Name: _____ Period: _____ Date: _____

AP STATISTICS SUMMER ASSIGNMENT

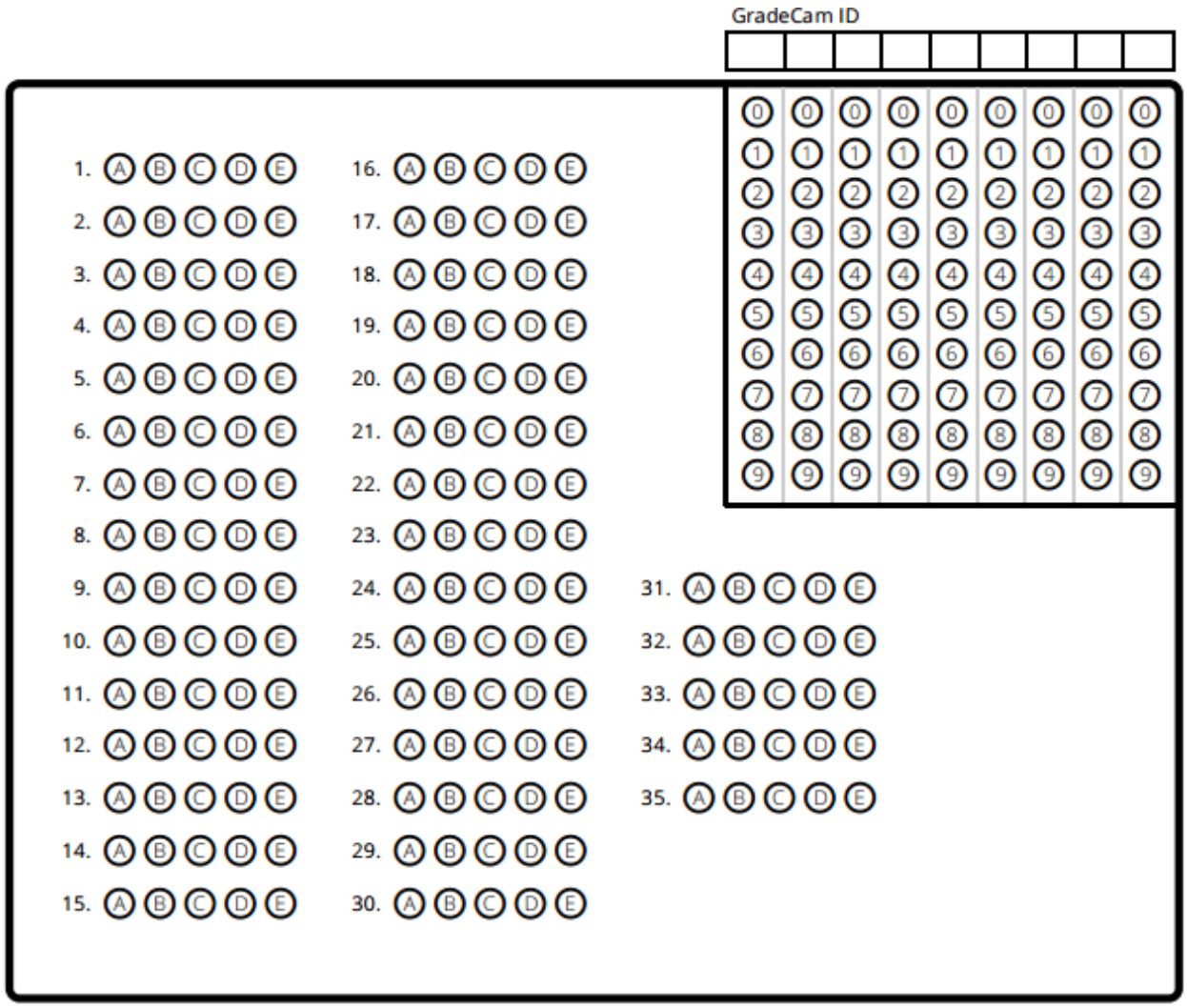
You have selected AP Statistics, a college level course, for your 2015-16 school year. In order to properly prepare you for the AP Statistics Exam in May 2016 we will aggressively work both independently and together as a group. Your first independent task is outlined below. This assignment in combination with an exam given within the first ten school days will be part of your MP1 grade.

Any issues with the assignment may be communicated to me via email; rieraci2@schools.nyc.gov or rieraci@taehs.org. Additionally, a Google classroom has been created (access code yv7oru2) where you can access the materials.

