Assignment-2

Mansi Dobariya

☆ Backward Elimination
→ highest P-value is Rejected

Step 1: Y V/s $x_1 x_2 x_3 x_4$
→ x_3 & x_4 are Insignificant
highest p-value 0.7658 of x_3

So x_3 will be eliminated.

Step 2: Y V/s $x_1 x_2 x_4$
→ x_4 is Insignificant
P-value 0.539711

So x_4 will be eliminated

Step 3: Y V/s $x_1 x_2$

None of these is Insignificant

x_1 and x_2 are significant
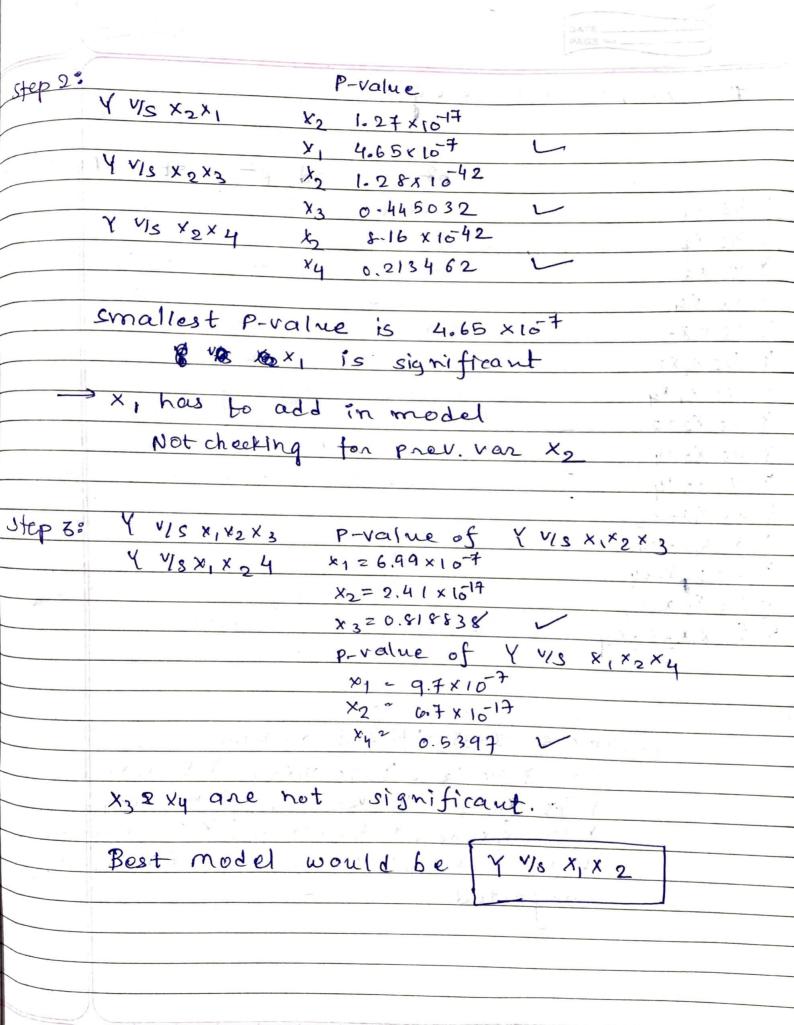So we stop the process and best model
is $\boxed{Y \text{ V/s } x_1 x_2}$

* stepwise Reg.

→ smallest P-value is selected

step 1: Reg. with single var.

	P-value
Y V/s x_1	4.22×10^{-33}
Y V/s x_2	1.26×10^{-43}
Y v/s x_3	0.07551
Y v/s x_4	0.004826

smallest P-value is of Y v/s x_2 which is significant

→ x_2 has to add in model

step 2:

		P-Value	taking in account
Y V/s $x_2 x_1$	x_2	1.27×10^{-17}	
~~Y v/s $x_2 x_2$~~	x_1	4.65×10^{-7}	✓
Y v/s $x_2 x_3$	x_2	1.28×10^{-42}	
	x_3	0.445032	✓
Y v/s $x_2 x_4$	x_2	8.16×10^{-42}	
	x_4	0.213462	✓

smallest P-value is of x_1 which is significant

checking for x_2's P-value 1.27×10^{-17} < 0.05

so its significant

→ x_1 has to add in model.

