Question 1

The weight loss of 24 candidates and days spent by them in the diet regime were observed and tabulated.

a. Find the best line fit based on Least Square method. Y = b0+b1x

b. Estimate the days that one should spend at diet regime if he or she aims at 17 kg of weight loss.

c. If the last entry is incorrectly entered as 3.1 , 27 instead of 4.1 , 28 then find out the line of best fit b0+b1x

For problems d to h, use the corrected entry in c (Without using regression tool in data analysis toolpack)

d. Find the coefficient of determination for the line fit. What does it tell you?

e. What is the estimate of Standard deviation of the regression model?

f. Determine Confidence interval for slope parameter with 98% confidence

g. Calculate MAD, MSD and MAPE in the line fit

h. Plot data and linear fit line

Answers	
Intercept (b0)	4.987279
Slope (b1)	4.612689
Estimated Days for 17 kg los	83.40298
Intercept (b0)	4.25291
Slope (b1)	4.682193
Coeff of Determination	0.912284
Std Dev of model	5.661073
CI Lower Limit	3.9058
CI Upper Limit	5.4586
MSD	29.3771
MAD	4.486468
MAPE	11.08699

Submission to include -
Attach snapshot of Answers Cells N2 to N16
Attach snapshot of the Plot(s)
Attach snapshot of your workings

Sl No	Weight loss in Kg (x)	Days Spent in diet regime (y)
1	5.50	26.00
2	10.10	49.00
3	11.10	53.00
4	6.70	30.00
5	1.50	11.00
6	10.50	51.00
7	11.90	67.00
8	0.00	5.00
9	6.80	40.00
10	7.10	41.00
11	8.50	36.00
12	11.40	55.00
13	10.10	42.00
14	13.90	75.00
15	11.80	48.00
16	5.40	28.00
17	4.80	25.00
18	10.40	53.00
19	10.30	60.00
20	3.70	29.00
21	10.30	60.00
22	11.00	55.00
23	14.70	79.00
24	3.10	27.00

Sl No	Weight loss (x)	days spent (y)	Weight loss in Kg X^2	Days Spent in diet regime Y^2	Weight loss in Kg (x)*Days Spent in diet regime (y) XY	SSxx	Ssyy	Ssxy
1	5.5	26	30.25	676	143	8.170069	307.7101	50.13993056
2	10.1	49	102.01	2401	494.9	3.033403	29.7934	9.506597222
3	11.1	53	123.21	2809	588.3	7.516736	89.46007	25.93159722
4	6.7	30	44.89	900	201	2.750069	183.3767	22.45659722
5	1.5	11	2.25	121	16.5	47.03674	1058.96	223.1815972
6	10.5	51	110.25	2601	535.5	4.586736	55.62674	15.97326389
7	11.9	67	141.61	4489	797.3	12.5434	550.2934	83.08159722
8	0	5	0	25	0	69.86174	1485.46	322.1440972
9	6.8	40	46.24	1600	272	2.428403	12.5434	5.519097222
10	7.1	41	50.41	1681	291.1	1.583403	6.460069	3.198263889
11	8.5	36	72.25	1296	306	0.020069	56.87674	-1.068402778
12	11.4	55	129.96	3025	627	9.251736	131.2934	34.85243056
13	10.1	42	102.01	1764	424.2	3.033403	2.376736	-2.685069444
14	13.9	75	193.21	5625	1042.5	30.71007	989.6267	174.3315972
15	11.8	48	139.24	2304	566.4	11.84507	19.87674	15.34409722
16	5.4	28	29.16	784	151.2	8.751736	241.5434	45.97743056
17	4.8	25	23.04	625	120	12.66174	343.7934	65.97743056
18	10.4	53	108.16	2809	551.2	4.168403	89.46007	19.31076389
19	10.3	60	106.09	3600	618	3.770069	270.8767	31.95659722
20	3.7	29	13.69	841	107.3	21.70007	211.4601	67.73993056
21	10.3	60	106.09	3600	618	3.770069	270.8767	31.95659722
22	11	55	121	3025	605	6.978403	131.2934	30.26909722
23	14.7	79	216.09	6241	1161.3	40.21674	1257.293	224.8649306
24	3.1	27	9.61	729	83.7	27.65007	273.6267	86.98159722
Total	200.6	1045	2020.72	53571	10321.4	344.04	8069.96	1586.9417
Mean	8.3583	43.54						
Slope	4.612688509							
Intercept	4.987278549							
Sxy	1586.941667							

a). Therefore, we find that the regression equation is:
Days Spent in diet regime (y) = 4.9873 + 4.6127 Weight loss in Kg (x)

b). Estimate the days that one should spend at diet regime if he or she aims at 17 kg of weight loss.
Days Spent in diet regime (y) = 4.9873 + 4.6127 Weight loss in Kg (x)
When, weight loss = 17 kgs
Days spent in diet regime:

