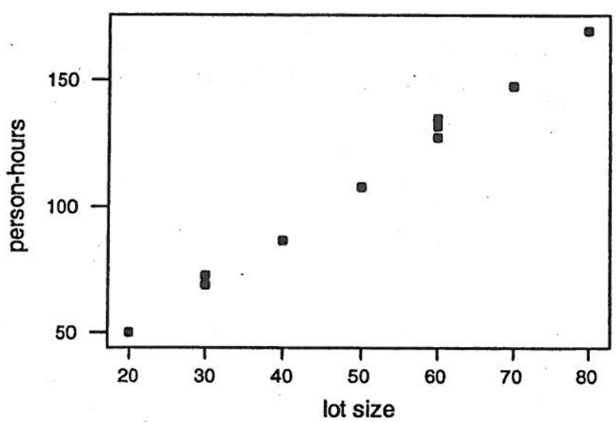
Answers

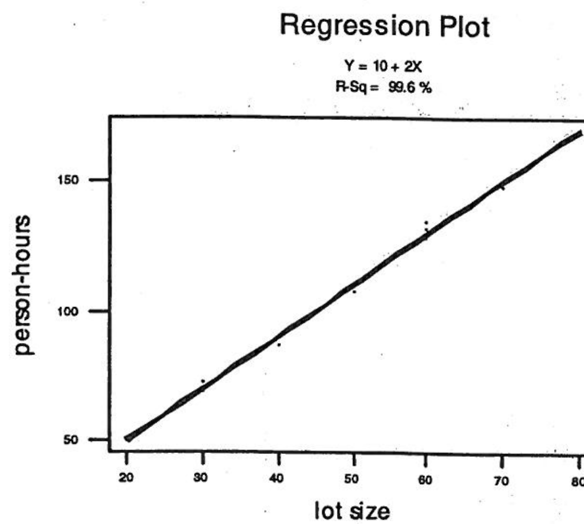
IE 3301: Simple Linear Regression

A company that manufactures a widget is interested in the relationship between:

Y = production person-hours
X = lot size

Production Run (i)	Person-Hours (Y_i)	Lot Size (X_i)	$x_i y_i$	x_i^2	$\hat{Y}_i$	e_i
1	73	30	2190	900	70	3
2	50	20	1000	400	50	0
3	128	60	7680	3600	130	-2
4	170	80	13600	6400	170	0
5	87	40	3480	1600	90	-3
6	108	50	5400	2500	110	-2
7	135	60	8100	3600	130	5
8	69	30	2070	900	70	-1
9	148	70	10360	4900	150	-2
10	132	60	7920	3600	130	2
Totals	1100	500	61,800	28,400	/	0





Solve for Least Squares Estimates:

$$\hat{\beta}_1 = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sum x_i^2 - \frac{\left(\sum x_i\right)^2}{n}} = \frac{61{,}800 - \frac{(500)(1100)}{10}}{28{,}400 - \frac{(500)^2}{10}} = \frac{\overset{S_{xy}}{6800}}{\underset{S_{xx}}{3400}} = 2.0$$

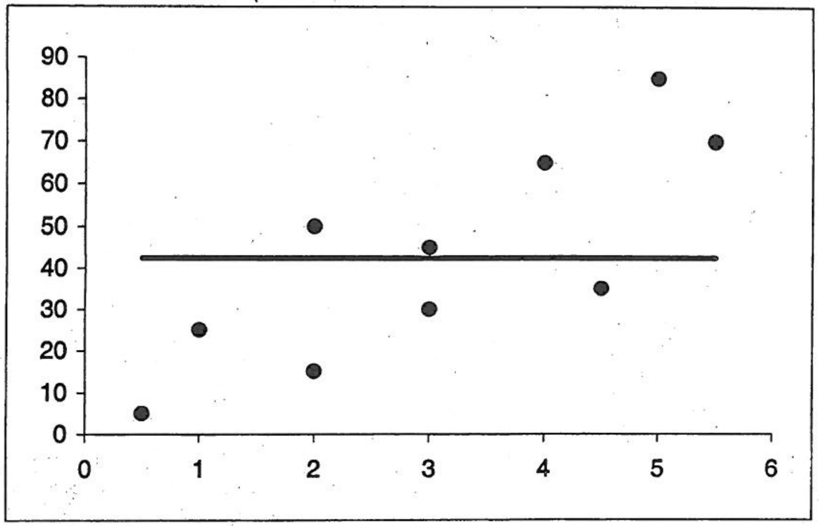
$$\hat{\beta}_0 = \frac{1}{n}\left(\sum y_i - \hat{\beta}_1 \sum x_i\right) = \frac{1}{10}\left(1100 - (2.0)(500)\right) = 10.0$$

