

PRINTED BY: jjewett@herzing.edu. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

- ownership: The ownership type of the provider including:
 - For-profit—corporation
 - For-profit—Individual
 - For-profit—Limited liability company
 - For-profit—Partnership
 - Government—City
 - Government—City/County
 - Government—County
 - Government—Hospital district
 - Government—State
 - Nonprofit—Church-related
 - Nonprofit—Corporation
 - Nonprofit—Other
- inhosp: Identifies whether the provider resides in a hospital; takes on a value of YES or NO
- adj_total: The adjusted total nurse staffing hours per resident per day, including certified nursing practitioner (CNA), license practicing nurse (LPN), and registered nurse (RN)
- adj_rn: The adjusted RN staffing hours per resident per day

Renaming Column Headers

Several data elements will be used in this chapter to answer the research question. The data elements that were extracted from MySQL Workbench have column names that are not particularly informative for understanding what variable we are working with for a specific analysis. To make the data easier to work with in RStudio, we will simply rename the column headers. Given the fact that RStudio is case sensitive, we will ensure that the revised names are all in lowercase. This will help to prevent errors that would otherwise occur if you tried to call a variable using the wrong case sensitivity. To rename column headers in Excel, simply click each individual cell, highlight the contents, delete the content, enter the new column name, and hit return. Repeat this process for each column header. The old and new column headers are listed in [table 13.2](#).

Reviewing Data Sets for Erroneous Values

One of the first steps in data preparation is to carefully examine if the data include erroneous values or blanks. The CSV that was exported from MySQL does not include any erroneous values or blanks, but it is still important to review the data set for errors and issues.

Recoding a Variable in Excel

After you have reviewed the data for accuracy and completeness, you need to create a new column of data. For the purpose of this chapter one of the objectives is to compare the nursing staffing measures and percentage of high-risk long-stay residents with pressure ulcers between the different provider ownership types. However, the ownership column of data includes 10 different types. Consequently, with such a large number of groups,

PRINTED BY: jjewett@herzing.edu. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

this will result in a small sample size for some of the ownership status. For example, there are only three providers in the Government—Federal group. When there are groups with very small sample sizes, this limits the quality of the analysis and also may limit the generalization of the findings. To improve the analysis, we are going to recode the owner column of data. The objective is to regroup the ownership status into three groups: for-profit, nonprofit, and government. We will group the providers based on the existing ownership status. In Excel, this task can be carried out using an IF statement.

Table 13.2. Original and revised column names within Excel

Original Name	Revised Name
provnum	provnum
msr_cd	msr_cd
msr_descr	msr_descr
avg_value	ulcer
ownership	owner
inhosp	hosp
adj_total	total
adj_rn	rn

The first step is to create a new column. The new column will be created to the right of the existing owner column. To insert a new column right click on the F column header. Next, select Insert. This will generate a blank column of data to the right of the existing owner column. Add a header to that column by typing owner_group in cell F1. In cell F2, we will enter the following formula to recode the owner column:

```
=IF(ISNUMBER(SEARCH("For profit",E2)), "For profit", IF(ISNUMBER(SEARCH("Non profit",E2)), "Non profit", IF(ISNUMBER(SEARCH("Government",E2)), "Government", "")))
```

This formula will search for specific text in part of cell E2. When the text is found, the function will return an output. For instance, if the ownership status contains the words "for profit" the ownership status will be recoded to just "for profit." If the ownership status contains the word "nonprofit" the ownership status will be recoded to just "nonprofit." If the ownership status contains the word "government" the ownership status will be recoded to just "government." **Figure 13.2** displays a screenshot of the Excel spreadsheet with this formula displayed in the formula toolbar.

After recoding the data, the data preparation phase is complete. Remember to save the file as a CSV. In Excel, click File > Save As. Under the Format drop-down window, select the Comma Separated Values option. When saving the file, name the document nurse_staff and save the file to your computer's desktop. When saving your file, you may receive a message:

Some features in your workbook might be lost if you save it as CSV (Comma defined). Do you want to keep using that format?

