3,4

TRABAJO CORTO 01 – ESTADISTICA II

ANDRÉS DUQUE RENDÓN

CRISTÒBAL HENAO RUEDA

MANUEL JOSÉ RAMÍREZ PINEDA

PEDRO ALEJANDRO PACHECO BOHORQUEZ

GRUPO 38

UNIVERSIDAD NACIONAL DE COLOMBIA - SEDE MEDELLÍN

30/MARZO/2023

Preguntas a resolver.

1. Estime un modelo de regresión lineal múltiple que explique el riesgo de infección en términos de las variables restantes (actuando como predictoras) Analice la significancia de la regresión y de los parámetros individuales. Interprete los parámetros estimados. Calcule e interprete el coeficiente de determinación múltiple R2 .

2. Use la tabla de todas las regresiones posibles, para probar la significancia simultánea del subconjunto de tres variables con los valores p más grandes del punto anterior. ¿Según el resultado de la prueba es posible descartar del modelo las variables del subconjunto? Explique su respuesta.

3. Plantee una pregunta donde su solución implique el uso exclusivo de una prueba de hipótesis lineal general de la forma H0 : Lβ = 0 (solo se puede usar este procedimiento y no SSextra). Especifique claramente la matriz L, el modelo reducido y la expresión para el estadístico de prueba (no hay que calcularlo).

4. Realice una validación de los supuestos en los errores y examine si hay valores atípicos, de balanceo e influenciales. ¿Qué puede decir acerca de la validez de éste modelo? Argumente su respuesta.

SOLUCIÓN

13pt

1. Con base en la tabla de parámetros estimados se obtiene la ecuación de regresión ajustada:

No va en ec. ajustada

$$\hat{Y}_i = -\ 1.5124 + 0.0313x_{i1} + 0.0455x_{i2} + 0.0315x_{i3} + 0.0249x_{i4} + 0.0026x_{i5} + \varepsilon_i$$

X

$i = 1, 2, \ldots$

Donde:
Y: Riesgo de infección Probabilidad promedio estimada (en porcentaje).
X1: Duración de la estadía (en días).
X2: Rutina de cultivos Razón del número de cultivos realizados en pacientes sin síntomas de infección hospitalaria, por cada 100.
X3: Número de camas
X4: Censo promedio diario
X5: Número de enfermeras

2pt

TABLA DE PARAMETROS ESTIMADOS				
	Estimado	error estd.	Estadístico t	Valor p
β0	-1,5124812	2,0046923	-0,7544705	0,4545860
β1	0,0313136	0,1177962	0,2658286	0,7916129
β2	0,0455480	0,0378258	1,2041533	0,2349682
β3	0,0315445	0,0187667	1,6808751	0,0998738
β4	0,0249176	0,0095867	2,5991873	0,0126700
β5	0,0026273	0,0009807	2,6789126	0,0103485

Para analizar la significancia de los parámetros del modelo, se plantea el siguiente juegos de hipótesis:
H0:$\beta_j = 0$ **vs** **H1:**$\beta_j \neq 0$ para j= 0,1,2,3,4,5.

Haciendo uso de los valores p proporcionados por la tabla de parámetros estimados a una significancia del 0.05, se concluye que los parámetros individuales β_0, β_1, β_2 y β_3 (para los cuales el valor p>0.05 y permiten rechazar Ho) son significativos cada uno en presencia de los demás parámetros, por otro lado se encuentra que β_4 y β_5 (para los cuales el valor p<0.05 y no permiten rechazar Ho) son individualmente no significativas en presencia de los demás parámetros.

0pt

Al revés.