Fitted values: $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i = 10 + 2x_i$ → $\hat{y}_1 = 10 + 2(30) = 70$, $\hat{y}_2 = 10 + 2(20) = 50$

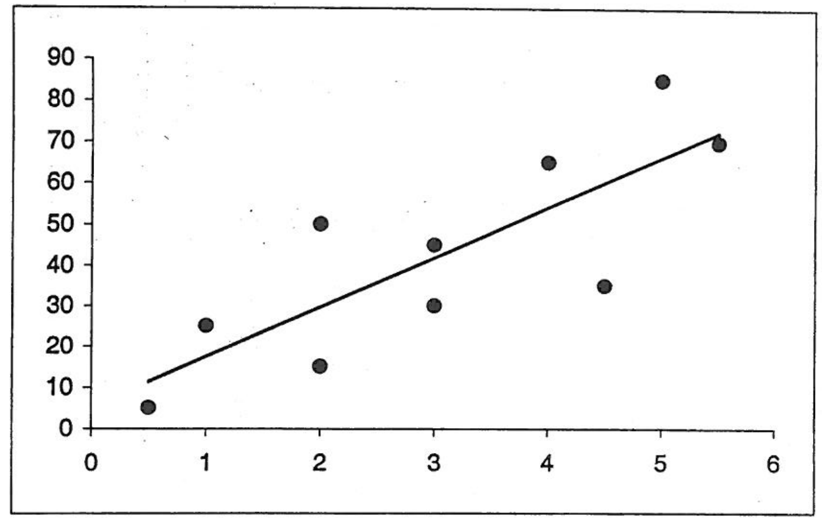
Residuals: $e_i = Y_i - \hat{Y}_i$ → $e_1 = y_1 - \hat{y}_1 = 73 - 70 = 3$, $e_2 = y_2 - \hat{y}_2 = 50 - 50 = 0$

3301: Graphical Illustration of Analysis of Variance (ANOVA)

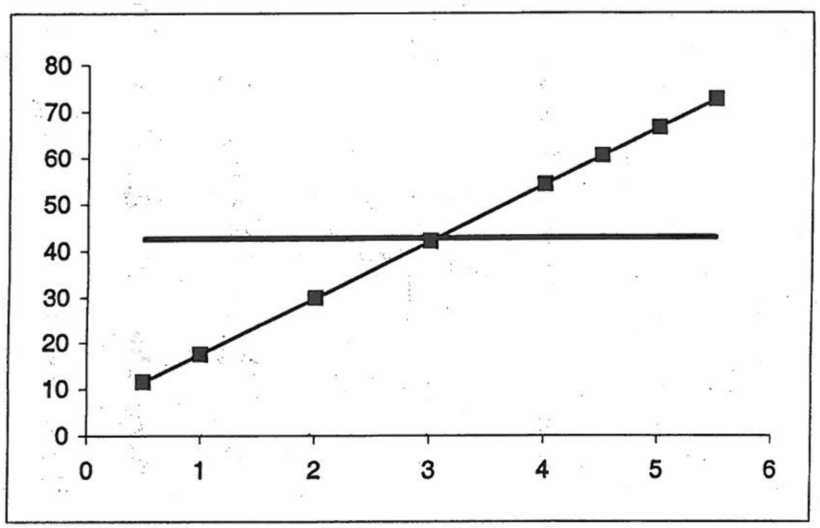
Vertical Deviations for Total Sum of Squares



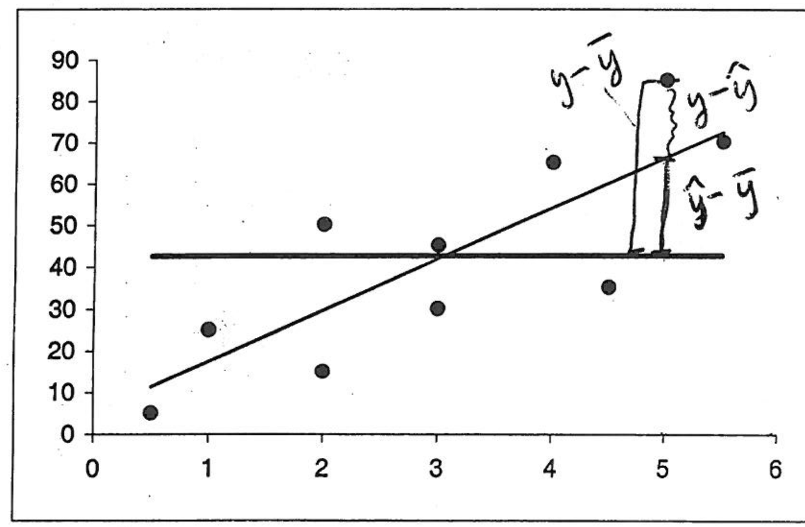
Vertical Deviations for Error Sum of Squares



