Quiz 11.2A AP Statistics Name:

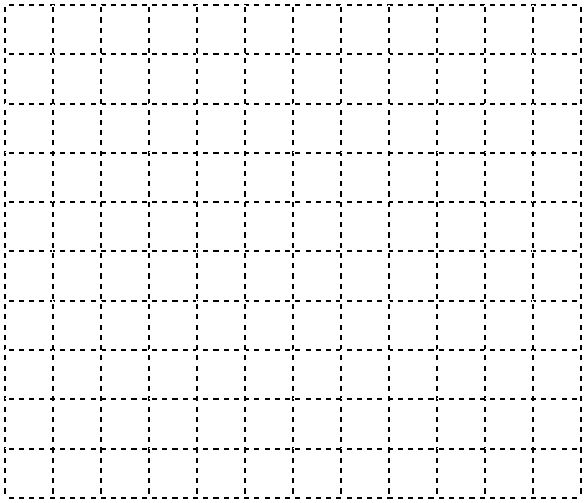
1. A radio station trying to determine what kind of music to play takes a simple random sample of 50 students at each of three locations: a local middle school, a high school, and a college. The students are asked to choose which of three different music genres they most enjoy hearing on the radio. Here are the results:

		Music Genre			
		Hip Hop	Alternative	Post-rock	Total
Age level	Middle School	28	18	4	50
	High School	22	22	6	50
	College	16	20	14	50
	Total	66	60	24	150

(a) In the table below, provide the appropriate conditional distributions based on the data collected for comparing the music-listening preferences of the three age levels, based on the data above.

		Music Genre		
		Hip Hop	Alternative	Post-rock
Age level	Middle School			
	High School			
	College			

(b) Make a graph that illustrates these conditional distributions effectively, and use the table from (a) and your graph to describe the relationship between age level and preferred music genre.



(c) Perform the appropriate statistical test to determine if there is a difference in the music preference of these three age groups.

(d) If you chose a chi-square test for homogeneity in part (c), explain how the data could have been obtained to make a chi-square test for independence appropriate. If you chose a test for independence, explain how the data could have been obtained to make a test for homogeneity appropriate.

Quiz 11.2B AP Statistics Name:

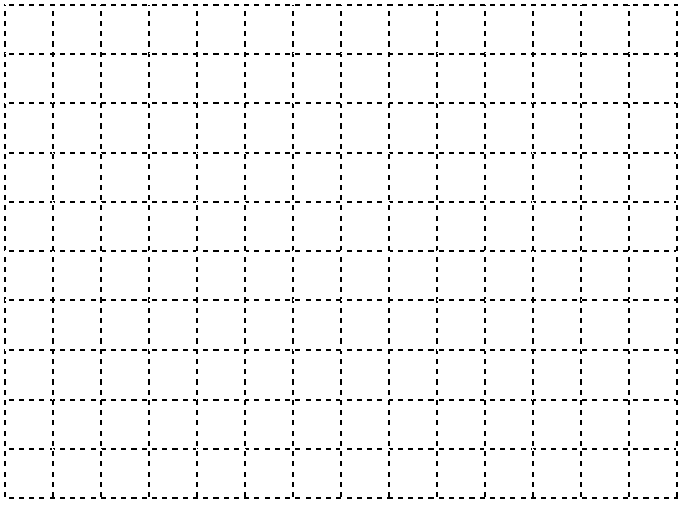
1. A bank wants to determine which age groups among its customers are using their on-line banking services. They select a simple random sample of customers from four age groups and ask each person if they do most of their banking on-line or at their local branch bank. Here are the results:

		Preferred banking method		
		Branch bank	On-line	Total
Age group (years)	Under 30	16	37	53
	31 – 45	18	38	56
	46 – 60	19	26	45
	Over 60	22	8	30
	Total	75	109	184

(a) In the table below, provide the appropriate conditional distributions for comparing the banking preferences of the four age groups, based on the data above.