Step 1: Read chapters 1 & 2 in *Probability and Statistics - Advanced* Second Edition by Lawsky, Ottman, Almukkahal, Meery and DeLancy. This on-line textbook may be found @ http://www.ck12.org/book/CK-12-Probability-and-Statistics-Advanced-Second-Edition/

Step 2: Answer the questions in the attached packet, indicating your answers on the answer sheet below (pen or pencil may be used – your OSIS # is your student ID).

Step 3: Review the material for the first week of school.



Name: _____ Class: _____ Date: _____ ID: A

AP STATISTICS SUMMER ASSIGNMENT

Multiple Choice
Identify the choice that best completes the statement or answers the question.

1. Data may be classified by one of the four levels of measurement. What is the name of the lowest level?
   A) simple B) ordinal C) ratio D) nominal E) interval

2. Identify the variable in the information below.
   What is the average miles per gallon (mpg) for all new cars? Using Consumer Reports, a random sample of 35 new cars gave an average of 21.1 mpg.
   A) cars B) miles C) number of new cars surveyed D) gallons E) miles per gallon

3. Identify the implied population in the information below.
   Government agencies carefully monitor water quality and its effect on wetlands (Reference: Environment Protection Agency Wetland Report EPA 832-R-93-005). Of particular concern is the concentration of nitrogen in water draining from fertilized lands. Too much nitrogen can kill fish and wildlife. Twenty-eight samples of water were taken at random from a lake. The nitrogen concentration (milligrams of nitrogen per liter of water) was determined for each sample. The variable in this information is nitrogen concentration (mg nitrogen/l water).
   A) nitrogen concentration (mg nitrogen/l water) in the entire lake B) samples of water taken at random
   C) samples of water taken at random D) none of these choices E) number of fish

4. Identify whether the variable in the information below is qualitative or quantitative.
   The archeological site of Tara is more than 4000 years old. Tradition states that Tara was the seat of the high kings of Ireland. Because of its archeological importance, Tara has received extensive study (Reference: Tara: An Archeological survey by Conor Newman, Royal Irish Academy, Dublin). Suppose an archeologist wants to estimate the density of ferromagnetic artifacts in the Tara region. For this purpose, a random sample of 55 plots, each of size 100 square meters, is used. The number of ferromagnetic artifacts for each plot is determined.
   A) neither qualitative nor quantitative B) quantitative C) Information does not have any variable.
   D) qualitative E) qualitative as well as quantitative

5. Identify the level of measurement corresponding to the data "Gross income for each of the past 5 years" associated with robotics.
   A) nominal B) interval C) ratio D) none of these choices E) ordinal

6. Identify the sampling technique used in the following information.
   An important part of employee compensation is a benefits package that might include health insurance, life insurance, child care, vacation days, retirement plan, parental leave, bonuses, etc. Suppose you want to conduct a survey of benefits packages available in private businesses in Hawaii. You want a sample size of 100. Sampling technique used to get the sample size of 100 is described below.
   Use the Island Business Directory. Number all the businesses. Select a starting place at random, and then use every 50th business listed until you have 100 businesses.
   A) Convenience sample B) Cluster sample C) Systematic sample D) Simple random sample
   E) Stratified sample

7. Find the technique for gathering data in the study below.
   An ecology class used binoculars to watch 23 turtles at Lowell Pond. It was found that 18 were box turtles, and 5 were snapping turtles.
   A) experiment B) none of these choices C) census D) observational study E) sampling

8. Find the technique for gathering data in the study below.
   A study of all league football scores attained through touchdowns and field goals was conduced by the National Football League to determine whether field goals account for more scoring events than touchdowns (USA Today).
   A) experiment B) census C) sampling D) simulation E) none of these choices

9. Finish times (to the nearest hour) for 10 dogsled teams are shown below.
   Find the class width. Use five classes. (Round your answer to the nearest integer.)

   234 271 339 361 354 263 236 290 315 254

   A) 121 B) 119 C) 26 D) 25 E) 27

10. Finish times (to the nearest hour) for 10 dogsled teams are shown below.
Make a frequency table showing class limits, class boundaries, midpoints, frequency, relative frequencies, and cumulative frequencies. Use three classes. (Round your answer for relative frequency to the nearest hundredth and for midpoint to the nearest tenth.)

