuando la serie no contiene componente estacionario alguno.

Los datos deben estar todos en la misma columna. Minitab permite elegir entre cuatro modelos diferentes: lineal, cuadrático, exponencial, y curva en forma de S. En el caso de elegir este último, es necesario eliminar de la columna todas las casillas que no contengan datos válidos (*missing data*).

El programa ofrece tres medidas para estimar la bondad del ajuste:

- $$\mathrm{MAPE} = \frac{\Sigma\left|(y_t - \hat{y}_t)/y_t\right|}{n} \times 100 \qquad (y_t \neq 0)$$

- $$\mathrm{MAD} = \frac{\Sigma\left|y_t - \hat{y}_t\right|}{n}$$

- $$\mathrm{MSD} = \frac{\Sigma(y_t - \hat{y}_t)^2}{n}$$

En las expresiones anteriores, y_t representa la observación, $\hat{y}_t$ representa el valor pronosticado, y n representa el número de predicciones a realizar.

Para las tres medidas, **cuanto menor sea su valor, mejor será el ajuste del modelo**.

Descomposición

El **método de descomposición** permite, dada una serie temporal, separarla en sus respectivos componentes: por un lado nos proporcionará la **tendencia** lineal y, por otro, su **estacionalidad**.

Usaremos el método de descomposición cuando:

(a) deseemos realizar predicciones y la serie tenga un componente estacional, o

(b) queramos examinar la naturaleza de los componentes de la serie.

El componente estacional de la serie puede tener, con respecto a la tendencia, un carácter aditivo o un carácter multiplicativo. Usaremos un **modelo multiplicativo** cuando la variación del patrón estacional aumente al desplazarnos hacia la derecha en el gráfico. Si, por el contrario, la variación del patrón estacional permanece constante, usaremos un **modelo aditivo**.

Modelo multiplicativo: y_t = Tendencia • Estacionalidad + Error

Modelo aditivo: y_t = Tendencia **+** Estacionalidad + Error

Generalmente, realizaremos la descomposición en un solo paso a partir de las observaciones colocadas en una única columna. Sin embargo, cuando las observaciones muestren una tendencia no lineal, suele ser conveniente realizar una descomposición de los residuos del modelo de tendencia previamente calculado. Esta alternativa suele mejorar el ajuste del modelo al combinar la información del análisis de tendencia con la información de la descomposición.

Modelo combinado

Cuando las observaciones muestren una tendencia no lineal, suele ser conveniente realizar una descomposición de los residuos del modelo de tendencia previamente calculado (**análisis combinado**). Esta alternativa suele mejorar el ajuste del modelo al combinar la información del análisis de tendencia con la información de la descomposición.

Medias móviles

El **método de las medias móviles** es un método dinámico que consiste en promediar observaciones consecutivas de una serie para “suavizar” el patrón que siguen los datos y realizar predicciones a corto plazo. Este procedimiento suele emplearse cuando las observaciones no muestran patrones de tendencia o estacionarios, si bien es posible emplearlo también con series que tengan componentes estacionarios o de tendencia.

Para calcular la media móvil se promedian grupos de observaciones consecutivas. Supongamos, por ejemplo, que una serie comienza con los números 4, 5, 8, 9, 10, … , y que usamos una longitud de 3 para calcular la media móvil. Entonces, los dos primeros valores de la media móvil serán desconocidos, mientras que el tercero será el promedio entre las observaciones 4, 5, y 8. Por su parte, el cuarto valor será el promedio entre 5, 8, y 9, etc.

Si la serie no tiene componente estacionario, suele ser habitual tomar medias móviles de poca longitud para “suavizar” la serie, si bien dicha longitud depende del nivel de ”ruido” (error) que contenga la serie: si tomamos una media móvil de longitud grande estaremos eliminando mucho “ruido”, pero el patrón resultante también será menos sensible a cambios en las series. Si la serie contiene un patrón estacionario se suele usar el período como longitud para la media móvil.

Autocorrelación y autocorrelación parcial

A la hora de realizar pronósticos, sería muy útil detectar la existencia, en las observaciones, de algún patrón que nos indicase cómo varían los datos de un instante temporal al siguiente. Por ejemplo, podría ocurrir que un valor por debajo de la media en el instante t propiciara obtener un valor alto en el instante t+1 (o viceversa).

La **función de autocorrelación** nos puede ayudar a identificar tales patrones. La idea es calcular el coeficiente de correlación entre el conjunto de observaciones y el conjunto de observaciones desplazadas en n instantes temporales (lag = n). Lógicamente, estaremos interesados en detectar niveles altos de autocorrelación.

Pero no debemos olvidar que los coeficientes de autocorrelación son dependientes entre sí. Por ejemplo: si la primera columna (observaciones) está fuertemente correlacionada con la segunda (lag = 1), y ésta a su vez con la tercera (lag = 2), entonces la primera estará también correlacionada con la tercera. Por este motivo, suele ser interesante calcular también la **función de autocorrelación parcial**, en la cual ya se eliminan las dependencias con columnas intermedias. En cierto sentido, podríamos decir que la autocorrelación parcial proporciona una visión más clara de las dependencias entre las observaciones.

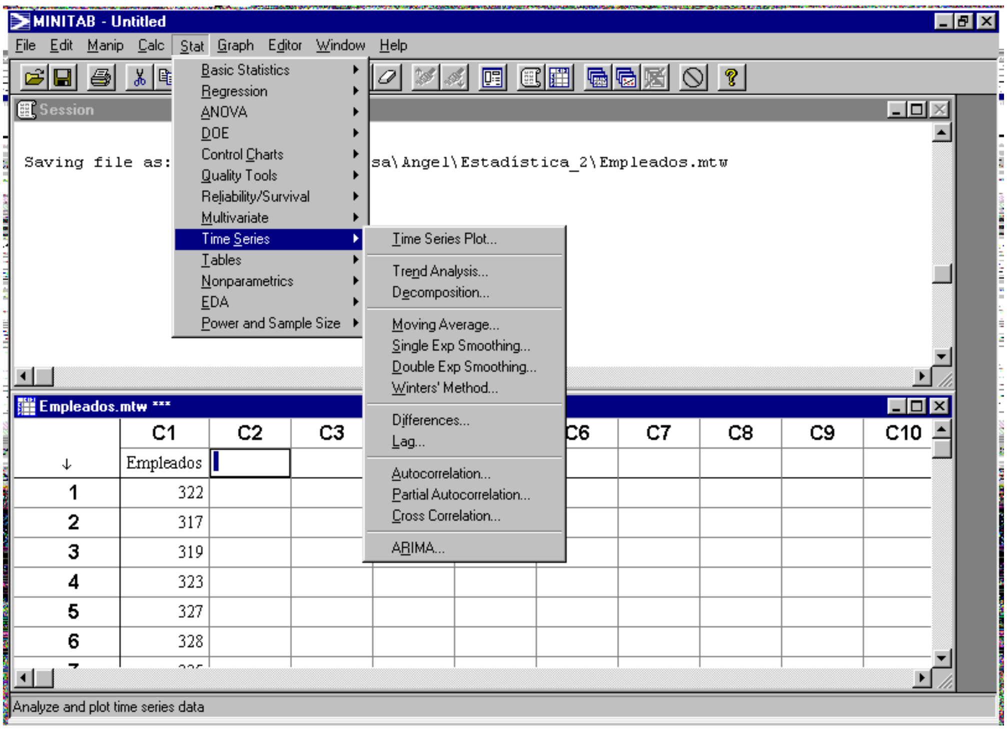
CASOS PRÁCTICOS CON SOFTWARE

Número de empleados de una empresa

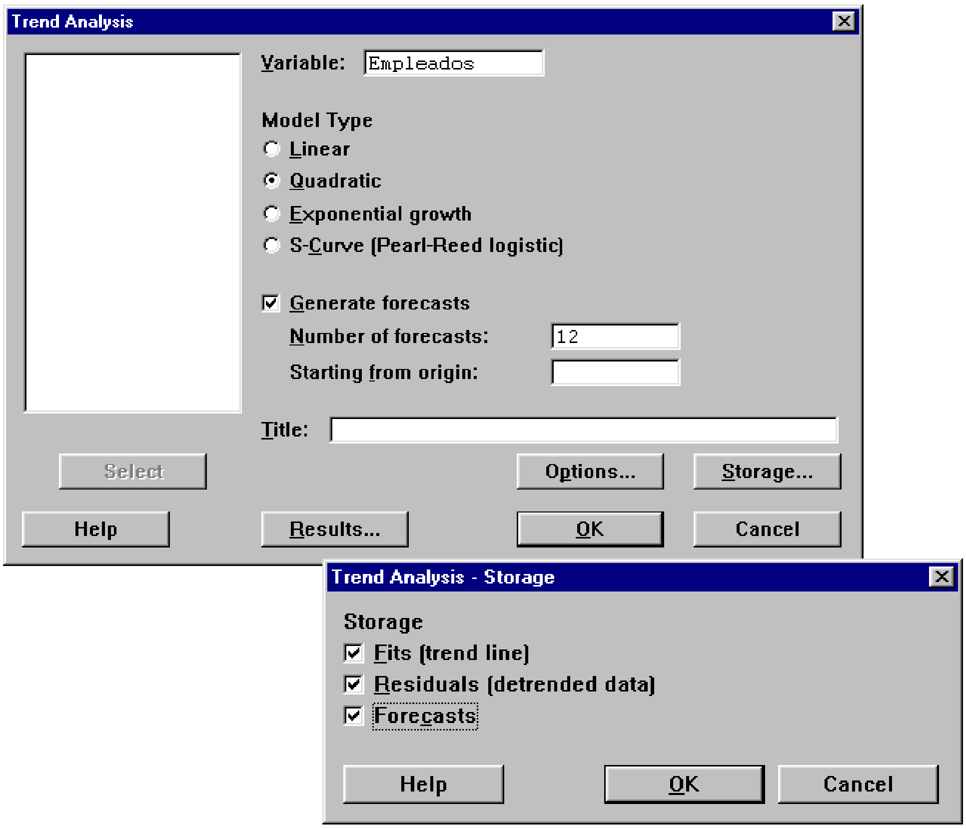
Durante los últimos 60 meses, hemos ido registrando el número de empleados de una gran empresa (fichero `Empleados.mtw`). Deseamos ahora hacer una predicción, sobre la evolución de este indicador, en los próximos 12 meses. Dado que se observa en los datos un patrón curvilíneo, usaremos un modelo cuadrático para ajustar las observaciones. Además, también se observa un componente estacionario, por lo que guardaremos los residuos a fin de realizar, más adelante, una descomposición de los mismos y poder así mejorar nuestro modelo.