Days 83.40298319

Sl No	Weight loss (x)	days spent (y)	Weight loss	Days Spent	Weight loss	SSxx	Ssyy	Ssxy
1	5.5	26	30.25	676	143	8.41	309.1736	50.99166667
2	10.1	49	102.01	2401	494.9	2.89	29.34028	9.208333333
3	11.1	53	123.21	2809	588.3	7.29	88.67361	25.425
4	6.7	30	44.89	900	201	2.89	184.5069	23.09166667
5	1.5	11	2.25	121	16.5	47.61	1061.674	224.825
6	10.5	51	110.25	2601	535.5	4.41	55.00694	15.575
7	11.9	67	141.61	4489	797.3	12.25	548.3403	81.95833333
8	0	5	0	25	0	70.56	1488.674	324.1
9	6.8	40	46.24	1600	272	2.56	12.84028	5.733333333
10	7.1	41	50.41	1681	291.1	1.69	6.673611	3.358333333
11	8.5	36	72.25	1296	306	0.01	57.50694	-0.758333333
12	11.4	55	129.96	3025	627	9	130.3403	34.25
13	10.1	42	102.01	1764	424.2	2.89	2.506944	-2.691666667
14	13.9	75	193.21	5625	1042.5	30.25	987.0069	172.7916667
15	11.8	48	139.24	2304	566.4	11.56	19.50694	15.01666667
16	5.4	28	29.16	784	151.2	9	242.8403	46.75
17	4.8	25	23.04	625	120	12.96	345.3403	66.9
18	10.4	53	108.16	2809	551.2	4	88.67361	18.83333333
19	10.3	60	106.09	3600	618	3.61	269.5069	31.19166667
20	3.7	29	13.69	841	107.3	22.09	212.6736	68.54166667

21	10.3	60	106.09	3600	618	3.61	269.5069	31.19166667		
22	11	55	121	3025	605	6.76	130.3403	29.68333333		
23	14.7	79	216.09	6241	1161.3	39.69	1254.34	223.125		
24	4.1	28	16.81	784	114.8	18.49	242.8403	67.00833333		
Total	201.6	1046	2027.92	53626.00	10352.50	334.48	8037.83	1566.10		
Mean	8.4	43.58333								
Slope	4.682193255									
Intercept	4.25290999									
Sxx	334.48									
Syy	8037.83									
Sxy	1566.10									

c). Therefore, we find that the regression equation is:
Days Spent in diet regime (y) = 4.6821 + 4.2529 Weight loss in Kg (x)

Slope bi	Ssxy/SSxx	4.612689								
Intercept (b0)		4.987279								
		4.25291								
Therefore, we find that the regression equation is: Days Spent in diet regime (y) = 4.2529 + 4.6822 Weight loss in Kg										

d). Find the coefficient of determination for the line fit. What does it tell you?
R-sq =Sxy/sqrt(Sxx*Syy).That means 91% of variation in the days spent is explained by the weight loss.

R-sq	0.955135339									
R2	0.912283516									

e)
As given above the error variance $\hat{\sigma}^2 = (Syy - \frac{Sxy^2}{Sxx}) * \frac{1}{n-2}$

sdv2			32.04775	(F91-((F92*F92)/F90))*(1/(E85-2))						
Std erro			5.661073							

f. Determine Confidence interval for slope parameter with 98% $(\hat{\beta} \mp t_{n-2,\alpha/2} \sqrt{\frac{\hat{\sigma}^2}{Sxx}})$