236 263 273 283 239 280 270 310 259 310

A)

Class Limits	Boundaries	Midpoint	Freq.	Relative Freq.	Cumulative Freq.
236 – 260	235.5 – 260.5	248.0	4	0.30	3
261 – 284	260.5 – 285.5	273.0	7	0.50	8
285 – 310	285.5 – 310.5	296.5	3	0.20	10

B)

Class Limits	Boundaries	Midpoint	Freq.	Relative Freq.	Cumulative Freq.
236 – 260	235.5 – 260.5	248.0	3	0.30	3
261 – 285	260.5 – 285.5	273.0	5	0.50	8
286 – 310	285.5 – 310.5	298.0	2	0.20	10

C)

Class Limits	Boundaries	Midpoint	Freq.	Relative Freq.	Cumulative Freq.
236 – 260	235.5 – 260.5	248.0	3	0.30	3
261 – 284	260.5 – 284.5	272.5	5	0.50	8
285 – 308	284.5 – 308.5	296.5	2	0.20	10

D)

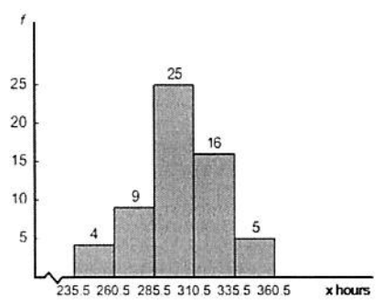
Class Limits	Boundaries	Midpoint	Freq.	Relative Freq.	Cumulative Freq.
236 – 260	235.5 – 260.5	248.0	4	0.30	3
261 – 284	260.5 – 284.5	272.5	7	0.50	8
285 – 308	284.5 – 308.5	296.5	3	0.20	10

E) none of these choices

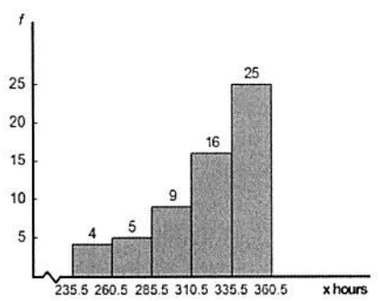
11. Finish times (to the nearest hour) for 59 dogsled teams are shown below. Draw a histogram. Use five classes.

261	269	236	244	280	296	284	299	288	288	249	256
338	360	341	333	261	266	287	296	313	311	306	306
299	303	277	283	304	305	288	290	288	289	297	299
332	330	309	327	306	327	285	291	295	298	306	315
310	318	318	320	333	321	323	324	327	239	358	

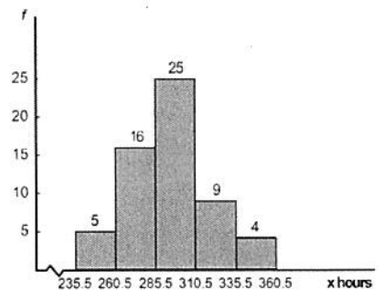
A)



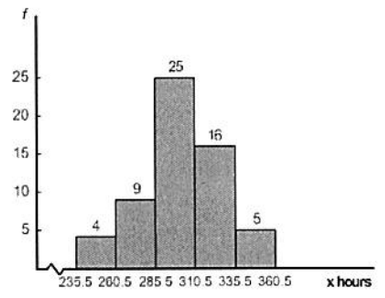
B)



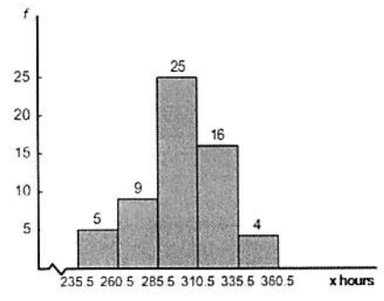
C)



D)



E)



12. Finish times (to the nearest hour) for 57 dogsled teams are shown below. Use five classes. Categorize the basic distribution shape as uniform or rectangular, mound-shaped symmetric, bimodal, skewed left, or skewed right.

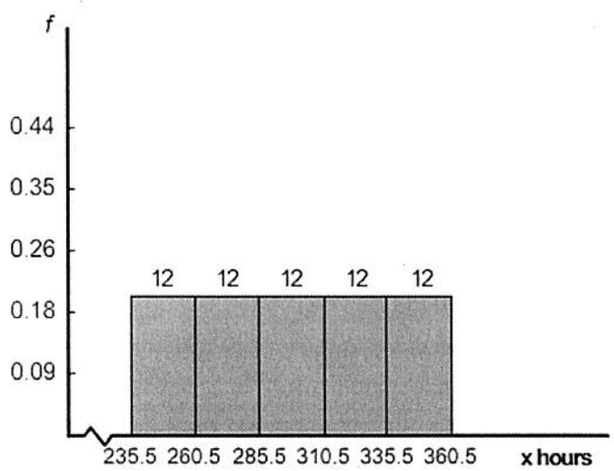
261	271	236	244	279	296	284	299	288	288	247	256
338	360	341	333	261	266	287	296	313	311	307	307
299	303	277	283	304	305	288	290	288	289	297	299
332	330	309	328	307	328	285	291	295	298	306	315
310	318	318	320	333	321	323	324	327			

A) approximately bimodal B) approximately mound-shaped symmetric C) approximately skewed left
D) approximately skewed right E) approximately uniform or rectangular

13. Finish times (to the nearest hour) for 60 dogsled teams are shown below. Use five classes. Categorize the basic distribution shape as uniform, mound-shaped symmetric, bimodal, skewed left, or skewed right.