<u>Interpretación de los parámetros estimados:</u> Los parámetros susceptibles de interpretación son aquellos que son significativos individualmente (β0 , β1 , β2 y β3) En el caso de **^β0**=-1.512 como $x_j = 0 \notin [x_{j,min}, x_{j,max}]$ $\forall j$ (evidenciado en la tabla de maximos y minimos para cada variable) entonces este valor no es interpretable

TABLA DE MINIMOS Y MAXIMOS						
	Y	X1	X2	X3	X4	X5
MIN	1,3	7,14	44,2	1,6	42,6	29
MAX	7,6	17,94	65,9	60,5	133,5	835

3pt

Para **^β1**=0.0313, indica que por cada dia que aumenta la duración promedio de la estadía de todos los pacientes en el hospital (x_1), el porcentaje de la probabilidad promedio estimada de adquirir infección en el hospital se incrementa en 0.0313 unidades.

Debian interpretar significativos, se los vuelgo porque son congruentes con el error

Para **^β2**=0.0455, indica que por cada punto de incremento en la razón del número de cultivos realizados en pacientes sin síntomas de infección hospitalaria (x_2), el porcentaje de la probabilidad promedio estimada de adquirir infección en el hospital se incrementa en 0.0455 unidades.

Para **^β3**=0.0315, indica que por cada unidad que se incrementa el número promedio de camas (x_3), el porcentaje de la probabilidad promedio estimada de adquirir infección en el hospital se incrementa en 0.0315 unidades.

Al interpretar x_1, x_2 y x_3 sin tener en cuenta sus coeficientes estimados sino únicamente la lógica y el sentido común, se puede pensar que al incrementar el número de cultivos realizados en pacientes sin síntomas de infección hospitalaria, el porcentaje de la probabilidad promedio estimada de adquirir infección en el hospital debería de disminuir ya que si se realizan cultivos con mayor frecuencia, se pueden detectar infecciones antes y prevenir su propagación, lo que reduciría el riesgo de infección. Pero se ve que según esta muestra no es así.

<u>Significancia de la regresión:</u> para verificar la significancia de la regresión se plantean el siguiente juego de hipótesis
H0:$\beta_1 = \beta_2 = \ldots = \beta_5 = 0$ **vs** **H1:** algún $\beta_j \neq 0$ para j= 1,2,3,4,5

5 pt

TABLA ANOVA					
	Suma de cuadrados	Grados de libertad	Cuadrado medio	Estadistico F	Valor P
Modelo	41,1881	5	8,23762	6,96077	7,26E-05
Error	~~520,711~~	44	1,18343		

De la tabla Anova se obtienen los valores del estadístico de prueba F0=6.96 y su valor p=7.26e-05, como vp<0.05=α se rechaza H0 y se concluye que el modelo de RLM planteado es significativo. Lo cual se traduce en que el porcentaje de la probabilidad promedio estimada de adquirir infecciones en el hospital depende significativamente de al menos una de las predictoras del modelo

3pt

<u>Coeficiente de determinación:</u> se sabe que $R^2 = \frac{SSR}{SST}$ de la tabla anova se obtiene que $R^2 = \frac{41.1881}{41.1881+52.0711} = \frac{41.1881}{93.2592} = 0.4417$ de lo que se puede concluir que el 44.17% de la variabilidad total en la probabilidad promedio de adquirir infecciones es explicado por el modelo propuesto.

Adicionalmente, se puede calcular el R^2 ajustado como una medidas de bondad de ajuste: $R^2_{adj} = 1 - \frac{(n-1)MSE}{SST} = 1 - \frac{(50-1)*1.18343}{41.1881+52.0711} = 1 - \frac{57.9881}{93.2592} = 0.3782$ ya que $R^2_{adj} = 0.3782$ es menor que $R^2 = 0.4417$ podemos afirmar que en el modelo existen variables que no dan un aporte significativo al modelo.

3pt

Modelo reducido?