Análisis de la tendencia

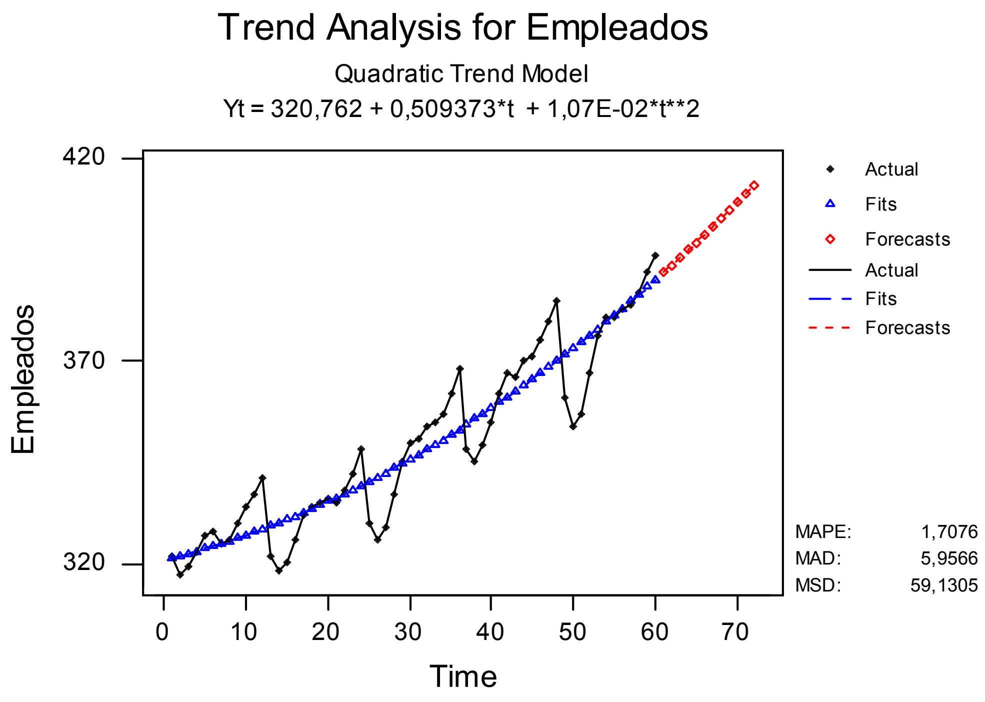
Seleccionamos ***Stat > Time Series > Trend Analisis*** :



Completamos la ventana "Trend Analisis" y la de "Storage" como se muestra a continuación:




```
Trend Analysis

Data         Empleados
Length       60,0000
NMissing     0

Fitted Trend Equation

Yt = 320,762 + 0,509373*t  + 1,07E-02*t**2

Accuracy Measures

MAPE:          1,70760
MAD:           5,95655
MSD:           59,1305

 Row   Period     FORE1

   1       61   391,818
   2       62   393,649
   3       63   395,502
   4       64   397,376
  ...     ...     ...
```

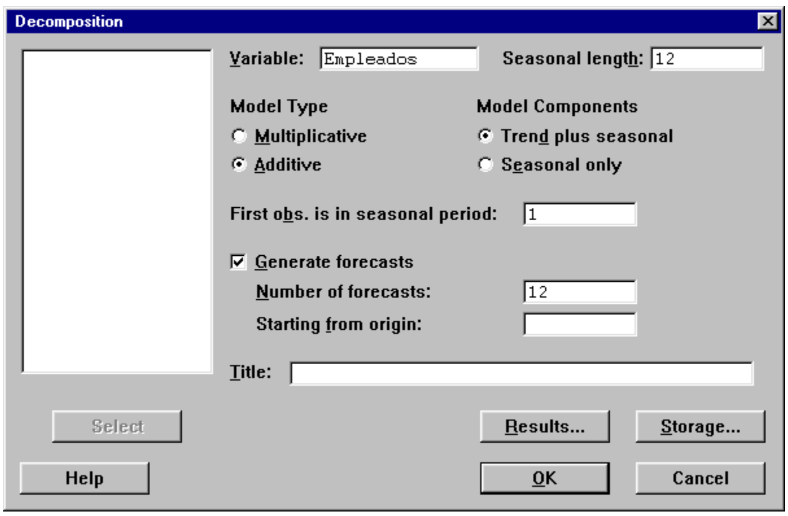


El gráfico anterior muestra las observaciones (Actual), la curva de tendencia que se ajusta a las mismas (Fits), y los valores pronosticados (Forecasts).

Las observaciones presentan una tendencia creciente, con un claro componente estacionario. La curva obtenida parece ajustarse bastante bien a la tendencia de las observaciones, pero el patrón estacionario no está siendo considerado en este modelo.

Análisis de descomposición

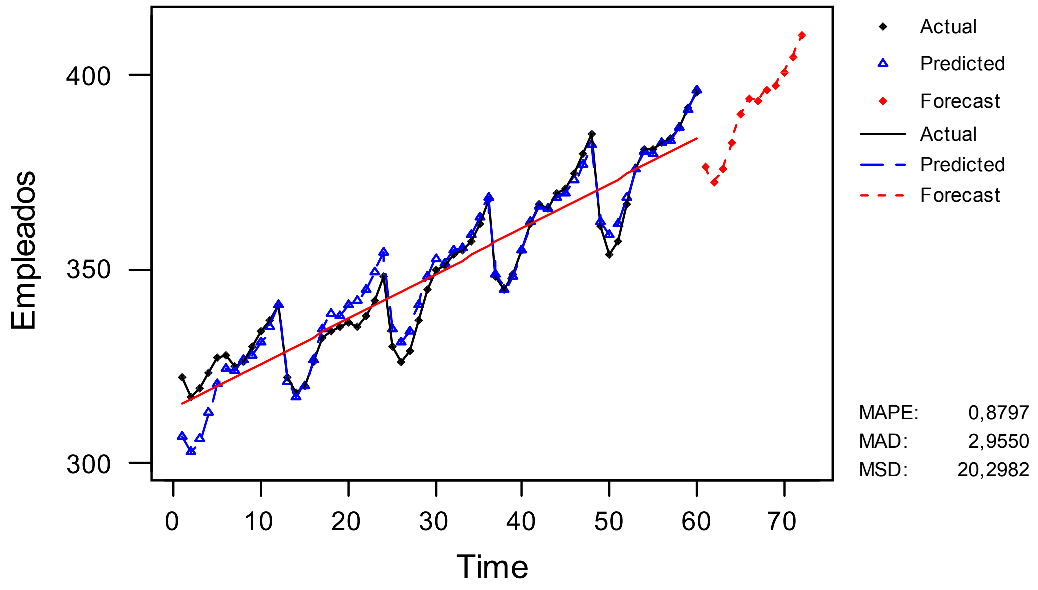
Seleccionamos ***Stat > Time Series > Decomposition*** y completamos las ventanas como se indica (elegiremos un modelo aditivo):



El "output" del programa nos ofrece información textual y gráfica. Observar que, con este modelo, obtenemos un valor de 20,30 para el MSD:

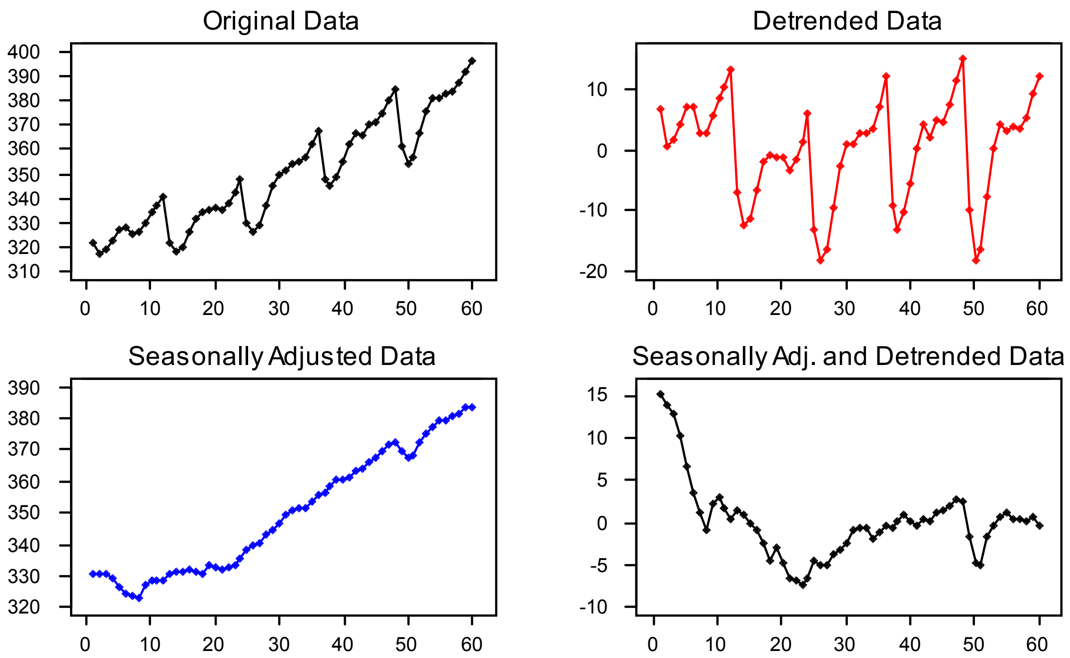
```
Time Series Decomposition

Data         Empleados
Length       60,0000
NMissing     0

Trend Line Equation

Yt = 313,989 + 1,16485*t

Seasonal Indices

Period     Index

   1       -8,48264
   2       -13,3368
   3       -11,4410
   4       -5,81597
   5       0,559028
   6        3,55903
   7        1,76736
   8        3,47569
   9        3,26736
  10        5,39236
  11        8,49653
  12        12,5590

Accuracy of Model

MAPE:       0,8797
MAD:        2,9550
MSD:       20,2982

Forecasts

 Row   Period   Forecast

   1       61    376,562
   2       62    372,873
   3       63    375,933
   4       64    382,723
   5       65    390,263
   6       66    394,428
   7       67    393,801
   8       68    396,674
   9       69    397,631
  10       70    400,921
  11       71    405,190
  12       72    410,417
```

