

UNIDAD 6: Práctica 27 – Modelos de Regresión Lineal

1. REGRESIÓN LINEAL SIMPLE

Los modelos de regresión lineal son modelos probabilísticos basados en una función lineal, expresamos el valor de nuestra variable de estudio (interés), a la que también llamamos variable dependiente, en función de una o más variables a quienes llamamos variables independientes o explicativas, y las cuales suponemos tienen un efecto sobre nuestra variable de estudio. Los pasos básicos a seguir en el estudio de un modelo lineal son:

- Escribir el modelo matemático con todas sus hipótesis.
- Estimación de los parámetros del modelo.
- Inferencias sobre los parámetros.
- Diagnóstico del modelo.

Ricardo Ríos <http://ricardorios.net> Informática Estadística

El modelo de regresión más simple que nos podemos encontrar es aquel en donde únicamente se considera a solamente una variable independiente, y se quiere estudiar su efecto sobre la variable dependiente; la ecuación del modelo es:

$$y_i = \beta_0 + \beta_1 x_i + u_i$$

Donde:

- y_i respondiente de la variable dependiente, es decir, el valor de la variable dependiente para el i – *simo* individuo de la muestra.
- x_i ; representa la observación i -ésima correspondiente de la variable independiente.
- β_0 ; representa el intercepto del modelo, es decir, valor de la variable dependiente cuando nuestra variable independiente toma el valor de cero. En muchos casos no tendrá interpretación, pues la variable independiente no puede tomar el valor de 0.

- β_1 ; representa la pendiente del modelo, es decir, el cambio esperado en la variable dependiente por cada cambio unitario realizado a la variable independiente.
- x_i ; representa el efecto de las demás variables omitidas en el modelo.

Las hipótesis básicas del modelo, son las mismas a las consideradas en el Análisis de Varianza, que como recordarán son las siguientes:

- El promedio de las perturbaciones es cero, es decir, se cumple que: $E[u_1] = 0; \forall i$
- La varianza de las perturbaciones es constante, es decir, se cumple que: $var(u_i) = \sigma^2; \forall i$
- La distribución de las perturbaciones debe ser normal, es decir se cumple que: $u_i \approx N(0; \sigma^2); \forall i$
- Las perturbaciones son independientes, es decir se cumple que: $cov(u_i; u_j) = 0; \forall i$

Las cuales pueden resumirse en: $u_i \sim NIID(0, \sigma^2); \forall i$

En R la función a utilizar para realizar o ajustar un modelo de regresión es `lm()` (de lineal model). Esta función no nos ofrece ninguna salida en pantalla si no que nos crea un objeto, o mejor dicho, nosotros creamos un objeto que va a ser un modelo de regresión lineal, y el cual podemos referenciarlo posteriormente en nuestro análisis.

La función `lm` tiene la siguiente sintaxis: `lm(formula, data, subset)`

- En formula escribimos: $y \sim x$, lo cual significa que a la izquierda del símbolo \sim especificamos quien es nuestra variable dependiente; mientras que a la derecha especificamos quien es nuestra variable independiente.
- En data especificamos el dataframe que contiene las variables del modelo, es recomendable que los datos se encuentren en un dataframe.
- En subset especificamos un subconjunto de observaciones para validar posteriormente el modelo. En caso que se desee utilizar conjuntos distinto para estimar y validar el modelo. Muy recomendado en muchas aplicaciones.

La función `lm` tiene muchas más opciones pero para conocer mejor su funcionamiento vamos a ver ejemplos.

- EJEMPLO 1

En el archivo “costes.dat” se encuentra la información correspondiente a 34 fábricas de producción en el montaje de placas para ordenador, el archivo contiene la información sobre el costo total (primera columna) y el número de unidades fabricadas (segunda columna). Suponga que deseamos ajustar un modelo de regresión simple a los datos para estimar el costo total en función del número de unidades fabricadas.

Ejecutamos lo siguiente. # Lectura de los datos.

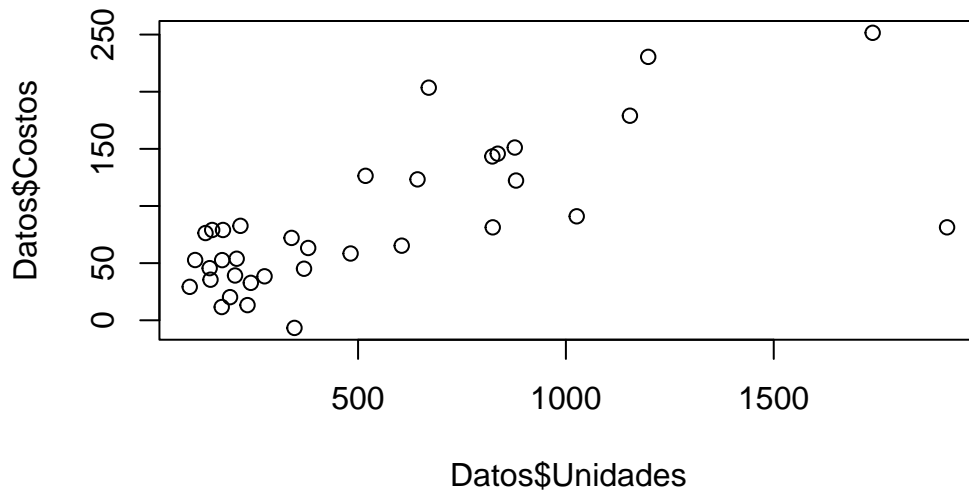
```
Datos=read.table("costes.dat")  
View(Datos)
```

Renombrando a las variables

```
names(Datos)= c ("Costos", "Unidades")
```

Realizando el diagrama de dispersión entres las dos variables

```
plot(Datos$Unidades,Datos$Costos)
```



Se aprecia una relación entre las variables por lo que se procede a ajustar el modelo de regresión

```
regresion <- lm(Datos$Costos ~ Datos$Unidades)
```

```
summary(regresion)
```

Call:

```
lm(formula = Datos$Costos ~ Datos$Unidades)
```

Residuals:

Min	1Q	Median	3Q	Max
-137.386	-24.496	-0.117	29.848	105.028

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	33.92200	11.57500	2.931	0.0061 **
Datos\$Unidades	0.09640	0.01665	5.789	1.8e-06 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 44.49 on 33 degrees of freedom

Multiple R-squared: 0.5039, Adjusted R-squared: 0.4888

F-statistic: 33.51 on 1 and 33 DF, p-value: 1.796e-06

En este caso el modelo resultante sería: $\text{costos} = 19.38 + 0.1345(\text{unidades})$

Se observa que el término constante no es significativo porque el p-valor correspondiente a la prueba de hipótesis $H_0 : \beta_0 = 0$ es 0.501; y además no tiene interpretación, pues en teoría si no se fabrican unidades no deberían existir costos asociados a la producción.