Vertical Deviations for Regression Sum of Squares



Decomposition of Vertical Deviations



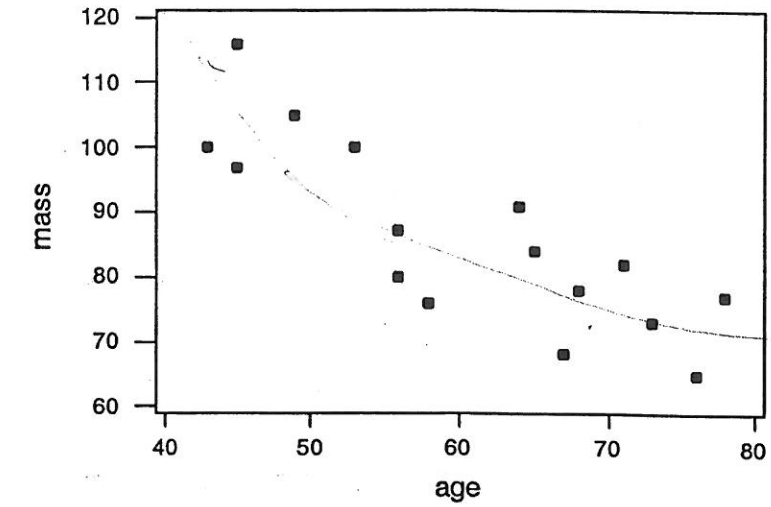
IE 3301: Simple Linear Regression Muscle Mass Example

Problem Description: A person's muscle mass is expected to decrease with age. To explore this relationship in women, a nutritionist randomly selected four women from each 10-year age group (40-49, 50-59, 60-69, 70-79).

Minitab

Data Display

Row	mass	age
1	82	71
2	91	64
3	100	43
4	68	67
5	87	56
6	73	73
7	78	68
8	80	56
9	65	76
10	84	65
11	116	45
12	76	58
13	97	45
14	100	53
15	105	49
16	77	78



Regression Analysis

```
The regression equation is
mass = 148 - 1.02 age
```

Predictor	Coef	StDev	T	P
Constant	148.05	11.56	12.80	0.000
age	-1.0236	0.1882	-5.44	0.000

S = 8.344   R-Sq = 67.9%   R-Sq(adj) = 65.6%

Analysis of Variance

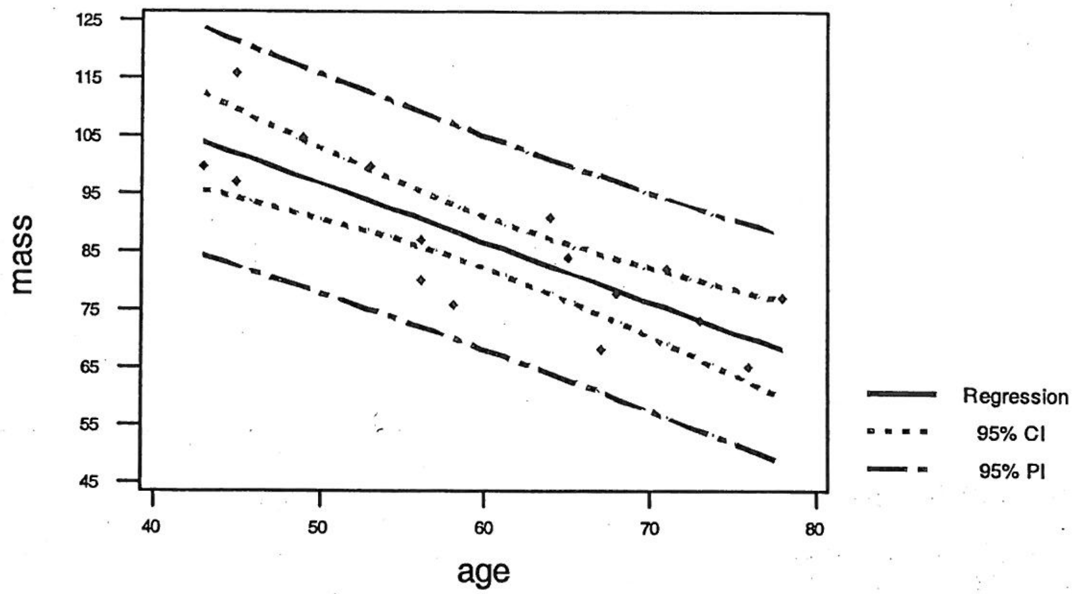
Source	DF	SS	MS	F	P
Regression	1	2059.8	2059.8	29.59	0.000
Residual Error	14	974.7	69.6		
Total	15	3034.4			

