

STAT 457 Fall, 2010 Lab #2

Topics:

- 2 way tables done by StatCrunch
- Segmented bar graphs of 2 way categorical tables done by StatCrunch

=====

[1] I want to do problem #2 of Chapter 3 in your text, located on page 41. I did this problem by hand, and the results are shown on the last pages of this lab. I would now like to have StatCrunch do the problem for me.

I first had to put the data set together, so that StatCrunch could construct my 2 way tables and my graph. I did my work in EXCEL, because it is a nice environment to construct spread sheet information. I have 2 columns, labeled "race" and "plans", where I have 198 rows of "white, 4 year", 44 rows of "minority, 4 year", 36 rows of "white 2 year", 6 rows of "minority, 2 year", etc., having a total of 325 rows of a value of "race" and a value of "plans". I could easily make these repetitive rows by using the copy and paste features of EXCEL. Then, I alphabetized rows by race, so all the white plans were shown, followed by all the minority plans. Then I made 4 more columns, labeled "race1", "plans1", "race2", and "plans2", where the race and plan 1's were the white data, with race and plan2 being the minority data. I just copied and pasted the respective white and minority data from the first 2 columns to make columns 3 – 6.

I then saved it as a type .csv (comma delimited file), calling it **2waytabl.csv** . I will use the first 2 columns to make the 2 way table in StatCrunch, and the last 4 to make segmented bar graphs of the conditional distribution of **plans** by **race**.

[1a] Open StatCrunch and upload the file **2waytabl.csv** into it. You should have a screen similar to that shown below.

StatCrunch logo: **StatCrunch** data analysis on the Web

Navigation: Home Explore My StatCrunch

Properties: 2waytabl.csv

Row	race	plans	race1	plans1	race2	plans2
1	white	4 year	white	4 year	minority	4 year
2	white	4 year	white	4 year	minority	4 year
3	white	4 year	white	4 year	minority	4 year
4	white	4 year	white	4 year	minority	4 year
5	white	4 year	white	4 year	minority	4 year
6	white	4 year	white	4 year	minority	4 year
7	white	4 year	white	4 year	minority	4 year
8	white	4 year	white	4 year	minority	4 year
9	white	4 year	white	4 year	minority	4 year
10	white	4 year	white	4 year	minority	4 year

[1b] Now click on **STAT TABLE CONTINGENCY WITH DATA** and click the **plans** variable as the “row” variable, with **race** variable being the “column” variable. Then click on **NEXT** and check all % boxes (there are 3 of them), then just click on **CALCULATE**. You should have a result similar to that shown below. Copy and paste your results in your WORD document.

Contingency table results:

Rows: plans
Columns: race

Cell format
Count (Row percent) (Column percent) (Total percent)

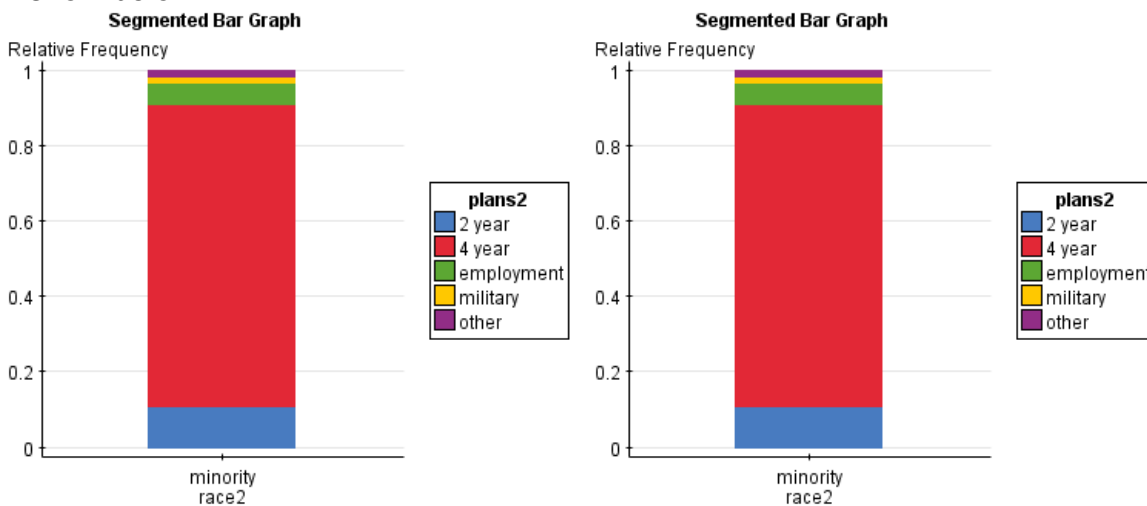
	minority	white	Total
2 year	6 (14.29%) (10.53%) (1.846%)	36 (85.71%) (13.43%) (11.08%)	42 (100.00%) (12.92%) (12.92%)
4 year	44 (18.18%) (77.19%) (13.54%)	198 (81.82%) (73.88%) (60.92%)	242 (100.00%) (74.46%) (74.46%)
employment	3 (17.65%) (5.263%) (0.9231%)	14 (82.35%) (5.224%) (4.308%)	17 (100.00%) (5.231%) (5.231%)
military	1 (20%) (1.754%) (0.3077%)	4 (80%) (1.493%) (1.231%)	5 (100.00%) (1.538%) (1.538%)
other	3 (15.79%) (5.263%) (0.9231%)	16 (84.21%) (5.97%) (4.923%)	19 (100.00%) (5.846%) (5.846%)
Total	57 (17.54%) (100.00%) (17.54%)	268 (82.46%) (100.00%) (82.46%)	325 (100.00%) (100.00%) (100.00%)