Como el término constante no es significativo se quitara del modelo, volvemos a realizar los cálculos con el R

Ejecutar lo siguiente:

```
regresion2 <- lm(Datos$Costos ~ Datos$Unidades -1)
```

```
summary(regresion2)
```

Call:

```
lm(formula = Datos$Costos ~ Datos$Unidades - 1)
```

Residuals:

Min	1Q	Median	3Q	Max
-174.579	-4.844	19.527	35.812	114.095

Coefficients:

```
              Estimate Std. Error t value Pr(>|t|)
Datos$Unidades  0.13350    0.01197   11.16 6.59e-13 ***
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Residual standard error: 49.21 on 34 degrees of freedom

Multiple R-squared: 0.7854, Adjusted R-squared: 0.7791

F-statistic: 124.5 on 1 and 34 DF, p-value: 6.591e-13

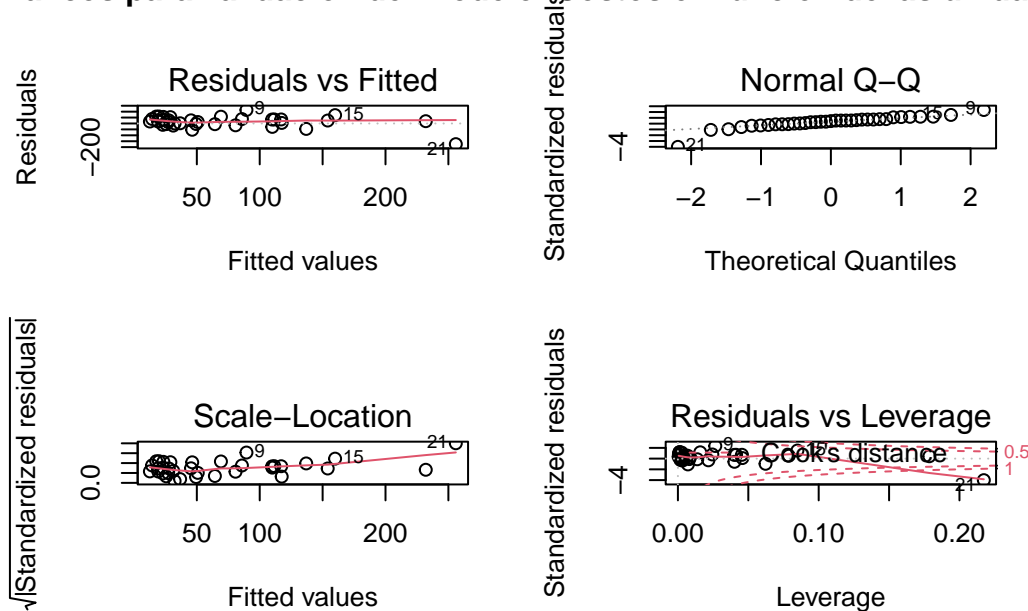
En este caso el modelo resultante sería: $\text{costos} = 0.1588(\text{unidades})$; el cual es un mejor modelo en términos de variabilidad explicada.

Una vez estimados los parámetros del modelo, el siguiente paso es validarlo, es decir verificar si se cumplen las cuatro hipótesis básicas del modelo (nulidad, normalidad, independencia y homocedasticidad de los residuos). Para verificar esto, podríamos realizar los siguientes pasos:

Efectúa un análisis gráfico de bondad de ajuste del modelo

```
par(mfrow = c(2, 2))
plot(regresion2)
par(oma=c(1,1,1,1), new=T, font=2, cex=0.5)
mtext(outer=T, "Gráficos para validación del modelo: Costos en función de las unidades",
      side=3)
```

Gráficos para validación del modelo: Costos en función de las unidades



En los gráficos que se muestra en la parte superior se contrasta los cuatro supuestos. En el de la izquierda se verifican: nulidad, independencia y homocedasticidad; a partir del gráfico mostrado parece existir indicios de falta de homocedasticidad, por su parte los residuos pueden considerarse constante pues no muestran ningún patrón; sin embargo, la media de los residuos no parece ser nula, lo cual indica falta de linealidad en el modelo (es decir, es necesario incorporar más variables o tal vez términos cuadráticos). En la figura de la derecha se contrasta la normalidad, y puede apreciarse que los residuos parecen seguir una distribución normal. # Por su parte, también es de mencionar que en el gráfico se muestran puntos que posiblemente sean observaciones atípicas, por lo que habría que estudiarlas.

Información sobre el modelo ajustado que proporciona la función `lm()` # Extrae la fórmula del modelo.

```
formula(regresion2)
```

```
Datos$Costos ~ Datos$Unidades - 1
```

Extrae el vector de coeficientes de regresión.

```
coef(regresion2)
```

```
Datos$Unidades
0.1334998
```

```
# Extrae el vector de residuos.
```

```
residuals(regresion2)
```

1	2	3	4	5	6
-5.9252241	24.9422285	57.2435787	16.1557654	29.5426287	55.6391079
7	8	9	10	11	12
37.4643315	26.5128603	114.0951533	16.5660408	34.0763027	53.6636345
13	14	15	16	17	18
26.6898859	-15.4751725	70.5629667	19.5271516	11.9533340	-5.3935511
19	20	21	22	23	24
-11.3620767	4.7973082	-174.5790615	-45.9982846	-17.9578900	-4.2938112
25	26	27	28	29	30
34.1602682	33.4958317	0.4420639	38.3225636	-53.0131662	26.0890974
31	32	33	34	35	
12.5185354	-28.6938846	58.5790516	1.7630220	59.1208689	

```
# Extrae un vector con los valores estimados.
```

```
#modelo2ted.values(regresion2)
```

```
# Extrae la matriz de covarianzas de los parámetros.
```

```
vcov(regresion2)
```

```
Datos$Unidades
Datos$Unidades 0.0001432011
```

```
# Calcula los residuales, errores estándar de los parámetros, distancias Cook.
```

```
ls.diag(regresion2)
```

```
$std.dev
[1] 49.20928
```

```
$hat
```

```

[1] 0.0137387236 0.0787524233 0.0158676214 0.0012433354 0.0017698828
[6] 0.0018110415 0.0244497405 0.0012092731 0.0265461726 0.0005337028
[11] 0.0454832506 0.0027846574 0.0068361273 0.0216452725 0.0848722948
[16] 0.1786289532 0.0024610058 0.0021799913 0.0017494809 0.0457949567
[21] 0.2173184179 0.0622511001 0.0032380535 0.0080957252 0.0413299485
[26] 0.0400545613 0.0034632436 0.0006897629 0.0071205126 0.0025584620
[31] 0.0085392456 0.0401519584 0.0010460576 0.0044721637 0.0013128794

```

\$std.res