261	271	236	244	279	296	284	299	288	288	247	256
338	360	341	333	261	266	287	296	313	311	307	307
279	283	277	283	285	275	259	239	288	289	297	299
341	358	257	328	244	328	245	258	259	259	319	315
339	359	347	348	333	321	323	324	327	349	351	355

The relative frequency histogram of the above data is given below.



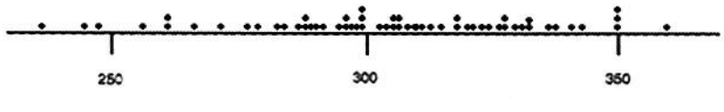
A) Skewed right B) none of these choices C) Uniform or rectangular D) mound-shaped symmetric
E) Bimodal

14. Finish times (to the nearest hour) for 57 dogsled teams are shown below. Make a dotplot for the data.

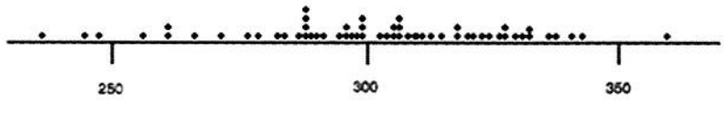
261	271	236	244	279	296	284	299	288	288	247	256
308	360	341	333	261	266	287	296	313	311	307	307
299	303	277	283	304	305	288	290	288	289	297	299
332	330	309	328	307	328	285	291	295	298	306	315
310	318	318	318	333	321	323	324	327			

A)

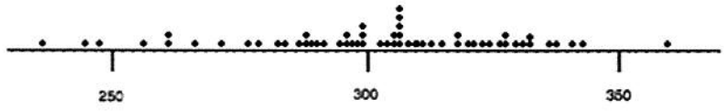
B)



C)



D)



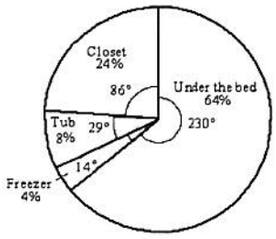
E) none of these choices

15. At Westgate Community College, a survey was done to determine when students are available for class. A questionnaire was given to a random sample of students. The instructions were to mark each of the time categories in which they could take classes. Many students marked more than one category. Responses from the students in the sample indicated that 52 would take early morning classes, 85 would take mid-morning classes, 41 would take afternoon classes and 37 would take evening classes. Would a circle graph be appropriate for this data? Give a reason for your answer.

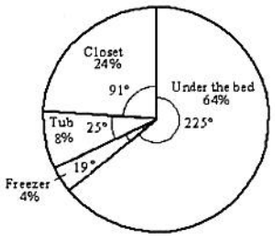
A) Yes. The categories represent all possible responses. B) No. There are too few categories for a circle graph to be useful. C) Yes. Each category represents a percentage of the total student population that could attend class at a certain time. D) No. Since there were multiple responses from some students this data does not represent parts of a whole. E) Yes. Circle graphs are most effective when the number of wedges is 10 or fewer.

16. A survey of 1000 adults uncovered some interesting housekeeping secrets. When unexpected company comes, where do we hide the mess? The survey showed that 64% of the adults toss their mess in the closet, 24% shove things under bed, 8% put things in the bathtub, and 4% put the mess in the freezer. Make a circle graph to display this information.

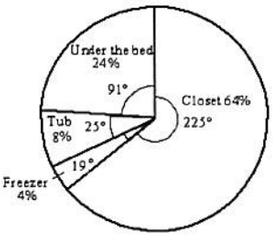
A)



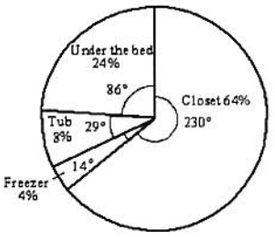
B)



C)



D)



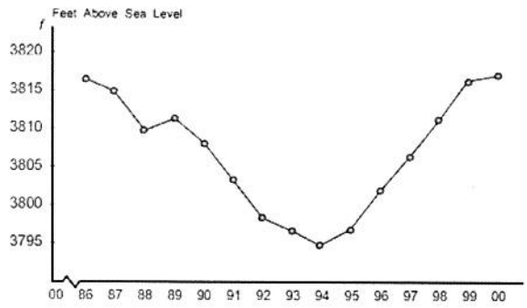
E) none of these choices

17. Pyramid Lake, Nevada, is described as the pride of the Paiute Indian Nation. It is a beautiful desert lake famous for very large trout. The elevation of the lake surface (feet above sea level) varies according to the annual flow of the Truckee River from Lake Tahoe. Assume that the U.S. Geological Survey provided the following data:

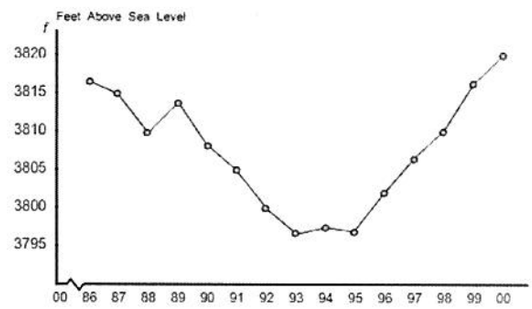
Year	1986	1987	1988	1989	1990	1991	1992	1993
Elevation	3817	3815	3810	3812	3808	3803	3798	3797
Year	1994	1995	1996	1997	1998	1999	2000	
Elevation	3795	3797	3802	3807	3811	3816	3817	

Make a time series graph.

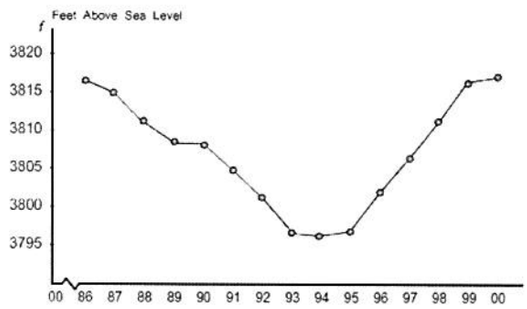
A)



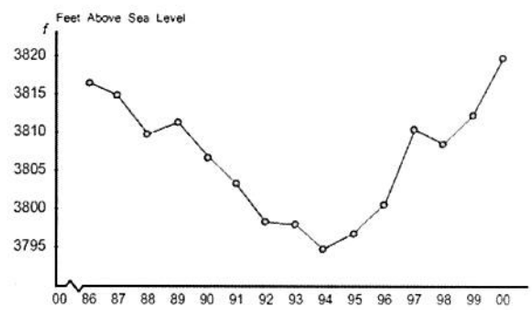
B)



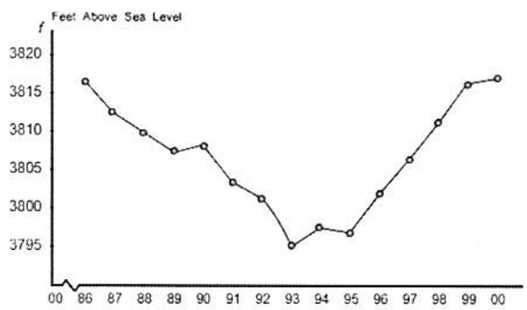
C)



D)



E)



18. Suppose the American Medical Association Center for Health Policy Research included data, by state, on the number of community hospitals and the average patient stay (in days) in its publication. The data (by state) are shown in the table.
Which two states have an unusually high number of hospitals?

State	Hospitals	State	Hospitals	State	Hospitals
Alabama	330	Colorado	79	Georgia	162
Alaska	16	Connecticut	35	Hawaii	19
Arizona	61	Delaware	8	Idaho	41
Arkansas	88	Dist. of Columbia	11	Illinois	279
California	236	Florida	289	Indiana	113
Iowa	123	Nebraska	90	Rhode Island	12
Kansas	133	Nebraska	21	S.Carolina	68
Kentucky	101	New Hampshire	28	S.Dakota	52
Louisiana	459	New Jersey	96	Tennessee	122
Maine	38	New Mexico	37	Texas	235
Maryland	51	New York	333	Utah	42
Mass.	101	N.Caroline	117	Vermont	15
Michigan	175	N.Dakota	47	Virginia	98
Minnesota	276	Ohio	193	Washington	92
Mississippi	102	Oklahoma	399	W.Virginia	59
Missouri	133	Oregon	66	Wisconsin	478
Montana	53	Pennsylvania	231	Wyoming	27

A) Wisconsin and Louisiana B) Maine and Iowa C) Alabama and Arkansas D) Florida and Wisconsin
E) none of these choices

19. Use the data given in the following table to make a stem – and – leaf display for milligrams of nicotine per cigarette smoked. In this case, truncate the measurements at the tenths position and use two lines per stem.

