Panel 1

Want to analyze 2 vars simultaneously:

Q: Does smoking cause cancer?
Do men make more money than women?
Do African Americans vote more Democrat than Republican?

smoking (not clear) vs cancer

gender (indep) vs income (dep)

race (indep) vs party affiliation (dep)

independent (explain in time)
vs dependent variable

1

Panel 2

The trick to analyzing 2 vars is Contingency Table

var A

var B

Ex: Residents of a town vote on zoning law (and their age was recorded).

Age vs. Opinion

opinion (dep) \ age (indep)	≤50	>50	total
For	97	82	179
Against	158	75	233
total	255	162	412

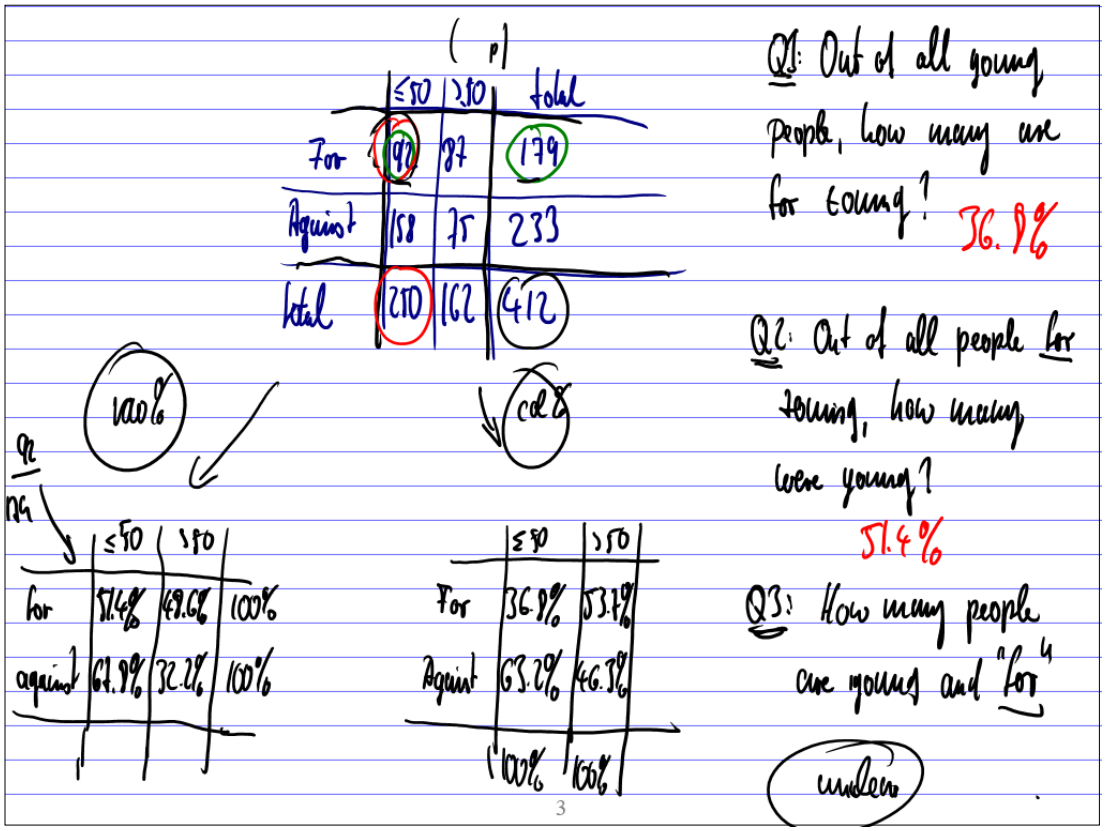
row %: row / row total in %

col %: col / col total in %

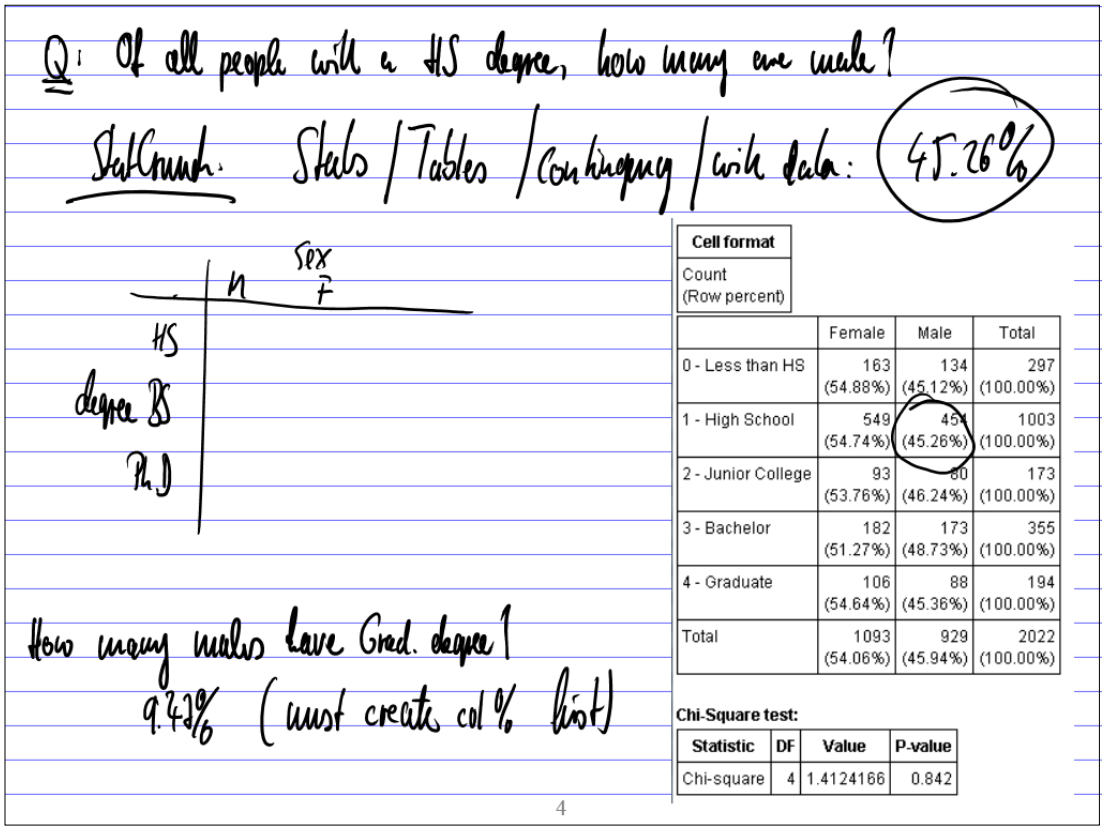
total %: cell / total in %

2

Panel 3



Panel 4



Cell format
Count (Row percent)

	Female	Male	Total
0 - Less than HS	163 (54.88%)	134 (45.12%)	297 (100.00%)
1 - High School	549 (54.74%)	454 (45.26%)	1003 (100.00%)
2 - Junior College	93 (53.76%)	80 (46.24%)	173 (100.00%)
