Tuesday, August 21, 2018

- **Warm-up**
  - **Copy and complete the following Contingency Table of AP scores from 2016 by calculating the marginal distributions (totals):**

		Grade					TOTAL
		5	4	3	2	1	
AP Test	Calculus BC	76,486	53,467	53,533	30,017	94,712	
	Statistics	29,647	44,884	51,367	32,120	48,565	
	TOTAL						

- **More with contingency tables and graphs**

- Content Objective: I will be able to use contingency tables to analyze data and prepare it for graphical representation.
- Social Objective: I will work with my group members to analyze data and make it easier to understand.
- Language Objective: I will read questions carefully so as to analyze data effectively.
- **Essential Question: What questions can I answer using data that is arranged in a contingency table?**

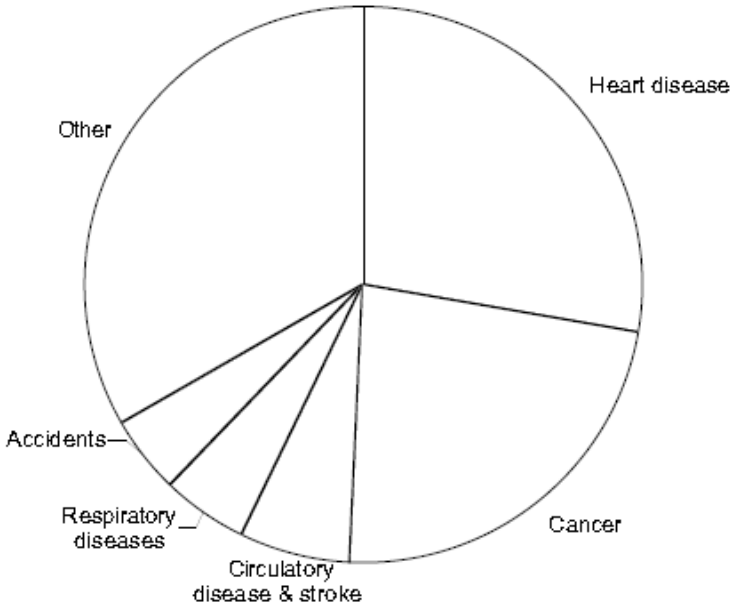
Check Homework: p 38 (11-13)

11. Causes of death 2004.

a) Yes, it is reasonable to assume that heart and respiratory disease caused approximately 33% of U.S. deaths in 2004, since there is no possibility for overlap. Each person could only have one cause of death.

b) Since the percentages listed add up to 66.4%, other causes must account for 33.6% of US deaths.

c) A pie chart is a good choice (with the inclusion of the "Other" category), since causes of US deaths represent parts of a whole. A bar chart would also be a good display.

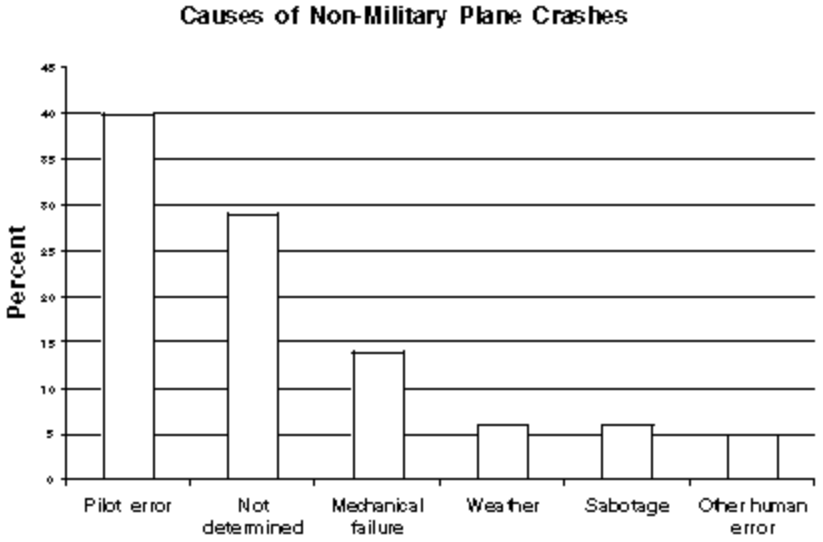


Check Homework: p 38 (11-13)

12. Plane crashes.

a) As long as each plane crash had only one cause, it would be reasonable to assume that weather or mechanical failures were the causes of about 20% of crashes.

b) Since the percentages listed add up to 71%, other causes must account for 29% of non-military plane crashes.



c) A relative frequency bar chart is a good choice (with the inclusion of the "Not determined" category). A pie chart might not be a good display, since several of the categories have nearly the same percentage. This makes it difficult to compare the percentages in a pie chart.