Decomposition Fit for Empleados

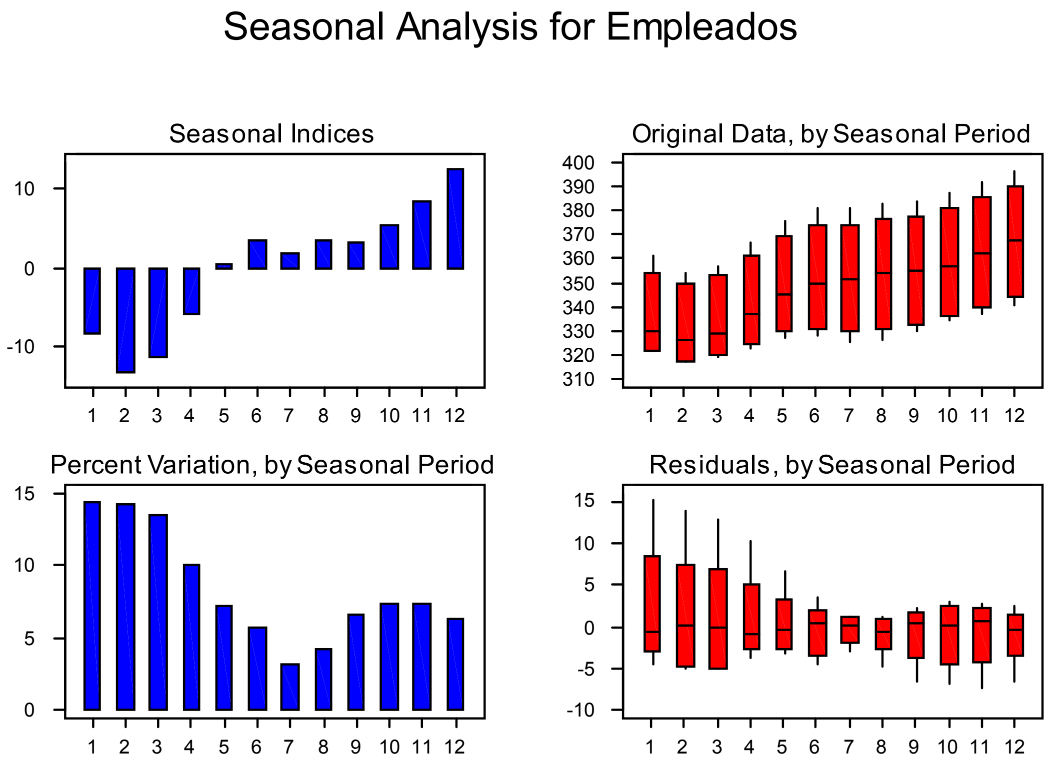


En las siguientes imágenes se muestran, por separado, la serie original de observaciones, los datos una vez eliminada la tendencia, los datos una vez extraído el patrón estacional, y los datos una vez extraídos el patrón estacional y la tendencia:

Component Analysis for Empleados



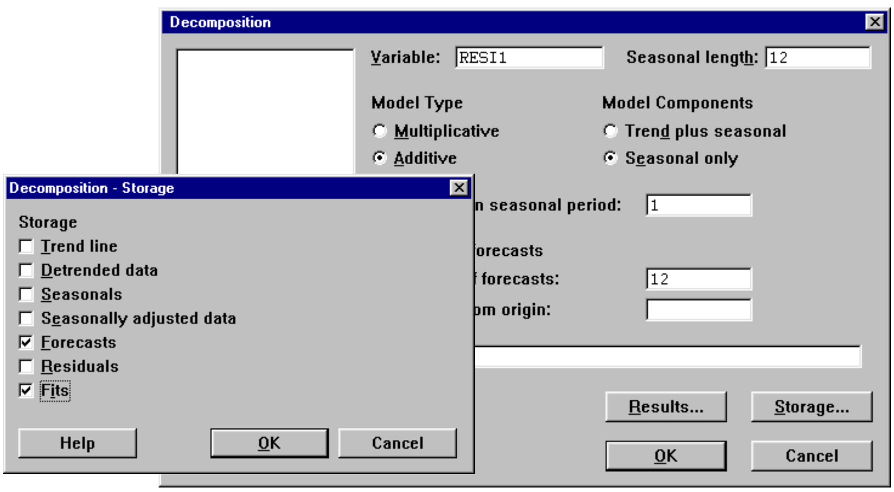
Finalmente, en el último conjunto de gráficos se muestra un análisis estacional: gráfico de índices estacionales, gráfico de variación porcentual por estaciones, gráfico de boxplots referidos a observaciones agrupadas por períodos estacionarios, y gráfico de boxplots de los residuos agrupados por períodos estacionarios.



Análisis combinado (tendencia+descomposición)

Usaremos los residuos obtenidos en el análisis de la tendencia (guardados en la columna RESI1) para combinarlo con el método de descomposición:

Seleccionamos ***Stat > Time Series > Decomposition*** y completamos las ventanas como se indica:

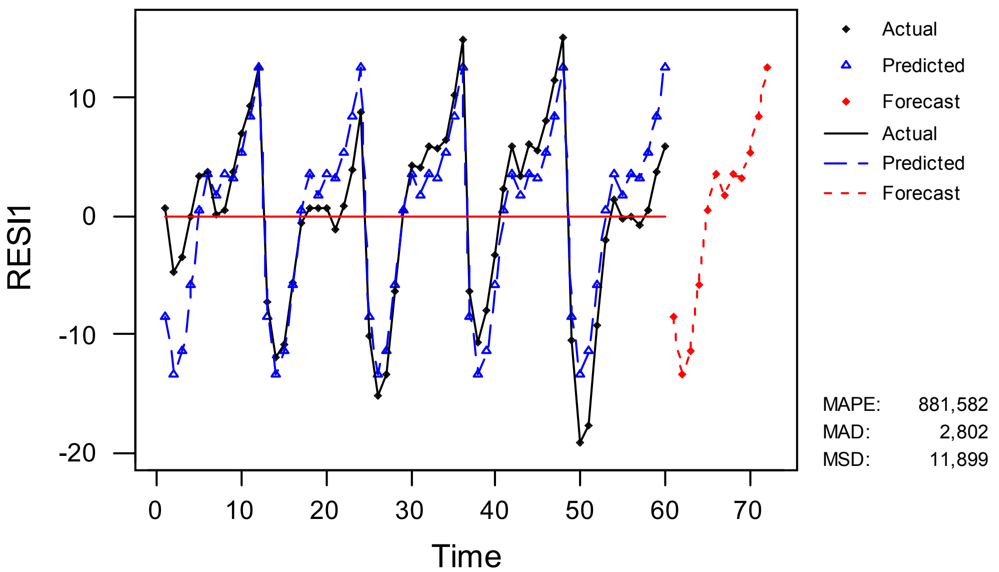


A continuación se muestra el "output" generado por el programa:

```
Time Series Decomposition

Data        RESI1
Length      60,0000
NMissing    0


Seasonal Indices

Period    Index

   1      -8,48264
   2      -13,3368
   3      -11,4410
   4      -5,81597
   5      0,559028
   6       3,55903
   7       1,76736
   8       3,47569
   9       3,26736
  10       5,39236
  11       8,49653
  12       12,5590



Accuracy of Model

MAPE:     881,582
MAD:        2,802
MSD:       11,899

Forecasts

 Row  Period     FORE2

   1      61   -8,4826
   2      62  -13,3368
   3      63  -11,4410
   4      64   -5,8160
   5      65    0,5590
   6      66    3,5590
   7      67    1,7674
   8      68    3,4757
   9      69    3,2674
  10      70    5,3924
  11      71    8,4965
  12      72   12,5590
```

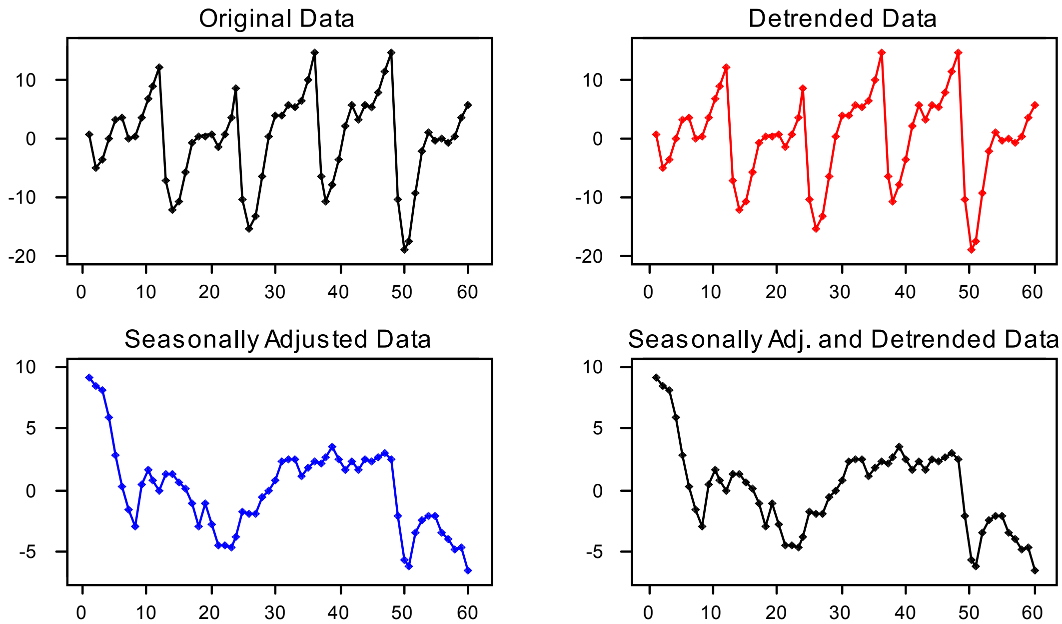
El gráfico siguiente nos proporciona la serie de residuos original (Actual), la línea de tendencia asociada (horizontal, ya que son los residuos), los valores estimados (Predicted), y los pronosticados (Forecasts):