3 - Bachelor	182 (51.27%)	173 (48.73%)	355 (100.00%)
4 - Graduate	106 (54.64%)	88 (45.36%)	194 (100.00%)
Total	1093 (54.06%)	929 (45.94%)	2022 (100.00%)

Chi-Square test:

Statistic	DF	Value	P-value
Chi-square	4	1.4124166	0.842

Panel 5

How many males have at least HS degree?

Cell format
Count (Row percent) (Column percent)

	Female	Male	Total
0 - Less than HS	163 (54.88%) (14.91%)	134 (45.12%) (14.43%)	297 (100.00%) (14.69%)
1 - High School	549 (54.74%) (50.23%)	454 (45.26%) (48.87%)	1003 (100.00%) (49.6%)
2 - Junior College	93 (53.76%) (8.509%)	80 (46.24%) (8.611%)	173 (100.00%) (8.556%)
3 - Bachelor	182 (51.27%) (16.65%)	173 (48.73%) (18.62%)	355 (100.00%) (17.56%)
4 - Graduate	106 (54.64%) (9.698%)	88 (45.36%) (9.473%)	194 (100.00%) (9.594%)
Total	1093 (54.06%) (100.00%)	929 (45.94%) (100.00%)	2022 (100.00%) (100.00%)

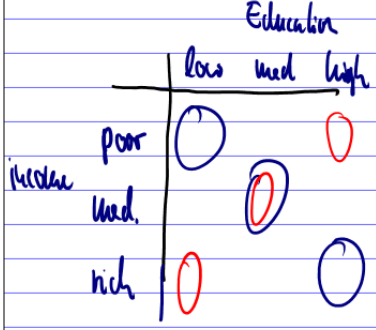
14.42%
+ 48.81%
63.19%

5

Panel 6

Big Question: Is there a relationship between two variables, and if there is one, which way does it go?