13. Oil spills 2006.

The bar chart shows that grounding is the most frequent cause of oil spillage for these 312 spills, and allows the reader to rank the other types as well. If being able to differentiate between these close counts is required, use the bar chart. The pie chart is also acceptable as a display, but it's difficult to tell whether, for example, there is a greater percentage of spills caused by grounding or hull failure. If you want to showcase the causes of oil spills as a fraction of all 312 spills, use the pie chart.

More with Contingency Tables

		Grade					TOTAL
		5	4	3	2	1	
AP Test	Calculus BC	76,486	53,467	53,533	30.017	94,712	**308,215**
	Statistics	29,647	44,884	51,367	32,120	48,565	**206,583**
	TOTAL	**106,133**	**98,351**	**104,900**	**62,137**	**143,277**	**514,798**

a) What percent of total scores were a 5? $\frac{106,133}{514,798} \approx 20.6\%$

b) Overall, did a higher percentage of Calc or Stats students score a 4?

$\approx 17.3\%$ $\approx 21.7\%$ 48.9%

c) What percent of total 3 scores are Statistics? $\frac{51,367}{104,900}$

d) What percent of Statistics tests were a score of 1? $\approx 23.5\%$

- Content Objective: I will be able to use contingency tables to analyze data and prepare it for graphical representation.
- Social Objective: I will work with my group members to analyze data and make it easier to understand.
- Language Objective: I will read questions carefully so as to analyze data effectively.

Even More with Contingency Tables

$\%M \stackrel{?}{=} \%C$

$P(A) \stackrel{?}{=} P(A|B)$

INDEPENDENT

AP Test	Grade 5	4	3	2	1	TOTAL
Calculus BC	76,486	53,467	53,533	30.017	94,712	**308,215**
Statistics	29,647	44,884	51,367	32,120	48,565	**206,583**
TOTAL	**106,133**	**98,351**	**104,900**	**62,137**	**143,277**	**514,798**

51% 53%

514,798 514,798 514,798 514,798

e) Calculate the marginal distribution of scores. → ÷ Grand total

→ of subtotals

f) Find the conditional distributions of 1's for each exam.

14.3% of stats are 5's $\frac{29,647}{206,583}$

g) Based on the data given, do you believe score is independent of test (is one easier)?

% of 5's $\frac{106,133}{514,798} \approx 20.6\%$

- Content Objective: I will be able to use contingency tables to analyze data and prepare it for graphical representation.
- Social Objective: I will work with my group members to analyze data and make it easier to understand.
- Language Objective: I will read questions carefully so as to analyze data effectively.

Another one…

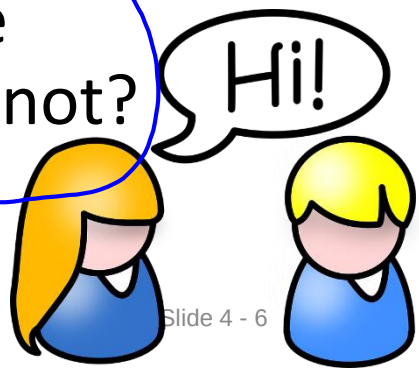
With Whom Is It Easiest to Make Friends?			
	Opposite Sex	Same Sex	No Difference
Females	58	16	63
Males	15	13	40