		Preferred banking method	
		Branch bank	On-line
Age group (years)	Under 30		
	31 – 45		
	46 – 60		
	Over 60		

(b) Make a graph that illustrates these conditional distributions effectively, and use the table from (a) and your graph to describe the relationship between age group and preferred banking method.



(c) Perform the appropriate statistical test to determine if there is a difference in the banking preferences of these four age groups.

(d) If you chose a chi-square test for homogeneity in part (c), explain how the data could have been obtained to make a chi-square test for independence appropriate. If you chose a test for independence, explain how the data could have been obtained to make a test for homogeneity appropriate.

Quiz 11.2C AP Statistics Name:

1. Have older cell-phone users caught on to texting? Simple random samples of people under 30 and over 30 were asked whether they made more phone calls or sent more text messages on their cell phone on a typical day. Here are the results:

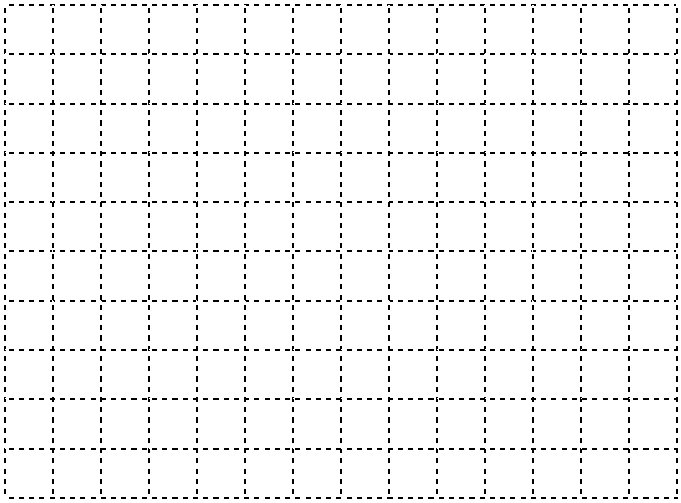
	Age group	
	Under 30	Over 30
More texts	146	32
More phone calls	34	98

Give the name of the appropriate chi-square procedure and state the null and alternative hypotheses for the test that will address this question. **Do not perform the test!**

2. A person studying fathers' involvement in their children's education interviews a simple random sample of fathers of school-age children. One question concerns regularly scheduled parent-teacher conferences. Here is a two-way table of the results:

	Attended all	Attended some	Attended none	Total
Fathers in two-parent families	109	132	203	444
Fathers in single-parent families	15	10	13	38
Non-resident fathers	11	5	82	98
Total	135	147	298	580

(a) Use appropriate graphical techniques to illustrate the relationship between these two variables.



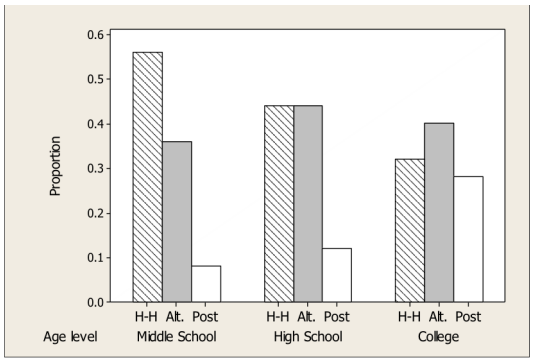
(b) Based on your graphical analysis, discuss the relationship between family type and how often fathers attended parent-teacher conferences. Support your conclusions with the appropriate statistical test.