```

[1] -0.12124442 0.52808038 1.17260833 0.32851156 0.60087863 1.13168807
[7] 0.77080784 0.53910368 2.34997086 0.33673450 0.70878358 1.09204005
[13] 0.54423845 -0.31793642 1.49895619 0.43784678 0.24320756 -0.10972401
[19] -0.23109519 0.09979981 -4.01007557 -0.96527557 -0.36552118 -0.08761148
[25] 0.70898925 0.69473704 0.00899894 0.77903567 -1.08115618 0.53084568
[31] 0.25548695 -0.59517005 1.19102962 0.03590740 1.20220642

```

\$stud.res

```

[1] -0.119473935 0.522403330 1.179328393 0.324159319 0.595144672
[6] 1.136532408 0.766111172 0.533401168 2.529690985 0.332300135
[11] 0.703499199 1.095240031 0.538526063 -0.313692652 1.528103030
[16] 0.432581090 0.239812980 -0.108117520 -0.227850384 0.098335619
[21] -5.441881509 -0.964279034 -0.360815373 -0.086323206 0.703706396
[26] 0.689354550 0.008865626 0.774436698 -1.083933598 0.525161706
[31] 0.251943705 -0.589430743 1.198655383 0.035376083 1.210400702

```

\$cooks

```

[1] 2.047755e-04 2.383898e-02 2.216993e-02 1.343476e-04 6.401583e-04
[6] 2.323641e-03 1.489076e-02 3.518799e-04 1.505953e-01 6.054895e-05
[11] 2.393841e-02 3.330121e-03 2.038767e-03 2.236389e-03 2.083829e-01
[16] 4.169239e-02 1.459274e-04 2.630304e-05 9.359475e-05 4.780083e-04
[21] 4.464949e+00 6.185333e-02 4.340279e-04 6.264812e-05 2.167080e-02
[26] 2.013939e-02 2.814313e-07 4.189037e-04 8.382848e-03 7.228166e-04
[31] 5.621878e-04 1.481789e-02 1.485440e-03 5.792050e-06 1.899999e-03

```

\$dfits

```

[1] -0.0141010149 0.1527389219 0.1497489923 0.0114372818 0.0250599080
[6] 0.0484105087 0.1212841638 0.0185600398 0.4177445106 0.0076788503
[11] 0.1535668861 0.0578762950 0.0446787837 -0.0466592443 0.4653661151
[16] 0.2017315204 0.0119114345 -0.0050535591 -0.0095385991 0.0215426354
[21] -2.8675090121 -0.2484465426 -0.0205651234 -0.0077986703 0.1461131657
[26] 0.1408138463 0.0005226419 0.0203463104 -0.0917931003 0.0265973773
[31] 0.0233816685 -0.1205549113 0.0387881801 0.0023710564 0.0438860875

```



```
$correlation
      Datos$Unidades
Datos$Unidades      1
```

```
$std.err
      [,1]
Datos$Unidades 0.01196667
```

```
$cov.scaled
      Datos$Unidades
Datos$Unidades 0.0001432011
```

```
$cov.unscaled
      Datos$Unidades
Datos$Unidades 5.913605e-08
```

Permite obtener el mejor conjunto de regresión y proporciona la estimación de los coeficientes (válido únicamente en modelos de regresión múltiple).

```
step(regresion2)
```

```
Start:  AIC=273.71
Datos$Costos ~ Datos$Unidades - 1
```

	Df	Sum of Sq	RSS	AIC
<none>			82333	273.71
- Datos\$Unidades	1	301376	383709	325.58

```
Call:
lm(formula = Datos$Costos ~ Datos$Unidades - 1)
```

```
Coefficients:
Datos$Unidades
0.1335
```

De todos los resultados anteriores nos concentraremos en la instrucción:

```
ls.diag(regresion2)
```

\$std.dev

[1] 49.20928

\$hat

[1] 0.0137387236 0.0787524233 0.0158676214 0.0012433354 0.0017698828
[6] 0.0018110415 0.0244497405 0.0012092731 0.0265461726 0.0005337028
[11] 0.0454832506 0.0027846574 0.0068361273 0.0216452725 0.0848722948
[16] 0.1786289532 0.0024610058 0.0021799913 0.0017494809 0.0457949567
[21] 0.2173184179 0.0622511001 0.0032380535 0.0080957252 0.0413299485
[26] 0.0400545613 0.0034632436 0.0006897629 0.0071205126 0.0025584620
[31] 0.0085392456 0.0401519584 0.0010460576 0.0044721637 0.0013128794

\$std.res

[1] -0.12124442 0.52808038 1.17260833 0.32851156 0.60087863 1.13168807
[7] 0.77080784 0.53910368 2.34997086 0.33673450 0.70878358 1.09204005
[13] 0.54423845 -0.31793642 1.49895619 0.43784678 0.24320756 -0.10972401
[19] -0.23109519 0.09979981 -4.01007557 -0.96527557 -0.36552118 -0.08761148
[25] 0.70898925 0.69473704 0.00899894 0.77903567 -1.08115618 0.53084568
[31] 0.25548695 -0.59517005 1.19102962 0.03590740 1.20220642

\$stud.res

[1] -0.119473935 0.522403330 1.179328393 0.324159319 0.595144672
[6] 1.136532408 0.766111172 0.533401168 2.529690985 0.332300135
[11] 0.703499199 1.095240031 0.538526063 -0.313692652 1.528103030
[16] 0.432581090 0.239812980 -0.108117520 -0.227850384 0.098335619
[21] -5.441881509 -0.964279034 -0.360815373 -0.086323206 0.703706396
[26] 0.689354550 0.008865626 0.774436698 -1.083933598 0.525161706
[31] 0.251943705 -0.589430743 1.198655383 0.035376083 1.210400702

\$cooks

[1] 2.047755e-04 2.383898e-02 2.216993e-02 1.343476e-04 6.401583e-04
[6] 2.323641e-03 1.489076e-02 3.518799e-04 1.505953e-01 6.054895e-05
[11] 2.393841e-02 3.330121e-03 2.038767e-03 2.236389e-03 2.083829e-01
[16] 4.169239e-02 1.459274e-04 2.630304e-05 9.359475e-05 4.780083e-04
[21] 4.464949e+00 6.185333e-02 4.340279e-04 6.264812e-05 2.167080e-02
[26] 2.013939e-02 2.814313e-07 4.189037e-04 8.382848e-03 7.228166e-04
[31] 5.621878e-04 1.481789e-02 1.485440e-03 5.792050e-06 1.899999e-03

\$dfits

[1] -0.0141010149 0.1527389219 0.1497489923 0.0114372818 0.0250599080
[6] 0.0484105087 0.1212841638 0.0185600398 0.4177445106 0.0076788503
[11] 0.1535668861 0.0578762950 0.0446787837 -0.0466592443 0.4653661151
[16] 0.2017315204 0.0119114345 -0.0050535591 -0.0095385991 0.0215426354