of grand total

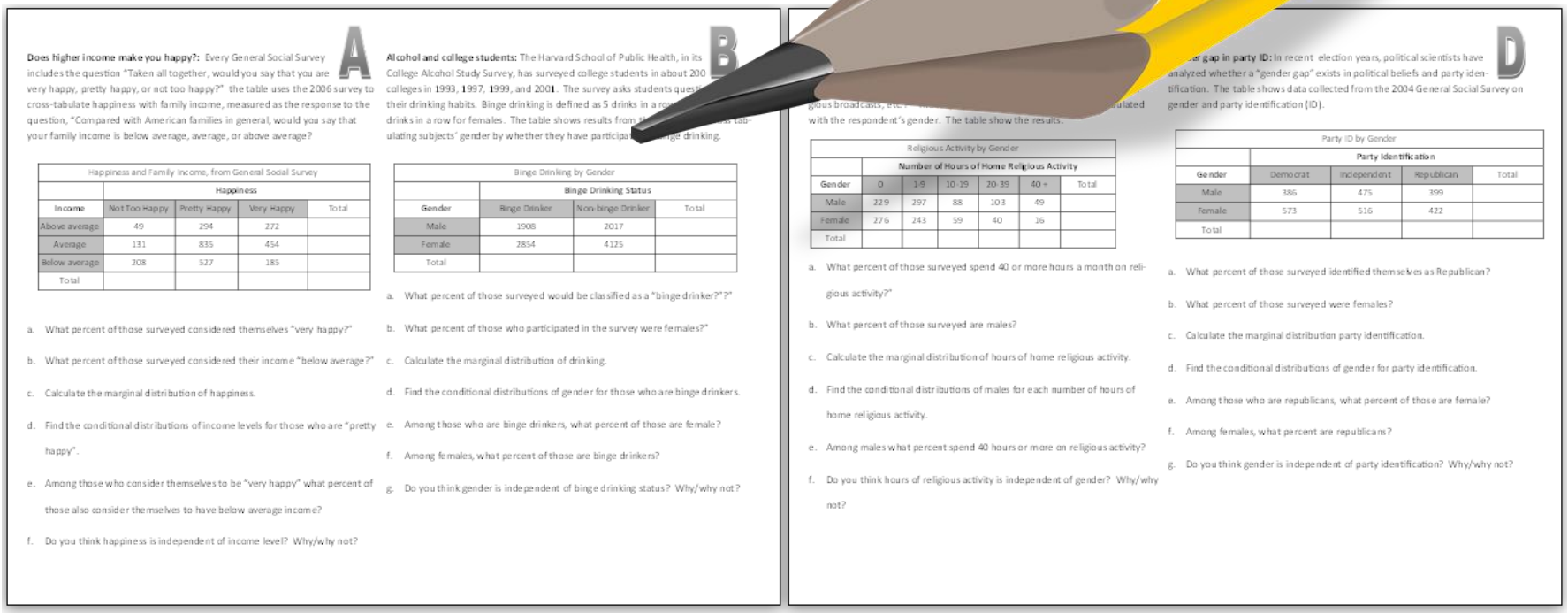
M% ?

C% ? of subtotal

- What percent of those surveyed found it easiest to make friends with the opposite sex?
- What percent of those surveyed were females?
- Calculate the marginal distribution of gender.
- Find the conditional distributions of friend types for females.
- Do you think "with whom it is easiest to make friends" is independent of gender? Why/why not?



Contingency table Practice



A

Does higher income make you happy?: Every General Social Survey includes the question "Taken all together, would you say that you are very happy, pretty happy, or not too happy?" the table uses the 2006 survey to cross-tabulate happiness with family income, measured as the response to the question, "Compared with American families in general, would you say that your family income is below average, average, or above average?

Happiness and Family Income, from General Social Survey

	Happiness			
Income	Not Too Happy	Pretty Happy	Very Happy	Total
Above average	49	294	272	
Average	131	835	454	
Below average	208	527	185	
Total				

a. What percent of those surveyed considered themselves "very happy?"

b. What percent of those surveyed considered their income "below average?"

c. Calculate the marginal distribution of happiness.

d. Find the conditional distributions of income levels for those who are "pretty happy".

e. Among those who consider themselves to be "very happy" what percent of those also consider themselves to have below average income?

f. Do you think happiness is independent of income level? Why/why not?

B

Alcohol and college students: The Harvard School of Public Health, in its College Alcohol Study Survey, has surveyed college students in about 200 colleges in 1993, 1997, 1999, and 2001. The survey asks students questi... their drinking habits. Binge drinking is defined as 5 drinks in a ro... drinks in a row for females. The table shows results from ... ulating subjects' gender by whether they have particip... nge drinking.

Binge Drinking by Gender

	Binge Drinking Status		
Gender	Binge Drinker	Non-binge Drinker	Total
Male	1908	2017	
Female	2854	4125	
Total			

a. What percent of those surveyed would be classified as a "binge drinker?"?

b. What percent of those who participated in the survey were females?"

c. Calculate the marginal distribution of drinking.

d. Find the conditional distributions of gender for those who are binge drinkers.

e. Among those who are binge drinkers, what percent of those are female?

f. Among females, what percent of those are binge drinkers?

g. Do you think gender is independent of binge drinking status? Why/why not?

C

...gious broadcasts, etc.) ... ulated with the respondent's gender. The table show the results.

Religious Activity by Gender

	Number of Hours of Home Religious Activity					
Gender	0	1-9	10-19	20-39	40 +	Total
Male	229	297	88	103	49	
Female	276	243	59	40	16	
Total						

a. What percent of those surveyed spend 40 or more hours a month on religious activity?"

b. What percent of those surveyed are males?

c. Calculate the marginal distribution of hours of home religious activity.

d. Find the conditional distributions of males for each number of hours of home religious activity.

e. Among males what percent spend 40 hours or more on religious activity?

f. Do you think hours of religious activity is independent of gender? Why/why not?

D

...er gap in party ID: In recent election years, political scientists have analyzed whether a "gender gap" exists in political beliefs and party identification. The table shows data collected from the 2004 General Social Survey on gender and party identification (ID).