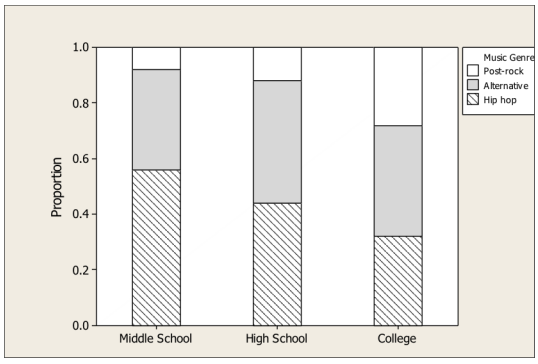
Quiz 11.2A

1. (a) & (b) Since age level is the explanatory variable, we calculate conditional distributions for each age level. See table and graph below (two options for graphs are shown: parallel and segmented bar graphs.) The data suggest that middle school students strongly prefer Hip Hop, most high school students are evenly divided between Hip Hop and Alternative, and college students are more evenly divided between all three genres. (c) State: We are testing the hypothesis H_o: The distribution of music genre is the same for the populations in all three age groups, against H_a: The distribution of music genre is not the same for the populations in all three age groups. We will use a significance level of $\alpha = 0.05$. Plan: The procedure is a chi-square test for homogeneity. Conditions: *Random*: the data come from SRSs of each population. *10%:* The radio station's listening audience is likely to be more than 10 times the sample size. *Large counts:* All expected counts are at least 5 (see expected counts table below). Do: Using a calculator, $\chi^2 = 10.673$; $df = 4$; $P - \text{value} = 0.0305$. (From Table C, P-value is between 0.025 and 0.05). Conclude: Since the P-value is smaller than $\alpha = 0.05$, we can reject H_0. There is convincing evidence that the distribution of music genre is different for the populations in the three age groups. (d) The test in (c) was a test of homogeneity. If we had treated all of the station's listeners as a single population from which we took one SRS, and if we treated age level and preferred music genre as two categorical variables, then a test of independence would have been appropriate.

Conditional distributions for part (a)

Age level	Hip Hop	Alternative	Post-rock
Middle School	0.56	0.36	0.08
High School	0.44	0.44	0.12
College	0.32	0.40	0.28





Expected counts for part (c)

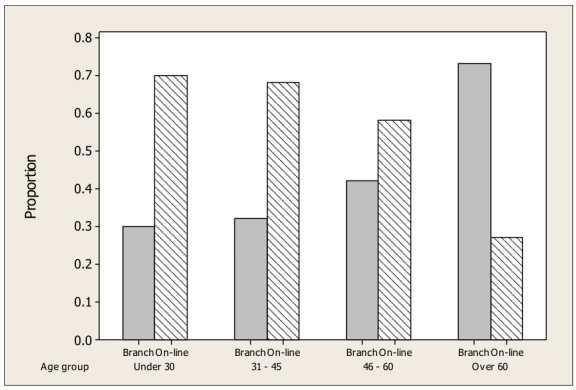
Age level	Hip Hop	Alternative	Post-rock
Middle School	22	20	8
High School	22	20	8
College	22	20	8