$(1-\alpha)\%$ is the confidence interval for population slope

Where alpha = 1 - 0.98 = 0.02

$=t_{\alpha/2,0.01}$	2.5083									
	98% CI : (3.9058, 5.4586)									
Upper	5.4586									
Lower	3.9058									

g) . MSD = Σ(y - Predicted y)^2 / n MAD = Σ | y - Predicted y | / n MAPE = Σ | y - Predicted y | / n * 100

x	y	yfit	Abs Dev	Square dev	Abs Percent Err					
5.5	26	30.005	4.005	16.04003	15.40385					
10.1	49	51.54312	2.54312	6.467459	5.190041					
11.1	53	56.22532	3.22532	10.40269	6.085509					
6.7	30	35.62364	5.62364	31.62533	18.74547					
1.5	11	11.2762	0.2762	0.076286	2.510909					
10.5	51	53.416	2.416	5.837056	4.737255					
11.9	67	59.97108	7.02892	49.40572	10.49093					
0	5	4.2529	0.7471	0.558158	14.942					
6.8	40	36.09186	3.90814	15.27356	9.77035					
7.1	41	37.49652	3.50348	12.27437	8.545073					
8.5	36	44.0516	8.0516	64.82826	22.36556					
11.4	55	57.62998	2.62998	6.916795	4.781782					
10.1	42	51.54312	9.54312	91.07114	22.72171					
13.9	75	69.33548	5.66452	32.08679	7.552693					
11.8	48	59.50286	11.50286	132.3158	23.96429					
5.4	28	29.53678	1.53678	2.361693	5.4885					
4.8	25	26.72746	1.72746	2.984118	6.90984					
10.4	53	52.94778	0.05222	0.002727	0.098528					
10.3	60	52.47956	7.52044	56.55702	12.53407					
3.7	29	21.57704	7.42296	55.10034	25.59641					
10.3	60	52.47956	7.52044	56.55702	12.53407					
11	55	55.7571	0.7571	0.5732	1.376545					
14.7	79	73.08124	5.91876	35.03172	7.492101					
3.1	27	23.44992	4.55008	20.70323	16.25029					
			4.486468	29.3771	11.08699					
			MAD	MSD	MAPE					

Testing two slopes

Weight loss(x)	days (y)
5.50	26.00
10.10	49.00
11.10	53.00
6.70	30.00
1.50	11.00
10.50	51.00
11.90	67.00
0.00	5.00
6.80	40.00
7.10	41.00
8.50	36.00
11.40	55.00

Weight loss (	days (y)
5.5	26
10.1	49
11.1	53
6.7	30
1.5	11
10.5	51
11.9	67
0	5
6.8	40
7.1	41
8.5	36
11.4	55

10.10	42.00
13.90	75.00
11.80	48.00
5.40	28.00
4.80	25.00
10.40	53.00
10.30	60.00
3.70	29.00
10.30	60.00
11.00	55.00
14.70	79.00
3.10	27.00

10.1	42
13.9	75
11.8	48
5.4	28
4.8	25
10.4	53
10.3	60
3.7	29
10.3	60
11	55
14.7	79
4.1	28

Comparison of two slopes

n	24	24	=COUNT(x)
b	4.612688509	4.682193	=SLOPE(y,x)
$s_{y\cdot x}$	5.838316786	5.661073	=STEYX(y,x)
s_x	3.867581726	3.813477	=STDEV(x)
s_b	0.314763408	0.309538	= $s_{y\cdot x}$ / (s_x * SQRT(n-1))

			s_{Res}^2	33.06685	= $((n_1-2)s_{y\cdot x1}^2+(n_2-2)s_{y\cdot x2}^2)/(n_1+n_2-4)$
s_{b1-b2}	0.44146308	= $SQRT(s_{b1}^2+s_{b2}^2)$	s_{b1-b2}	0.441559	= s_{Res}*SQRT(1/(s_{x1}^2(n_1-1))+1/(s_{x2}^2(n_2-1)))
t	-0.157441811	= $(b_1-b_2)/(s_{b1-b2})$	t	-0.15741	= $(b_1-b_2)/(s_{b1-b2})$
df	44	= n_1+n_2-4	df	44	= n_1+n_2-4
alpha	0.02		alpha	0.05	
p-value	0.875616954	= TDIST(|t|,df,2)	p-value	0.875644	= TDIST(|t|,df,2)
t-crit	2.414134368	= TINV(α,df)	t-crit	2.015368	= TINV(α,df)
sig	no	= yes if p-value < α	sig	no	= yes if p-value < α

h)