Party ID by Gender

	Party Identification			
Gender	Democrat	Independent	Republican	Total
Male	386	475	399	
Female	573	516	422	
Total				

a. What percent of those surveyed identified themselves as Republican?

b. What percent of those surveyed were females?

c. Calculate the marginal distribution party identification.

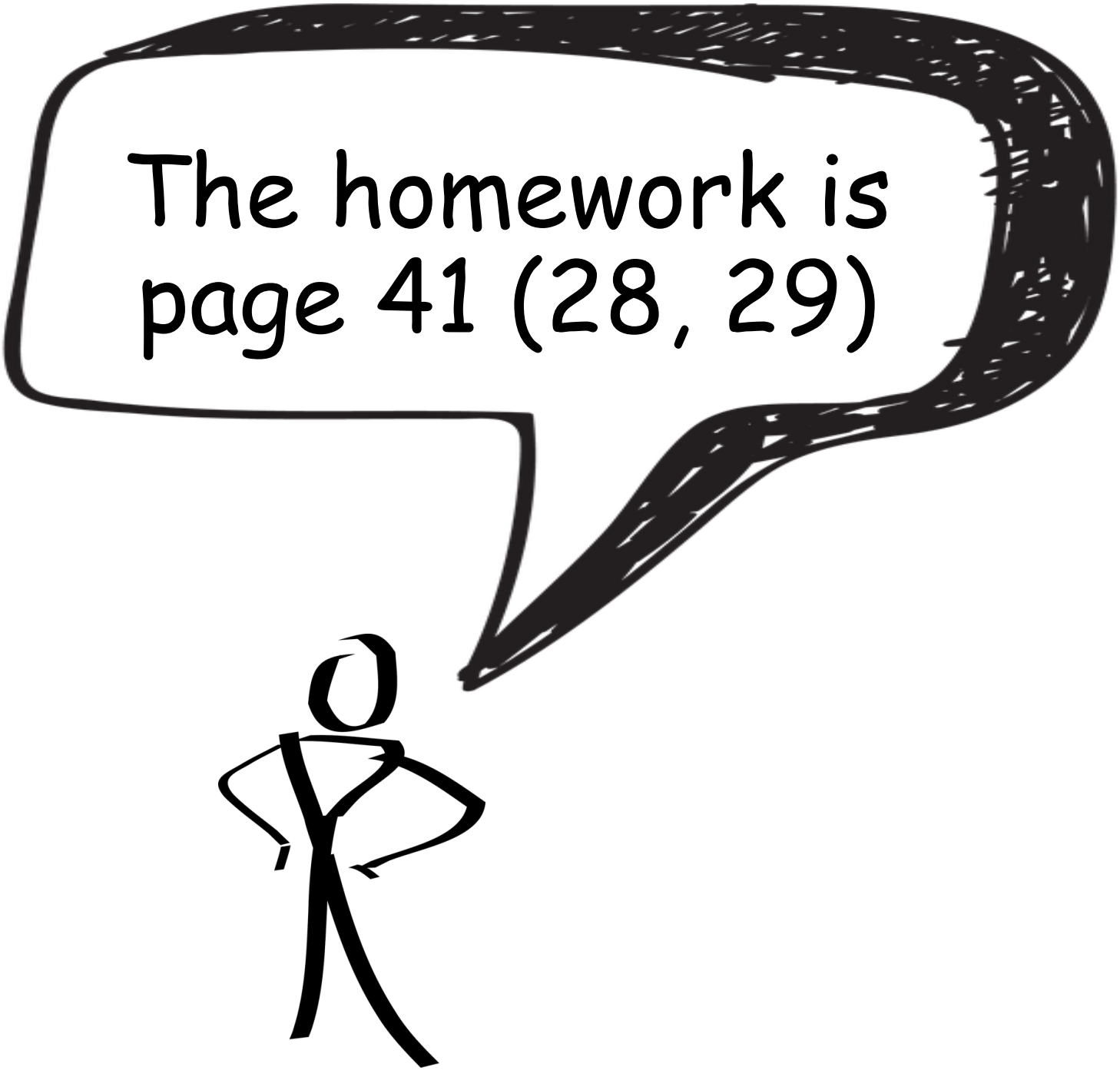
d. Find the conditional distributions of gender for party identification.

e. Among those who are republicans, what percent of those are female?

f. Among females, what percent are republicans?

g. Do you think gender is independent of party identification? Why/why not?

- Content Objective: I will be able to use contingency tables to analyze data and prepare it for graphical representation.
- Social Objective: I will work with my group members to analyze data and make it easier to understand.
- Language Objective: I will read questions carefully so as to analyze data effectively.
- **Essential Question: What questions can I answer using data that is arranged in a contingency table?**


The homework is page 41 (28, 29)