```
[21] -2.8675090121 -0.2484465426 -0.0205651234 -0.0077986703 0.1461131657
[26] 0.1408138463 0.0005226419 0.0203463104 -0.0917931003 0.0265973773
[31] 0.0233816685 -0.1205549113 0.0387881801 0.0023710564 0.0438860875
```

```
$correlation
          Datos$Unidades
Datos$Unidades          1
```

```
$std.err
          [,1]
Datos$Unidades 0.01196667
```

```
$cov.scaled
          Datos$Unidades
Datos$Unidades 0.0001432011
```

```
$cov.unscaled
          Datos$Unidades
Datos$Unidades 5.913605e-08
```

Con esta instrucción obtenemos para cada observación en el conjunto de datos, medidas que nos ayudarán a identificar observación atípicas (tienen un impacto únicamente en las medidas resumen del modelo) y observaciones influyentes (tienen un efecto marcado en la estimación de los parámetros). Al digitar la instrucción anterior en R se mostrará los siguientes resultados (cada uno de ellos en un vector).

- $\$ \hat{H}$. Corresponde a los elementos de la diagonal de la matriz $H = X(X^t X)^{-1} X^t$, y se examina H_{ii} que mide la distancia de X_i (observación i -ésima) al centro de los datos (medida estandarizada). Los elementos grandes indican observaciones potencialmente influyentes. Si se cumple que $H_{ii} > 2 \left(\frac{k+1}{n} \right)$ se trata de una observación influyente.
- $\$std.res$. Son los residuos estandarizados (la varianza de los residuos se supone es la misma) del modelo. Una observación se considera influyente si su residuo estandarizado es mayor en valor absoluto a 3.
- $\$stud.res$. Son los residuos estudentizados del modelo (se considera que la varianza de los residuos es diferente); estos residuos siguen una distribución t de Student para $n - 3$ grados de libertad. Por lo que si para una observación su residuo estandarizado es mayor en valor absoluto al percentil 95 de la distribución t de Student se considera como punto influyente.
- $\$cooks$. Es la distancia de Cook (mide el efecto de eliminar una observación en la estimación de cada uno de los parámetros, el efecto se mide en desviaciones típicas). Si dicha distancia es mayor a 1 el punto se considera como influyente.

- *\$dfits*. Es el valor del DDFITS (mide el cambio ocurrido en la estimación de una observación cuando esta observación es descartada y luego incluida en el modelo).

Se considera que una observación es influyente si su correspondiente DDFITS es mayor, en valor absoluto, a $2\sqrt{\frac{k+1}{n}}$. Donde k es el número de variables en el modelo (en regresión simple es igual a 1).

2. REGRESIÓN LINEAL MÚLTIPLE

Al igual que en el modelo de regresión simple, el modelo de regresión múltiple trata de ajustar una ecuación matemática en la que se relacione a una única variable dependiente en función de dos o más variables independientes. La forma general del modelo es la siguiente: $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + u_i$

Como siempre debe cumplirse que: $u_i \sim NIID(o, \sigma^2); \forall i$

La función para estimar cada uno de los parámetros del modelo, a partir de la información suministrada por la muestra, los datos disponibles, es como siempre `lm()`, sin embargo, en la expresión fórmula debemos escribir $y \sim x_1 + x_2 + \dots + x_k$. Todas las instrucciones utilizadas en regresión simple son válidas también para regresión múltiple (diagnosis de los residuos e identificación de puntos influyentes).

Veamos el siguiente ejemplo. - EJEMPLO 2.

En el archivo “precioscasas.dat” tienen la información sobre 100 datos de precios de viviendas y sus características, el archivo se encuentra estructurado de la siguiente forma:

- Primera columna: precios de viviendas en euros.
- Segunda columna: superficie en metros cuadrados.
- Tercera: numero de cuartos de baño.
- Cuarta: número de dormitorios.
- Quinta: número de plazas de garaje.
- Sexta: edad de la vivienda .
- Séptima: 1 =buenas vistas y 0 =vistas corrientes

Suponga que deseamos estimar un modelo de regresión en el cual relacionemos el precio de una vivienda en función de sus características.

Ejecutar lo siguiente: `# leyendo los datos`

```

datos <- read.table(file="preciocasas.dat")

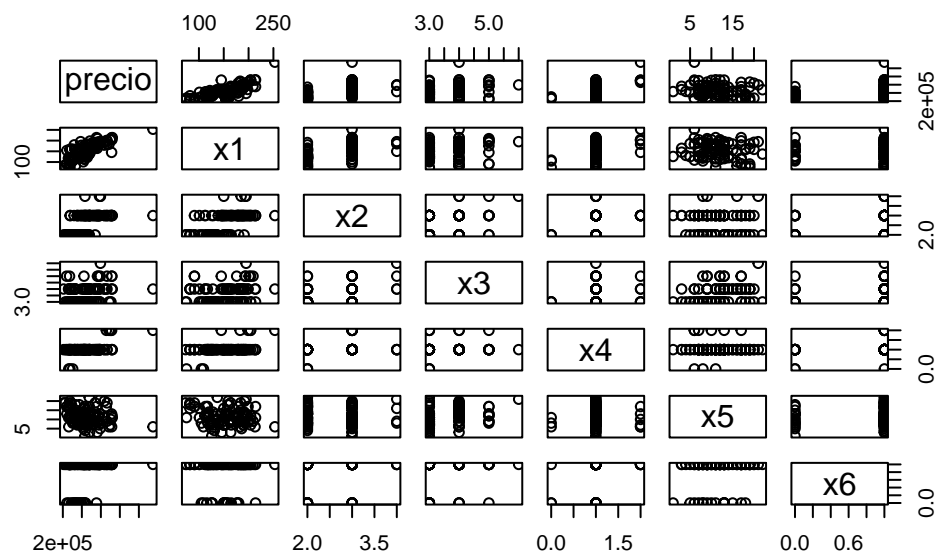
# nombrando a las columnas

names(datos) <- c("precio", "x1", "x2", "x3", "x4", "x5", "x6" )