Method 1: Create contingency table, put ind. var. into columns
dep var. into rows
Compute row % and circle highest value in each row.



ind dep
Education vs. income

if they line up along major diagonal, there is a
positive relationship
negative relationship

6

Panel 7

Does money make you happy? indep. dep. Both are ordinal

money

happy

Contingency table results:

Rows: GENERAL HAPPINESS

Columns: SATISFACTION WITH FINANCIAL SITUATION

Cell format
Count (Row percent)

	1 - satisfied	2 - more or less	3 - not at all satisfied	Total
1 - very happy	267 (44.8%)	236 (39.6%)	93 (15.6%)	596 (100.00%)
2 - pretty happy	273 (24.89%)	476 (43.39%)	348 (31.72%)	1097 (100.00%)
3 - not too happy	31 (9.841%)	102 (32.38%)	182 (57.78%)	315 (100.00%)
Total	571 (28.44%)	814 (40.54%)	623 (31.03%)	2008 (100.00%)

7

Panel 8

Our analysis so far works only for ordinal variables. Want to expand to nominal vars. (and later to numeric vars).

Need: Expected Value = (row total) × (col. total) / (total)

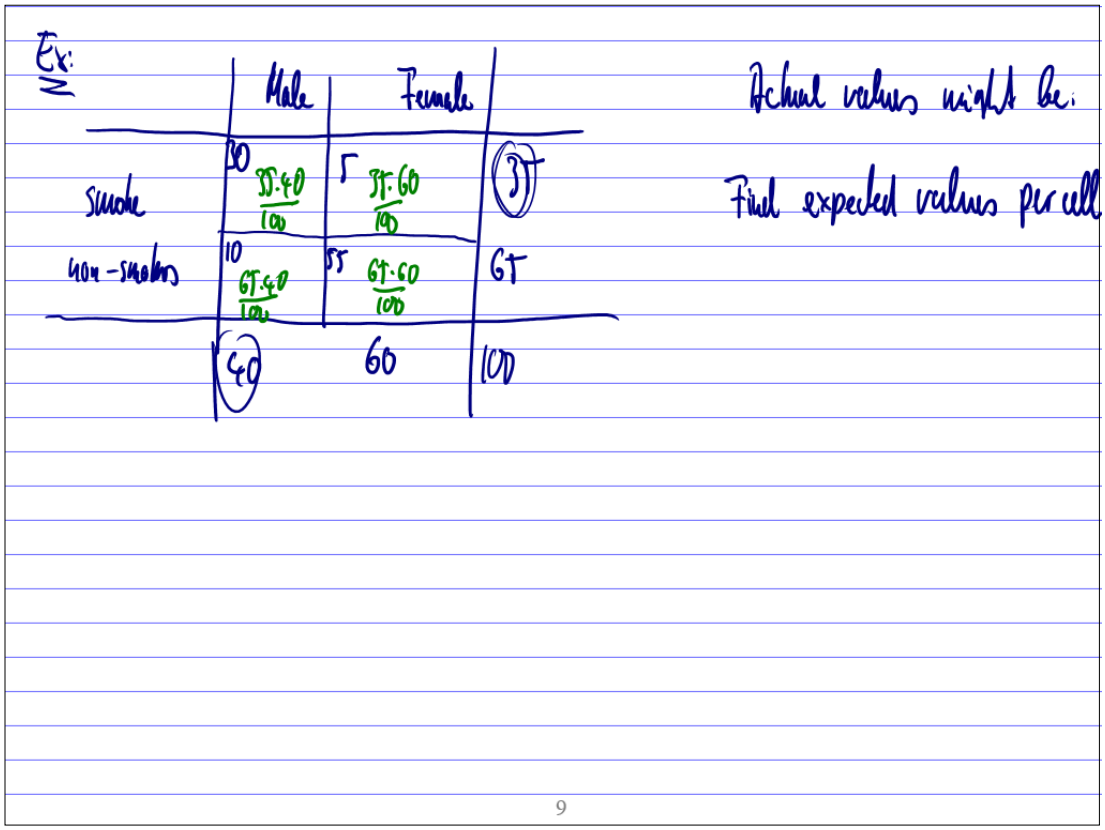
Ex.

	Male	Female	
smoke	$\frac{40}{100} \cdot 30 = 12$	$\frac{60}{100} \cdot 30 = 18$	30
non smoke	$\frac{40}{100} \cdot 70 = 28$	$\frac{60}{100} \cdot 70 = 42$	70
	40	60	100

are the values you'd expect if there was no relationship between the row/col variable.

8

Panel 9


Ex:
Male
Female
Smoke
non-smoker
Actual values might be:
Find expected values per cell
9