1	provnum	msr_cd	msr_descr	avg_value	ownership	owner_group	inhosp	adj_total	adj_in
2	15009	403	Percent of	5.7	For profit - Corporation	For profit	NO	3.80644	0.61047
3	335392	403	Percent of	5.7	For profit - Corporation	For profit	NO	2.58548	0.2412
4	335393	403	Percent of	4.3	For profit - Corporation	For profit	NO	3.92595	0.44411
5	335394	403	Percent of	2.4	Non profit - Corporation	Non profit	NO	4.16761	0.47048
6	335396	403	Percent of	19.9	For profit - Corporation	For profit	NO	3.1949	0.34541
7	335397	403	Percent of	10.7	For profit - Individual	For profit	NO	2.69132	0.15481
8	335398	403	Percent of	8.2	For profit - Corporation	For profit	NO	3.32308	0.27157
9	335399	403	Percent of	9.9	For profit - Partnership	For profit	NO	3.99555	0.22652
10	335400	403	Percent of	8.3	For profit - Corporation	For profit	NO	2.95724	0.33862
11	335401	403	Percent of	9.3	Non profit - Other	Non profit	NO	3.54686	0.47761
12	335402	403	Percent of	6.4	For profit - Corporation	For profit	NO	4.63343	0.67497
13	15010	403	Percent of	4.5	Non profit - Other	Non profit	YES	8.96123	1.54976
14	335403	403	Percent of	4.8	Non profit - Corporation	Non profit	YES	5.39035	0.74719
15	335404	403	Percent of	12.3	For profit - Corporation	For profit	NO	3.68712	0.36513
16	335405	403	Percent of	3.9	Non profit - Corporation	Non profit	NO	2.90712	0.2528
17	335406	403	Percent of	6.1	Government - County	Government	NO	3.90435	0.33929
18	335407	403	Percent of	3.4	For profit - Corporation	For profit	NO	3.04702	0.48998
19	335408	403	Percent of	9.4	For profit - Corporation	For profit	NO	3.87411	0.43312
20	335409	403	Percent of	9.3	For profit - Corporation	For profit	NO	2.40703	0.33602
21	335410	403	Percent of	10.3	For profit - Corporation	For profit	NO	3.19519	0.37431
22	335411	403	Percent of	13.9	For profit - Corporation	For profit	NO	2.96032	0.56168
23	335412	403	Percent of	3.1	Government - County	Government	NO	4.16762	0.38446
24	335413	403	Percent of	5.6	For profit - Partnership	For profit	NO	6.4233	0.32995
25	335416	403	Percent of	5	For profit - Corporation	For profit	NO	2.96459	0.39394
26	335418	403	Percent of	6.4	Non profit - Corporation	Non profit	NO	4.65384	0.48979
27	335419	403	Percent of	16.3	Non profit - Church related	Non profit	YES	3.56626	0.48167
28	335421	403	Percent of	1.9	For profit - Corporation	For profit	NO	3.2433	0.28211
29	335422	403	Percent of	7.9	For profit - Corporation	For profit	NO	2.96987	0.27954
30	335423	403	Percent of	7.1	Government - County	Government	NO	3.89124	0.57257
31	335424	403	Percent of	8.9	Non profit - Corporation	Non profit	NO	2.87422	0.27237
32	335425	403	Percent of	6.3	Government - County	Government	NO	4.37749	0.62455
33	15014	403	Percent of	4.9	For profit - Corporation	For profit	NO	4.64293	0.72849
34	335426	403	Percent of	14.9	For profit - Corporation	For profit	NO	3.08391	0.37253

Figure 13.2. A screenshot of Excel after recoding the owner column
Source: Used with permission from Microsoft.

PRINTED BY: jjewett@herzing.edu. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.