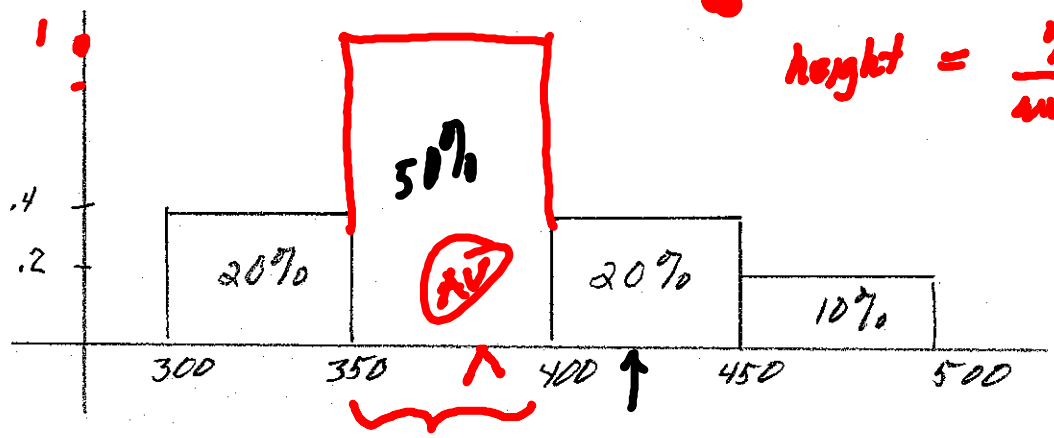
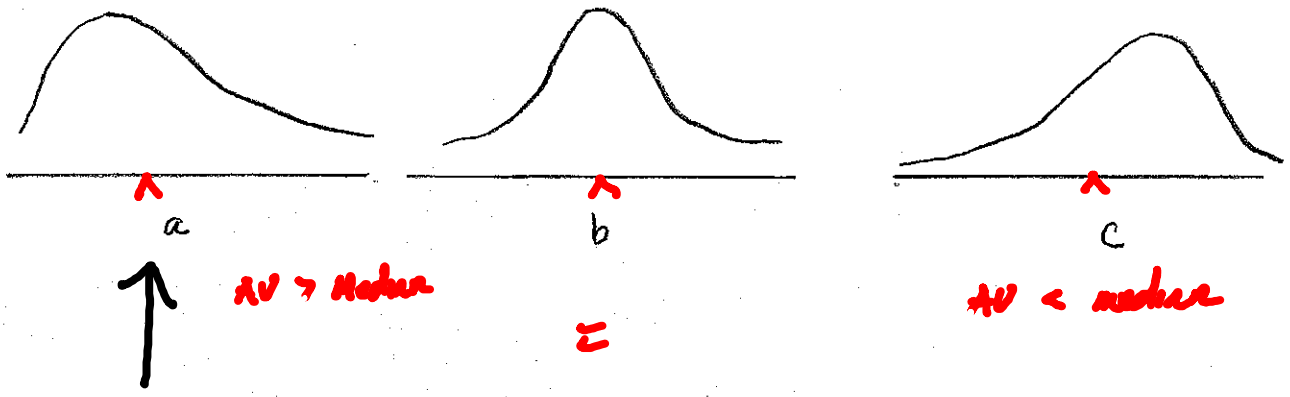
Review for Quiz 2:

1. Below is a histogram for rents paid in Logan. Find the height of the missing block from 350 to 400. Estimate the percentage of rents less than $400 per month. Estimate the percentage of rents less than $425 per month. Approximately, what is the average rent? Estimate the median rent.