Decomposition Fit for RESI1



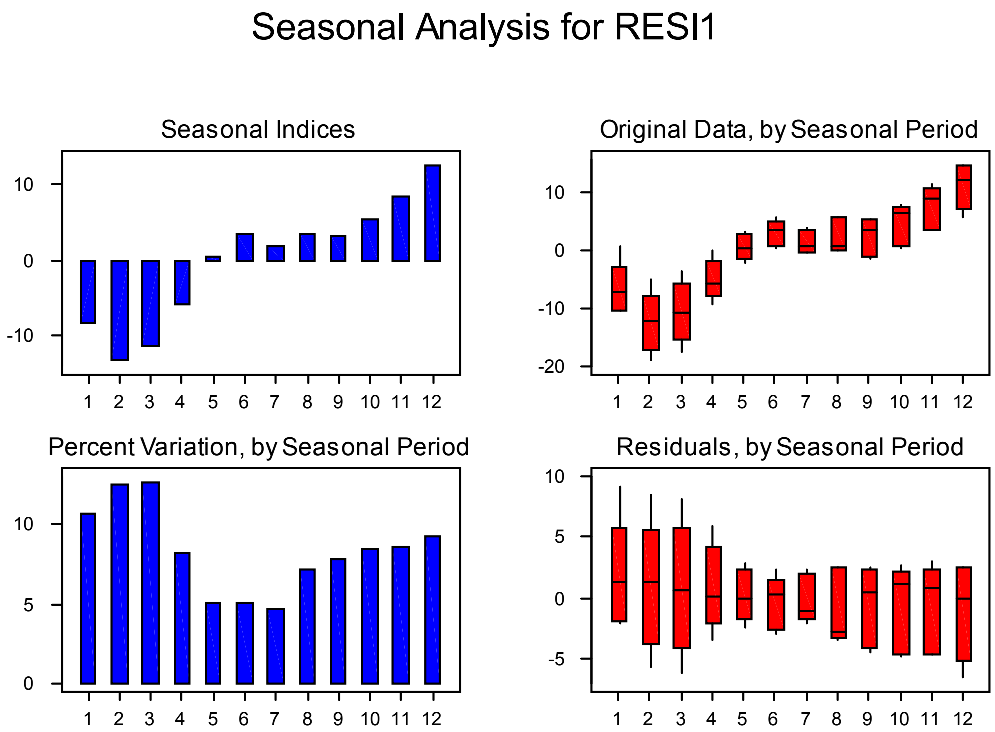
En las siguientes imágenes se muestran, por separado, la serie original de observaciones (residuos), los datos una vez eliminada la tendencia (en este caso queda igual, por ser la tendencia horizontal), los datos una vez extraído el patrón estacional, y los datos una vez extraídos el patrón estacional y la tendencia:

Component Analysis for RESI1



En el primero de los gráficos (Decomposition FIT) se observa que los residuos obtenidos en el análisis de tendencia (ver ejemplo anterior) se ajustan bastante bien por el modelo generado usando el método de descomposición. Si acaso, vemos que el valor estimado en el primero de los ciclos es considerablemente menor que el valor real, mientras que ocurre todo lo contrario en el último de los ciclos. Esto también se puede apreciar claramente en el gráfico de nombre *Seasonally Adj. And Detrended Data*.

Finalmente, en el último conjunto de gráficos se muestra un análisis estacional: gráfico de índices estacionales, gráfico de variación porcentual por estaciones, gráfico de boxplots referidos a observaciones (residuos) agrupadas por períodos estacionarios, y gráfico de boxplots de los residuos (de las observaciones) agrupados por períodos estacionarios.



Veamos ahora cómo podemos calcular los valores estimados (Predicted) y los pronosticados (Forecasted):

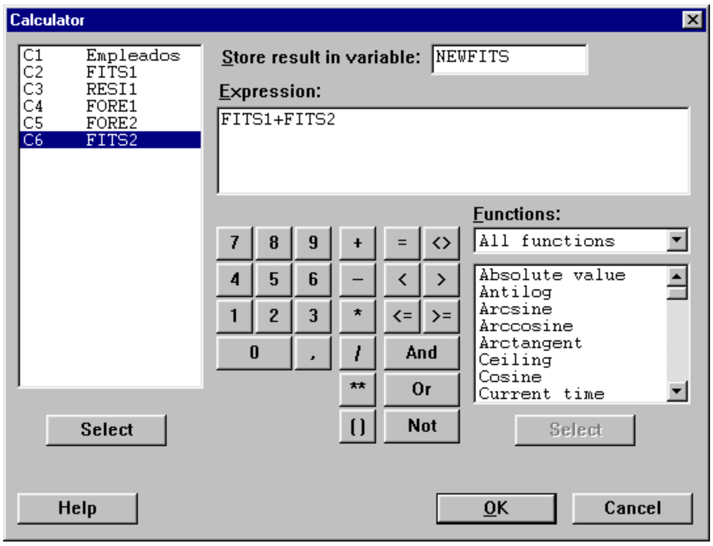
Seleccionamos ***Calc > Calculator***.

Guardaremos los nuevos valores estimados, obtenidos como suma (por ser un modelo aditivo) de:

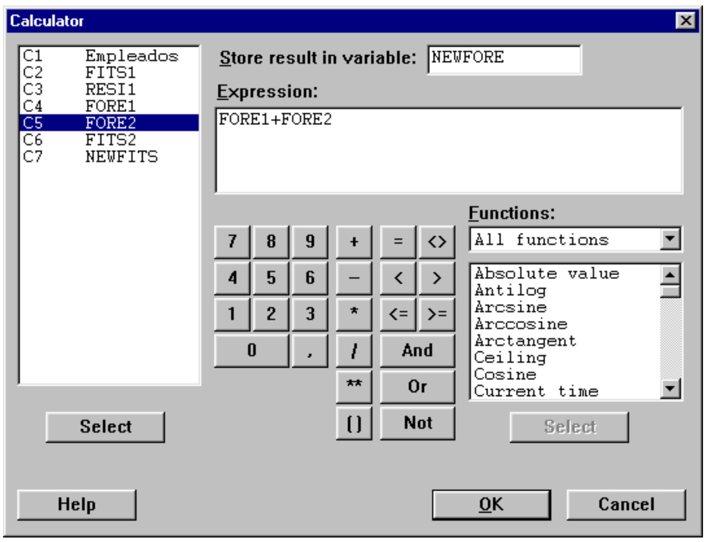
(a) los valores estimados provenientes del análisis de la tendencia (FITS1), y

(b) los provenientes de la descomposición de los residuos (FITS2)

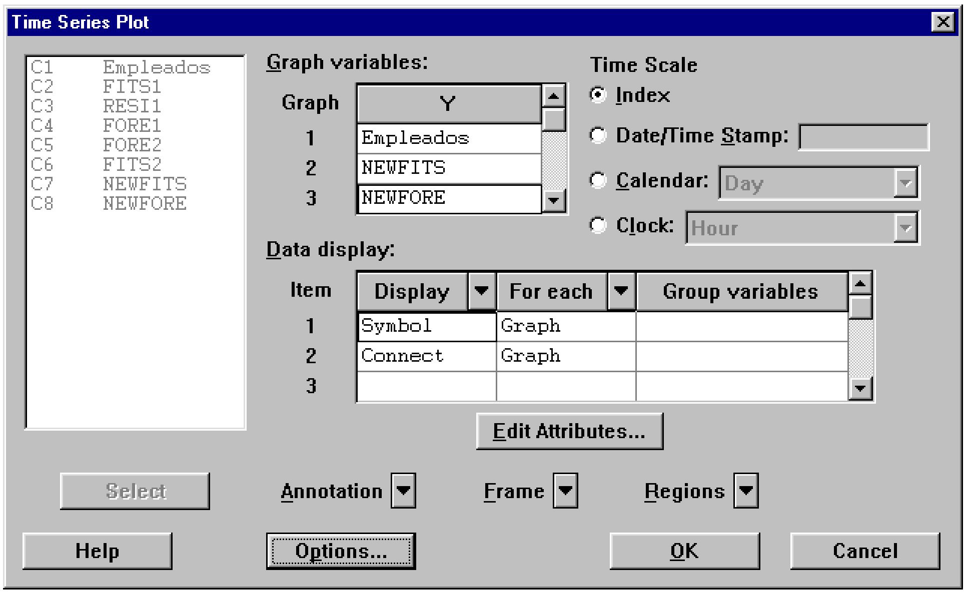
Completaremos, pues, la ventana como se muestra a continuación:

