

## Supplementary exercises 9.36 and 9.38 of IPS7e

Data: Enrollment status of US high school graduates aged 16 to 24 years, separated by gender. As the purpose of the two exercises is purely descriptive, we will not introduce notation to describe the methods. For use in both exercises, we use Minitab to display a table of counts together with marginal and conditional proportions. For a start, we relabel the gender groups more informatively using the Data-Recode-To Text menu.

```
MTB > Code ("g1") "Men" ("g2") "Women" 'gender';
SUBC> TSummary;
SUBC> Original.
MTB > XTabs 'gender' 'status';
SUBC> Layout 1 1;
SUBC> Frequencies 'count';
SUBC> Counts;
SUBC> RowPercents;
SUBC> ColPercents;
SUBC> DMissing 'gender' 'status'.
```

### Tabulated Statistics: gender, status

Using frequencies in count

**Rows: gender Columns: status**

	s1	s2	s3	s4	s5	s6	All
Men	890	340	2897	249	306	160	4842
	18.38	7.02	59.83	5.14	6.32	3.30	100.00
	47.88	45.76	46.59	39.40	45.54	53.87	46.46
Women	969	403	3321	383	366	137	5579
	17.37	7.22	59.53	6.87	6.56	2.46	100.00
	52.12	54.24	53.41	60.60	54.46	46.13	53.54
All	1859	743	6218	632	672	297	10421
	17.84	7.13	59.67	6.06	6.45	2.85	100.00
	100.00	100.00	100.00	100.00	100.00	100.00	100.00

*Cell Contents*  
 Count  
 % of Row  
 % of Column

### Exercise 9.36

- (a) There were 890 thousand male students in a two-year college, full-time program. (Men × s1 – Count, or directly from the IPS data table)
- (b) The marginal gender distribution is the totals for Men and Women. There were 4842 thousand male students and 5579 thousand female students, and the proportions of male and female

students were therefore  $4842/(4842 + 5579) = 46.46\%$ , and  $53.54\%$ , respectively.  
 (Men and Women  $\times$  All – Count and % of Column)

- (c) The marginal distribution of status among all students is in the row for All, redisplayed here for clarity:

	s1	s2	s3	s4	s5	s6	All
All	1859	743	6218	632	672	297	10421
	17.84	7.13	59.67	6.06	6.45	2.85	100.00

The by far most common enrolment status is Four-year college, full-time (category s3).

*Exercise 9.38*

The conditional distributions of enrolment for each gender are in the rows for Men and Women; we also redisplay those here for clarity.

	s1	s2	s3	s4	s5	s6	All
Men	890	340	2897	249	306	160	4842
	18.38	7.02	59.83	5.14	6.32	3.30	100.00

	s1	s2	s3	s4	s5	s6	All
Women	969	403	3321	383	366	137	5579
	17.37	7.22	59.53	6.87	6.56	2.46	100.00

The proportions (or percentages) look quite similar across all status categories, with no difference larger than 0.02 (or 2%). Note that despite the visually similar distributions, a Pearson chi-square test for homogeneity between genders is strongly significant; this is because of the large sample size, even without converting the counts from thousands. Still it seems most meaningful to describe the common patterns across gender, and as noted before the by far most common status category is s3 with close to 60% of students. The second-largest category is s1 (Two-year college, full-time), with around 18% of students. The remaining four status categories all have much lower representation, ranging from about 2.5% to 7%.