Statistic	DF	Value	P-value
Chi-square	4	0.43779305	0.9793

For now, don't worry about the last table (with P value, DF, etc.). We are just interested in the main table results.

[1c] Now, in StatCrunch, click on **GRAPHICS BAR PLOT WITH DATA** . Put **race1** In "column" and **plans1** in "groups", then click **STACKED BARS** and **NEXT** . Now click **RELATIVE FREQUENCY** for "type" then **NEXT** .Now you may give it a title, if you choose, then **CREATE GRAPH** . Your result should look like the one below. Copy and paste this to your WORD document.

[1d] Repeat the **GRAPHICS** procedure for race2 and plans2, then copy/paste your resulting segmented bar graph for minority next to your graph for white. My results show below.



[1e] Comment on whether the two variables, race and plans, are independent or not independent. Give reasons for your decision. Be sure to use proper sentence structure, English, and spelling.

=====

[2] EXTRA CREDIT: Download the data file **bloodpres.csv** from the web site. Repeat problem [1] with this data set I built up in the same way I constructed the **2waytab1.csv** file for problem [1]. Below are some of the screens I got when I did this problem. Be sure to show your work in WORD, and label this part as EXTRA CREDIT, should you decide you have the time to do this in StatCrunch. Shown below is what I did on the blood pressure problem. You do not have to do exactly what I did, but, I wanted you to see what you should sort of end up with. Have fun with this!!

Set Properties

[Delete]

File name: m.olear

Created: Aug 21, 2010

Size: 1.4KB

Format: no

Rows: 0

Results for this data:

Feature is new to StatCrunch 5.0 and does not

bloodpres.csv

Row	pres	age	pres1	age1	pres2	age2
1	low	under 30	normal	under 30	high	unc
2	low	under 30	normal	under 30	high	unc
3	low	under 30	normal	under 30	high	unc
4	low	under 30	normal	under 30	high	unc
5	low	under 30	normal	under 30	high	unc
6	low	under 30	normal	under 30	high	unc
7	low	under 30	normal	under 30	high	unc
8	low	under 30	normal	under 30	high	unc
9	low	under 30	normal	under 30	high	unc
10	low	under 30	normal	under 30	high	unc

Contingency table results:

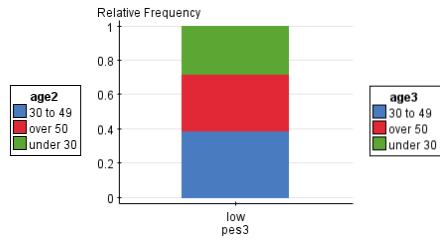
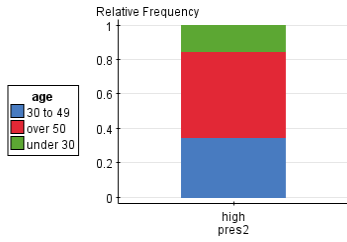
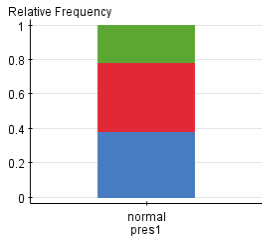
Rows: pres

Columns: age

Cell format
Count
(Row percent)
(Column percent)
(Total percent)

	30 to 49	over 50	under 30	Total
high	51 (34.69%) (28.49%) (10.76%)	73 (49.66%) (37.06%) (15.4%)	23 (15.65%) (23.47%) (4.852%)	147 (100.00%) (31.01%) (31.01%)
low	37 (38.95%) (20.67%) (7.806%)	31 (32.63%) (15.74%) (6.54%)	27 (28.42%) (27.55%) (5.696%)	95 (100.00%) (20.04%) (20.04%)
normal	91 (39.22%) (50.84%) (19.2%)	93 (40.09%) (47.21%) (19.62%)	48 (20.69%) (48.98%) (10.13%)	232 (100.00%) (48.95%) (48.95%)
Total	179 (37.76%) (100.00%) (37.76%)	197 (41.56%) (100.00%) (41.56%)	98 (20.68%) (100.00%) (20.68%)	474 (100.00%) (100.00%) (100.00%)

Statistic	DF	Value	P-value
Chi-square	4	9.352049	0.0529



041 (ch 3)

23)

23. Class of 2000. Prior to graduation, a high school class of 2000 was surveyed about their plans. The following table displays the results for white and minority students. (The "Minority" group included African-American, Asian, Hispanic, and Native American students.)

