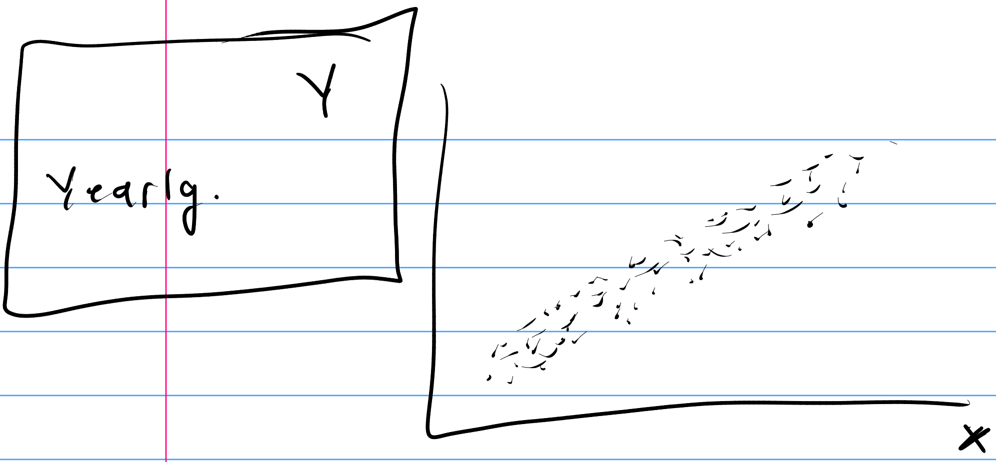
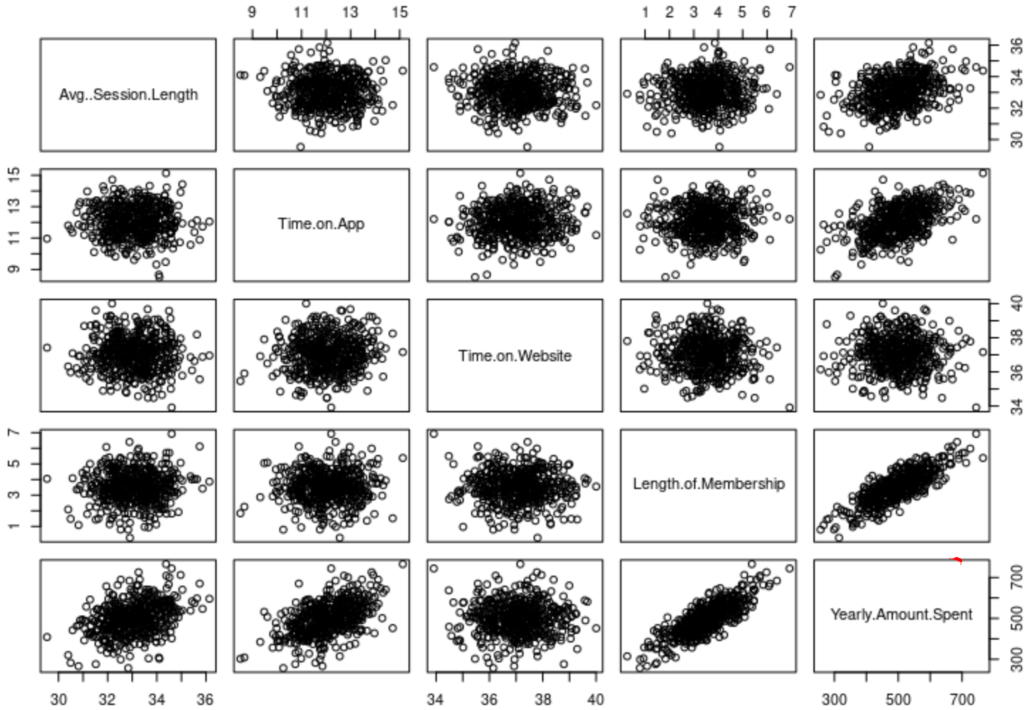


