

Exercise 3A.1

In the New Hampshire student survey data, one of the variables collected was where the student lives. Complete the table below with the appropriate summaries (as labeled at the top of the columns).

ANSWER: (see below)

Guessed			
grades	Freq.	Percent	Cum.
As & Bs	110	$110/242=0.45$	$110/242=0.45$
Bs & Cs	120	$120/242=0.50$	$(110+120)/242=0.95$
Cs & Ds	12	$12/242=0.050$	$(110+120+12)/242=1$
Total	242	100	-

```
. tab grades  
  
Guessed |  
grades this |  
semester | Freq. Percent Cum.  
-----+-----  
As & Bs | 110  
Bs & Cs | 120  
Cs & Ds | 12  
-----+-----  
Total | 242 100.00 --
```

Exercise 3A.2

In a study of asthma in minority youth in Detroit, Michigan, two pieces of information collected about the children were their asthma severity (3 levels: intermittent, mild persistent, moderate/severe persistent) and their family income. These data were obtained and summarized in the table below:

Income	Asthma Severity		
	Intermittent	Mild Persistent	Moderate/Severe Persistent
<\$15,000	71 (43%)	41 (53%)	34 (49%)
\$15,000-\$40,000	66 (40%)	27 (35%)	29 (41%)
\$40,000+	28 (17%)	10 (13%)	7 (10%)

- (a) Are the percentages in this table overall percentages, row percentages, or column percentages?

ANSWER: Column percentages

They sum to 100% down the row, e.g., $43\% + 40\% + 17\% = 100\%$ (Note: Column 2 percentages sum to 101% due to rounding)

- (b) Why might the investigators have chosen to use these percentages (instead of the other options)?

ANSWER: The column percentages allow us to easily compare the distribution of income across the three asthma severity groups. So we could see whether kids with more severe asthma tended to have lower incomes (for example).

- (c) If one child with intermittent asthma is randomly selected from this study, what is the probability that he has the highest level of household income (\$40,000+)?

ANSWER: 0.17

Read the column percent right off the table, or calculate as $28/(71 + 66 + 28) = 28/156 = 0.17$

