Your χ^2 project will involve constructing a statistic for the info we collected in class.

You will be assigned a question, and given the distribution of answers provided by 2 or more subgroups of the class.

In our case the there were four ways that we formed subgroups: class standing (F,So, Ju, Sen); gender (Female or Male); self identified ethnicity (white or not); employment (yes/no).

Your goal is to see if the differences seen in a subgoup in the response rate from these subgroups are statisitically significant or not.

Example with other data

1. **Our Sample Problem Introduction**
   Here is our set up for our sample investigation.

   158 people are asked if they agree or disagree with the following statement:

   > textitI get my money's worth at NIU.

   The posible answers are - **Yes, NO, Don't Know**.

   They are also asked if they are religious, and the possible answers are **Yes, No, Somewhat**.

   I form my NULL HYPOTHESIS: The response rate to the question is independent of the person's religious classification (Yes, No, Sometimes).

   The alternative hypothesis is the negation: The respose rate is dependent upon one's religious classification.

   In general, we compute our χ^2 statistic, with given α level, and know d, and if

   - If my χ^2 statistic is smaller than the table entry for our values of α and d, then my hypothesis is SUPPORTED (not proven),
   - if my statistic is larger than the table entry my hypothesis is REJECTED (the differences are staistically significant).

	Yes	No	Don't Know
Religous	34	25	14
Not Religous	9	17	1
Sometimes Religous	25	25	8

   or, note that I totaled all the rows and columns here - those totals are the

same for the observed and expected.

	Yes	No	Don't Know	
Religous	34	25	14	73
Not Religous	9	17	1	27
Sometimes Religous	25	25	8	58
	68	67	23	158

Recall, d is a feature of the table!!!
BUT, alpha (α) is chosen by the researcher (you) or given to you.

We set $\alpha = 0.01$ for our investigation.

2. **Computing the Expected**:
For the question, I counted 68 Yes, 67 No and so there are 23 Don't knows (since there are a total of 158 people, I need only count 2 of the 3 responses, I can get the 3rd from knowing the totals).

I compute the proportion of each type of answer:

Answer	Frequency	Proportion
Yes	68	$68/158 = 0.43$
No	67	$67/158 = 0.42$
Don't Know	23	$23/158 = 0.15$

To fill out the expected, we assume that the proportion of each type of answer is the same in each of the three categories of Religousness = Yes, No or Sometimes.

Out of the people surveyed 73 answered *Yes, they are religious*, 27 answered *No* and 58 said *Sometimes*.

Of the 73 who answered *Yes to Question 1* we would expect to see the same proportion of answers Yes/No/Don't Know as in the whole population. Thus we

> expect that of these 73 people, $73 \times 0.43 = 31.4$
> would answer Yes to question 7: $73 \times 0.42 = 30.7$ would say No;,
> and
> $73 \times 0.15 = 10.9$ would have no opinion.

We could similarly find that we would expect 11.6, 11.3 and 4.1 of the 27 who answered *No* to the religion question to answer Yes, No or Don't know, respectively on question 7.

I can use the fact that I know the totals in the columns to find the numbers for Sometimes. For instance, we saw a total of 68 Yes's, but we've listed 31.4 in the Religous category and 11.6 in the Not religous category, this

leaves $68 - (31.4 + 11.6) = 68 - 43 = 25$ people who said Yes to #7, but are only Sometimes Religious.

Our table will eventually look like

	Yes	No	Don't Know
Religous	31.4	30.7	10.9
Not Religous	11.6	11.3	4.1
Sometimes Religous	25	25	8

3. **The degree of freedom** is 4. Note: you used this if you computed your values like I did. The **Alpha** level is $\alpha = .01$.

4. **Computing the statistic**

   We compute the ratio $(E-F)^2/E$, for each pair of corresponding **E**xpected and observed **F**requencies. Use 3 decimal places here.

   My statistic is

$$\begin{aligned}\chi^2 = & \frac{(31.4-34)^2}{31.4} + \frac{(30.7-25)^2}{30.7} + \frac{10.9-14)^2}{10.9} \\ & + \frac{(11.6-9)^2}{11.6} + \frac{(11.3-17)^2}{11.3} + \frac{(4.1-1)^2}{4.1} \\ & + \frac{(25-25)^2}{25} + \frac{(25-25)^2}{25} + \frac{(8-8)^2}{8} \\ = & \ 0.215 + 1.058 + 0.882 + 0.583 + 2.875 + 2.344 + 0 + 0 + 0 \\ = & \ 7.957\end{aligned}$$

5. **Interpreting the statistic.**

   At $\alpha = 0.01$, and $d = 4$, the table entry is 13.277. The statistic I obtained is SMALLER than the entry in the χ^2 table. Thus my null hypothesis is supported (not proven). That is, there is no statistically significant difference in the opinions of people on whether or not NIU is a good value, when we examine this responses with respect to religiousness.

   Note also that our χ^2 value is so small that it would not have been rejected at .025 or 0.05 either.