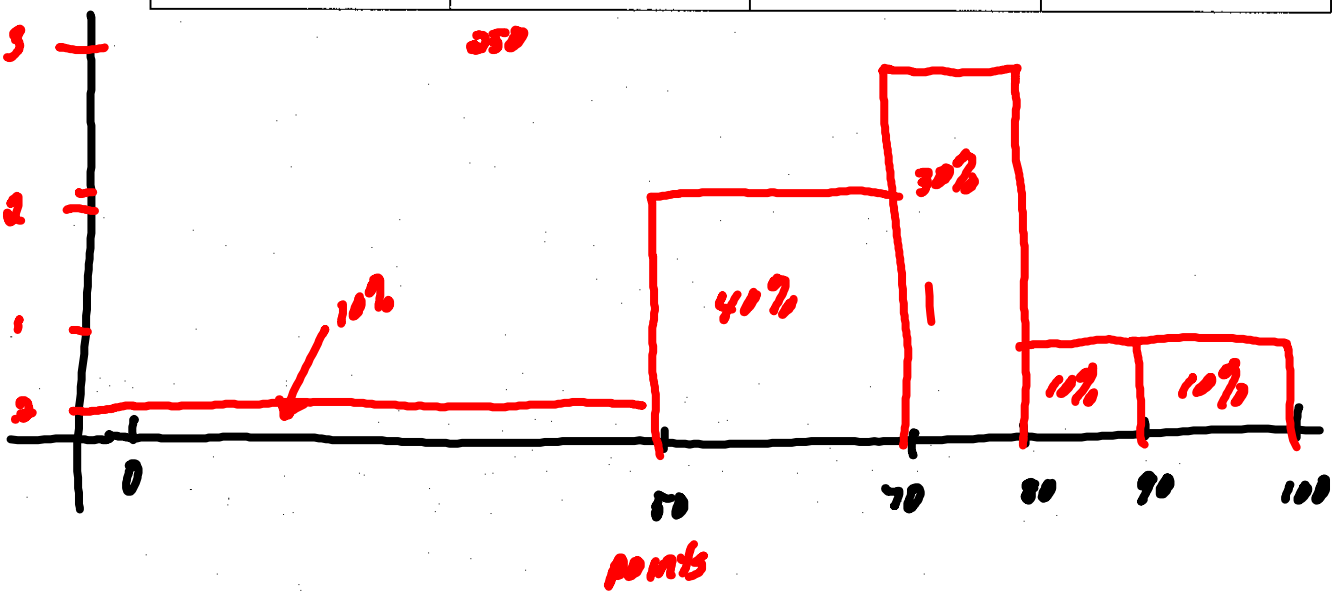


2. One question on the 2000 census form asks how far you have to commute to work. When the data is analyzed, it is found that the average distance is 13.3 miles, with an SD of 16.8 miles.
Which of the following sketches do you think best approximates the histogram for the distance traveled? Do you think the median distance traveled would be smaller than 13.3 miles, larger than 13.3 miles, or about 13.3 miles?



3. A table for the scores on a statistics midterm is given below. The class intervals include the left endpoint, but not the right. Draw the histogram.

Points	Number of Students	%	height
0-50	25	10	.2
50-70	100	40	2
70-80	75	30	3
80-90	25	10	1
90-100	25	10	1



4. The heights (inches) for the starting five on a women's basketball team are 66, 67, 69, 71, and 72.

a) Find their average height.

$\frac{66+67+69+71+72}{5} = \boxed{69}$ ↑ median

b) Find the SD of the list.

deviations from AV: $-3, -2, 0, 2, 3$

SD = RMS of deviations from (mean size of the deviations)

$9, 4, 0, 4, 9$

$\frac{9+4+0+4+9}{5} = \frac{26}{5}$

$\sqrt{\frac{26}{5}} = 2.28$

5. The average of a list of numbers is 1943 and the SD is 68. If you add 2012 to each number on the list, find the average and SD of the new list.

New AV = $1943 + 2012$

SD is still 68

6. The average of a list of 12 numbers is 9 and the SD is equal to 3. If the numbers 6 and 12 are included with this list of numbers, what is the average and SD of the new list?

New AV is still 9.

SD = ?

