

PHPM 631

Health Information Management Systems

Lab 3: Data to Decision III

Michelle Mellers (mmellers@tamu.edu)
Postdoctoral Research Associate

Population Informatics Lab (<https://pinformatics.org/>)

Course URL: <http://pinformatics.org/phpm631>

License:
Health Information Technology by Hye-Chung Kum is licensed under a
Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License



1

1

Lab 1 Feedback

- Great Job!!
 - Everyone had nice looking graphs!
- A few Comments
 - Pay attention to which elements are selected (ie, which years of data)
 - Try to arrange graphs in dashboard so all elements are easily readable

2

2

Background for Lab 3: Data to Decision

- Reminder: Submit lab by end of Sunday, 2/9
- Builds on lab 1 and 2
 - Create new calculated variables
 - If...else (repeat from lab 1)
 - Use sets to create graphs
 - Examine distributions and handle missing data
- Goal: learn to explore and make decisions using data

3

3

Create New Calculated Variables

- Goal: How do we create a variable to group aspects of the data?
- If...else structure
 - Classic programming structure to create new variables based on conditions
 - Form: if (statement) then {action 1} else {action 2}
 - Can be read as: if statement is true then "action 1" is performed, and if statement is not true then "action 2" is performed
 - The "statement" is a statement that can be evaluated to be either true or false

4

4

Create New Calculated Variables

- Create variable in lab to determine if hospitals contain specialty beds:
 - if ((int(psych+accute+peds+ob) = 0)AND(other = 0) then 0
 - elseif (int(psych+accute+peds+ob) = 0)AND(other != 1) then 1
 - else 2
 - end
- Can be interpreted as:
 - Assigning 0 if no specialty or other beds
 - Assigning 1 if no specialty, and other beds
 - Assigning 2 if specialty beds

5

5

Create New Calculated Variables

- The new variable can be used to easily segment the dataset based on characteristics
- Example from other sources:
 - Grades are reported as A, B, C, D, etc although each of these commonly represent a percent
 - Can allow to ask if there are differences between those getting high grades and low grades.

6

6

Use Sets to Create Graphs

- Divide datasets based on desired conditions
- Can be used to easily segment datasets and determine if differences within data
- In lab:
 - Use the previously created variable determining if “specialty” beds are present to divide the dataset
 - Hospitals may be different based on what type of beds they contain
 - Creation of variable lets us easily graph the differences between the two types

7

7

Examine Distributions and Handle Missing Data

- After we create a new parameter showing proportion of general beds to specialty beds.
 - How can we interpret the parameter showing the ratio of general beds to specialty beds?
- Can we report a general mean and is this meaningful?
 - This question can be answered by plotting a distribution of the data
 - A distribution of the data is a plot showing counts for each value
 - If values are “extreme” reporting the mean may not be meaningful

8

8

Examine Distributions and Handle Missing Data

- Missing data from dataset should be evaluated.
 - Different software packages can make different assumptions with missing data so shown calculations may differ
 - Most simply, missing data is excluded from calculations
- Tableau indicates if there is missing data, go through process within lab

9

9

Lab 3

- Ending Questions
 - Goal is to learn to interpret numbers from the data and graphs
 - Always ask: How is this meaningful? What am I showing?
 - Determining basic set of descriptions from datasets using Tableau
 - Ask for calculation of numbers
 - Know basic numbers from dataset to “learn” the data
 - Asks for interpretation of graphs
 - Graphs allow for easily seeing differences that may otherwise be hidden in data

10

10