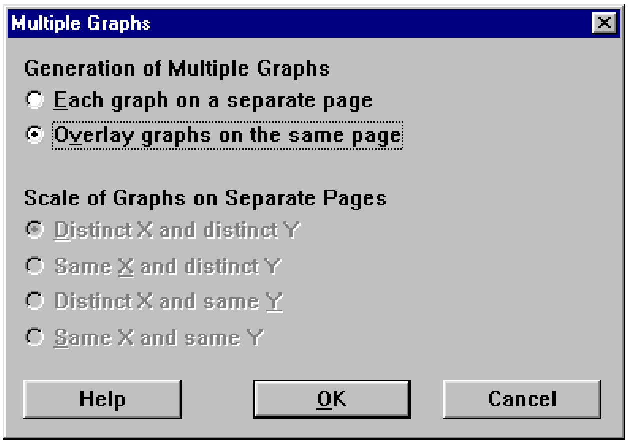


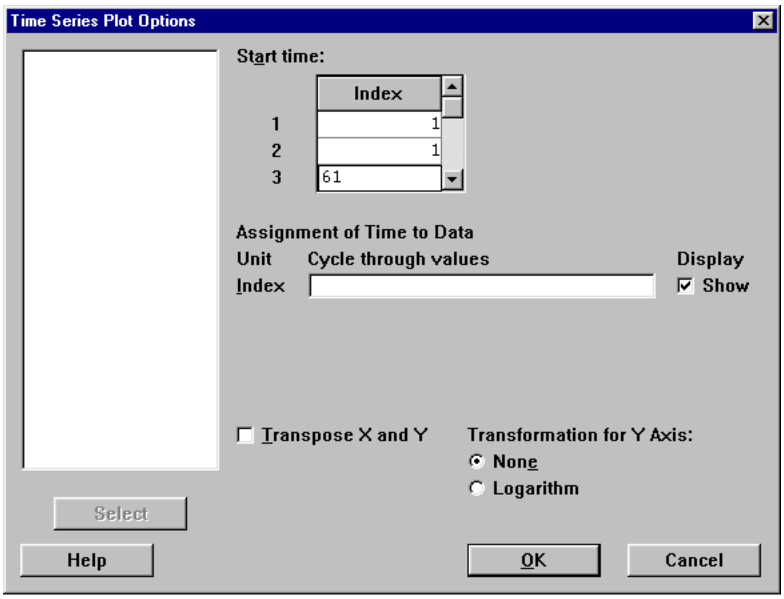
Ahora haremos lo mismo con los valores pronosticados:

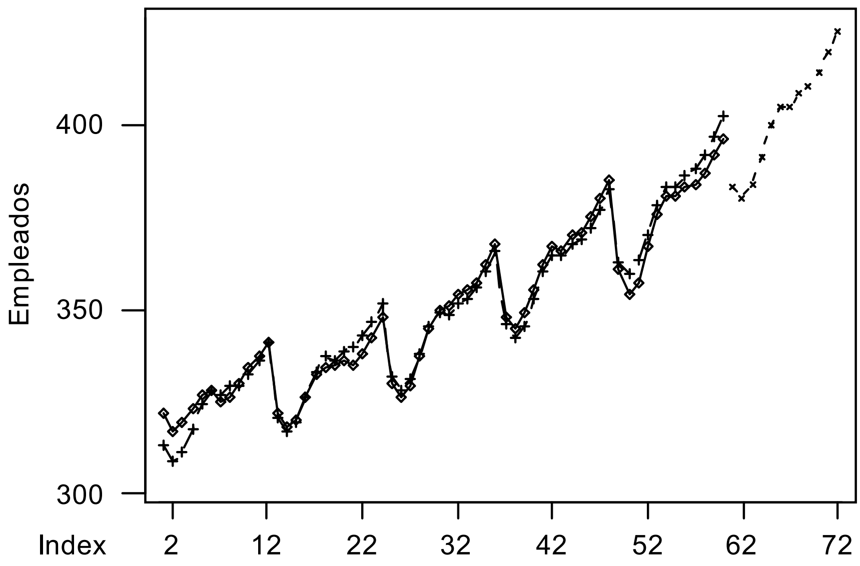


Comprobemos cómo quedan todos los datos anteriores en un gráfico de series temporales. Usaremos la opción ***Graph > Time Series Plot*** :



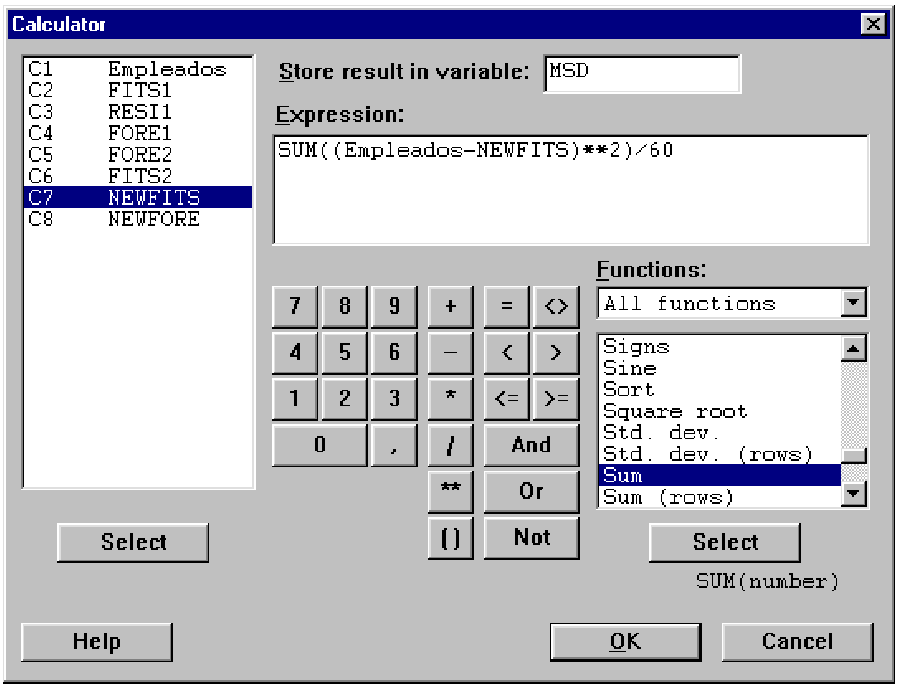






En el gráfico anterior, los círculos representan las observaciones, los símbolos + representan las estimaciones. Los pronósticos se representan con otro símbolo.

Calculemos ahora el valor del MSD. Haremos uso para ello de la fórmula y de la opción ***Calc > Calculator*** :



Como se observa en la siguiente pantalla, el valor del MSD que se obtiene con este método combinado es de 11,90:

Empleados.mtw ***

	C1	C2	C3	C4	C5	C6	C7	C8	C9
↓	Empleados	FITS1	RESI1	FORE1	FORE2	FITS2	NEWFITS	NEWFORE	MSD
1	322	321,282	0,7179	391,818	-8,4826	-8,4826	312,799	383,335	11,8989
2	317	321,824	-4,8237	393,649	-13,3368	-13,3368	308,487	380,312	
3	319	322,387	-3,3868	395,502	-11,4410	-11,4410	310,946	384,061	
4	323	322,971	0,0286	397,376	-5,8160	-5,8160	317,155	391,560	
5	327	323,577	3,4225	399,271	0,5590	0,5590	324,137	399,830	
6	328	324,205	3,7949	401,188	3,5590	3,5590	327,764	404,747	

Podemos comparar la bondad del ajuste de los diferentes modelos usando el MSD obtenido. El valor del MSD para el modelo de tendencia cuadrática era de 59,13. Los modelos de descomposición aditiva y multiplicativa con tendencia lineal darían un MSD de 20,30 y 18,54 respectivamente. El valor del MSD para la combinación de tendencia cuadrática y descomposición de residuos es de 11,90, lo que indica que este método combinado es el que proporciona un mejor ajuste. Probablemente sea también interesante calcular el valor MSD para el modelo multiplicativo.

Evolución de la temperatura de un río

En el archivo Rio.mtw se ha registrado (en la variable ***Temp***) la temperatura del agua de un río en las últimas 90 horas.

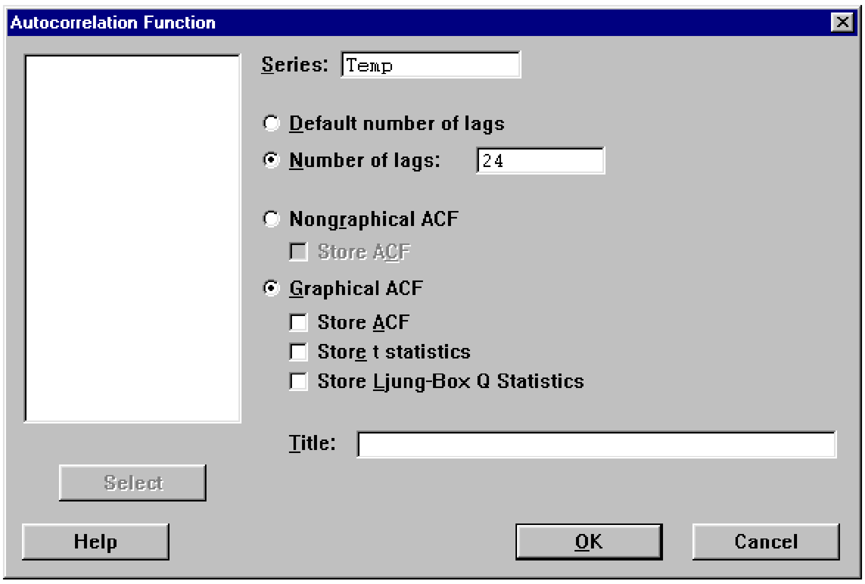
En primer lugar, a fin de determinar si la temperatura en una hora concreta está correlacionada con la temperatura registrada una hora antes, dos horas antes, etc., calcularemos la función de autocorrelación con lags = 24

Posteriormente realizaremos con Minitab un análisis de la tendencia y un análisis de descomposición (no combinado) de la serie temporal que origina la variable ***Temp***.