SD is still 3

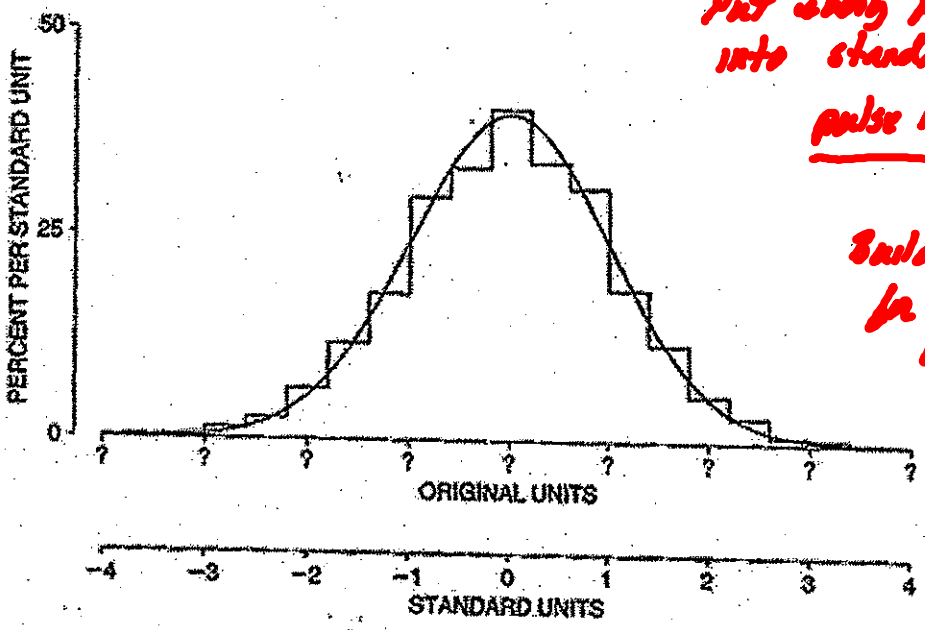
Normal Approximation for Data

1. A group of men working at an auto assembly plant have an average pulse rate of 70 beats per minute and an SD equal to 11. What does it mean to say that their pulse rates "follow the normal curve"?

Put every pulse rate into standard units

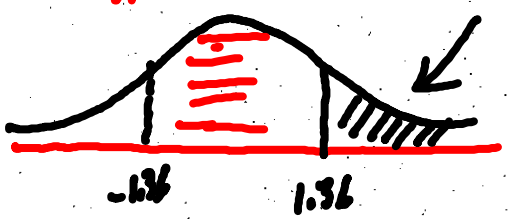
$$\frac{\text{pulse rate} - 70}{11}$$

Build a histogram for the standard unit values.



2. What percentage has a pulse rate greater than 85 beats per minute?

$$\frac{85-70}{11} = 1.36$$

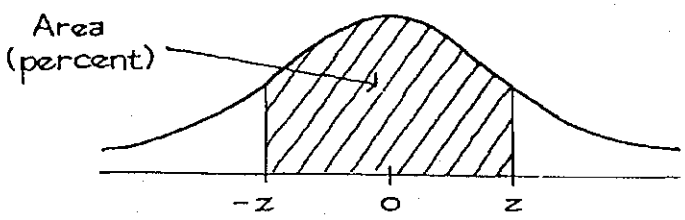


$\Phi(1.36) = 82\%$

$100 - 82 = 18\%$

$\frac{18\%}{2} = 9\%$

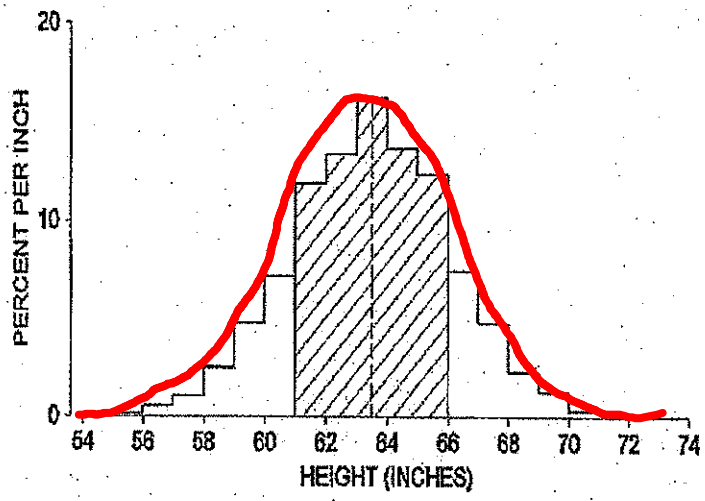
A NORMAL TABLE



z	Area	z	Area	z	Area
0.00	0	1.50	86.64	3.00	99.730
0.05	3.99	1.55	87.89	3.05	99.771
0.10	7.97	1.60	89.04	3.10	99.806
0.15	11.92	1.65	90.11	3.15	99.837
0.20	15.85	1.70	91.09	3.20	99.863
0.25	19.74	1.75	91.99	3.25	99.885
0.30	23.58	1.80	92.81	3.30	99.903
0.35	27.37	1.85	93.57	3.35	99.919
0.40	31.08	1.90	94.26	3.40	99.933
0.45	34.73	1.95	94.88	3.45	99.944
0.50	38.29	2.00	95.45	3.50	99.953
0.55	41.77	2.05	95.96	3.55	99.961
0.60	45.15	2.10	96.43	3.60	99.968
0.65	48.43	2.15	96.84	3.65	99.974
0.70	51.61	2.20	97.22	3.70	99.978
0.75	54.67	2.25	97.56	3.75	99.982
0.80	57.63	2.30	97.86	3.80	99.986
0.85	60.47	2.35	98.12	3.85	99.988
0.90	63.19	2.40	98.36	3.90	99.990
0.95	65.79	2.45	98.57	3.95	99.992
1.00	68.27	2.50	98.76	4.00	99.9937
1.05	70.63	2.55	98.92	4.05	99.9949
1.10	72.87	2.60	99.07	4.10	99.9959
1.15	74.99	2.65	99.20	4.15	99.9967
1.20	76.99	2.70	99.31	4.20	99.9973
1.25	78.87	2.75	99.40	4.25	99.9979
1.30	80.64	2.80	99.49	4.30	99.9983
1.35	82.30	2.85	99.56	4.35	99.9986
1.40	83.85	2.90	99.63	4.40	99.9989
1.45	85.29	2.95	99.68	4.45	99.9991

Example: The histogram of heights of 6,588 women aged 18-74 in NHANES II.

