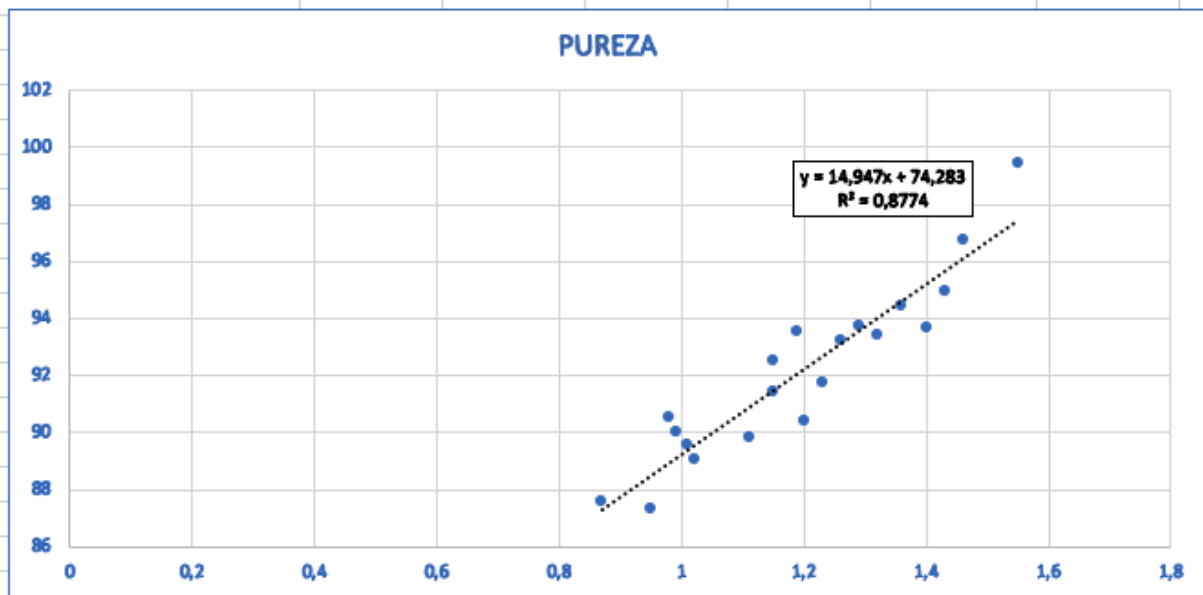


## Lista de Exercício - Regressão Linear

### **Bootstrap Paramétrico (Excel): Nível de Hidrocarboneto x Pureza**

Plotando um gráfico de relação dispersão entre as variáveis Nível de Hidrocarboneto e pureza com uma linha de tendência observamos na Figura abaixo um relação linear entre as duas variáveis.

A	B	C	D	E	F	G	
NIVEL_HIDROCARBONETO	PUREZA	B0^	B1^	PUREZA^	RESÍDUO	R2	dp^
0,99	90,01	74,2833142	14,9474797	89,0813192	0,928680825	0,87743571	1,1
1,02	89,05			89,5297436	-0,479743567		
1,15	91,43			91,4729159	-0,042915932		
1,29	93,74			93,5655631	0,174436905		
1,46	96,73			96,1066346	0,623365351		
1,36	94,45			94,6118867	-0,161886676		
0,87	87,59			87,2876216	0,302378393		
1,23	91,77			92,6687143	-0,898714311		
1,55	99,42			97,4519078	1,968092175		
1,4	93,65			95,2097859	-1,559785865		
1,19	93,54			92,0708151	1,469184878		
1,15	92,52			91,4729159	1,047084068		
0,98	90,56			88,9318444	1,628155622		
1,01	89,54			89,3802688	0,15973123		
1,11	89,85			90,8750167	-1,025016743		
1,2	90,39			92,2202899	-1,830289919		
1,26	93,25			93,1171387	0,132861297		
1,32	93,41			94,0139875	-0,603987487		
1,43	94,98			95,6582103	-0,678210257		
0,95	87,33			88,48342	-1,153419986		



Logo:

$$H_0 : \beta_1 = 0$$

$$H_0 : \beta_1 \neq 0$$

E considerando que os nossos dados seguem um modelo de regressão linear, temos:

$$PUREZA = \beta_0 + \beta_1 * NIVEL\_HIDROCARBONETO + ERRO$$

Onde erro é:

$$ERRO \sim N(0; \sigma^2)$$

Para estimar as variáveis  $\hat{\beta}_0$  e  $\hat{\beta}_1$  iremos utilizar o método dos mínimos quadrados de modo a minimizar a soma dos quadrados dos desvios verticais.