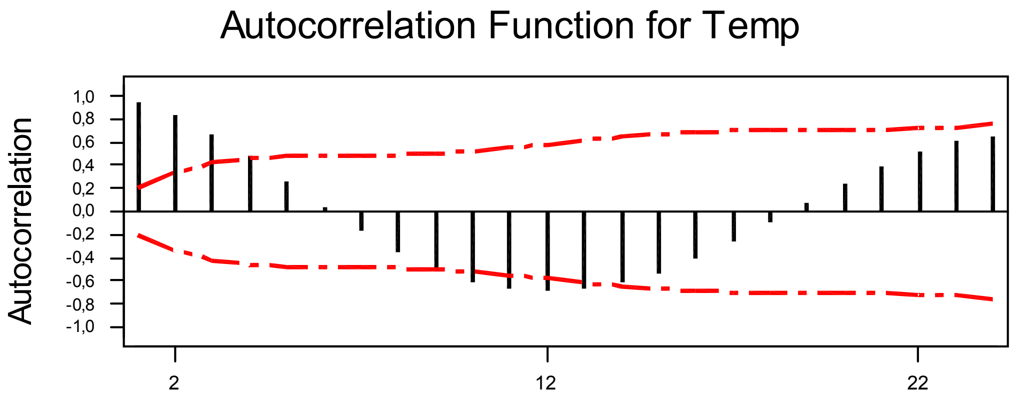
Por otra parte, usando medias móviles de longitud 4, intentaremos predecir la temperatura de las próximas 12 horas

Estudio de la autocorrelación y la autocorrelación parcial

Stat > Time Series > Autocorrelation :



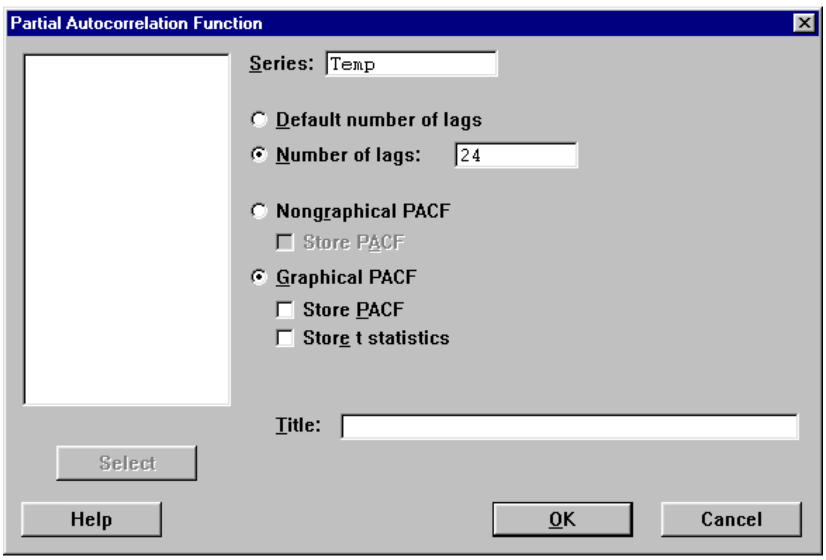
Observar en el gráfico siguiente que la función de correlación tiene una forma senosuidal, lo cual sugiere que las temperaturas de horas cercanas estarán positivamente correlacionadas (de hecho, el valor obtenido para la correlación entre la columna de las observaciones y la columna desplazada en una unidad es de 0,95), mientras que temperaturas separadas por 12 horas estarán negativamente correlacionadas (-0,70 en este ejemplo). Notar también la existencia de un componente estacionario de período 24 horas.

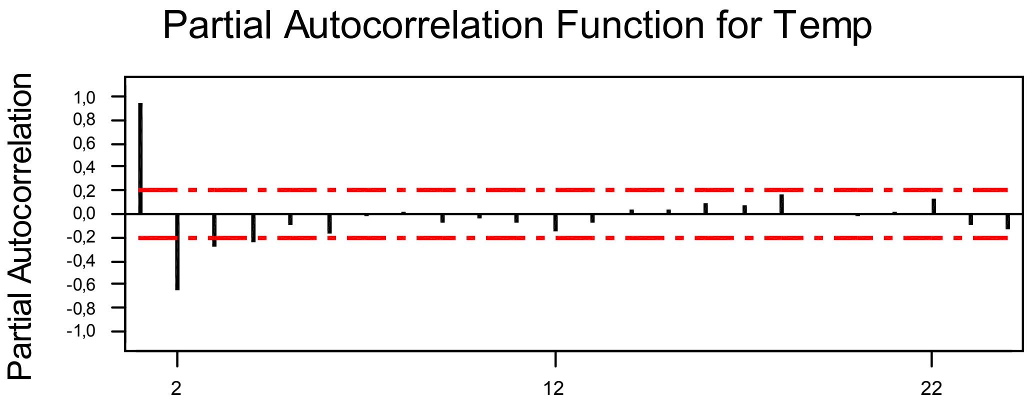


Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	0,95	9,25	88,34	8	-0,36	-1,44	249,83	15	-0,53	-1,57	549,45	22	0,52	1,45	640,21
2	0,84	4,87	157,57	9	-0,50	-2,00	277,09	16	-0,41	-1,18	569,00	23	0,62	1,68	688,98
3	0,67	3,21	203,14	10	-0,61	-2,33	317,54	17	-0,26	-0,75	577,17	24	0,66	1,75	745,79
4	0,48	2,06	226,19	11	-0,67	-2,43	367,27	18	-0,10	-0,27	578,29				
5	0,26	1,08	233,18	12	-0,70	-2,37	420,99	19	0,08	0,21	578,99				
6	0,04	0,16	233,34	13	-0,68	-2,19	472,74	20	0,24	0,69	586,23				
7	-0,17	-0,70	236,44	14	-0,62	-1,91	516,86	21	0,39	1,11	605,55				

Al analizar gráficos como el anterior, no debemos olvidar que los coeficientes de autocorrelación son dependientes entre sí. Por ejemplo: si la primera columna (observaciones) está fuertemente correlacionada con la segunda (lag = 1), y esta a su vez con la tercera (lag = 2), entonces la primera estará también correlacionada con la tercera. Por este motivo, suele ser interesante calcular también la **función de autocorrelación parcial**, en la cual ya se eliminan las dependencias con columnas intermedias. En cierto sentido, podríamos decir que la autocorrelación parcial proporciona una visión más clara de las dependencias entre las .

Stat > Time Series > Partial Autocorrelation:



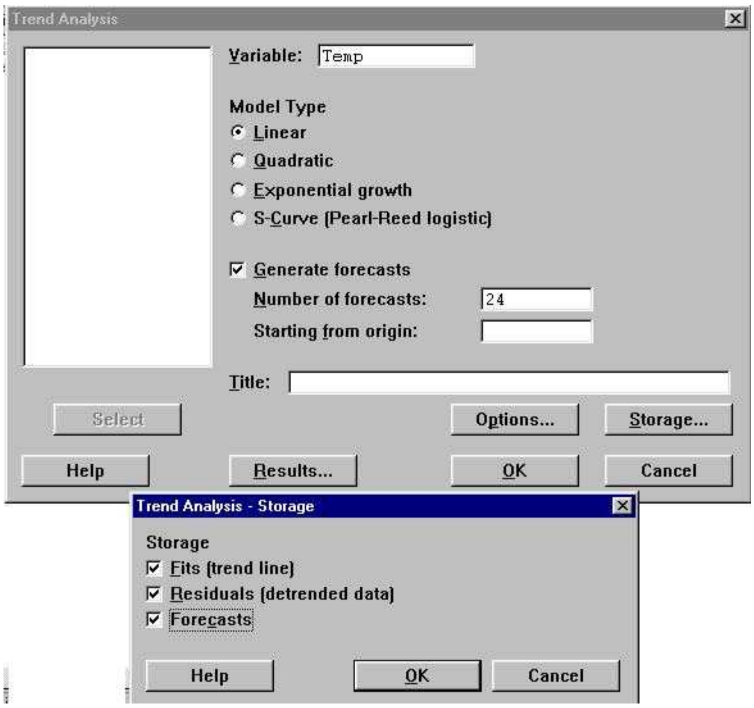


Lag	PAC	T	Lag	PAC	T	Lag	PAC	T	Lag	PAC	T
1	0,95	9,25	8	0,02	0,24	15	0,04	0,40	22	0,14	1,36
2	-0,66	-6,44	9	-0,08	-0,77	16	0,10	0,95	23	-0,10	-0,93
3	-0,29	-2,80	10	-0,03	-0,30	17	0,07	0,72	24	-0,13	-1,25
4	-0,23	-2,28	11	-0,08	-0,77	18	0,16	1,56			
5	-0,09	-0,84	12	-0,14	-1,36	19	0,01	0,06			
6	-0,16	-1,60	13	-0,07	-0,70	20	-0,02	-0,16			
7	-0,01	-0,11	14	0,04	0,35	21	0,02	0,19			

Análisis de la tendencia

Dado que se observa en los datos un patrón lineal, usaremos un modelo lineal para ajustar las observaciones. Además, también se observa un claro componente estacionario, por lo que guardaremos los residuos a fin de realizar, más adelante, una descomposición de los mismos y poder así mejorar nuestro modelo.