# haciendo la matriz de diagramas de dispersión

plot(datos)

```



se observa gráficamente que las variables independientes parecen influir en el comportamiento de nuestra variable dependiente.

ajustamos el modelo de regresión

```

modelo1 <- lm( precio ~ x1 + x2 + x3 + x4 + x5 + x6 , data = datos)

#resumen del modelo

summary(modelo1)

```

```
Call:
lm(formula = precio ~ x1 + x2 + x3 + x4 + x5 + x6, data = datos)
```

Residuals:

Min	1Q	Median	3Q	Max
-101248	-23050	-345	18036	141928

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	29844.7	26365.3	1.132	0.26056
x1	1159.3	142.9	8.112	1.98e-12 ***
x2	13284.5	9286.2	1.431	0.15591
x3	8695.2	6708.7	1.296	0.19814
x4	59777.1	14604.0	4.093	9.06e-05 ***
x5	-3198.4	974.3	-3.283	0.00145 **
x6	34312.9	10963.6	3.130	0.00234 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 38920 on 93 degrees of freedom

Multiple R-squared: 0.7505, Adjusted R-squared: 0.7344

F-statistic: 46.61 on 6 and 93 DF, p-value: < 2.2e-16

de los resultados anteriores puede apreciarse que el intercepto, y las variables x_2 (número de cuarto de baño) y x_3 (número de dormitorios) no parecen influir en la estimación del precio de la vivienda por lo podrían descartarse de la ecuación.

una forma alternativa y mucho más eficiente para seleccionar el mejor conjunto de variables independientes es utilizar la instrucción `step()`, con la cual se utilizan los algoritmos conocidos para seleccionar variables (selección hacia adelante -“forward”-, hacia atrás -“backward”- o selección por pasos -“both”-).

```
step(modelo1, direction="both")
```

Start: AIC=2120.58

precio ~ x1 + x2 + x3 + x4 + x5 + x6

	Df	Sum of Sq	RSS	AIC
- x3	1	2.5444e+09	1.4340e+11	2120.4

```

<none>                1.4086e+11 2120.6
- x2      1 3.0996e+09 1.4395e+11 2120.8
- x6      1 1.4835e+10 1.5569e+11 2128.6
- x5      1 1.6322e+10 1.5718e+11 2129.6
- x4      1 2.5376e+10 1.6623e+11 2135.2
- x1      1 9.9664e+10 2.4052e+11 2172.1

```

Step: AIC=2120.37

precio ~ x1 + x2 + x4 + x5 + x6

	Df	Sum of Sq	RSS	AIC
<none>			1.4340e+11	2120.4
+ x3	1	2.5444e+09	1.4086e+11	2120.6
- x2	1	5.3488e+09	1.4875e+11	2122.0
- x5	1	1.3780e+10	1.5718e+11	2127.6
- x6	1	1.6460e+10	1.5986e+11	2129.2
- x4	1	2.4664e+10	1.6806e+11	2134.2
- x1	1	1.0510e+11	2.4850e+11	2173.4

Call:

```
lm(formula = precio ~ x1 + x2 + x4 + x5 + x6, data = datos)
```

Coefficients:

(Intercept)	x1	x2	x4	x5	x6
42224	1182	16724	58864	-2686	35913

- EJERCICIO 1.

Se deja como ejercicio al estudiante, elegir el mejor conjunto de variables a incluir en el modelo, y para el modelo resultante (llamarlo modelo2), realizar el diagnóstico de los residuos y el estudio de las observaciones atípicas e influyentes.

```
coefficients(modelo1) # coeficientes del modelo
```

(Intercept)	x1	x2	x3	x4	x5
29844.731	1159.255	13284.517	8695.202	59777.054	-3198.381
	x6				
	34312.868				

```
confint(modelo1, level=0.95) # Intervalos de confianza para los parámetros
```

	2.5 %	97.5 %
(Intercept)	-22511.5521	82201.014
x1	875.4687	1443.041
x2	-5156.1158	31725.150
x3	-4626.8649	22017.269
x4	30776.4939	88777.614
x5	-5133.1468	-1263.615
x6	12541.3783	56084.359

```
fitted(modelo1) # valores estimados
```

1	2	3	4	5	6	7	8
279222.7	266544.8	279358.4	189667.8	257849.6	296943.2	339872.9	314353.6
9	10	11	12	13	14	15	16
235171.0	290392.2	178634.1	318058.5	328027.0	263914.7	283839.6	295909.1
17	18	19	20	21	22	23	24
294267.0	270922.5	281152.3	374349.5	287379.2	329024.5	271191.4	303744.5
25	26	27	28	29	30	31	32
335401.2	315658.1	301569.6	348909.3	302264.1	352862.4	377290.2	353524.7
33	34	35	36	37	38	39	40
379091.6	393779.0	380187.5	434665.3	411467.3	532447.2	219574.1	200445.6
41	42	43	44	45	46	47	48
245099.4	229169.3	245099.4	168574.1	243291.9	258491.8	290025.0	222772.5
49	50	51	52	53	54	55	56
287794.3	290350.5	264479.6	274721.4	299945.6	288891.8	275192.3	365697.6
57	58	59	60	61	62	63	64
295629.7	265042.8	325082.0	361924.8	376089.2	411489.0	393554.2	272961.6
65	66	67	68	69	70	71	72
308219.7	301037.1	328088.8	294749.8	328300.4	298428.6	334567.6	325423.2
73	74	75	76	77	78	79	80
363896.2	356423.6	346248.0	358762.2	349529.8	389720.9	330018.5	386584.3
81	82	83	84	85	86	87	88
365655.9	327917.5	356401.9	349269.0	367435.7	420763.0	379008.2	387743.5
89	90	91	92	93	94	95	96
397530.2	389503.3	418697.8	388132.6	370893.4	357178.3	351102.6	442604.2
97	98	99	100				
463750.1	454155.0	426300.0	405371.6				