Step3:

$$Y \text{ V/s } X_1 X_2 X_3$$
$$Y \text{ V/s } X_1 X_2 X_4$$

P-value of Y V/s $X_1 X_2 X_3$

$X_1 = 6.99 \times 10^{-07}$
$X_2 = 2.41 \times 10^{-17}$
$X_3 = 0.818438$ ✓

P-value of Y V/s $X_1 X_2 X_4$

$X_1 = 9.71 \times 10^{-07}$
$X_2 = 6.7 \times 10^{-14}$
$X_4 = 0.539711$ ✓

X_3 & X_4 are Insignificant we have to stop process here

And our best model would be

$$\boxed{Y \text{ V/s } X_1 X_2}$$

✲   Forward selection
⟶ Smallest p-value is selected

Step 1: Reg. with single var

	P-value
Y V/s X_1	4.22×10^{-33}
Y V/s X_2	1.26×10^{-43}
Y V/s X_3	0.07551
Y V/s X_4	0.004826

Y V/s X_2 has smallest p-value is significant

⟶ X_2 has to add in model.

Step 2:

		P-value	
Y v/s $x_2 x_1$	x_2	1.27×10^{-17}	
	x_1	4.65×10^{-7}	⌐
Y v/s $x_2 x_3$	x_2	1.28×10^{-42}	
	x_3	0.445032	⌐
Y v/s $x_2 x_4$	x_2	8.16×10^{-42}	
	x_4	0.213462	⌐

smallest P-value is 4.65×10^{-7}

∴ v/s x_1 is significant

→ x_1 has to add in model

Not checking for prev. var x_2

Step 3:   Y v/s $x_1 x_2 x_3$ P-value of   Y v/s $x_1 x_2 x_3$

Y v/s $x_1 x_2 4$ $x_1 = 6.99 \times 10^{-7}$

$x_2 = 2.41 \times 10^{-17}$

$x_3 = 0.818838$ ✓

P-value of Y v/s $x_1 x_2 x_4$

$x_1 = 9.7 \times 10^{-7}$

$x_2 = 6.7 \times 10^{-17}$

$x_4 = 0.5397$ ✓

x_3 & x_4 are not significant.

Best model would be | Y v/s $x_1 x_2$ |

✦ All possible reg.

		R^2	adj R^2	SE	No. of Signi
Y v/s X_1	✓	0.7706	0.7683	10.9447	1/1
Y v/s X_2	✓	0.8599	0.8585	8.6519	1/1
Y v/s X_3	✗	0.03275	0.02288	22.4763	0/1
Y v/s X_4	✓	0.0782	0.06883	21.94147	1/1
Y v/s $X_1 X_2$	✓	0.8923	0.89016	7.5358	2/2
Y v/s $X_1 X_3$	✗	0.77198	0.7672	10.96907	1/2
Y v/s $X_1 X_4$	✗	0.7775	0.7729	10.83547	1/2
Y v/s $X_2 X_3$	✗	0.8608	0.8579	8.5699	1/2
Y v/s $X_2 X_4$	✗	0.8623	0.8594	8.5272	1/2
Y v/s $X_3 X_4$	✗	0.1093	0.0909	21.67913	1/2
Y v/s $X_1 X_2 X_3$	✗	0.89243	0.88907	7.5729	2/3
Y v/s $X_1 X_2 X_4$	✗	0.8928	0.8894	7.5601	2/3
Y v/s $X_1 X_3 X_4$	✗	0.7795	0.7725	10.8433	1/3
Y v/s $X_2 X_3 X_4$	✗	0.86311	0.8568	8.543	1/3
Y v/s $X_1 X_2 X_3 X_4$	✗	0.8929	0.8883	7.5962	2/4

1. Remove Insignificant values

so Y v/s X_3, Y v/s $X_1 X_3$, Y v/s $X_1 X_4$,
Y v/s $X_2 X_3$, Y v/s $X_2 X_4$, Y v/s $X_3 X_4$,
Y v/s $X_1 X_2 X_3$, Y v/s $X_1 X_2 X_4$, Y v/s $X_1 X_3 X_4$,
Y v/s $X_1 X_3 X_4$, Y v/s $X_2 X_3 X_4$, Y v/s $X_1 X_2 X_3 X_4$

Rejected

among

	R^2
Y v/s x_1	0.7706
Y v/s x_2	0.8599
Y v/s x_4	0.0782
Y v/s $x_1 x_2$	0.8923

highest R^2 value will be 0.8923 so

best model would be [Y v/s $x_1 x_2$]