Seleccionamos ***Stat > Time Series > Trend Analisis***



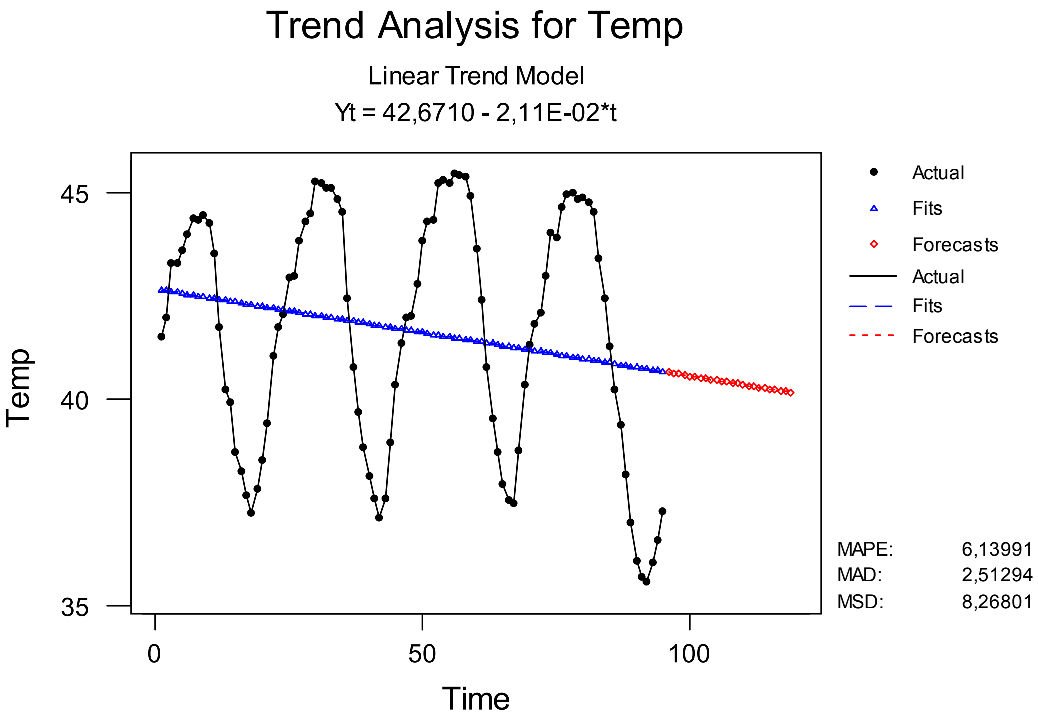
La salida que ofrece Minitab es la siguiente:

Trend Analysis

```
Data        Temp
Length      95,0000
NMissing    0

Fitted Trend Equation
Yt = 42,6710 - 2,11E-02*t


Accuracy Measures

MAPE:         6,13991
MAD:          2,51294
MSD:          8,26801

 Row  Period     FORE1

   1      96   40,6501
   2      97   40,6291
   3      98   40,6080
   4      99   40,5869
   5     100   40,5659
   6     101   40,5448
   7     102   40,5238
   8     103   40,5027
   9     104   40,4817
  10     105   40,4606
  11     106   40,4396
  12     107   40,4185
  13     108   40,3975
  14     109   40,3764
  15     110   40,3554
  16     111   40,3343
  17     112   40,3133
  18     113   40,2922
  19     114   40,2712
  20     115   40,2501
  21     116   40,2291
  22     117   40,2080
  23     118   40,1870
  24     119   40,1659
```

Observar que las medidas de bondad del ajuste (MAPE, MAD y MSD) mantienen unos valores bastante buenos (pequeños).

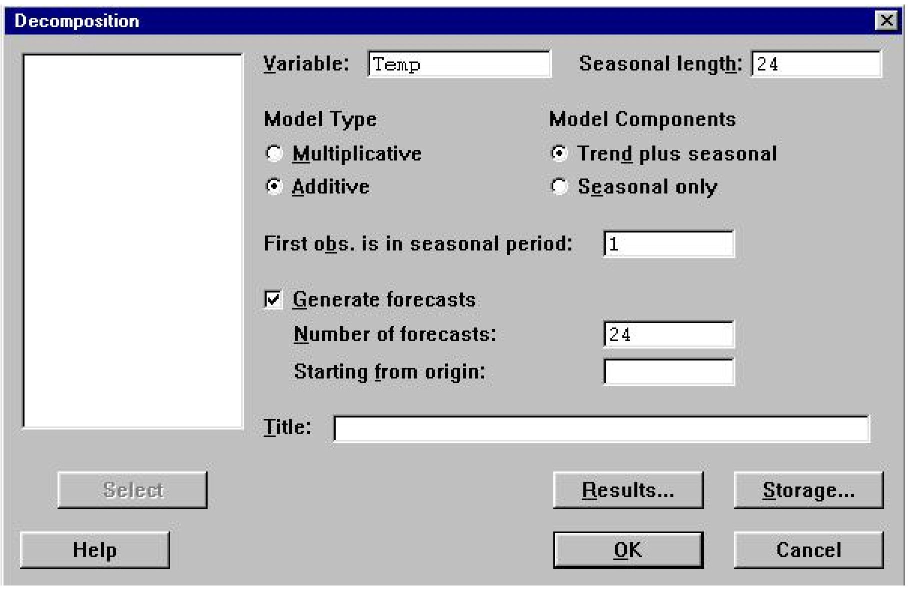
El siguiente gráfico muestra las observaciones (Actual), la curva de tendencia que se ajusta a las mismas (Fits), y los valores pronosticados (Forecasts).
Las observaciones presentan una tendencia decreciente, con un claro componente estacionario. La recta obtenida no ajusta mal la tendencia de las observaciones, pero el patrón estacionario no está siendo considerado en este modelo.



Análisis de descomposición

Utilizamos un modelo aditivo, al no observarse que la variación del patrón estacional crezca con el tiempo.

Seleccionamos ***Stat > Time Series > Decomposition***



La salida que ofrece Minitab es la siguiente:

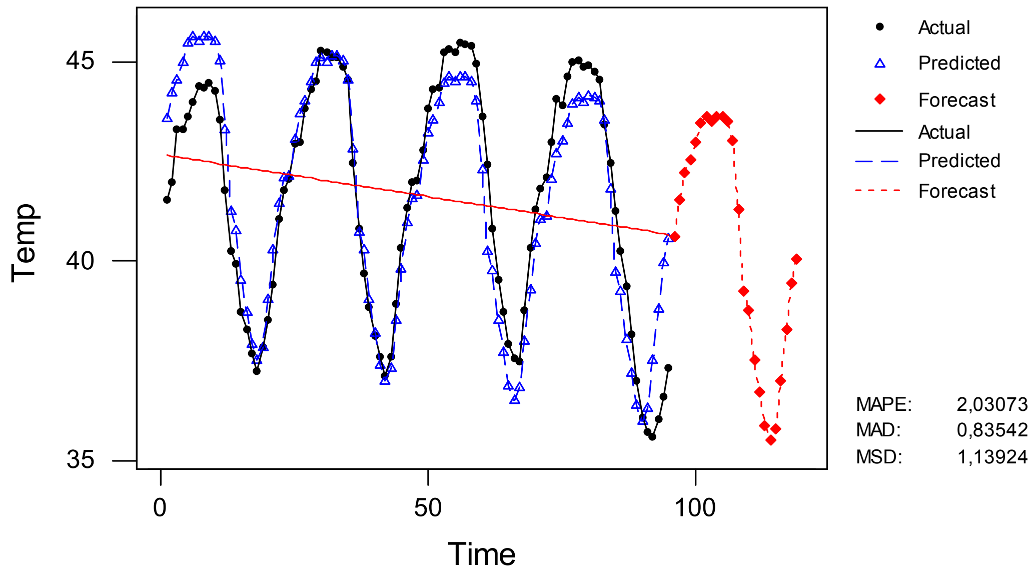
```
Time Series Decomposition

Data         Temp
Length       95,0000
NMissing     0

Trend Line Equation

Yt = 42,6710 - 2,11E-02*t

Seasonal Indices

Period    Index

   1      0,919957
   2       1,60121
   3       1,94391
   4       2,41391
   5       2,91829
   6       3,09621
   7       2,99412
   8       3,15204
   9       3,16475
  10       3,07621
  11       2,61058
  12      0,905998
  13      -1,14609
  14      -1,59484
  15      -2,81734
  16      -3,60754
  17      -4,40838
  18      -4,77692
  19      -4,44046
  20      -3,22254
  21      -1,92629
  22      -0,73692
  23      -0,10129
  24      -0,01859

 Accuracy of Model

MAPE:     2,03073
MAD:      0,83542
MSD:      1,13924

Forecasts

 Row  Period  Forecast

   1      96   40,6315
   2      97   41,5490
   3      98   42,2092
   4      99   42,5309
   5     100   42,9798
   6     101   43,4631
   7     102   43,6200
   8     103   43,4969
   9     104   43,6337
  10     105   43,6254
  11     106   43,5158
  12     107   43,0291
  13     108   41,3035
  14     109   39,2304
  15     110   38,7606
  16     111   37,5170
  17     112   36,7057
  18     113   35,8839
  19     114   35,4943
  20     115   35,8097
  21     116   37,0065
  22     117   38,2817
  23     118   39,4501
  24     119   40,0646
```

Notar que todas las medidas de bondad del ajuste (MAPE, MAD y MSD) son menores que en el caso anterior, lo cual indica que este modelo se ajusta mejor a la serie. Esto también se ve claramente en el primero de los gráficos siguientes, donde se observa que ahora el componente estacional sí se está teniendo en cuenta.