```
residuals(modelo1) # residuos
```


1	2	3	4	5
-57889.68926	-36565.81290	-49379.40675	41695.15707	-22682.61106
6	7	8	9	10
-60047.16363	-101247.91267	-68984.63691	10370.98785	-38999.18751
11	12	13	14	15
73135.93784	-57967.51228	-65194.03290	-1081.71872	-15818.63830
16	17	18	19	20
-19588.07789	-17945.95009	9202.48178	5197.68004	-80736.51621
21	22	23	24	25
6578.80514	-31607.52031	37568.64742	7159.52240	-24151.21283
26	27	28	29	30
2508.86990	18115.36629	-21592.29001	34207.88296	-13600.37389
31	32	33	34	35
-37858.16771	9600.34471	1325.36583	-6791.04628	10604.48751
36	37	38	39	40
-5832.26646	48490.68128	141927.79018	-15532.13765	12242.35614
41	42	43	44	45
-25495.39355	2365.71917	-9932.39355	68992.85235	-4666.94695
46	47	48	49	50
-14844.82285	-30649.96225	43173.48129	-752.29527	13982.53546
51	52	53	54	55
42893.36326	34270.61190	10266.39228	22358.24265	38824.68839
56	57	58	59	60
-49259.62111	32912.30503	73874.20103	19022.01500	8117.19902
61	62	63	64	65
17295.80992	20803.02791	62945.75914	-10128.64462	-31898.72900
66	67	68	69	70
-24716.08965	-34130.82548	6671.17677	-26733.43567	18009.35116
71	72	73	74	75
-18129.63248	-5527.16050	-40542.15465	-31340.58382	-11827.02143
76	77	78	79	80
-21920.16275	-8883.84843	-47345.86309	12356.54383	-32451.27461
81	82	83	84	85
-5988.89813	42124.46281	16752.06955	27689.03797	12981.28878
86	87	88	89	90
-38617.00938	8324.81178	2702.47073	-4663.21304	6475.72725
91	92	93	94	95
-22026.83322	9575.40941	47564.59441	61279.71336	81189.36959
96	97	98	99	100
62.81828	-14167.14854	-1113.00535	30200.03063	52857.40711

```
influence(modelo1) # puntos de influencia
```

```

$hat
      1      2      3      4      5      6      7
0.07644959 0.04329746 0.03101799 0.16470221 0.04905149 0.10054929 0.01897510
      8      9     10     11     12     13     14
0.06554873 0.07650649 0.03416588 0.15576025 0.10836388 0.04982851 0.07485435
     15     16     17     18     19     20     21
0.09354194 0.03757592 0.08757059 0.03427498 0.13159339 0.05890256 0.07085183
     22     23     24     25     26     27     28
0.07729138 0.09235552 0.04200637 0.12319478 0.06365060 0.07744857 0.04052721
     29     30     31     32     33     34     35
0.03219950 0.03809646 0.02541698 0.03275262 0.07081194 0.06528581 0.04193972
     36     37     38     39     40     41     42
0.13289981 0.04986938 0.16786143 0.07359808 0.09044749 0.11550283 0.06575711
     43     44     45     46     47     48     49
0.11550283 0.15413191 0.10208614 0.03911696 0.05499194 0.06973092 0.03927547
     50     51     52     53     54     55     56
0.03295164 0.07257119 0.03206266 0.03181964 0.03757695 0.05922658 0.02390075
     57     58     59     60     61     62     63
0.04319111 0.10588291 0.02262057 0.09015279 0.02702758 0.08352555 0.20328452
     64     65     66     67     68     69     70
0.03199632 0.08829534 0.04839577 0.04638077 0.03773781 0.03545056 0.15835318
     71     72     73     74     75     76     77
0.12184146 0.04217239 0.02977390 0.02462946 0.03748933 0.03614752 0.06572773
     78     79     80     81     82     83     84
0.03282231 0.03668504 0.04376268 0.02187867 0.04667088 0.03649125 0.15511531
     85     86     87     88     89     90     91
0.03576599 0.09019921 0.02406875 0.04441708 0.08172344 0.04900424 0.10789185
     92     93     94     95     96     97     98
0.16871732 0.02092219 0.07365580 0.05656898 0.14285883 0.13476353 0.13227248
     99    100
0.06338635 0.07303966

```