No **EXCEL** usaremos as seguintes fórmulas:

$$\hat{\beta}_0 = INTERCEPT(B2 : B21; A2 : A21) = 74,28331424 \quad \hat{\beta}_1 = SLOPE(B2 : B21; A2 : A21) = 14,94747973$$

Logo:

$$\widehat{PUREZA} = 74,28331424 + 14,94747973 * NIVEL\_HIDROCARBONETO$$

E

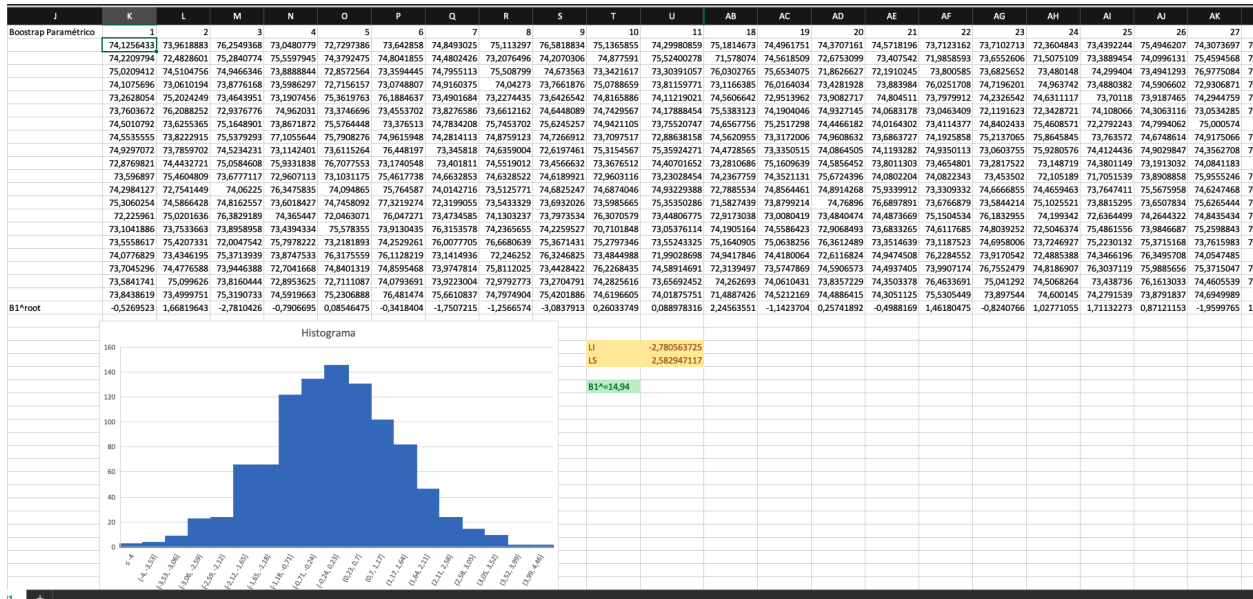
$$RESÍDUO = Y - \hat{Y}$$

E

$$\sigma^2 = (\sum r_i^2)/n - 2 = 1,180545382 \quad \sigma = \sqrt{\sigma^2} = 1,086529053$$

Onde  $r_i$  é o resíduo de cada amostra e  $n$  é a quantidade de amostras.





A fórmula usada para realizar as simulações foi:

$\$C\$2 + NORM.INV(RAND();0; \$I\$2)$ , onde  $\$C\$2$  é  $\hat{\beta}_0 = 74,28331424$  e  $\$I\$2$  é o desvio padrão do resíduo ( $\sigma = 1,086529053$ ).

Em seguida foi calculado o valor do  $\hat{\beta}_{1root}$  usando a fórmula  $SLOPE(K2 : K21; \$A\$2 : \$A\$21)$

E por fim, plotamos o histograma desses valores e calculamos os Limites inferiores e superiores, com as fórmulas:

- LI:  $PERCENTILE.EXC(K22 : ALV22; 0,025) = -2,780563725$
- LS:  $PERCENTILE.EXC(K22 : ALV22; 0,975) = 2,582947117$

Como  $\hat{\beta}_1 = 14,94747973 > LS(2,582947117)$  **rejeito**  $H_0 : \beta_1 = 0$

ao nível de confiança de 95%, ou seja, existe uma relação linear entre o **nível de hidrocarboneto** e a **pureza** do oxigênio.

Usualmente ao invés de usarmos  $\hat{\beta}_1$  para comparar com os percentis 2,5% e 97,5% (confiança de 95%) optamos por usar

$$T_{obs} = \hat{\beta}_1 / desvio(\hat{\beta}_1) = 14,94747973 / 1,31675827 = 11,35172649$$

Quando dividimos todos os  $\hat{\beta}_1 Boot$  pelo desvio (cada amostra pelo seu desvio) o histograma terá a forma de uma distribuição T com  $n - 2$  graus de liberdade.