2. Se busca probar la significancia simultánea de las 3 variables con valor p más alto las cuales son β_1, β_2 y β_3 aunque β_0 tiene mayor valor p que alguno de los coeficientes anteriores este no se incluye ya que no está directamente relacionado con la contribución de las variables independientes especificadas que se están examinando. Para probar esto se tiene el siguiente juego de hipótesis

2pt

H0: $\beta_1 = \beta_2 = \beta_3 = 0$ **vs** **H1:** algún $\beta_j \neq 0$ para j= 1,2,3

Para esta prueba de hipótesis el estadístico de prueba es $F_0 = \frac{MSextra}{MSE} = \frac{MSR(\beta_1,\beta_2,\beta_3|\beta_0,\beta_4,\beta_5)}{MSE} = \frac{SSR(\beta_1,\beta_2,\beta_3|\beta_0,\beta_4,\beta_5)}{3*MSE} = \frac{SSE(\beta_0,\beta_4,\beta_5)-SSE(MF)}{3*MSE}$ de la tabla de todas las posibles regresiones se saca el SEE de la regresión que incluye β_0, β_4 y β_5 y de la Anova se obtiene SSE y MSE, reemplazando se tiene que $F_0 = \frac{56.868-52.0711}{3*1.18343}$ entonces $F_0 = 1.3511$ para el criterio de decisión se requiere obtener el valor crítico de una distribución $f_{\boxed{3,50-3}} = \cancel{f_{3,47}}$ a un nivel de significancia α=0.05, esto es $f_{0.05,\cancel{3,47}} = \cancel{2.8024}$, como $F_0 = 1.3511 < f_{0.05,3,47} = 2.8024$, entonces con α=0.05 no se rechaza H0 y no se puede concluir que al menos una variable del subconjunto sea significativa, sin embargo, estas variables no se pueden eliminar del modelo sin antes considerar otras pruebas de significancia individual para evaluar la relevancia de cada variable en el modelo, si se determina que alguna de las variables es importante para explicar la variable respuesta, entonces, debe mantenerse en el modelo, incluso si el subconjunto completo no es significativo.

$F_{3,44}$

1pt

Esta prueba se hace precisamente porque ninguna es significativa y se quiere saber si vale la pena sacarlas

TABLA DE TODAS LAS POSIBLES REGRESIONES						
	k	R^2	R^2 ajustado	SSE	Est. de Mallows	Variables
1	1	0,219	0,203	72,808	15,522	X1
2	1	0,195	0,179	75,033	17,402	X3
3	1	0,168	0,151	77,593	19,566	X4
4	1	0,166	0,149	77,755	19,702	X5
5	1	0,001	-0,02	93,206	32,759	X2
6	2	0,39	0,364	56,868	4,054	X4 X5
7	2	0,311	0,282	64,234	10,278	X1 X3
8	2	0,298	0,268	65,493	11,342	X1 X4
9	2	0,289	0,259	66,328	12,047	X3 X5
10	2	0,262	0,231	68,789	14,127	X3 X4
11	2	0,257	0,226	69,252	14,517	X1 X5
12	2	0,224	0,191	72,356	17,141	X2 X3
13	2	0,223	0,189	72,504	17,265	X1 X2
14	2	0,179	0,144	76,602	20,728	X2 X5
15	2	0,168	0,133	77,589	21,563	X2 X4
16	3	0,415	0,377	54,581	4,12	X3 X4 X5
17	3	0,401	0,362	55,845	5,189	X2 X4 X5
18	3	0,4	0,361	55,911	5,245	X1 X4 X5
19	3	0,349	0,306	60,737	9,322	X1 X3 X4
20	3	0,338	0,295	61,692	10,129	X1 X3 X5
21	3	0,335	0,291	62,039	10,423	X2 X3 X5
22	3	0,316	0,271	63,799	11,91	X1 X2 X3
23	3	0,301	0,255	65,218	13,109	X1 X2 X4
24	3	0,278	0,231	67,369	14,926	X2 X3 X4
25	3	0,258	0,209	69,238	16,506	X1 X2 X5
26	4	0,441	0,391	52,155	4,071	X2 X3 X4 X5
27	4	0,423	0,372	53,787	5,45	X1 X3 X4 X5
28	4	0,406	0,353	55,415	6,825	X1 X2 X4 X5
29	4	0,356	0,299	60,066	10,756	X1 X2 X3 X5
30	4	0,351	0,293	60,564	11,177	X1 X2 X3 X4
31	5	0,442	0,378	52,071	6	X1 X2 X3 X4 X5

Presenten sólo lo que se necesita, no saturen el reporte

3pt

3. ¿Existen relaciones lineales entre las variables: duración de la estadía, número de camas, censo promedio diario y riesgo de infección en un hospital, tales que si se suman los coeficientes de regresión de duración de la estadía y números de camas, y si se suman los coeficientes de regresión de censo promedio diario y número de enfermeras el resultado es cero?

