$$AIC = \underbrace{2(k+1)}_{\text{penalización}} - 2log(L)$$

↳ Verosimilitud

↓ AIC : mejor

M1: $Y \sim X_1 + X_2 + X_3$ AIC = 999

M2: $Y \sim X_1 + X_2$ AIC = 1000

} dif < 2 ⇒ Modelo 2 ✓

$$SBC = BIC = \underbrace{(k+1) \times log(n)}_{\text{penalización}} - 2log(L)$$

↳ Verosimilitud

$$SBIC = n \times \log\left(\frac{SCE}{n}\right) + 2(p+3)q - 2q^2$$

Modelos	C_p	N° var.	$\lvert C_p - N°var \rvert$
$y \sim x_1 + x_2$	5	2	3
$y \sim x_1 + x_3$	6	2	4
$y \sim x_2 + x_3 + x_5$	4	3	1
$y \sim x_1 + x_2 + x_3 + x_6$	2	4	2

Mejores subconjuntos

Y, X_1, X_2, X_3

- Subconjuntos de 1 variable
  - $Y \sim X_1$
  - $Y \sim X_2$
  - $Y \sim X_3$

  mejor

- Subconjuntos de 2 variables
  - $Y \sim X_1 + X_2$
  - $Y \sim X_1 + X_3$
  - $Y \sim X_2 + X_3$

  mejor

- Subconjuntos de 3 variables — $Y \sim X_1 + X_2 + X_3$ ✓ mejor

mejor

Subsets Regression Summary

Model	R-Square	Adj. R-Square	Pred R-Square	C(p)	AIC	SBIC	SBC
~~1~~	0.4725	0.4709	0.4659	18.1803	1724.8932	816.5992	1736.1981
2	0.4998	0.4967	0.4906	2.8855	1709.8884	801.8265	1724.9617
3	0.5034	0.4987	0.4909	2.6108	1709.5830	801.5988	1728.4246
4	0.5056	0.4994	0.4906	3.2048	1710.1497	802.2451	1732.7596
5	0.5068	0.4989	0.4888	4.4810	1711.4093	803.5763	1737.7876
~~6~~	0.5075	0.4965	0.4832	8.0000	1714.9164	805.1534	1748.8313

AIC: Akaike Information Criteria

menor dif.

2 - 2.8855

3 - 2.6108

4 - 3.2048

5 - 4.4810

```
> modelo |> ols_step_backward_p()
```

modelo completo

```
                                 Stepwise Summary
----------------------------------------------------------------------------
Step    Variable        AIC          SBC         SBIC        R2        Adj. R2
----------------------------------------------------------------------------
 0      Full Model    1714.916     1748.831    805.153    0.50755    0.49650
 1      region        1711.409     1737.788    803.576    0.50679    0.49894
 2      x3            1710.150     1732.760    802.245    0.50565    0.49937
----------------------------------------------------------------------------
```

→ Quitó region y X_3

1° variable que se retira

```
Final Model Output
------------------

                         Model Summary
-----------------------------------------------------------------
R                         0.711       RMSE                   3.436
R-Squared                 0.506       MSE                   11.809
Adj. R-Squared            0.499       Coef. Var          10415.875
Pred R-Squared            0.491       AIC                 1710.150
MAE                       2.746       SBC                 1732.760
-----------------------------------------------------------------

                                Parameter Estimates
----------------------------------------------------------------------------
      model      Beta    Std. Error    Std. Beta       t       Sig      lower    upper
----------------------------------------------------------------------------
(Intercept)    -0.340         0.283                 -1.202    0.230    -0.897    0.217
      chas1     0.628         0.389        0.064     1.611    0.108    -0.139    1.394
         x1     3.118         0.180        0.686    17.284    0.000     2.763    3.473
         x8     0.809         0.190        0.169     4.261    0.000     0.436    1.183
  x_noise_1    -0.227         0.190       -0.047    -1.189    0.235    -0.601    0.148
----------------------------------------------------------------------------
```

Cuando se realizó la hipótesis de los coeficientes de región, estos resultaron estadísticamente iguales a cero, se obtuvo el **pvalor más alto.**

Luego, se volvió a correr el modelo, y resultó que la variable x3 tenía un coeficiente estadísticamente igual a cero (H0: beta = 0 no se rechazó), pues el **pvalor fue alto.**

Luego, todas las variables que restaron, tuvieron pvalores bajos, por lo que ninguna otra tuvo que ser retirada.

```
> modelo |> ols_step_backward_aic()


                              Stepwise Summary
--------------------------------------------------------------------------
Step    Variable        AIC         SBC        SBIC        R2       Adj. R2
--------------------------------------------------------------------------
 0      Full Model    1714.916    1748.831    805.153    0.50755    0.49650
 1      region        1711.409    1737.788    803.517    0.50679    0.49894
 2      x3            1710.150    1732.760    802.187    0.50565    0.49937
 3      x_noise_1     1709.583    1728.425    801.563    0.50343    0.49871
--------------------------------------------------------------------------
```

A cada paso, se retira una variable, porque al hacerlo disminuye el AIC.
Luego de retirar x_noise_1

```
                                 Parameter Estimates
------------------------------------------------------------------------------------
      model        Beta    Std. Error    Std. Beta      t         Sig      lower    upper
------------------------------------------------------------------------------------
(Intercept)      -0.305         0.282                -1.084     0.279    -0.860    0.249
      chas1       0.587         0.388        0.060    1.512     0.132    -0.177    1.351
         x1       3.130         0.180        0.688   17.364     0.000     2.775    3.484
         x8       0.801         0.190        0.167    4.220     0.000     0.428    1.175
------------------------------------------------------------------------------------
```

```
> modelo |> ols_step_forward_p()
```

Modelo Nulo : $Y \sim \beta_0$

```
                              Stepwise Summary
------------------------------------------------------------------------------
Step    Variable        AIC          SBC         SBIC         R2       Adj. R2
------------------------------------------------------------------------------
 0     Base Model    1927.592     1935.128    1017.991     0.00000     0.00000
 1     x1            1724.893     1736.198     816.599     0.47254     0.47088
 2     x8            1709.888     1724.962     801.827     0.49984     0.49668
 3     chas          1709.583     1728.425     801.599     0.50343     0.49871
 4     x_noise_1     1710.150     1732.760     802.245     0.50565     0.49937
------------------------------------------------------------------------------
```

x1 es la variable que arroja el pvalor más bajo.

Luego, se tiene Y ~ X1 y comienzan a competir las restantes. La siguiente variable con el pvalor más bajo fue X8, de modo que el modelo queda como Y ~ X1 + X8.

Luego, ...

```
                                Parameter Estimates
---------------------------------------------------------------------------------------
      model       Beta    Std. Error    Std. Beta       t        Sig      lower    upper
---------------------------------------------------------------------------------------
(Intercept)     -0.340         0.283                 -1.202    0.230    -0.897    0.217
         x1      3.118         0.180        0.686    17.284    0.000     2.763    3.473
         x8      0.809         0.190        0.169     4.261    0.000     0.436    1.183
      chas1      0.628         0.389        0.064     1.611    0.108    -0.139    1.394
  x_noise_1     -0.227         0.190       -0.047    -1.189    0.235    -0.601    0.148
---------------------------------------------------------------------------------------
```