Relación lineal → Correlación > 0

< 0



Y : Yearly

Se plantea un modelo de RLS, donde
Y: Yearly.amount.spent y X : Length of membership

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i ; \quad \varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$$

$$i \le i \le n ; \quad n = 400$$

Ajustamos

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$$

$[1, 2, 3, 4, \ldots, 500]$

Sample(1:500, 400)

$[1, 3, 4, \ldots, 53, 98]$

$$\boxed{\text{Yearly} = \beta_0 + \beta_1 \text{Length.amount}}$$

↓

R

$\boxed{\text{Yearly.Amount.Spent} \sim \text{Length...}}$

Yearly... ~ $\boxed{1}$ + Length

↳ β_0

$x=0$ Yearly... ~ 0 + Length

[ininterpretable?

β_0 ↲ $\hat{\beta}_0$ → $se(\hat{\beta}_0)$ → $T = \frac{\hat{\beta}_0}{se(\hat{\beta}_0)}$

Coefficients:

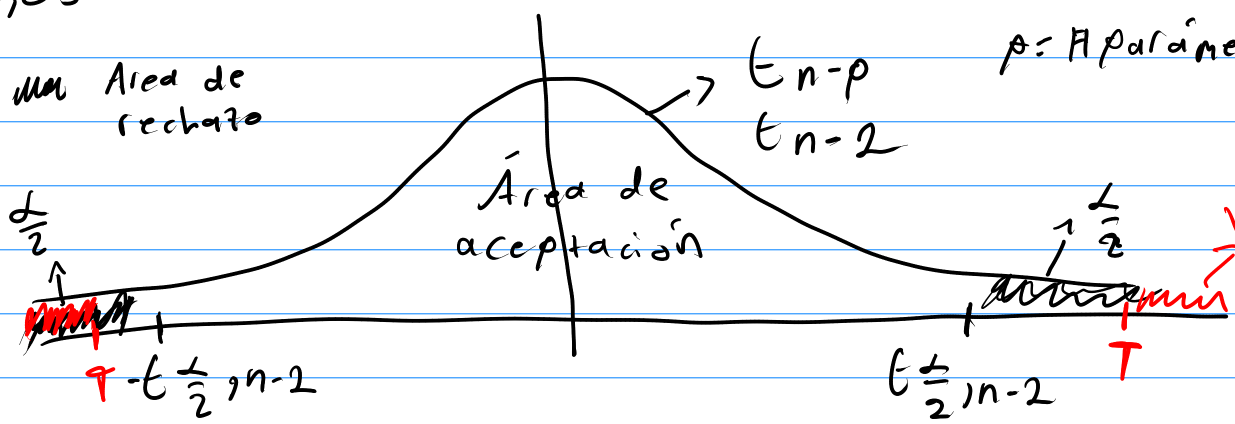
	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	271.417	8.461	32.08	<2e-16	***
Length.of.Membership	64.491	2.298	28.06	<2e-16	***

β_1 ↳ interpretable

$\hat{\beta}_1$ $se(\hat{\beta}_1)$ $T = \frac{\hat{\beta}_1}{se(\hat{\beta}_1)}$

$\alpha = 0{,}05$

p: # parámetros

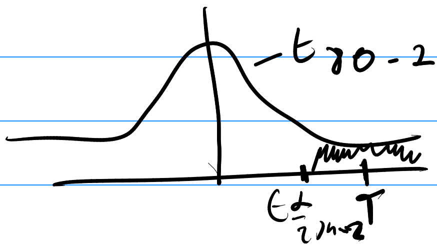


Val-P $< \alpha \Rightarrow |T| > |t_{\frac{\alpha}{2}, n-2}| \Rightarrow \beta_1$ signific.