```

$coefficients
      (Intercept)      x1      x2      x3      x4
1    -4227.090617 -16.01194059 1278.2081179  508.0538595 -718.264482
2     -332.494918   3.98687049  872.7083389 -457.9104431 -242.979377
3    -3908.905468   5.28021271  795.6853947  376.5560717 -278.410450
4     8976.007651   0.28821773 -303.3697850  223.3498552 -5199.729746
5    -1216.589486   0.65829628  264.2962209  423.0197697  -77.191809
6    -1299.786890 -39.40326282 1883.5157231 1749.5181453 1228.339847
7     2561.333249  13.30258539 -1437.9884477 -625.6021869  761.253504
8     1096.128770 -37.36592369 3106.4306263 -506.0976325 1120.074014
9      942.764318  -2.88554528  -36.4772476  -58.2867279  332.353187

```

10	-3384.699795	3.39207970	752.2125420	169.4473776	-186.340610
11	14961.049526	-1.19844346	-265.3043156	111.0311345	-8950.546768
12	-899.432006	-34.41247374	2852.4321459	-1275.2585546	-141.163489
13	-560.062942	-26.98974686	3032.8785222	-1081.2036829	793.605962
14	-118.457082	0.22734899	13.1385971	-3.1745651	-31.565079
15	-268.862409	-4.18879688	555.5886876	-313.3694953	-219.271434
16	-1778.493995	1.51023181	409.8337411	52.9185261	-85.290138
17	-764.811605	5.58523647	-730.4544316	315.7125092	-318.225001
18	600.197396	-0.36818940	-135.7843200	-121.1663049	32.353534
19	26.402059	-0.33488035	226.0398350	-173.0608562	51.770712
20	370.758161	-43.26778083	-707.3502575	2442.6161836	2583.786005
21	615.213007	0.94068318	-148.7481259	-5.6584646	109.669589
22	837.293817	-1.75376691	-478.3271520	-274.5887413	-167.059882
23	4779.254733	-11.29577510	-502.9110188	336.0131435	1222.764085
24	673.158315	-0.27973335	-168.4983576	-8.6474396	21.013987
25	1814.605579	1.44565814	-85.7446243	-1043.5213258	-246.593185
26	274.611008	-0.16600133	-67.6931907	10.5540804	9.238930
27	1491.220860	6.91744621	-533.5215263	-72.6476356	153.160246
28	1.551722	6.65268041	-306.3758583	-343.6647552	41.883955
29	2606.650211	3.83328604	-844.3714076	-245.4265045	-74.926075
30	817.549220	-4.15727643	-80.3133598	65.1139889	310.475991
31	853.094863	-5.56392430	-50.1038242	-367.8672604	669.672775
32	257.988257	0.44319746	187.6687045	-191.0461372	-140.505672
33	15.870421	-0.23109060	6.3130521	30.1244078	-9.840414
34	610.793070	-2.32689387	111.9952416	-238.1972509	166.144315
35	177.084349	3.92702358	94.4303937	-232.5818271	-278.190164
36	860.341166	-0.49080311	7.6718913	19.7866902	-738.461206
37	-2348.979556	33.14787811	-678.2769422	167.6729744	-1786.452824
38	-17509.859560	104.43325943	-5485.1736125	2074.6245717	15324.677776
39	-252.346540	9.08597918	97.7894629	-187.2106487	-368.866033
40	824.636286	-7.89507797	134.3398235	-240.3775799	311.929220
41	643.027149	15.32354165	658.8295344	-1259.3722135	-630.887687
42	71.753121	-1.50267197	-19.3669430	41.5678216	59.739210
43	250.507947	5.96968413	256.6641776	-490.6211950	-245.778704
44	12859.103660	-17.91620546	202.7578138	-581.9568627	-9327.174331
45	145.564223	4.53082820	-244.5106356	6.6394321	-101.126000
46	-1305.864335	5.35046872	115.7619502	92.8245612	-217.121124
47	198.920350	22.98590910	-1081.1019959	-303.5891837	-426.371251
48	904.028520	-25.94718296	-298.7749563	599.4203339	1045.655779
49	-70.807036	0.13201508	13.2948856	1.6154793	-5.897439
50	886.780879	1.92267393	-306.1086109	-174.0800409	-39.850186
51	-881.827993	-36.46221935	1883.1211446	148.2263485	752.702850
52	2777.349945	-5.57821899	-488.9285676	-250.7575080	261.023386

53	791.210187	0.86402111	-243.7434823	-72.0010396	-12.356352
54	399.681737	-0.17375459	-687.9158038	380.6737777	62.683887
55	-1184.798030	-25.14941681	1481.8704365	22.0594298	413.406711
56	889.005642	-0.42348439	-292.7710852	-516.6255151	629.148112
57	3202.965438	-4.53429834	-666.9137207	-21.3299700	211.637547
58	1666.700393	27.09217012	-1410.9857587	-2314.9260394	-740.321103
59	-484.623970	-4.79880558	370.7542475	98.8484886	-60.989735
60	-90.734880	7.51204788	-517.7547060	103.0733660	-244.171601
61	-782.505070	6.16354378	-14.2375961	30.2835067	-428.995091
62	-1675.769908	8.24354240	-462.8007549	851.1603095	-558.802336
63	-10194.781335	-63.24790612	714.8487483	2995.5731700	11124.379201
64	-709.929744	0.72497562	150.7100521	113.9350625	-46.068046
65	-1965.072247	11.86595157	-1198.9229267	319.4147321	-623.605163
66	-747.636421	10.46105690	-971.8142915	562.5564726	-49.648958
67	-2897.206164	-8.79506156	1146.0406708	99.0714794	258.072388
68	608.821969	-0.60791358	-136.4741485	-17.4938312	32.361160
69	-826.498074	5.38744373	-778.4039674	541.5818297	155.722947
70	2002.766694	-16.67375545	896.8733786	70.3450012	704.765662
71	1641.382809	11.94962102	-1559.6553919	254.0801955	110.652321
72	-243.244939	1.94224636	-175.9076660	89.0221566	4.269793
73	2118.718654	-13.10571793	-111.3706229	59.0223603	953.032241
74	454.340658	3.19652063	-302.3818383	-348.6592668	277.902092
75	-145.470068	-1.51277298	-243.5559963	302.9040269	205.222094
76	-802.432378	0.07185263	-412.7403683	356.4117075	286.969187
77	-582.444543	3.11902241	-231.0628019	68.6059122	13.591751
78	2129.539109	-22.07377830	267.6009254	-135.2556784	1365.683343
79	229.887446	-0.84227737	335.2663535	-302.4253466	-128.525829
80	-843.498695	-10.87520015	-248.4450250	594.3257750	816.934374
81	245.581544	-1.37992322	-19.9572446	-14.6582042	121.463579
82	-2639.114680	4.97593939	623.4366526	-303.8937894	-679.734450
83	237.497649	3.24873012	287.0345841	-406.1752877	-330.947720
84	-507.771005	22.35065995	389.8587108	-1031.0384107	-575.042764
85	279.735321	2.77954932	181.6670702	-271.5273030	-267.683679
86	3281.126424	-19.99358454	1018.6029701	-1565.5639562	1206.982587
87	-288.825569	2.31975508	-5.6116758	46.3473386	-184.598013
88	69.063954	0.94441704	19.4282174	-49.7449892	-69.428810
89	323.559041	1.15285259	-281.9396702	13.9310759	93.194749
90	238.434755	1.67292249	50.8252425	-93.8551284	-146.948493
91	2594.208849	5.25972213	-976.5664023	-694.0141846	457.237133
92	-2370.692904	-0.45656606	400.4425037	402.2675913	-267.761011
93	-1497.161858	8.62605653	121.4633213	290.0465319	-889.029156
