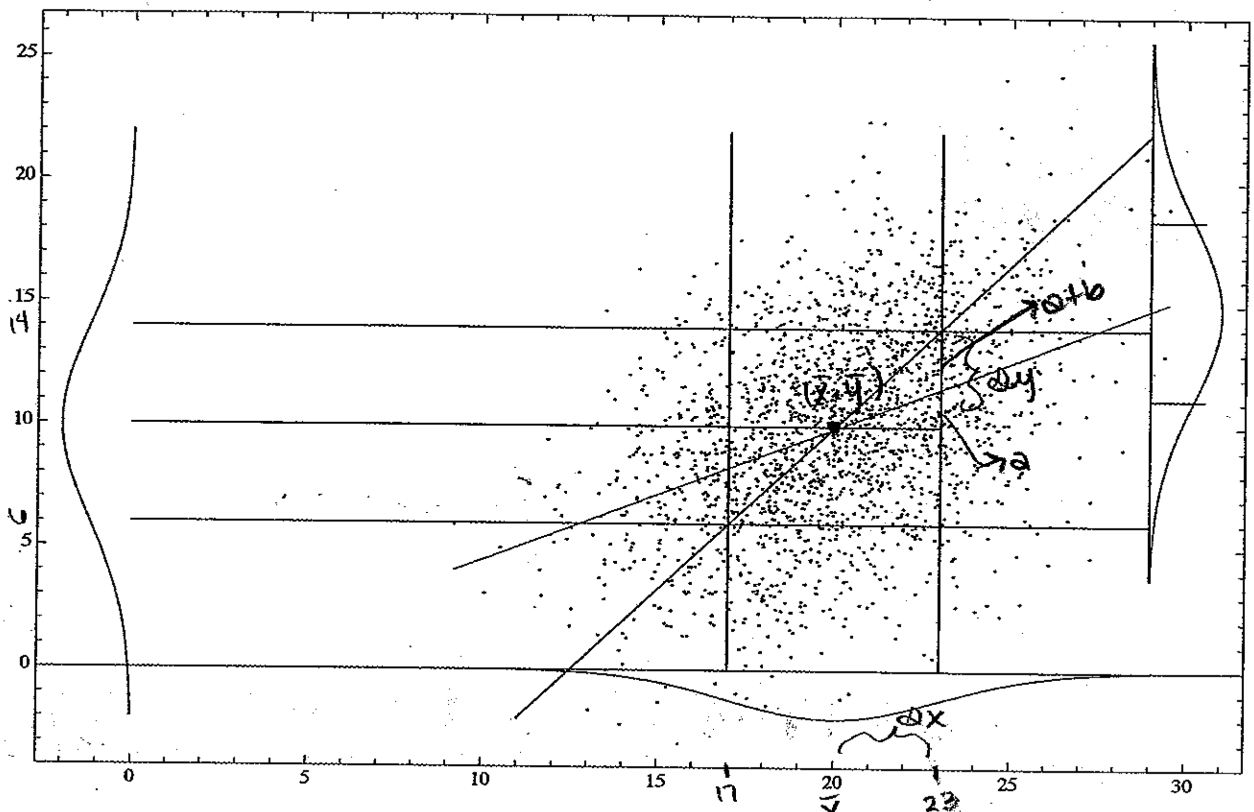
1 - 7. Refer to the (x, y) plot below (y is vertical).
Closest answers. Pay attention to scales on axes and origin.



1. μ_x (approx $\bar{x}$) (a) 17 (b) 18 (c) 19 (d) 20 (e) 21

2. S_x (a) 1 (b) 2 (c) 3 (d) 4 (e) 5 $= \frac{23-17}{2} = 3$

3. S_y (a) 2 (b) 3 (c) 4 (d) 5 (e) 6 $= \frac{14-6}{2} = 8/2 = 4$

4. r (a) 1 (b) .75 (c) .4 (d) .0 $r = a/(a+b) = 1.5/4 = .375$

5. sd of y for x = 15 (a) 0 (b) 1.6 (c) 2.9 (d) 3.5

std. dev. $= \sqrt{1-r^2}\,(\Delta y)$
$= \sqrt{1-(.4)^2}\,(4)$
$= 3.666$

6. naive line slope
(a) 1.3 (b) 2.1 (c) 3.0 (d) 3.6 (e) 4.5

naive slope $= \Delta y / \Delta x = 4/3 = 1.33$

7. regrline slope (a) .2 (b) .5 (c) .7 (d) 1.1 (e) 1.5

regr. slope $= r(\Delta y/\Delta x) = .4(1.3) = .52$

.52

8-12. For the data whose averages are reported below, choose closest answers.