No se está mirando correlaciones,

Para responder la pregunta se tienen que verificar simultáneamente las 2 ecuaciones, planteando entonces el siguiente juego de hipótesis

H0: $\beta_1 + \beta_3 = 0$ y $\beta_4 + \beta_5 = 0$ **vs** **H1:** $\beta_1 + \beta_3 \neq 0$ $\boxed{\text{y}}$ $\beta_4 + \beta_5 \neq 0$ ✗

ó

Lo cual se puede reescribir como

se está probando si los efectos llevan relación.

Nunca hagan esto - Escriban las matrices

$$H_0: \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{vs} \quad H_1: \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

1 pt

donde L= $\begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$

se observa que L tiene 2 filas linealmente independientes, por lo tanto r=2, el modelo el reducido puede escribirse como: $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 - \beta_1 X_3 + \beta_4 X_4 - \beta_4 \beta_5 + \boxed{\varepsilon}$ 0,5 pt

$Y = \beta_0 + \beta_1(X_1 - X_3) + \beta_2 X_2 + \beta_4(X_4 - X_5) + \varepsilon$ sin supuestos?

RM: $Y = \beta_0 + \beta_1 X_{1,3} + \beta_2 X_2 + \beta_4 X_{4,5} + \varepsilon$, donde $X_{1,3} = (X_1 - X_3)$ y $X_{4,5} = (X_4 - X_5)$

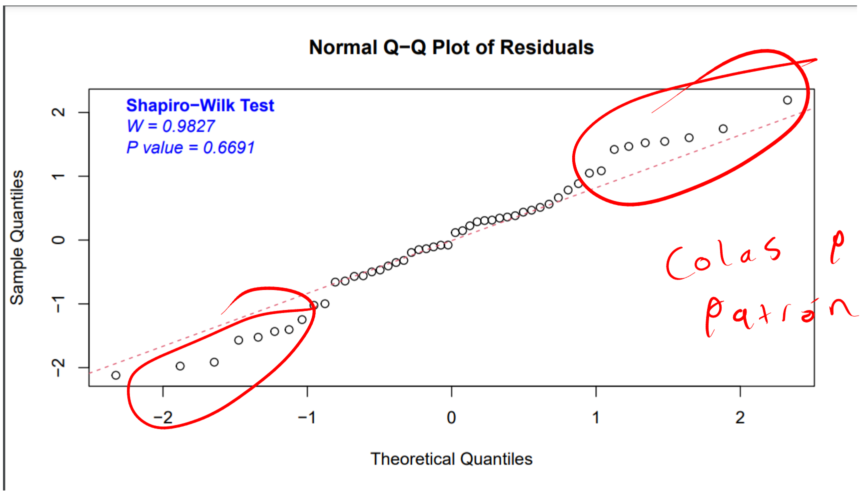
¿cómo distribuye?

La expresión del estadístico de prueba está dado por

$F_0 = \frac{MSH}{MSE} = \frac{SSH}{2*MSE} = \frac{SSE(RM)-SSE(FM)}{2*MSE} = \frac{SSE(RM)-52.0711}{2*1.18343} = \frac{SSE(RM)-52.0711}{2.3669} = F_0$ Con la información suministrada no podemos obtener el valor de SSE(RM) ya que de la tabla de todas las posibles regresiones no se puede sacar este valor, dado que esta no admite restas entre variables, por lo tanto, no se puede rechazar o aceptar alguna hipótesis. 1,5 pt

15 pt

4. <u>Validación de los supuestos sobre los errores:</u> se quiere probar 3,5 pt
H0: $\varepsilon_i \sim normal$ **vs** **H1:** $\varepsilon_i \nsim normal$



Colas pesadas y patrón irregular

Y ya! Análisis muy poco profundo.