Average = 63.5"
SD = 2.5"



Average – 1 SD = 61"
Average + 1 SD = 66"
} 67% of the women were between 61" and 66" tall (shaded)

Average – 2 SDs = 58.5"
Average + 2 SDs = 68.5"
} 94% of the women were between 58.5" and 68.5" tall

1. Find the percentage of women who are between 5'2" and 5'7".

2. What percentage of women are taller than 5'8"?

3. What height represents the 90-th percentile?

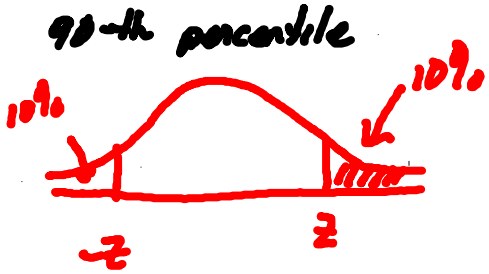
4. What percentile is represented by a height of 60"?

$\frac{62 - 63.5}{2.5} = -.6$

$\frac{67 - 63.5}{2.5} = 1.4$

$\frac{A(1.4)}{2} + \frac{A(.6)}{2}$

$\approx \frac{84}{2} + \frac{45}{2} \approx 65\%$



$A(z) = 80\%$

$z = 1.3$

Chapter 6: Measurement Error

If the same thing is measured several times, several different measurements are likely:

Measurement = exact value + chance error

seen

not seen

To estimate the exact value, we take the average of the measurements.

The SD of the measurements gives the likely size of the chance error in any single measurement.

Sometimes, measurements also have bias (they are systematically too small or too big):

Measurement = exact value + bias + chance error

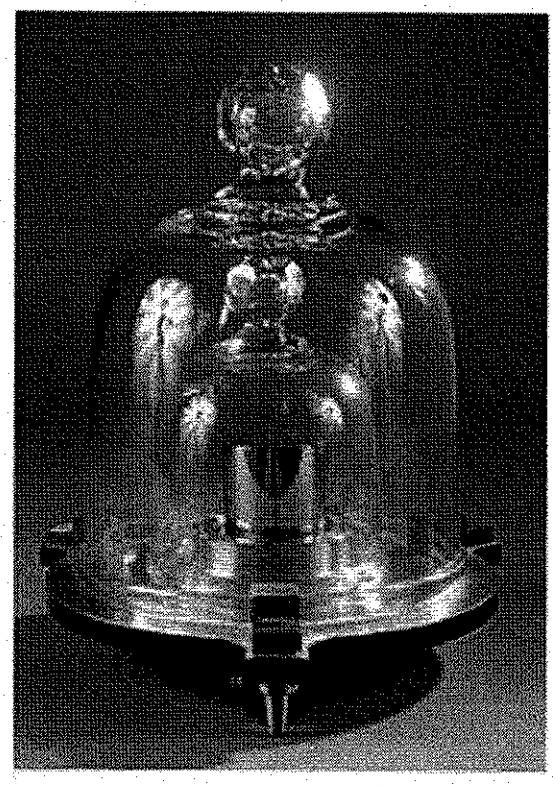
For example, if we measure someone's height when they are wearing shoes, the measurements will have some bias.

Statistics can tell us about the chance error but we can only estimate the bias using other information, e.g. we could measure the height of the shoes.

Measurement Error

$$m_1 = w + b + e_1$$
$$m_2 = w + b + e_2$$
$$m_3 = w + b + e_3$$
$$m_4 = w + b + e_4$$
$$m_5 = w + b + e_5$$

$$\frac{m_1 + m_2 + m_3 + m_4 + m_5}{5} = w + b + \frac{e_1 + e_2 + e_3 + e_4 + e_5}{5}$$