x	y	x^2	y^2	xy
1	0	1	0	0
2	8	4	64	16
3	10	9	100	30
6	6	36	36	36
—	—	—	—	—
3.	6.	12.5	50.	20.5

$\hat{\sigma}_x = \sqrt{12.5 - (3^2)}$

8. s_x (a) 1 (b) 2 (c) 3 (d) 4 (e) 5

$\hat{\sigma}_x = \sqrt{\overline{x^2} - \bar{x}^2} = \sqrt{12.5 - (3^2)} = 1.87$

9. s_y (a) 1.7 (b) 2.7 (c) 3.7 (d) 4.7 (e) 5.7

$\hat{\sigma}_y = \sqrt{\overline{y^2} - \bar{y}^2} = \sqrt{50 - (36)} = 3.74$

$20.5 - 18 = \frac{2.5}{6.919} = .36$

10. r (a) .26 (b) .36 (c) .46 (d) .56 (e) .66

$r = \frac{\overline{xy} - \bar{x}\bar{y}}{\sqrt{\overline{x^2} - \bar{x}^2}\sqrt{\overline{y^2} - \bar{y}^2}} = \frac{20.5 - (3 \cdot 6)}{(1.87)(3.7)} =$

11. naive line slope (a) 0 (b) 1 (c) 2 (d) 3 (e) 4

naive slope $= \Delta y / \Delta x$ $3.7 / 1.87 = 1.97$

12. regr line slope (a) .6 (b) .7 (c) .8 (d) .9 (e) 1

regr. slope $= r(\Delta y / \Delta x)$ $.36(1.97) = .7092$

13-15. A multiple linear regression on males at risk for stroke employs variables

y = probability of stroke within five years

x1 = age (years)   x2 = years since last incident

The MLR gives R = 0.86, s_y = 0.07, $\hat{b}_0 = 0.17$, $\hat{b}_1 = 0.003$, $\hat{b}_2 = -0.02$.

13. Fraction of s_y explained by MLR. (a) .70 (b) .74 (c) .78 (d) .82 (e) .86

R= 0.86 → R^2 = 0.86^2 = .7396

14. $\hat{y}$ (predicted probability of a stroke within five years) for a subject 74 yrs of age, with 5 years since his last incident.

(a) 0.07 (b) 0.14 (c) 0.29 (d) 0.33 (e) 0.36

$\hat{y} = \hat{b}_0 + \hat{b}_1 x_1 + \hat{b}_2 x_2$ $\hat{y}$ = 0.17+ 0.003(74) + -0.02 (5)

$\hat{y}$ = .292

15. If the multidimensional plot is normal (elliptical) what is the standard deviation of y for all subjects 74 years of age with 5 years since their last incident?

(a) .006 (b) .036 (c) .066 (d) .096 (e) .126

$\sqrt{1-r^2}$ (s_y) ~~1.96 (0.07/√n)~~ $\sqrt{1-.74}$ (.07) = .0356

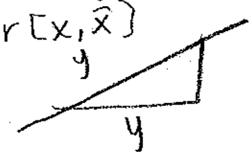
16-19 These concern MLR and straight line regression in general.

16. What is the range of R in general? (a) [0, 1] (b) [-1, 1] (c) [-0.5, 0.5]

17. If in a MLR we have r[y, $\hat{y}$] = 0.7 what is R?

(a) 0.49 (b) 0.7 (c) 1 - 0.49 (d) 1 - 0.7

r_{MLR} = |r| r[y, $\hat{y}$] = 0.7   r[x, $\bar{x}$]



18. For straight line regression what is r[-3x+4, 2y-6]?

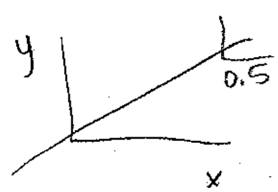
(a) -6 r[x, y] (b) -6 r[x, y] - 24 (c) r[x, y] (d) -r[x, y] (e) r[x, y] - 24

r[-3x+4, 2y-6] = -r[x, y]

19. If in a straight line regression we have r[x, y] = -0.5 what is R?

(a) -0.5 (b) -0.25 (c) 0.25 (d) 0.5 (e) .75

b/c r_{MLR} = |r|



correlation of y to $\hat{y}$ = .70