Bootstrap Não Paramétrico	1	2	3	4	5	6	7	8	9	10	11	18	
	75,75249912	74,240398	73,258297	73,605104	73,605104	74,416176	75,330398	75,91147	74,443045	74,283314	74,58569263	74,90668	74,45
	74,24039831	73,605104	73,129894	72,453024	74,121428	74,240398	72,723528	72,453024	72,723528	73,679327	74,41617554	74,121428	73,80
	74,45775115	74,416176	74,443045	74,585693	74,240398	73,3846	73,803571	75,752499	74,457751	72,453024	74,45775115	75,330398	75,9
	74,58569263	74,283314	74,585693	75,752499	74,283314	74,416176	74,121428	74,283314	72,453024	74,90668	75,75249912	75,91147	74,41
	73,2582975	74,283314	76,251406	73,258297	73,3846	74,283314	72,453024	73,3846	73,129894	74,90668	72,72352838	74,240398	73,25
	73,80357067	75,91147	73,258297	75,752499	74,416176	76,251406	72,453024	74,416176	73,258297	73,803571	76,25140642	75,91147	74,45
	76,25140642	73,679327	73,129894	74,416176	74,416176	75,330398	74,416176	73,129894	74,121428	74,457751	75,33039831	72,723528	74,45
	74,28331424	73,129894	74,283314	74,457751	74,90668	74,443045	73,803571	76,251406	75,330398	72,453024	72,72352838	74,121428	75,9
	72,45302432	73,129894	73,679327	75,91147	75,752499	74,416176	74,585693	74,457751	76,251406	76,251406	75,75249912	73,679327	74,28
	73,38459993	73,605104	73,3846	74,443045	75,91147	73,258297	75,330398	74,443045	74,457751	75,330398	72,72352838	72,453024	74,24
	73,80357067	72,453024	73,605104	74,283314	73,679327	73,258297	74,585693	74,90668	74,585693	74,121428	75,33039831	74,90668	73,12
	74,58569263	72,453024	74,585693	74,283314	76,251406	74,457751	74,283314	76,251406	73,803571	75,752499	73,38459993	72,723528	74,28
	73,12989425	76,251406	74,240398	74,121428	75,752499	74,443045	74,240398	74,585693	73,129894	75,752499	73,2582975	75,91147	73,
	75,33039831	73,3846	74,121428	74,585693	75,91147	74,585693	76,251406	73,129894	75,752499	73,679327	76,25140642	75,91147	74,24
	74,28331424	74,585693	73,679327	74,283314	74,457751	73,679327	73,258297	73,803571	73,605104	73,129894	74,24039831	74,240398	75,33
	73,80357067	76,251406	75,91147	74,240398	72,453024	74,416176	74,283314	72,723528	73,258297	74,443045	75,33039831	72,723528	73,60
	74,45775115	74,90668	73,679327	74,283314	73,258297	75,330398	72,723528	74,121428	76,251406	75,91147	74,90667959	73,605104	73,25
	76,25140642	74,457751	73,129894	74,457751	76,251406	73,3846	75,752499	73,605104	73,258297	73,803571	74,28331424	76,251406	74,44
	75,33039831	74,121428	75,91147	74,457751	74,121428	74,585693	72,453024	75,330398	74,283314	74,416176	76,25140642	76,251406	74,44
	73,67932675	74,121428	73,605104	75,330398	75,330398	73,3846	74,90668	74,90668	74,416176	74,585693	75,91146986	75,330398	72,72
B1*boot	-2,065860398	-0,172286	1,7913811	1,3385145	-0,090487	0,0970226	-1,801211	0,4819098	0,6258498	1,4038205	-0,35089737	-0,013256	0,331
Desvio B1*boot	1,175155349	1,3436279	1,1217441	0,9689764	1,3597397	0,9594151	1,3487871	1,4042174	1,3663057	1,2927894	1,514711702	1,5977904	1,057
Tcal=B1*boot/Desvio B1*boot	-1,757946641	-0,128225	1,5969605	1,3813695	-0,066547	0,1011269	-1,33543	0,3431875	0,4580599	1,0858849	-0,23165951	-0,008296	0,312
	LI	-2,211615											
	LS	2,1140701											
	Tobs=11,35172649												

Chart Title

Bin Range	Frequency
-4.4 to -4.2	5
-4.2 to -4.0	10
-4.0 to -3.8	15
-3.8 to -3.6	20
-3.6 to -3.4	25
-3.4 to -3.2	30
-3.2 to -3.0	40
-3.0 to -2.8	50
-2.8 to -2.6	60
-2.6 to -2.4	80
-2.4 to -2.2	100
-2.2 to -2.0	120
-2.0 to -1.8	140
-1.8 to -1.6	150
-1.6 to -1.4	140
-1.4 to -1.2	120
-1.2 to -1.0	100
-1.0 to -0.8	80
-0.8 to -0.6	60
-0.6 to -0.4	40
-0.4 to -0.2	20
-0.2 to 0.0	10
0.0 to 0.2	5
0.2 to 0.4	2
0.4 to 0.6	1