Predicted Values

Fit	StDev Fit	95.0% CI	95.0% PI
75.38	2.88	(69.20, 81.56)	(56.44, 94.31)
82.54	2.19	(77.84, 87.24)	(64.04, 101.04)
104.04	3.89	(95.70, 112.38)	(84.29, 123.78)
79.47	2.42	(74.27, 84.67)	(60.83, 98.11)
90.73	2.25	(85.91, 95.55)	(72.20, 109.26)
73.33	3.15	(66.57, 80.09)	(54.20, 92.46)
78.45	2.53	(73.03, 83.86)	(59.75, 97.14)
90.73	2.25	(85.91, 95.55)	(72.20, 109.26)
70.26	3.60	(62.55, 77.97)	(50.77, 89.74)
81.52	2.26	(76.68, 86.36)	(62.98, 100.06)
101.99	3.58	(94.32, 109.66)	(82.52, 121.46)
88.68	2.14	(84.10, 93.26)	(70.21, 107.16)
101.99	3.58	(94.32, 109.66)	(82.52, 121.46)
93.80	2.51	(88.41, 99.19)	(75.11, 112.49)
97.89	3.00	(91.47, 104.32)	(78.88, 116.91)
68.21	3.91	(59.83, 76.59)	(48.45, 87.97)

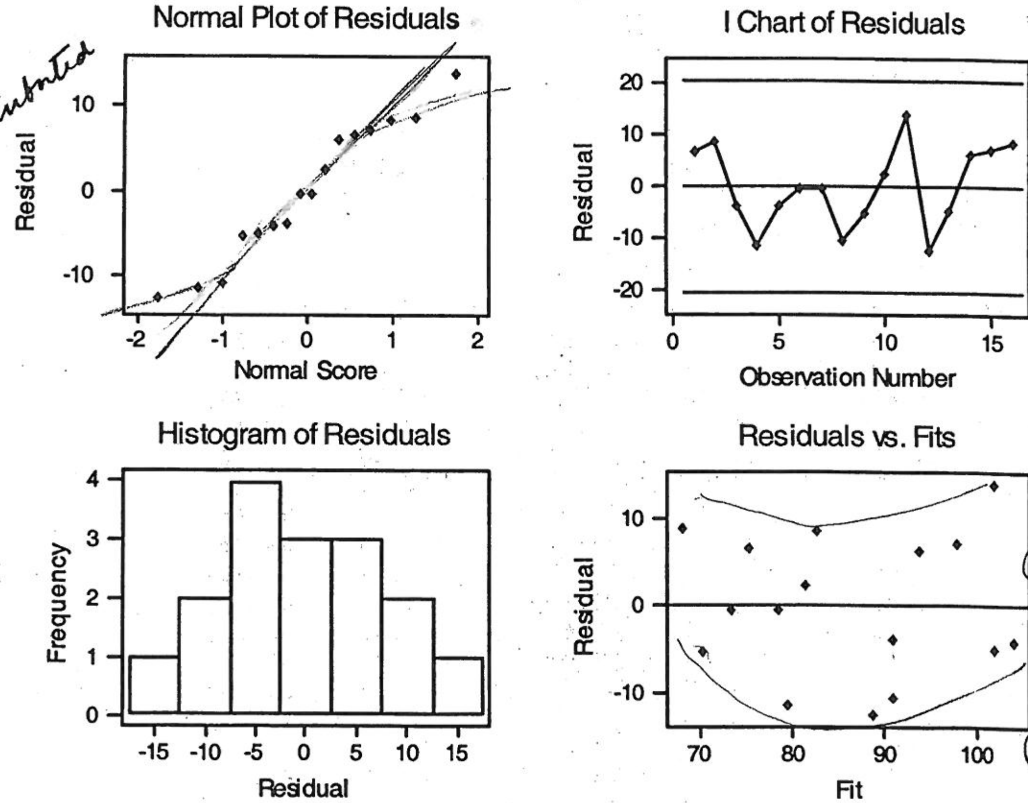
Regression Plot

Y = 148.051 - 1.02359X
R-Sq = 67.9 %



Residual Model Diagnostics

not quite normally distributed



① very slight curvature ⇒ maybe transformati

② no funnel constant variance OK