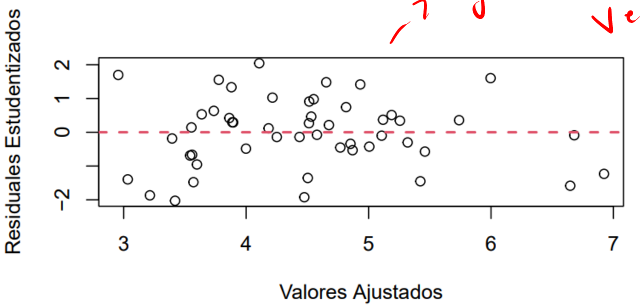
Si bien según la prueba de shapiro-wilk el valor permite aceptar la hipótesis nula, por lo que según este criterio los errores distribuyen normal, ahora bien, si vemos la gráfica, vemos que los puntos no se ajustan muy bien a la recta, entonces según la gráfica los errores no se distribuyen normal. Esa contradicción se puede deber a la presencia de observaciones influenciales. Entonces se acepta el criterio de la gráfica y se concluye que el supuesto de normalidad no se cumple

y los de balanceo?

<u>Supuesto de varianza constante:</u> se quiere probar **H0:** $V[\varepsilon_i] = \sigma^2$ **vs** **H1:** $V[\varepsilon_i] \neq \sigma^2$

1,5 pt

→ gráfica mal pegada, es difícil ver algo ahí.



podría ser un patrón de aumento constante y aún así decir que var cte!

Del análisis de la gráfica se tiene que el patrón de los puntos indica un comportamiento constante. De lo cual podemos concluir que el supuesto de varianza constante se cumple. ✓

Análisis de la presencia de observaciones extremas: para identificar si en el modelo hay observaciones extremas se busca calcular los estadísticos que nos permitan aplicar criterios a estos puntos, los cuales incluyen: residuales estudentizados,los h_{ii}, las distancias de cook(d_i) y los DFFITS.

Una observación i es atípica cuando $|r_i| > 3$ de acuerdo a la columna res.stud de residuales estudentizados se tiene que ningun valor es menor a -3 o mayor que 3, por lo tanto se concluye que no tiene observaciones atípicas ✓ 3pt

Puntos de balanceo: se asume que la observación i es un punto de balanceo si $h_{ii} > \frac{2P}{n}$ en la práctica tenemos que $h_{ii} > \frac{2P}{n} = 2 * \frac{6}{50} = 0.24$, de acuerdo a las columnas $h_{ii}.value$ de valores de la diagonal de la matriz H se tiene que las observaciones 17, 22, 40 y 45 son puntos de balanceo 2pt

¿Qué causan?

Observaciones influenciales: se dice que la observación y será influencial si $D_i > 1$ según este criterio no hay ninguna observación influencial. Adicionalmente si $|DFFITS| > 2\sqrt{\frac{P}{n}} = 2\sqrt{\frac{6}{50}} = 0.693$. De acuerdo a la columna DFFITS tenemos que las observaciones 12, 17, 39, 45 y 47 son influenciales. 3pt

¿Qué causan

En resumen, para el análisis de observaciones extremas se tiene que:

- no hay valores atípicos
- las observaciones 17, 22, 40 y 45 son puntos de balanceo
- las observaciones 12, 17, 39, 45 y 47 son influenciales

No es válidos si no se cumplen supuestos

Como el supuesto de normalidad no se cumple y se verificó la existencia de puntos de balanceo e influenciales significa que la validez del modelo puede estar comprometida y se deben tomar medidas para corregir estas limitaciones. 2pt
En general es recomendable realizar una validación exhaustiva de los supuestos y examinar cuidadosamente los datos antes de tomar decisiones basados en los resultados del modelo. ✓