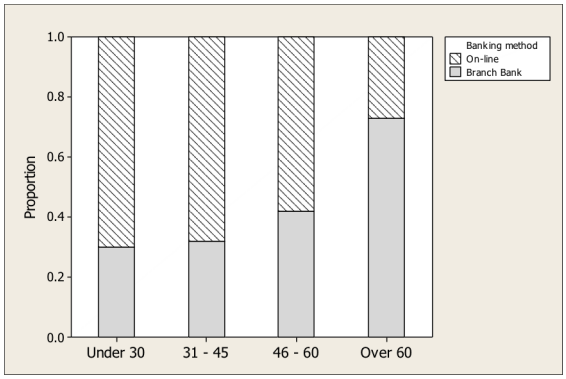
Quiz 11.2B

1. (a) & (b) Since age group is the explanatory variable, we calculate conditional distributions for each age group. See table and graph below (two options for graphs are shown: parallel and segmented bar graphs.) The data suggest that younger customers are more likely to bank online, and older customers are more likely to use the branch bank. This difference is especially strong for the over 60 age group. (c) State: We are testing the hypothesis H_o: The distribution of preferred banking method is the same for all age group populations, against H_a: The distribution of preferred banking method is not the same for all age group populations. We will use a significance level of $\alpha = 0.05$. Plan: The procedure is a chi-square test for homogeneity. Conditions: *Random*: the data come from SRSs of each population. *10%*: The populations are finite, but the bank is likely to have more customers than 10 times each sample size. *Large counts*: All expected counts are at least 5 (see expected counts table below). Do: Using a calculator, $\chi^2 = 17.397$; $df = 3$; $P-\text{value} = 0.00059$. (From Table C, P-value is between 0.0005 and 0.001). Conclude: Since the P-value is smaller than $\alpha = 0.05$, we can reject H_0. There is convincing evidence that the distribution of preferred banking method is different for the four age group populations. (d) The test in (c) was a test of homogeneity. If we had treated all of the bank's customers as a single population from which we took one SRS, and if we treated age group and preferred banking method as two categorical variables, then a test of independence would have been appropriate.

Conditional distributions for part (a)

Age group	Branch bank	On-line
Under 30	0.30	0.70
31 – 45	0.32	0.68
46 – 60	0.42	0.58
Over 60	0.73	0.27



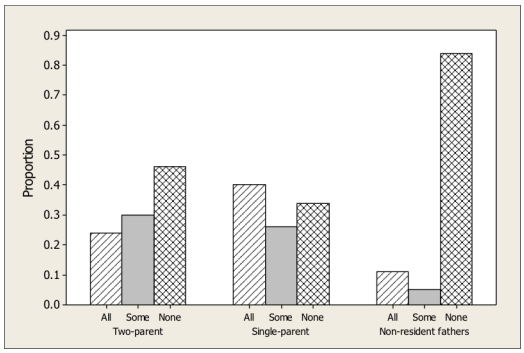


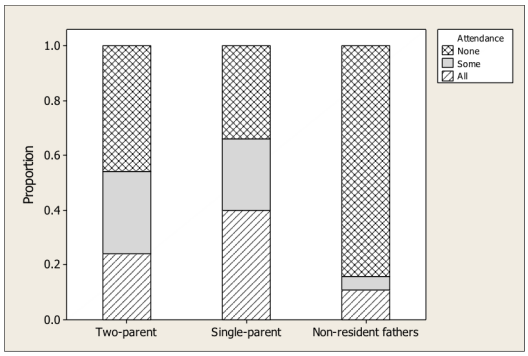
Expected counts for part (c)

Age group	Branch bank	On-line
Under 30	21.6	31.4
31 – 45	22.8	33.2
46 – 60	18.3	26.7
Over 60	12.2	17.8

Quiz 11.2C

1. Chi-square test for homogeneity; H_0: The distribution of texts and phone calls for the two age group populations are the same. H_a: The distributions of texts and phone calls for the two age group populations are different. **2.** (a) See graph below (two options for graphs are shown: parallel and segmented bar graphs.) (b) The more direct responsibility fathers have for the daily care of their children, the more likely they are to attend parent-teacher conferences: fathers in one-parent families are most likely to attend all conferences; non-resident fathers rarely attend any conferences. This is supported by the following chi-square test for independence. State: We are testing the hypothesis H_o: There is no association between attendance rate and family type, against H_a: There is an association between attendance rate and family type. We will use a significance level of $\alpha = 0.05$. Plan: Since the problem refers to a single sample of fathers, family type and attendance should be considered two variables measured in a single population, which makes this a chi-square test for independence. Conditions: *Random*: the data come from a SRS of fathers of school-age children. *10%*: The population of fathers with school-age children is at least 5800. *Large counts*: All expected counts are at least 5 (see expected counts table below). Do: Using a calculator, $\chi^2 = 54.774$; $df = 4$; $P-\text{value} \approx 0$. (From Table C, P-value is well below 0.0005). Conclude: Since the P-value is much smaller than $\alpha = 0.05$, we can reject H_0. There is convincing evidence of an association between family type and attendance at teacher conferences.





Expected counts for part (b)

	Attended all	Attended some	Attended none
Fathers in two-parent families	103.3	112.5	228.1
Fathers in single-parent families	8.8	9.6	19.5
Non-resident fathers	22.8	24.8	50.4