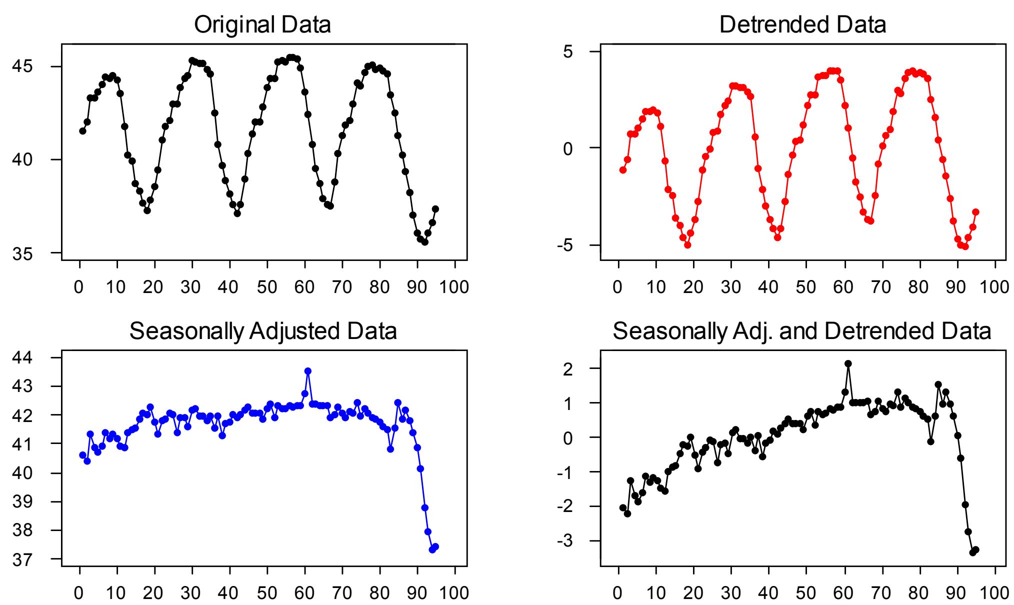
En el segundo bloque de imágenes se muestran, por separado, la serie original de observaciones, los datos una vez eliminada la tendencia, los datos una vez extraído el patrón estacional, y los datos una vez extraídos el patrón estacional y la tendencia

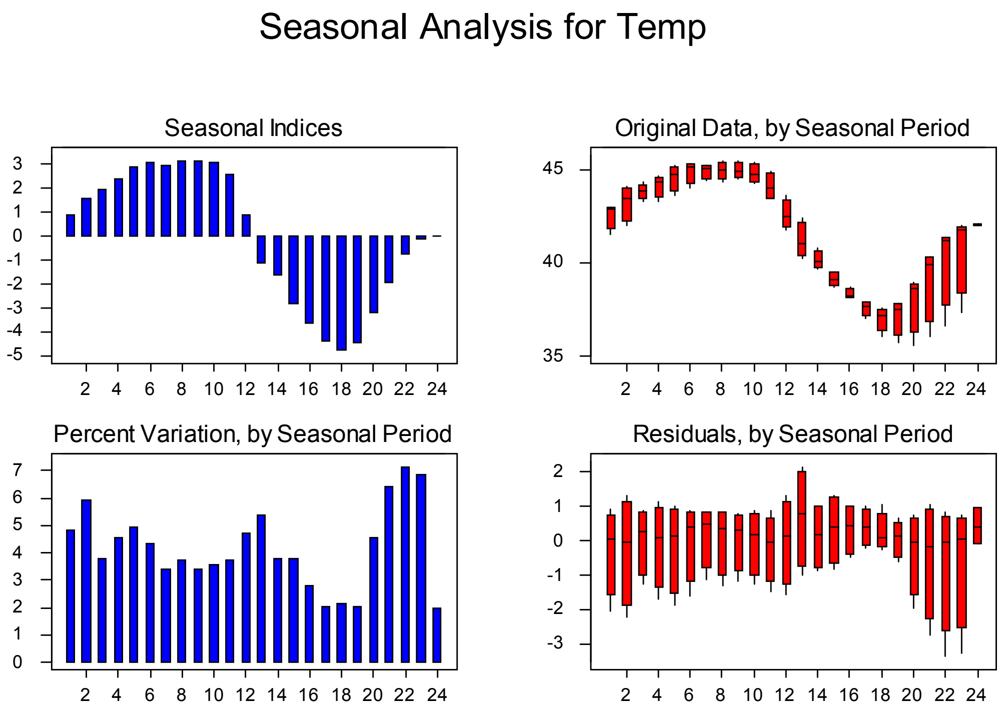
Finalmente, en el último conjunto de gráficos se muestra un análisis estacional: gráfico de índices estacionales, gráfico de variación porcentual por estaciones, gráfico de boxplots referidos a observaciones agrupadas por períodos estacionarios, y gráfico de boxplots de los residuos agrupados por períodos estacionarios.

Decomposition Fit for Temp



Component Analysis for Temp



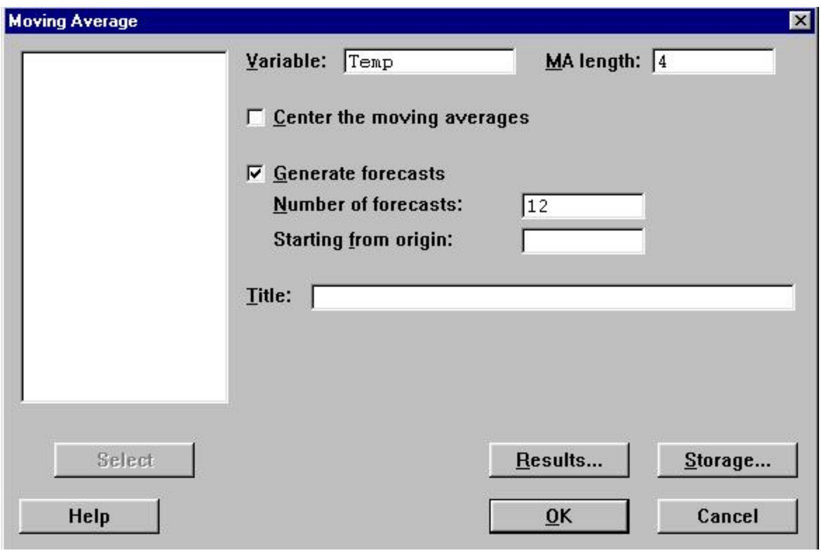


A pesar de que una transformación de la variable suele tener efectos positivos en el análisis, en este caso no parece que sea necesario a la vista del gráfico de residuos, ya que presentan un comportamiento errático alrededor del 0, sin ningún patrón claro.

Medias móviles

Utilizamos un modelo aditivo, al no observarse que la variación del patrón estacional crezca con el tiempo.

Seleccionamos ***Stat > Time Series > Moving Average***



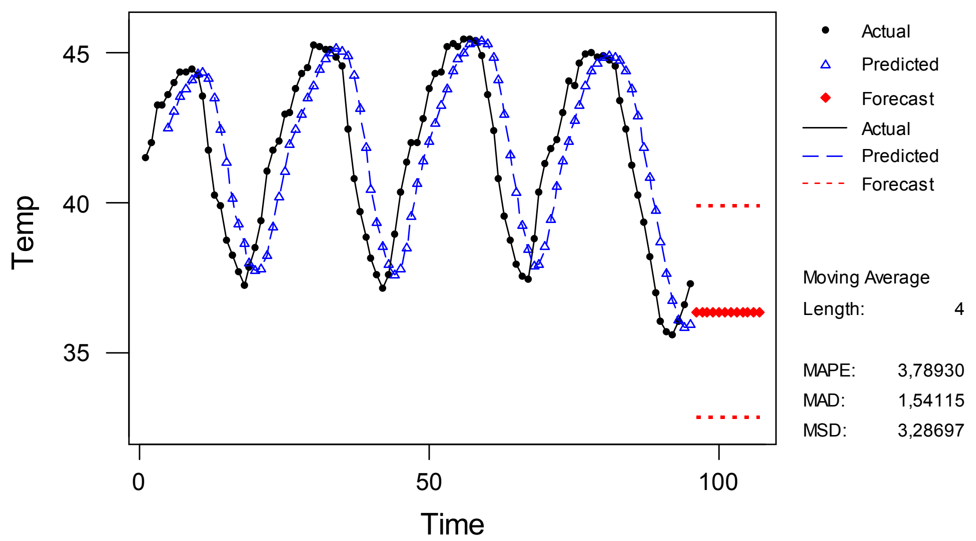
Moving average

```
Data        Temp
Length      95,0000
NMissing    0

Moving Average
Length: 4

Accuracy Measures
MAPE: 3,78930
MAD:  1,54115
MSD:  3,28697
```

Row	Period	Forecast	Lower	Upper
1	96	36,37	32,8165	39,9235
2	97	36,37	32,8165	39,9235
3	98	36,37	32,8165	39,9235
4	99	36,37	32,8165	39,9235
5	100	36,37	32,8165	39,9235
6	101	36,37	32,8165	39,9235
7	102	36,37	32,8165	39,9235
8	103	36,37	32,8165	39,9235
9	104	36,37	32,8165	39,9235
10	105	36,37	32,8165	39,9235
11	106	36,37	32,8165	39,9235
12	107	36,37	32,8165	39,9235

La predicción que nos ofrece el análisis es que la temperatura en las próximas 12 horas estará entre 32,82 y 39,92; siendo el valor medio 36,37.

En el siguiente gráfico se muestran las observaciones (Actual), los valores estimados (Predicted), y los valores pronosticados (Forecast). Observar que el patrón de los valores estimados está ligeramente desplazado a la derecha con respecto al patrón de las observaciones (ello se debe a que el valor estimado en el instante t es el valor de la media móvil en t-1).



BIBLIOGRAFÍA

[1] Baró, J. y Alemany, R. (2000): “Estadística II”. Ed. Fundació per a la Universitat Oberta de Catalunya. Barcelona.

[2] Peña Sánchez de Rivera, D. (1987): “Estadística. Modelos y Métodos. Volumen 2”. Alianza Editorial. Madrid. ISBN: 84-206-8110-5

[3] Johnson, R. R. (1996): “Elementary statistics”. Belmont, etc. : Duxbury, cop

[4] Kvanli, A. (????): “Introduction to Business Statistics”. South-Western.

[5] Martín-Guzmán, P. (1991): “Curso básico de estadística económica”. AC, DL. Madrid. ISBN: 84-7288-142-3

[5] Uriel, E. (????): “Análisis de datos”. ¿¿¿???. Valencia. ISBN: ¿¿¿¿-????

[7] Pepió, M. (2001): “Series Temporales”. Edicions UPC. Barcelona. ISBN: 84-8301-526-9

ENLACES

- http://perso.wanadoo.es/bledatobias/series.html
  Curso "Análisis, Regresión y Predicción de Series Temporales Epidemiológicas"

- http://www.ii.uam.es/~asuarez/docencia/doctorado/TS2001.html
  Curso de doctorado de la Universidad Autónoma de Madrid: "Series Temporales"