Brand		Brand	
Alpine	0.82	Multifilter	0.78
Benson & Hedges	1.11	Newport Lights	0.73
Bull Durham	2.07	Now	0.24
Camel Lights	0.67	Old Gold	1.26
Carlton	0.38	Pall Mall Lights	1.08
Chesterfield	1.04	Raleigh	0.92
Golden Lights	0.76	Salem Ultra	0.42
Kent	0.95	Tareyton	1.01
Kool	1.19	True	0.61
L&M	1.02	Viceroy Rich Light	0.69
Lark Lights	1.01	Virginia Slim	1.02
Marlboro	0.90	Winston Lights	0.82
Merit	0.57		

A) 0 2 = 0.2 milligram
0 4 2 3 7
0 6 6 6 7 7 8 8 9 9 9
1 0 0 0 0 0 1 1 1 2
2 0

B) 0 2 = 0.2 milligram
0 2 3 4
0 6 6 6 7 7 7 8 8 9 9 9
1 0 0 0 0 0 0 0 1 1 2
2 0

C) 0 2 = 0.2 milligram
0 4 2 3 6 7
0 7 6 6 8 7 7 8 8 9 9 9
1 0 0 0 0 0 1 1 1 2
2 0

D) 0 2 = 0.2 milligram
0 4 2 3
0 7 6 6 6 7 8 8 9 9 9
1 0 0 0 0 0 1 1 1 2
2 0

E) none of these choices

20. How hot does it get in Death Valley? Assume that the following data are taken from a study conducted by the National Park System, of which Death Valley is a unit. The ground temperatures ($°F$) were taken from May to November in the vicinity of Furnace Creek. Compute the median for these ground temperatures. Round your answer to the nearest tenth.

148	153	167	172	196	178	193
196	178	178	167	162	153	145

A) 194.5 B) 169.5 C) 167.0 D) 160.0 E) 339.0

21. The Grand Canyon and the Colorado River are beautiful, rugged, and sometimes dangerous. Assume there is a physician at the park clinic in Grand Canyon Village. Suppose the physician has recorded (for a 5-year period) the number of visitor injuries at different landing points for commercial boat trips down the Colorado River in both the Upper and Lower Grand Canyon.

Upper Canyon: Number of injuries per Landing Point Between North Canyon and Phantom Ranch
4 5 3 3 5 6 7 10 5 3 5

Lower Canyon: Number of injuries per Landing Point Between Bright Angel and Lava Falls
9 3 3 0 7 8 4 12 5 0 3 11 4 3

The mean, median, and mode for Upper Canyon are 5.091, 5.0, and 5, respectively.
The mean, median, and mode for Lower Canyon are 5.214, 4.0, and 3, respectively.
Compare the mean, median, and mode found in Upper Canyon and Lower Canyon, respectively.
A) Lower Canyon mean is greater than Upper Canyon, Lower Canyon median is smaller than Upper Canyon, and Lower Canyon mode is smaller than Upper Canyon. B) Lower Canyon mean is greater than Upper Canyon, Lower Canyon median is greater than Upper Canyon, and Lower Canyon mode is greater than Upper Canyon. C) Lower Canyon mean is smaller than Upper Canyon, Lower Canyon median is smaller than Upper Canyon, and Lower Canyon mode is smaller than Upper Canyon. D) Lower Canyon mean is greater than Upper Canyon, Lower Canyon median is greater than Upper Canyon, and Lower Canyon mode is smaller than Upper Canyon. E) Lower Canyon mean is greater than Upper Canyon, Lower Canyon median is smaller than Upper Canyon, and Lower Canyon mode is greater than Upper Canyon.

22. In your biology class, your final grade is based on several things: a lab score, score on two major tests, and your score on the final exam. There are 100 points available for each score. However, the lab score is worth 20% of your total grade, each major test is worth 27.5%, and the final exam is worth 25%. Compute the weighted average for the following scores: 91 on the lab, 83 on the first major test, 92 on the second major test, and 84 on the final exam. Round your answer to the nearest hundredth.
A) 84.85 B) 85.93 C) 89.08 D) 87.68 E) 87.33

23. Find the range for the following sample data.

x:	24	18	15	31	29

A) 16 B) 11 C) 2 D) 7 E) 3

24. Find the sample standard deviation s for the following sample data. Round your answer to the nearest hundredth.

x:	24	17	15	33	25

A) 9.58 B) 8.57 C) 7.16 D) 6.40 E) 9.60

25. Do bonds reduce the overall risk of an investment portfolio? Let x be a random variable representing annual percent return for the Vanguard Total Stock Index (all Stocks). Let y be a random variable representing annual return for the Vanguard Balanced Index (60% stock and 40% bond). For the past several years, assume the following data. Compute $\sum x$.