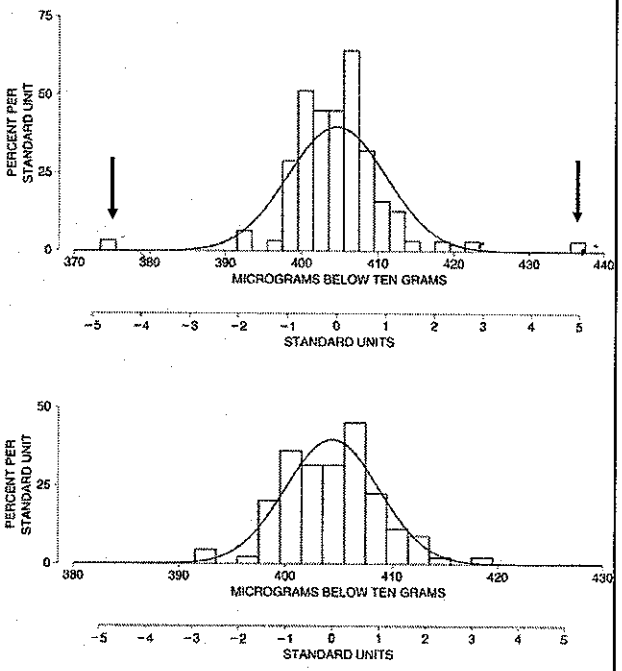
The U.S. National Prototype Kilogram, K20

Even careful measurements have occasional outliers.

Table 1. One hundred measurements on NB 10. Almer and Jones, National Bureau of Standards. Units are micrograms below 10 grams.

No.	*Result*	*No.*	*Result*	*No.*	*Result*	*No.*	*Result*
1	409	26	397	51	404	76	404
2	400	27	407	52	406	77	401
3	406	28	401	53	407	78	404
4	399	29	399	54	405	79	408
5	402	30	401	55	411	80	406
6	406	31	403	56	410	81	408
7	401	32	400	57	410	82	406
8	403	33	410	58	410	83	401
9	401	34	401	59	401	84	412
10	403	35	407	60	402	85	393
11	398	36	423	61	404	86	437
12	403	37	406	62	405	87	418
13	407	38	406	63	392	88	415
14	402	39	402	64	407	89	404
15	401	40	405	65	406	90	401
16	399	41	405	66	404	91	401
17	400	42	409	67	403	92	407
18	401	43	399	68	408	93	412
19	405	44	402	69	404	94	373
20	402	45	407	70	407	95	409
21	408	46	406	71	412	96	406
22	399	47	413	72	406	97	398
23	399	48	409	73	409	98	406
24	402	49	404	74	400	99	403
25	399	50	402	75	408	100	404

Figure 2. Outliers. The top panel shows the histogram for all 100 measurements on NB 10; a normal curve is drawn for comparison. The curve does not fit well. The second panel shows the data with 3 outliers removed. The curve fits better. Most of the data follow the normal curve, but a few measurements are much further away from average than the curve suggests.



Removing Outliers:

$\bar{x} = 404, \quad SD = 4$

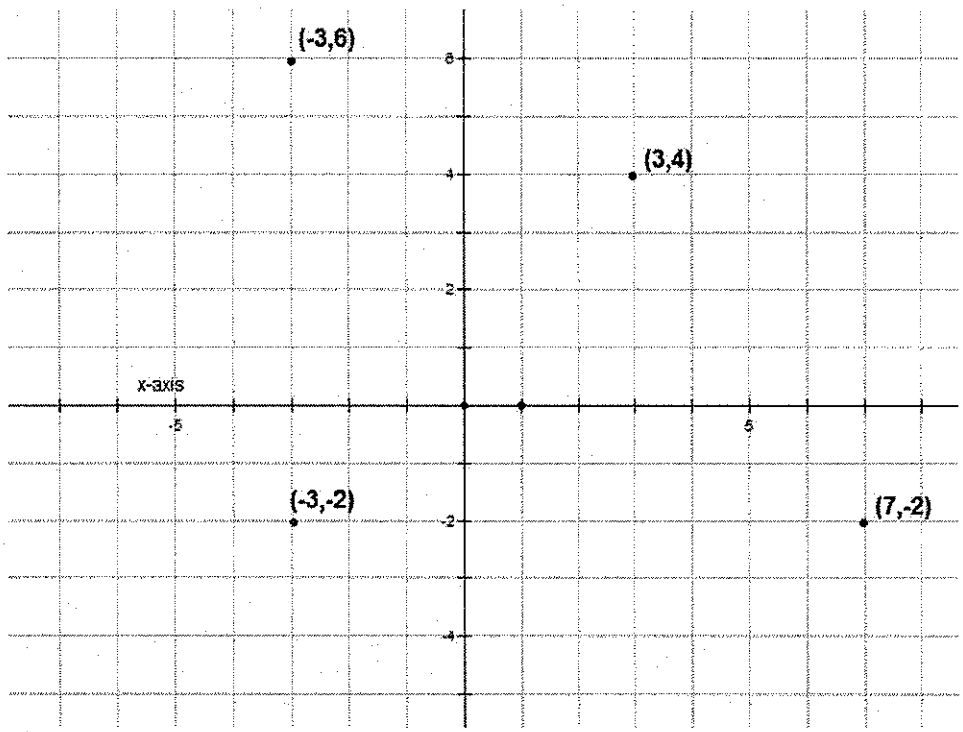
Table 1. One hundred measurements on NB 10. Almer and Jones, National Bureau of Standards. Units are micrograms below 10 grams.