Tabla para el nálisis de la presencia de observaciones extremas													
i	**Y**	**X1**	**X2**	**X3**	**X4**	**X5**	**y gorro**	**se,y gorro**	**residuales**	**res,stud**	**Cooks,D**	**valor hii**	**Dffits**
1	4,3	9,89	45,2	11,8	108,7	190	4,4359	0,4859	-0,1359	-0,1397	0,0008	0,1995	-0,0689
2	5,6	8,95	53,7	18,9	122,8	147	5,256	0,4118	0,344	0,3416	0,0033	0,1433	0,1383
3	5	10,33	55,8	21,2	104,3	266	5,3191	0,2537	-0,3191	-0,3016	0,0009	0,0544	-0,0716
4	5,3	8,15	54,9	12,3	79,8	99	3,8798	0,2215	1,4202	1,3334	0,0128	0,0415	0,2798
5	2,6	9,76	53,2	6,9	80,1	64	3,598	0,2994	-0,998	-0,9543	0,0124	0,0758	-0,2729
6	4,3	9,23	51,6	11,6	42,6	620	4,1832	0,4136	0,1168	0,1161	0,0004	0,1445	0,0472
7	4,2	7,39	51	14,6	88,4	72	3,8943	0,2876	0,3057	0,2914	0,0011	0,0699	0,079
8	6,4	11,62	53,9	25,5	99,2	133	4,9321	0,3281	1,4679	1,4153	0,0334	0,0909	0,4529
9	4,8	9,36	54,1	18,3	90,6	165	4,5131	0,1896	0,2869	0,2679	0,0004	0,0304	0,0469
10	4,7	8,77	54,5	5,2	47	143	2,9554	0,3551	1,7446	1,6966	0,0572	0,1065	0,5991
11	5,7	11,2	56,5	34,5	88,9	180	5,1881	0,4204	0,5119	0,5102	0,0076	0,1493	0,2119
12	7,6	11,4	61,1	16,6	97,9	535	5,9965	0,4264	1,6035	1,6022	0,0777	0,1537	0,6955
13	5,4	7,9	64,1	7,5	98,1	68	4,5151	0,4857	0,8849	0,9091	0,0343	0,1994	0,4527
14	4,2	9	56,3	14,6	76,4	72	3,8871	0,2384	0,3129	0,2948	0,0007	0,048	0,0655
15	6,1	13,6	54	24,2	111,7	312	5,7391	0,4028	0,3609	0,3572	0,0034	0,1371	0,1409
16	3,7	7,1	59	2,6	75,8	70	3,5531	0,3544	0,1469	0,1428	0,0004	0,1061	0,0487
17	5,9	17,9	56,2	26,4	91,8	835	6,9231	0,7031	-1,0231	-1,2325	0,1816	0,4177	-1,0502
18	4,3	7,6	47,1	16,4	65,7	318	3,8623	0,319	0,4377	0,4209	0,0028	0,086	0,1279
19	5,5	10,9	57,2	10,6	71,9	593	5,1181	0,3424	0,3819	0,3698	0,0025	0,0991	0,1214
20	1,3	8,2	60,9	1,9	58	73	3,2139	0,3667	-1,9139	-1,8687	0,0746	0,1136	-0,6893
21	6,2	10,2	51,9	16,4	59,2	568	4,6541	0,3046	1,5459	1,4803	0,0311	0,0784	0,4378
22	6,6	13,95	65,9	15,6	133,5	356	6,6799	0,6186	-0,0799	-0,0893	0,0006	0,3233	-0,061
23	5,3	9,77	50,2	15,7	89,7	154	4,2149	0,2502	1,0851	1,0249	0,0098	0,0529	0,242
24	4,5	9,31	47,2	30,2	101,3	170	4,8524	0,3531	-0,3524	-0,3424	0,0023	0,1053	-0,1163
25	2	8,93	56	6,2	72,5	95	3,5695	0,2387	-1,5695	-1,4788	0,0184	0,0481	-0,3373