x:	14	0	36	24	31	26	27	−14	−14	−24
y:	8	−1	26	17	25	18	17	−1	−2	−8

A) 130 B) 134 C) 106 D) 75 E) 154

26. Do bonds reduce the overall risk of an investment portfolio? Let x be a random variable representing annual percent return for the Vanguard Total Stock Index (all Stocks). Let y be a random variable representing annual return for the Vanguard Balanced Index (60% stock and 40% bond). For the past several years, assume the following data. Compute $\sum y$.

x:	12	0	39	25	33	27	28	−12	−12	−25
y:	7	−1	29	15	26	18	15	−1	−2	−7

A) 107 B) 101 C) 99 D) 106 E) 100

27. Base your answer on the chart below and your knowledge of social studies.

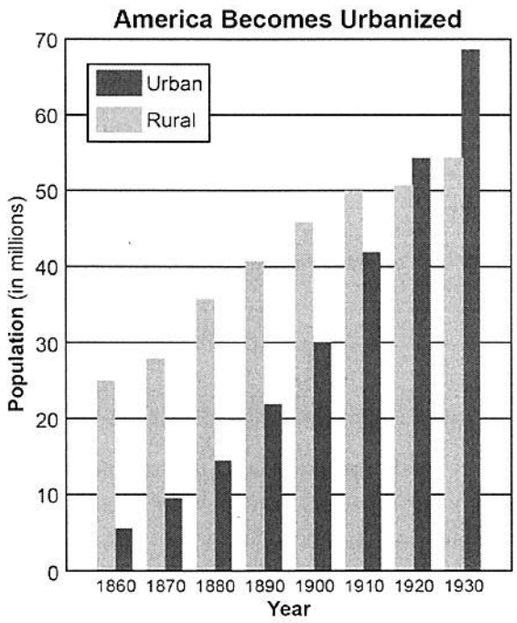
MEDIAN EARNINGS OF MEN AND WOMEN IN THE UNITED STATES, 1960–1990

Year	Women	Men	Women's Earnings as a Percent of Men's	Earnings Gap in Constant 1990 Dollars
1960	$ 3,257	$ 5,368	60.7	$ 8,569
1970	5,323	8,966	59.4	11,529
1980	11,197	18,612	60.2	11,776
1990	19,822	27,678	71.6	7,856

Source: Bureau of the Census

The data in this chart support the conclusion that between 1960 and 1990
A) government failed to pass laws that granted women equal access to jobs B) the earnings gap between men and women was only slightly improved C) women's earnings consistently increased faster than those of men
D) most higher paying jobs were still not legally open to women

28. Base your answer on the graph below and your knowledge of social studies.



Source: U. S. Census (adapted)

According to the graph, which was the first year in which more Americans lived in urban areas than in rural areas?
A) 1860 B) 1890 C) 1920 D) 1930

29. Base your answer on the chart below and your knowledge of social studies.

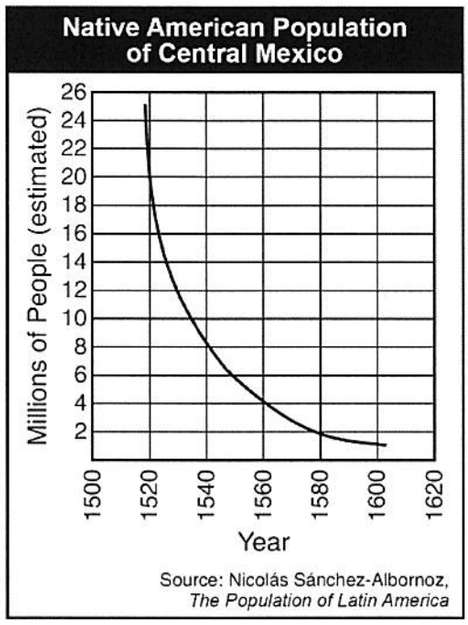
ESTIMATED PERCENTAGE OF VOTING-AGE AFRICAN AMERICANS REGISTERED IN 1965 AND 1988		
STATE	**March 1965**	**November 1988**
Alabama	19.3	68.4
Georgia	27.4	56.8
Louisiana	31.6	77.1
Mississippi	6.7	74.2
N. Carolina	46.8	58.2
S. Carolina	37.3	56.7
Virginia	38.3	63.8

Source: U.S. Department of Justice, Civil Rights Division (adapted)

Which state had the largest increase in the percentage of African Americans registered to vote between 1965 and 1988?