No.	*Result*	*No.*	*Result*	*No.*	*Result*	*No.*	*Result*
1	409	26	397	51	404	76	404
2	400	27	407	52	406	77	401
3	406	28	401	53	407	78	404
4	399	29	399	54	405	79	408
5	402	30	401	55	411	80	406
6	406	31	403	56	410	81	408
7	401	32	400	57	410	82	406
8	403	33	410	58	410	83	401
9	401	34	401	59	401	84	412
10	403	35	407	60	402	85	393
11	398	36	423	61	404	86	437
12	403	37	406	62	405	87	418
13	407	38	406	63	392	88	415
14	402	39	402	64	407	89	404
15	401	40	405	65	406	90	401
16	399	41	405	66	404	91	401
17	400	42	409	67	403	92	407
18	401	43	399	68	408	93	412
19	405	44	402	69	404	94	375
20	402	45	407	70	407	95	409
21	408	46	406	71	412	96	406
22	399	47	413	72	406	97	398
23	399	48	409	73	409	98	406
24	402	49	404	74	400	99	403
25	399	50	402	75	408	100	404

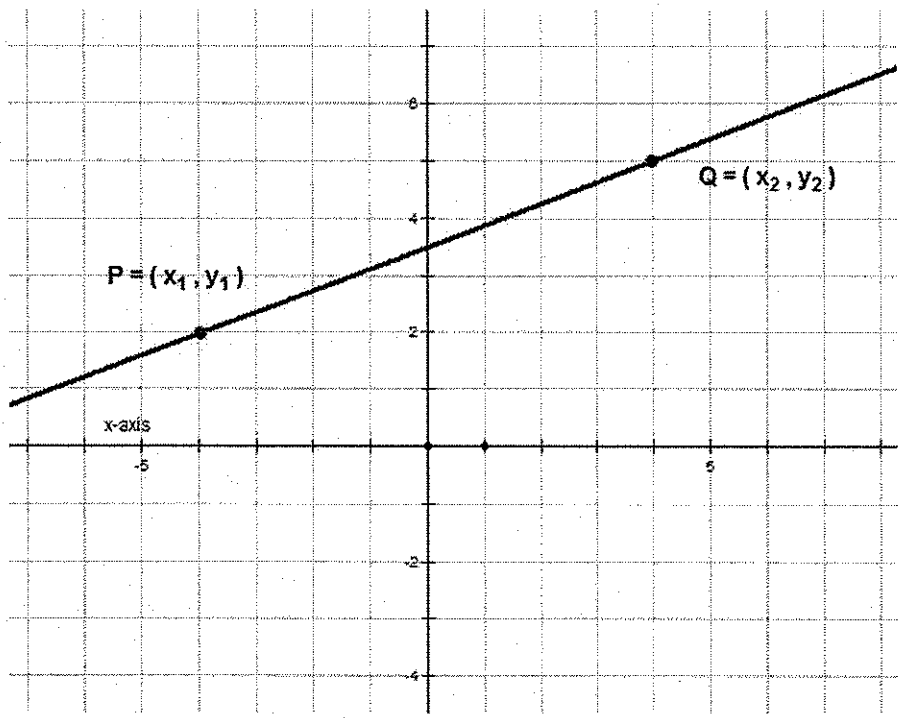
AV = 405 SD = 6

Review of Linear Equations:

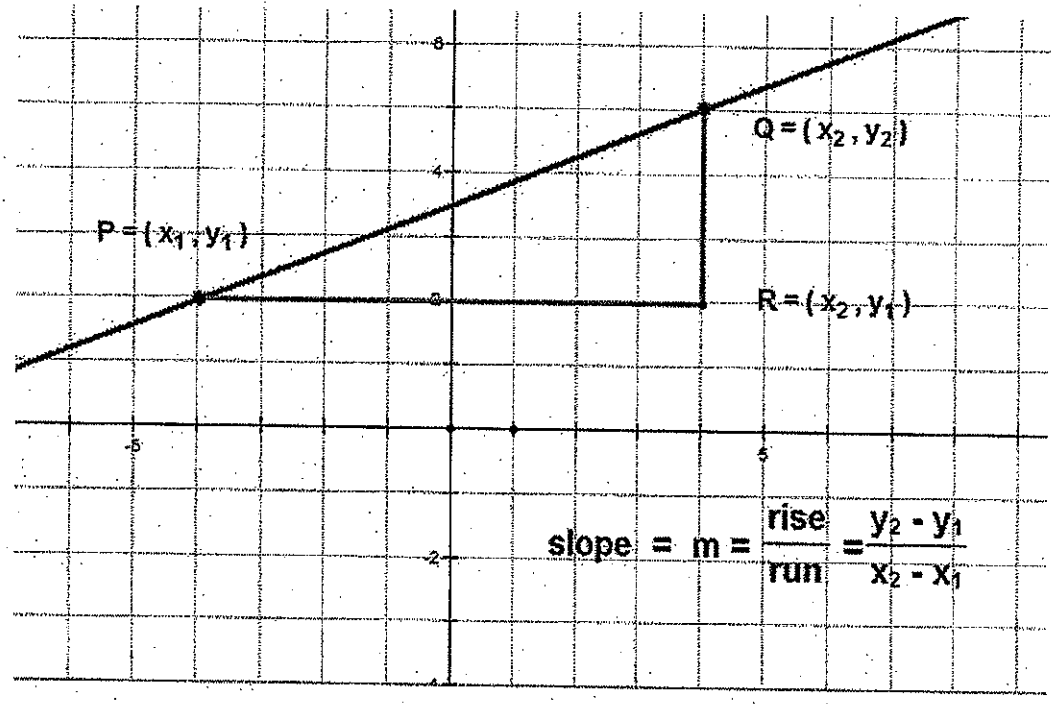
1. Cartesian Coordinate System

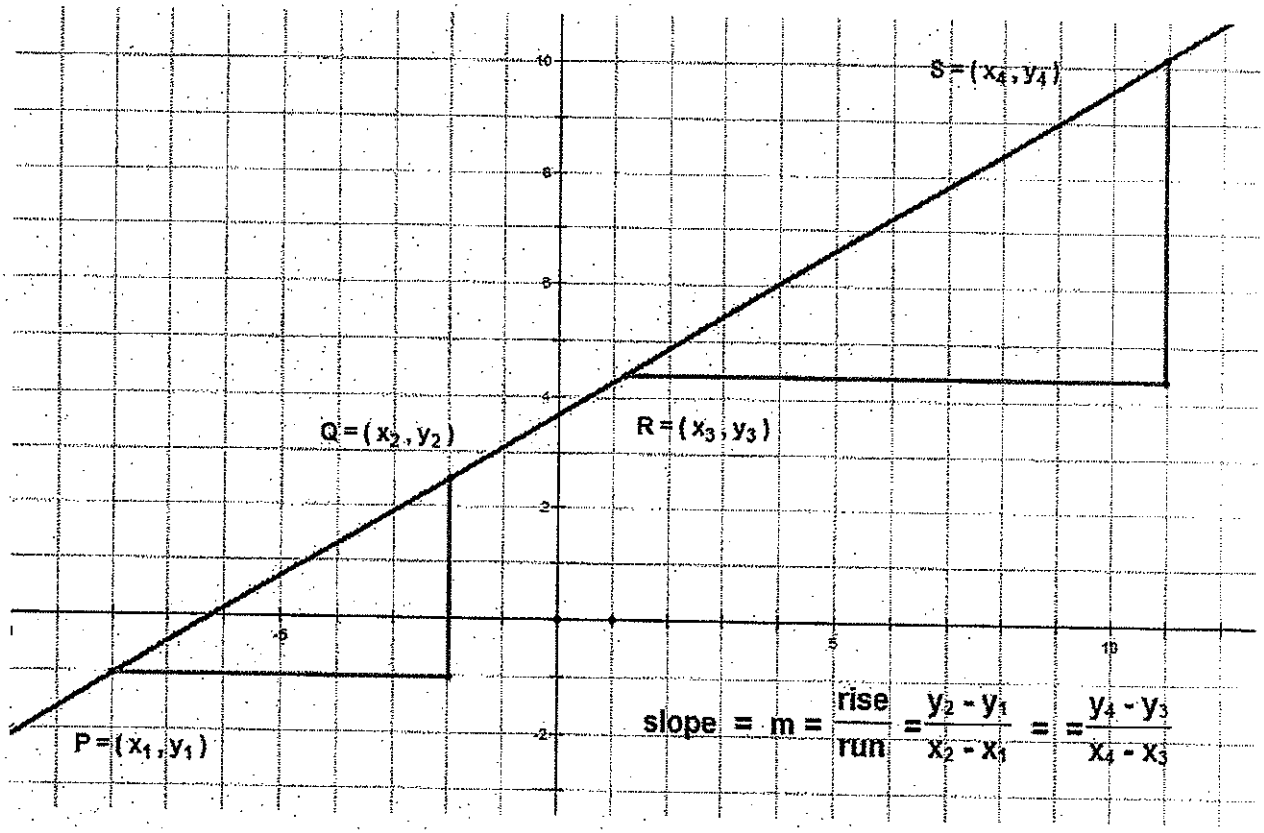


2. Two Points Determine a Line



3. Slope of a Line



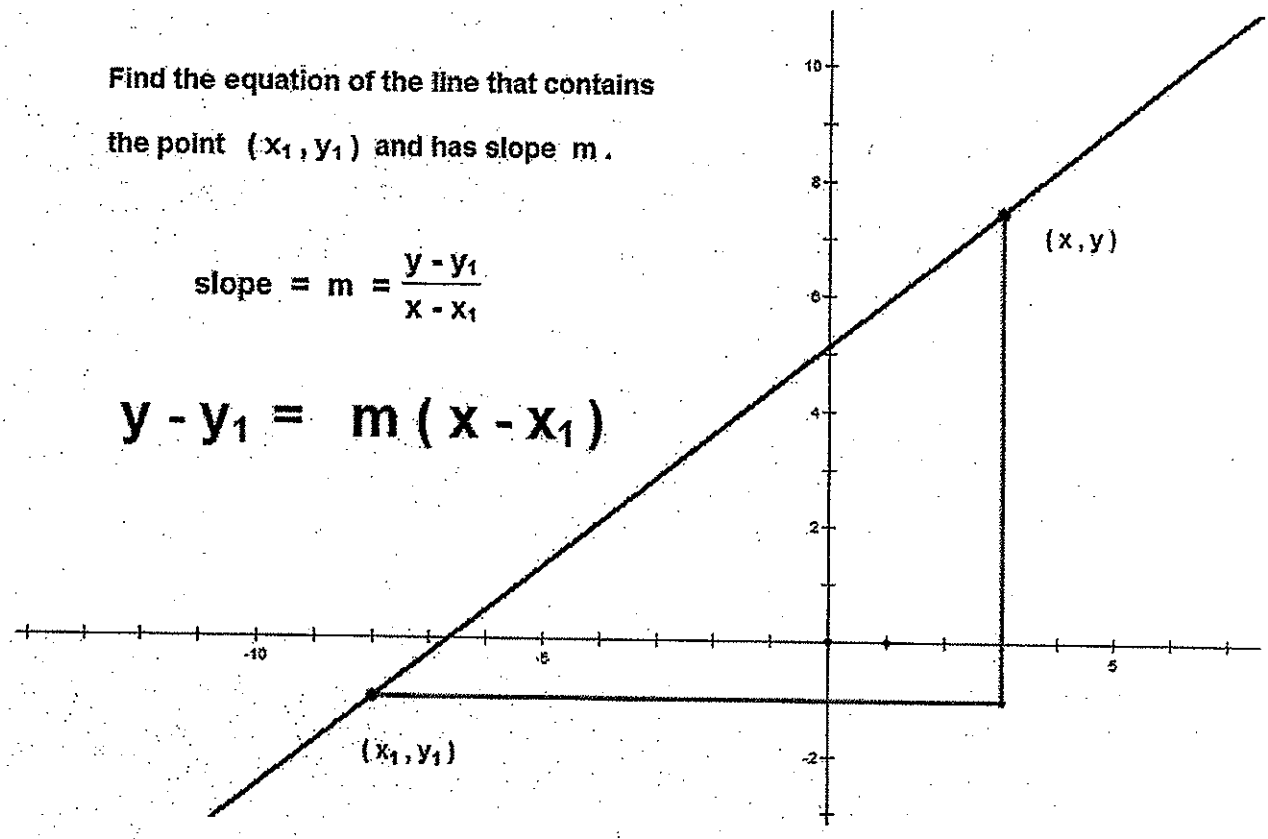


4. Equation of a Line: Point-Slope Form

Find the equation of the line that contains the point (x_1, y_1) and has slope m.

$$\text{slope} = m = \frac{y - y_1}{x - x_1}$$

$$y - y_1 = m(x - x_1)$$



5. Find the equation of the line containing the points (-3 , 5) and (1 , -7). Find the y-intercept and the x-intercept.