26	4,4	7,7	56,9	12,2	67,9	129	3,736	0,2885	0,664	0,6331	0,0051	0,0703	0,1729
27	1,3	8,92	53,9	2,2	79,5	56	3,4194	0,3036	-2,1194	-2,0288	0,058	0,0779	-0,6123
28	6,3	9,74	54,4	11,4	76,1	221	4,1068	0,1767	2,1932	2,0432	0,0189	0,0264	0,3496
29	1,6	8,82	58,2	3,8	51,7	80	3,0329	0,3599	-1,4329	-1,3958	0,0399	0,1094	-0,4948
30	4,2	9,06	52,8	6,9	75,9	134	3,6371	0,2396	0,5629	0,5304	0,0024	0,0485	0,1188
31	3,2	8,19	52,1	10,8	59,2	176	3,3952	0,263	-0,1952	-0,185	0,0004	0,0585	-0,0456
32	5,6	11,48	57,6	20,3	82	252	4,8163	0,2691	0,7837	0,7436	0,006	0,0612	0,1889
33	2,9	10,8	63,9	1,6	57,4	130	3,5585	0,4568	-0,6585	-0,667	0,0159	0,1763	-0,3066
34	4,6	10,16	54,2	8,4	51,5	831	5,0059	0,5244	-0,4059	-0,4259	0,0092	0,2324	-0,2321
35	4,9	9,89	50,5	17,7	103,6	167	4,676	0,2782	0,224	0,213	0,0005	0,0654	0,0557
36	5,6	10,12	51,7	14,9	79,1	362	4,5513	0,1839	1,0487	0,978	0,0047	0,0286	0,1676
37	3,9	10,73	50,6	19,3	101	445	5,4229	0,2883	-1,5229	-1,4518	0,0265	0,0702	-0,4042
38	4,9	10,23	53,2	9,9	77,9	752	5,4601	0,4767	-0,5601	-0,5728	0,013	0,192	-0,2771
39	5,3	11,77	54,1	17,3	56	196	3,7763	0,4696	1,5237	1,5528	0,0921	0,1864	0,7557
40	2,9	10,79	44,2	2,6	56,6	461	3,5422	0,5605	-0,6422	-0,6888	0,0286	0,2655	-0,4116
41	4,5	8,28	48,1	26	101,8	108	4,5782	0,3489	-0,0782	-0,0759	0,0001	0,1029	-0,0254
42	4,3	10,39	54,6	14	88,3	353	4,8691	0,1918	-0,5691	-0,5315	0,0015	0,0311	-0,0944
43	4,1	9,05	51,2	20,5	79,8	195	4,2504	0,2111	-0,1504	-0,1409	0,0001	0,0377	-0,0276
44	5	7,78	45,5	20,9	71,6	489	4,5317	0,4017	0,4683	0,4632	0,0056	0,1364	0,1824
45	5,4	11,18	45,7	60,5	85,8	640	6,647	0,7507	-1,247	-1,5838	0,38	0,4762	-1,5373
46	4,3	9,42	50,6	24,8	62,8	508	4,769	0,3314	-0,469	-0,4527	0,0035	0,0928	-0,1435
47	2,5	8,54	56,1	27	82,5	98	4,4751	0,3613	-1,9751	-1,9248	0,0766	0,1103	-0,7002
48	5	11,03	49,9	19,7	102,1	318	5,1068	0,2814	-0,1068	-0,1016	0,0001	0,0669	-0,0269
49	3,1	9,41	59,5	20,6	91,7	29	4,5032	0,3298	-1,4032	-1,3536	0,0309	0,0919	-0,4348
50	3,5	7,94	49,5	6,2	92,3	195	3,9986	0,3611	-0,4986	-0,4859	0,0049	0,1102	-0,1695

¿Para qué todo eso?
Saturan el reporte.
Por ejemplo, no ve en el reporte dónde usan el se.yhat
y no todos los datos son de interés al lector del reporte estadístico.