	White	Minority	Total
4-year college	198	44	242
2-year college	36	6	42
Military	4	1	5
Employment	14	3	17
Other	16	3	19

Total 268 57 325

- What percent of the graduates are white?
- What percent of the graduates are planning to attend a 2-year college?
- What percent of the graduates are white and planning to attend a 2-year college?
- What percent of the white graduates are planning to attend a 2-year college?
- What percent of the graduates planning to attend a 2-year college are white?
- Create a graph comparing the plans of white and minority students.
- Do you see any important differences in the post-graduation plans of white and minority students? Write a brief summary of what these data show, including comparisons of conditional distributions.

$$(a) \% \text{ white} = \frac{268}{325} = 0.8246$$

$$(b) \% \text{ 2-year} = \frac{42}{325} = 0.1292$$

$$(c) \% \text{ white-2-year} = \frac{36}{325} = 0.1108$$

$$(d) \% \text{ white are 2-year} = \frac{36}{268} = 0.1343$$

$$(e) \% \text{ 2-year are white} = \frac{36}{42} = 0.8571$$

$$(f) \text{ white} = 198/268 = 0.7388 \text{ (4-year college)} \quad 0.7719 = 4/57 = \text{minority}$$

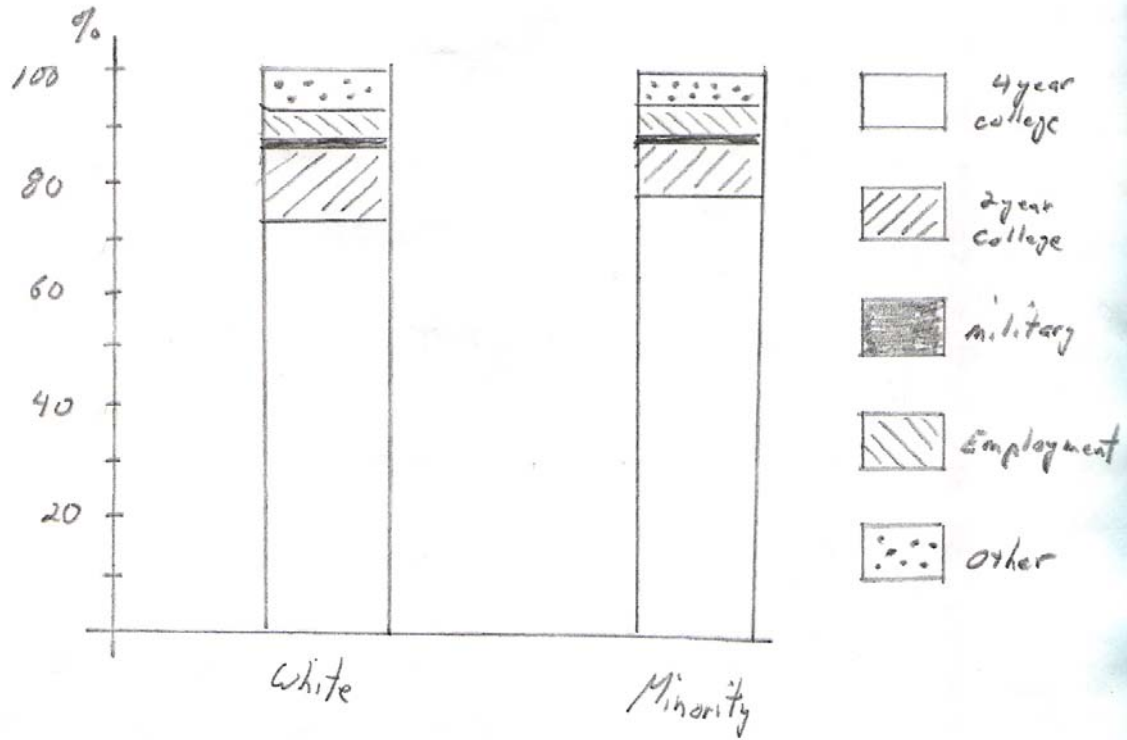
$$36/268 = 0.1343 \text{ (2-year college)} \quad 0.1053 = 6/57$$

$$4/268 = 0.0149 \text{ (military)} \quad 0.0175 = 1/57$$

$$14/268 = 0.0522 \text{ (employment)} \quad 0.0526 = 3/57$$

$$16/268 = 0.0597 \text{ (other)} \quad 0.0526 = 3/57$$

segmented bar graph



I do not see any striking differences in the 2 graphs, so there are no striking differences in the conditional distributions. Therefore, the post-high school plans for both groups are very similar