Val-p > α => $|T| < |t_{\frac{\alpha}{2}, n-p}|$ -> β_1 no es signif.

$$\begin{cases} H_0: \beta_1 = 0 \\ H_a: \beta_1 \neq 0 \end{cases} \rightarrow$$

$$T = \frac{\hat{\beta}_1}{se(\hat{\beta}_1)} \sim t_{n-2}$$



$\boxed{n = 80}$

$$T = \frac{\hat{\beta}_1}{se(\hat{\beta}_1)} = 28{,}06$$

① $t_{0,01,\,70} = 1{,}32$	$t_{0,03,\,70} = 1{,}99$ ②
③ $t_{0,05,\,98} = 1{,}20$	$t_{0,025,\,80} = 1{,}24$ ④
⑤ $\boxed{t_{0,025,\,78} = 1{,}27}$	$t_{0,05,\,78} = 1{,}96$ ⑥

Con una significancia $\alpha = 0{,}05$, diga cuál es correcta

$\rightarrow t_{\frac{0{,}05}{2}, 78} = t_{0{,}025, 78} = 1{,}27$ T vs $t_{\frac{\alpha}{2}, n-2}$

$28{,}06 > 1{,}27$

c) β_1 es significativo

d) el cuantil $t_{\frac{\alpha}{2}, n-2} = 1{,}96$

¿β_1? Cambio en Y por aumento X

k: # covariables

Reg $\qquad$ $\underbrace{\text{anova}}$

analysis of variance

	SS	g.l	MS	F	Val-P
Reg	SSR	1	$MSR = \frac{SSR}{1}$	$\frac{MSR}{MSE}$	
error	SSE	n-2	$\boxed{MSE} = \frac{SSE}{n-2} \to \hat{\sigma}^2$		
	SST	n-1			

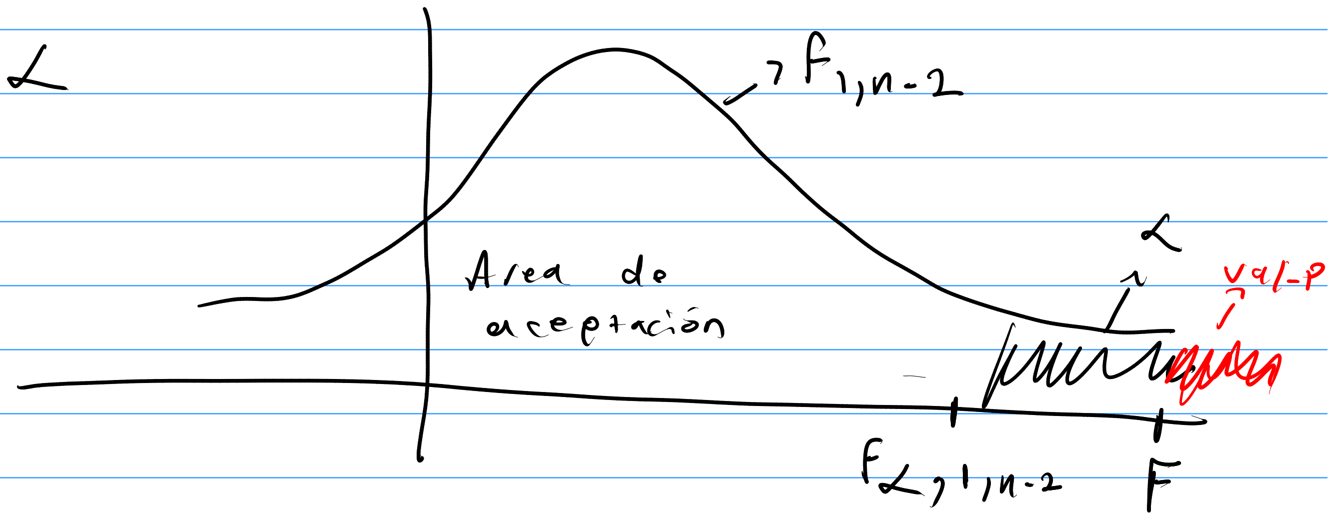
$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