94	-5048.827659	33.56096586	137.7286855	-712.8092850	-1936.717903
95	-5887.791028	32.89919767	437.6922346	-723.8958103	-2145.274714

96	-2.194293	-0.01867873	0.8189008	-0.8927769	8.804631
97	712.157425	-0.74438811	-37.7357436	254.4893601	-1794.307591
98	72.828313	-0.12436032	-5.2849082	26.6692121	-138.519974
99	-756.120135	17.74050493	-511.5506005	390.5122026	-1027.782614
100	-5041.585863	25.95155590	-1136.2541923	1825.7481625	-1576.126894
	x5	x6			
1	-112.4132787	3514.8068219			
2	-60.9311326	-423.7930175			
3	19.4753790	-956.9771960			
4	-141.5620527	-1831.3785561			
5	-79.7634794	-394.9183079			
6	-396.8977434	-1453.7171024			
7	29.1178563	-875.8766750			
8	-219.7113410	-1273.0043561			
9	12.8154798	-692.5395758			
10	63.2741991	-807.9964274			
11	-143.0361923	-3285.0110313			
12	-31.8460010	3752.2022530			
13	13.1684692	-1273.1845264			
14	2.1090654	68.3326815			
15	-23.2345029	1080.0570220			
16	43.9507797	-419.8013988			
17	-21.1717404	1294.6073968			
18	14.4389585	171.1633782			
19	33.5237888	-400.3785981			
20	-238.5814513	-1674.8019070			
21	-5.5387982	-392.0983161			
22	-36.3349657	2322.8883766			
23	-153.9085457	-2391.2758980			
24	-20.3093843	160.9507127			
25	19.9266467	2026.5723119			
26	-12.2274189	61.0764985			
27	2.0018738	-1050.3511605			
28	79.1899425	-207.2656486			
29	-25.3477614	763.5726894			
30	-46.3862783	-149.3112233			
31	67.9971488	-504.3545467			
32	-2.2209447	164.1512440			
33	-8.2279044	18.1197954			
34	9.8994703	-63.6388056			
35	1.8928601	220.5353017			
36	-20.6935779	69.9722591			
37	5.2972360	892.7797500			

38	-493.0549498	184.3314797
39	-37.6726504	-88.7456534
40	51.7164144	127.0176326
41	17.3493946	-51.6576396
42	0.9285681	15.8445083
43	6.7589078	-20.1245768
44	-49.5192857	2941.3795616
45	-15.6270975	16.5482210
46	10.6772479	-249.1828026
47	33.0882768	-69.9512778
48	75.1930984	260.5664722
49	1.7850319	-15.3444253
50	16.3212300	294.1595545
51	60.0329324	-34.5861986
52	-15.9618439	644.0363022
53	-7.9683363	225.9498833
54	-2.2594154	317.6008356
55	87.9331587	28.7788092
56	96.9998673	-580.9385131
57	-97.6106500	705.3607373
58	563.5726227	1478.4579135
59	3.3665566	130.2661406
60	8.7794434	204.7068904
61	16.4998690	234.1200937
62	-70.8812255	244.0959688
63	-166.4756666	-1780.4460196
64	-9.1171521	-189.7681137
65	49.4522043	2252.8593792
66	-26.5294862	-263.0507020
67	84.3622500	-881.4447198
68	-15.0899764	141.9766524
69	-2.1305767	-371.9539239
70	-119.8919720	-1415.2840764
71	-56.6552769	68.8832005
72	7.5821381	-74.3541828
73	-87.5790767	-492.7314059
74	66.1727186	-332.0478063
75	-20.7762586	-194.9919234
76	33.9867023	-387.3888571
77	41.1500436	-147.5986488
78	-22.3185448	-725.3095029
79	18.6060238	177.1166843
80	36.7359623	-697.9657928

81	-4.7373806	-72.5284163
82	188.2179428	334.1514606
83	20.2909967	296.8482299
84	194.5198924	-1911.5797297
85	-0.2395637	246.4357207
86	126.7148694	-506.4295173
87	-3.2141404	114.0597066
88	-3.0128437	58.5811746
89	5.2928652	-20.6427892
90	-16.2574128	141.4172255
91	81.0234177	-2.3554635
92	36.2930253	-106.6175722
93	-24.1422034	599.5361548
94	355.0357185	751.6703116
95	396.1132770	904.1556496
96	-0.1716774	-0.2752113
97	22.1913599	16.1309102
98	-0.6733397	2.5000967
99	-98.2511653	611.1052056
100	-67.9086896	583.4102407

\$sigma

	1	2	3	4	5	6	7	8
38621.12	38933.81	38777.32	38838.26	39053.19	38567.60	37649.06	38414.54	
9	10	11	12	13	14	15	16	
39112.23	38909.07	38238.28	38601.42	38502.10	39128.24	39090.05	39073.00	
17	18	19	20	21	22	23	24	
39079.36	39116.23	39124.09	38154.24	39121.94	38977.74	38911.83	39120.98	
25	26	27	28	29	30	31	32	
39035.90	39127.48	39078.97	39060.86	38960.11	39101.69	38923.61	39115.18	
33	34	35	36	37	38	39	40	
39128.15	39121.56	39112.11	39122.96	38783.15	35607.78	39092.23	39105.52	
41	42	43	44	45	46	47	48	
39026.20	39127.58	39112.92	38338.83	39125.04	39096.55	38990.09	38849.11	
49	50	51	52	53	54	55	56	
39128.33	39100.32	38851.89	38959.51	39113.29	39056.20	38905.23	38781.59	
57	58	59	60	61	62	63	64	
38970.85	38271.25	39076.96	39118.35	39085.69	39062.77	38431.46	39113.69	
65	66	67	68	69	70	71	72	
38973.09	39039.15	38958.37	39121.99	39025.36	39074.85	39076.39	39123.98	
73	74	75	76	77	78	79	80	
38892.40	38988.29	39108.22	39059.11	39116.68	38805.16	39106.39	38975.15	
81	82	83	84	85	86	87	88	

39123.32	38869.02	39087.94	39002.17	39104.13	38900.08	39118.55	39127.35
89	90	91	92	93	94	95	96
39125.12	39122.29	39052.80	39113.09	38806.13	38561.25	38145.61	39128.41
97	98	99	100				
39096.18	39128.21	38992.93	38707.51				

\$wt.res

1	2	3	4	5
-57889.68926	-36565.81290	-49379.40675	41695.15707	-22682.61106
6	7	8	9	10
-60047.16363	-101247.91267	-68984.63691	10370.98785	-38999.18751
11	12	13	14	15
73135.93784	-57967.51228	-65194.03290	-1081.71872	-15818.63830
16	17	18	19	20
-19588.07789	-17945.95009	9202.48178	5197.68004	-80736.51621
21	22	23	24	25
6578.80514	-31607.52031	37568.64742	7159.52240	-24151.21283
26	27	28	29	30
2508.86990	18115.36629	-21592.29001	34207.88296	-13600.37389
31	32	33	34	35
-37858.16771	9600.34471	1325.36583	-6791.04628	10604.48751
36	37	38	39	40
-5832.26646	48490.68128	141927.79018	-15532.13765	12242.35614
41	42	43	44	45
-25495.39355	2365.71917	-9932.39355	68992.85235	-4666.94695
46	47	48	49	50
-14844.82285	-30649.96225	43173.48129	-752.29527	13982.53546
51	52	53	54	55
42893.36326	34270.61190	10266.39228	22358.24265	38824.68839
56	57	58	59	60
-49259.62111	32912.30503	73874.20103	19022.01500	8117.19902
61	62	63	64	65
17295.80992	20803.02791	62945.75914	-10128.64462	-31898.72900
66	67	68	69	70
-24716.08965	-34130.82548	6671.17677	-26733.43567	18009.35116
71	72	73	74	75
-18129.63248	-5527.16050	-40542.15465	-31340.58382	-11827.02143
76	77	78	79	80
-21920.16275	-8883.84843	-47345.86309	12356.54383	-32451.27461
81	82	83	84	85
-5988.89813	42124.46281	16752.06955	27689.03797	12981.28878
86	87	88	89	90
-38617.00938	8324.81178	2702.47073	-4663.21304	6475.72725

91	92	93	94	95
-22026.83322	9575.40941	47564.59441	61279.71336	81189.36959
96	97	98	99	100
62.81828	-14167.14854	-1113.00535	30200.03063	52857.40711