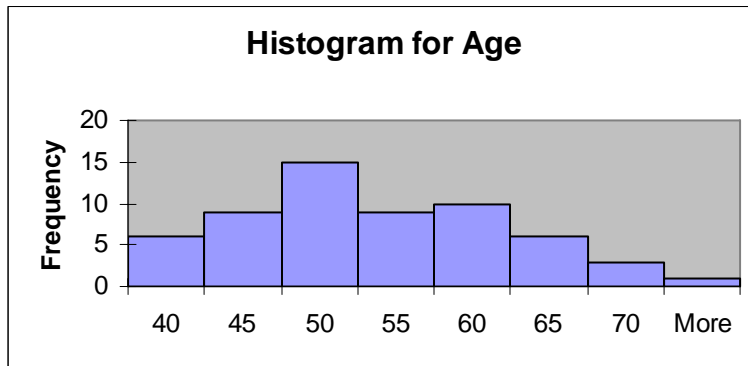
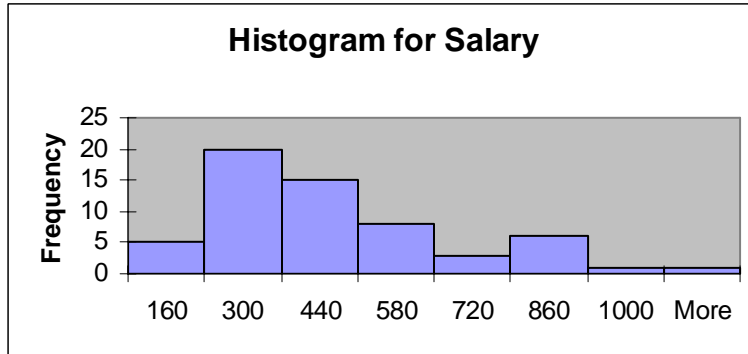


Answer Key Sample 2

1)
a)



From the graphs, we can see that both distributions are positively (right) skewed. However, distribution of age is more symmetrical. Both data sets seem to have outliers. Also there is more variation in Salary.

b)

AGE		SALARY	
Mean	51.5424	Mean	404.1695
Standard Error	1.1690	Standard Error	28.7110
Median	50	Median	350
Mode	50	Mode	291
Standard Deviation	8.9795	Standard Deviation	220.5335
Sample Variance	80.6318	Sample Variance	48635.0397
Kurtosis	0.0606	Kurtosis	0.6585
Skewness	0.1756	Skewness	0.9397
Range	42	Range	1082
Minimum	32	Minimum	21
Maximum	74	Maximum	1103
Sum	3041	Sum	23846

There is more skewness in “Salary”. “Age” is more symmetrical, mean is equal to median for age. Also by comparing coefficient of variation we can see that there is more dispersion in “Salary”. Also “Salary” has small and large outliers.

c) Best measure of center for salary is the median, because mean is inflated by large outliers. Median is better than mean, because it is not sensitive to outliers.

d)

	AGE	SAL
AGE	79.26515	
SAL	248.3149	47810.72

Covariance is 248.3149. This shows that both variables move or covary in the same direction. We can also say that there is a positive relationship or positive correlation between these two variables.

e)

Lower Limit	Upper Limit	Frequency, fi	Midpoint, Xi	fi*Xi	Xi-mu	(Xi-mu)^2	fi*(Xi-mu)^2
20	160	5	90	450	-306.1017	93698.2476	468491.2381
160	300	20	230	4600	-166.1017	27589.7731	551795.4611
300	440	15	370	5550	-26.1017	681.2985	10219.4772
440	580	8	510	4080	113.8983	12972.8239	103782.5912
580	720	3	650	1950	253.8983	64464.3493	193393.0480
720	860	6	790	4740	393.8983	155155.8747	930935.2485
860	1000	1	930	930	533.8983	285047.4002	285047.4002
1000	1140	1	1070	1070	673.8983	454138.9256	454138.9256
Total		59		23370			2997803.3898
				mean=	396.1017		
					23370/59		
						Var=	51686.2269 (=2997803.39/58)
						Std. Dev.=	225.4112

Mean = 396.1017

Median class is one where the $(59+1)/2 = 30^{\text{th}}$ observation lives. If we look at the frequency column, 30^{th} observation lies within the class 300-440. So this is the median class.

Modal class is the one with the highest frequency, i.e., 160-300.

f) Standard deviation is 225.4112 (See above)

We applied the formula for standard deviation of the grouped data.

2) 

a)

Since there are 3 possible answers for each question, the probability of guessing an answer correctly is $1/3$, i.e., $\pi=1/3$.

$$= \text{binomdist}(10,20,1/3,0) = 0.054$$

b)

At least $20 \cdot (60/100) = 12$ questions should be answered correctly to solve the exam. We are looking for the probability that the student guesses 12 or more questions correctly.

$$= 1 - \text{binomdist}(11,20,1/3,1)$$

c) $20 \cdot (25/100) = 5$

$$= 1 - \text{binomdist}(4,20,1/3,1) = 0.848$$

d)

$$= \text{binomdist}(10,20,1/3,1) = 0.962 \text{ or}$$

$$= 1 - \text{binomdist}(9,20,2/3,1) = 0.962$$

e)

$$E(X) = n \cdot \pi = 20 \cdot (1/3) = 6.667$$

$$\text{Var}(X) = n \cdot \pi \cdot (1 - \pi) = 20 \cdot (1/3) \cdot (1 - 1/3) = 4.444$$

$$\text{Std.Dev.} = \sqrt{\text{variance}} = 2.108$$

3)

a)

Interest in Finance	Ability in Mathematics			Total
	Low	Average	High	
Low	60	15	20	95
Average	15	35	10	60
High	10	10	25	45
Total	85	60	55	200

b) $45/200$

c) We use the following rule

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

$$60/200 + 60/200 - 35/200 = 85/200$$

d) We just look at the “Average Ability in Mathematics” column.

$$(15+10)/60$$

e) We look at “Average and High Ability in Mathematics” columns.

$$(15+20)/(60+55)=35/115$$