$$SSE = SST - SSR$$

$$SST = SSR + SSE$$

$$\begin{cases} H_0: \beta_1 = 0 \\ H_a: \beta_1 \neq 0 \end{cases}$$

$$F = \frac{SSR/1}{SSE/n-2} = \frac{MSR}{MSE} \sim F_{1, n-2}$$



$$F = T^2$$

F = 92

```
Analysis of Variance Table

Response: Yearly.Amount.Spent
                     Df  Sum Sq Mean Sq F value    Pr(>F)    
Length.of.Membership   1 1651518 1651518  787.47 < 2.2e-16 ***
Residuals            398  834701    2097
```

gl SS MS F

$900 - 2 = 398$

$F_{\alpha, 1, 398} = 3,8649$

F vs $F_{\alpha, 1, 398}$

$787,47 > 3,8649$

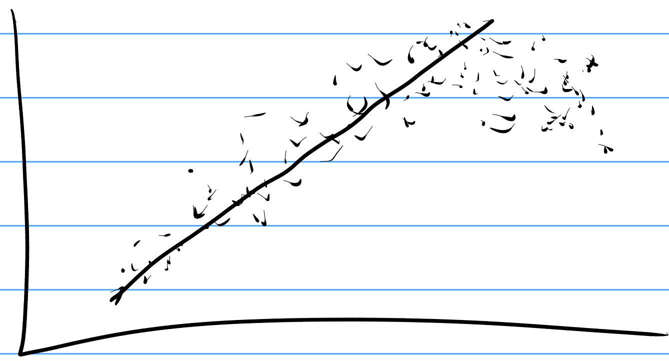
Reg es significativa

R^2 : variabilidad de la respuesta explicada por la regresión → proporción

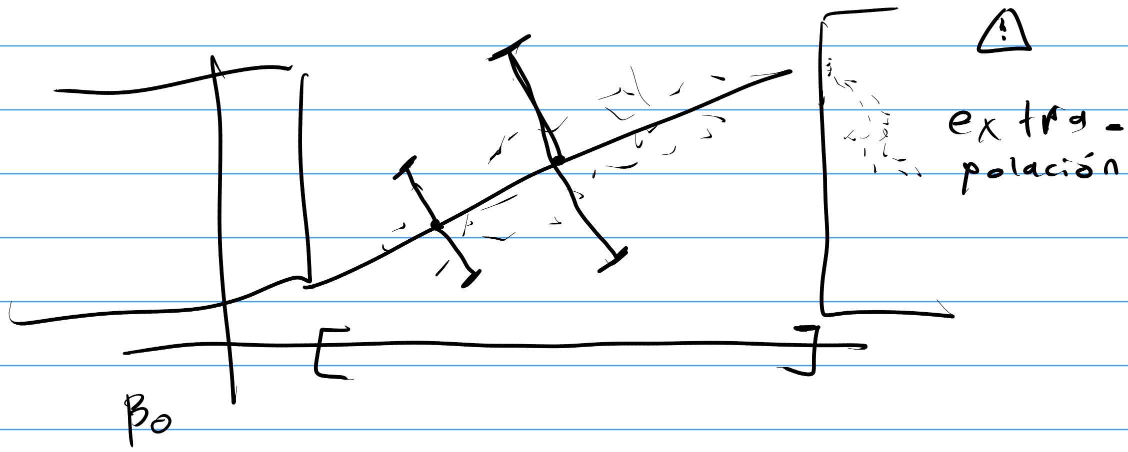
$$R^2 = \frac{SSR}{SST} = \frac{SSR}{SSR + SSE} = 0,6693$$

66,93% de la va

R^2 alto



Inferencias media	inferencias valor futuro
- interpolación	- interpolación
- Usados en el ajuste	- No usados en el ajuste



$$\hat{y}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_0$$

Media

$$\hat{y}_0 \pm t_{\frac{\alpha}{2}, n-2} \cdot se(\hat{y}_0)$$

Futuro

$>>$

$$\hat{y}_0 \pm t_{\frac{\alpha}{2}, n-2} \cdot se(\hat{y}_0 - y_0)$$