Ejercicio demostrar que SST = SSR + SSD

Vamos a demostrar formalmente la igualdad SST = SSR + SSD que hemos estado utilizando en la teoría de los vídeos anteriores.

Utilizando la notación del ejercicio anterior, es decir:

* y\_i = el valor actual i-ésimo
* y\_m = media de los valores actuales
* z\_i = valor predictor i-ésimo (el que nos da la recta de regresión).

Podemos escribir

y\_i - y\_m = (y\_i - z\_i) + (z\_i - y\_m)

Y si ahora elevamos al cuadrado, y aplicamos la identidad notable obtenemos que :

(y\_i - y\_m) ^2 = (y\_i - z\_i) ^2 + (z\_i-y\_m) ^2 + 2(y\_i-z\_i)(z\_i-y\_m)

De aquí, aplicando sumatorio a cada uno de los factores anteriores se deduce que para demostrar que SST = SSR + SSD, tenemos que probar que:

sum(y\_i-z\_i)(z\_i-y\_m) = 0

Notemos que podemos obviar el 2 porque.

Para ello expresamos  (y\_i-z\_i)(z\_i-y\_m) de la siguiente forma:

(y\_i-z\_i)(z\_i-y\_m) =(y\_i-z\_i)z\_i - (y\_i-z\_i)y\_m

Ahora bien, en la demostración que hicimos en el vídeo de la clase 74, la primera ecuación normal (en el minuto 5:26) vemos que:

sum(y\_i-z\_i)=0

Por lo tanto será suficiente ver que

sum(y\_i-z\_i)z\_i = 0

Para ello, basta escribir que es z\_i, no es otra cosa que:

z\_i = a + bx\_i

y si entonces sustituimos en la suma anterior y operamos, tenemos que:

sum(y\_i-z\_i)z\_i = sum(y\_i-z\_i)a + bsum(y\_i-z\_i)x\_i

De nuevo el primer sumatorio es cero por la primera ecuación. El segundo sumatorio también es cero por la segunda ecuación normal que aparece en el minuto 5:26 de vídeo.