A) Alabama B) Georgia C) Louisiana D) Mississippi

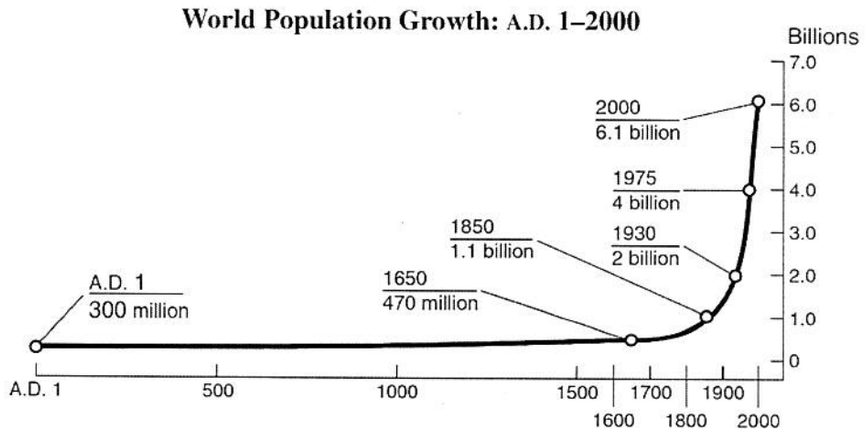
30. Base your answer on the chart below and your knowledge of social studies.



Between 1540 and 1580, what happened to the native population of Central Mexico?

A) It decreased by approximately 6 million people. B) It increased to a high of 26 million people. C) It increased by approximately 2 million people per year. D) It decreased to 8 million people.

31. Base your answer on the graph below and on your knowledge of social studies.



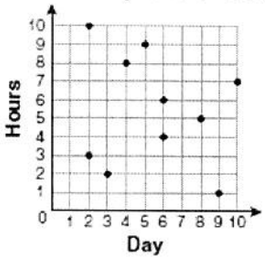
Source: Population Reference Bureau and de Blij and Murphy, *Human Geography: Culture, Society, and Space*, John Wiley & Sons, 1999 (adapted)

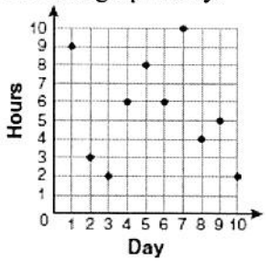
Which statement can be supported by the information in the graph?
A) The population of the world remained the same from A.D. 1 to A.D. 1650. B) Most of the world's population growth took place during the period from 1000 to 1500. C) The population growth rate decreased during the period from 1650 to 1800. D) The world's population tripled between 1930 and 2000.

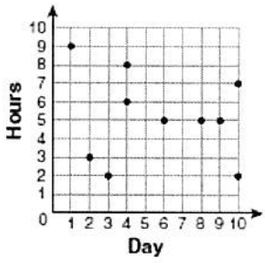
32. For 10 days, Romero kept a record of the number of hours he spent listening to music. The information is shown in the table below.

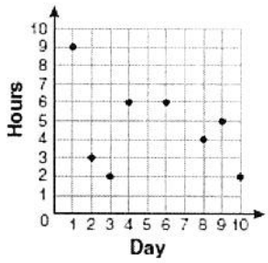
Day	1	2	3	4	5	6	7	8	9	10
Hours	9	3	2	6	8	6	10	4	5	2

Which scatter plot shows Romero's data graphically?

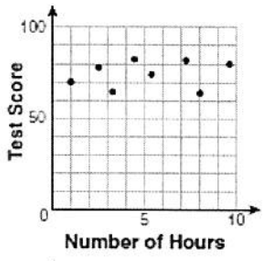
A) 

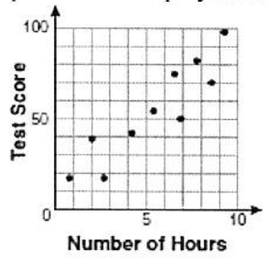
B) 

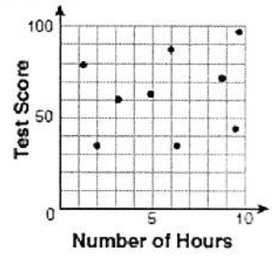
C) 

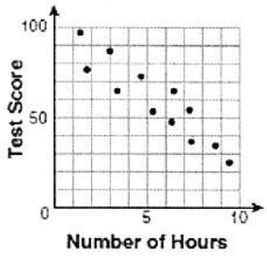
D) 

33. There is a negative correlation between the number of hours a student watches television and his or her social studies test score. Which scatter plot below displays this correlation?

A) 

B) 

C) 

D) 

34. Which statement is true about the data set 3, 4, 5, 6, 7, 7, 10?
A) mean = mode B) mean > mode C) mean = median D) mean < median

35. The mean of a normally distributed set of data is 52 and the standard deviation is 4. Approximately 95% of all the cases will lie between which measures?
A) 44 and 52 B) 44 and 60 C) 48 and 56 D) 52 and 64