Multivariate Analysis

Mosaic Plots

Ira A. Fulton Schools of Engineering
Arizona State University

Objective

Objective

Apply methods of visualizing discrete data values along two axes

Introduction to Mosaic Plots

Graphical display that allows you to examine the relationship among two or more categorical variables

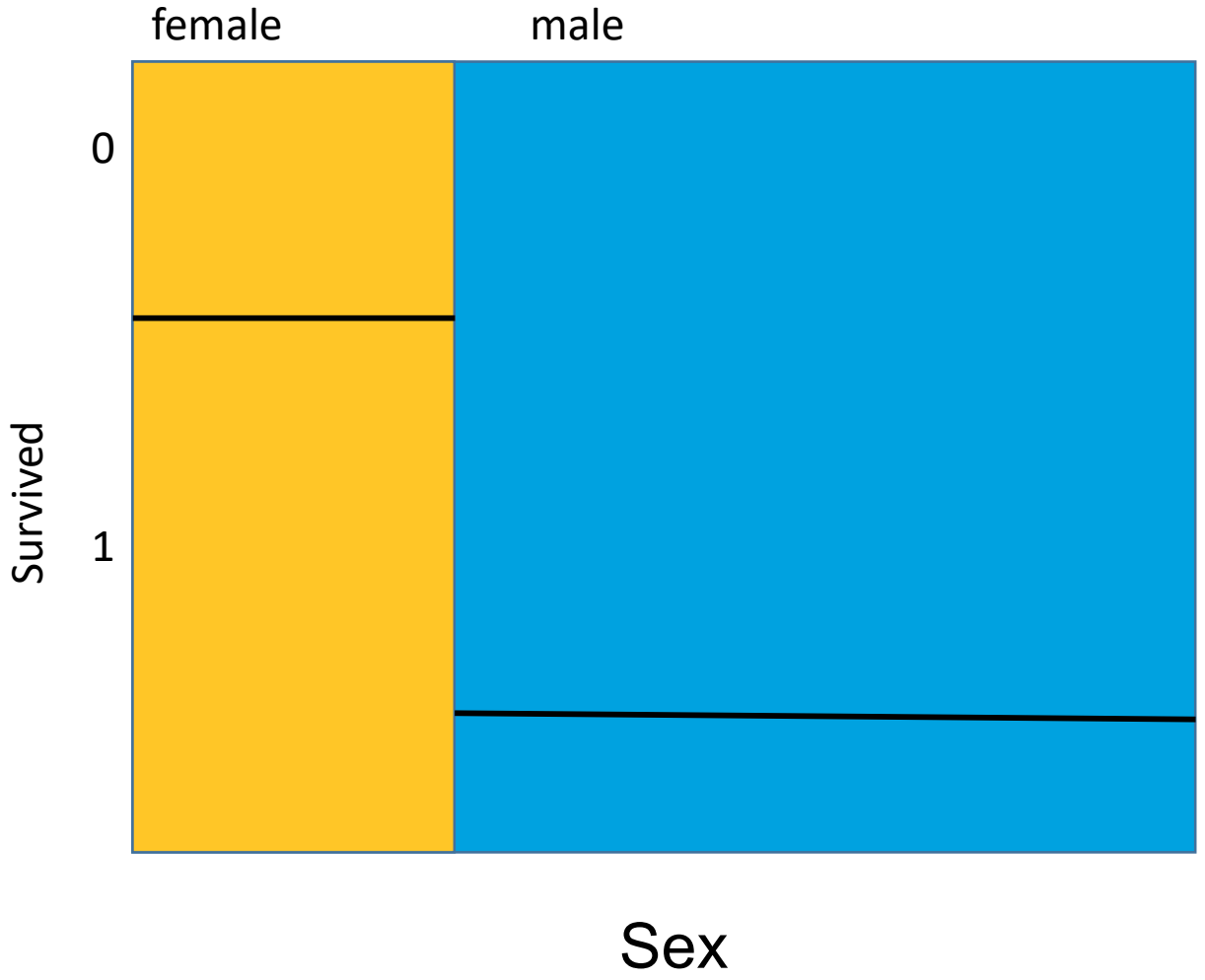
To create:

- Start as a square with length one
- Divide first into horizontal bars whose widths are proportional to the probabilities associated with the first categorical variable
- Next each bar is split vertically by the conditional probability of the second categorical variable

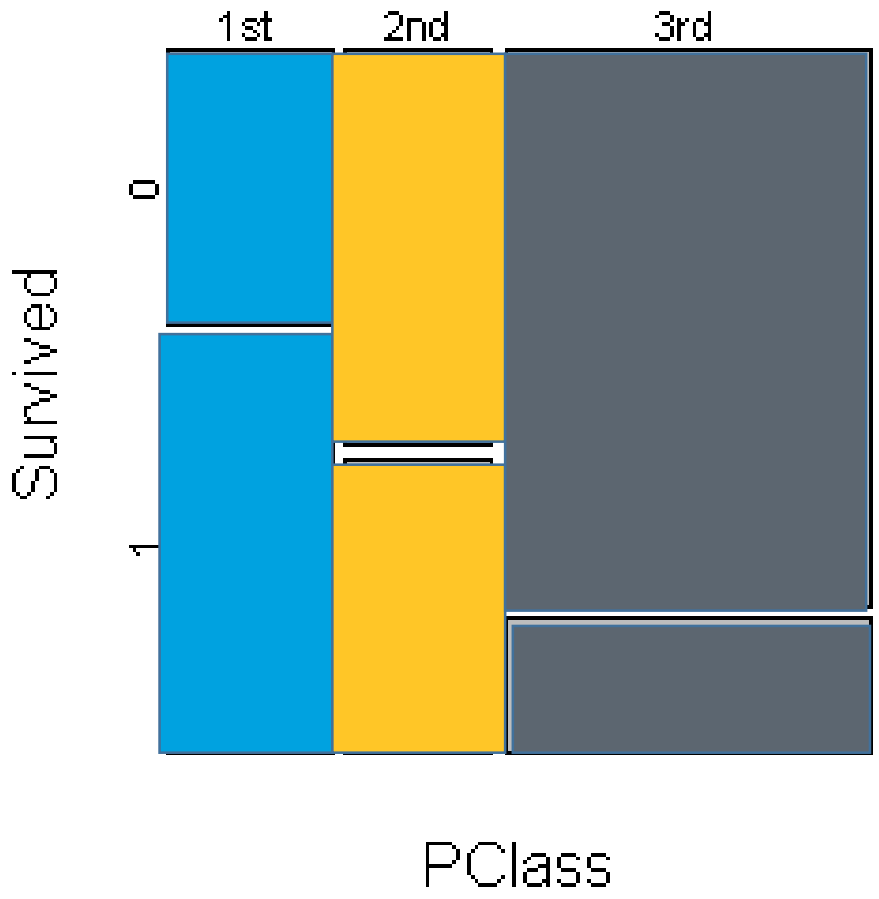
Example: Mortality rates

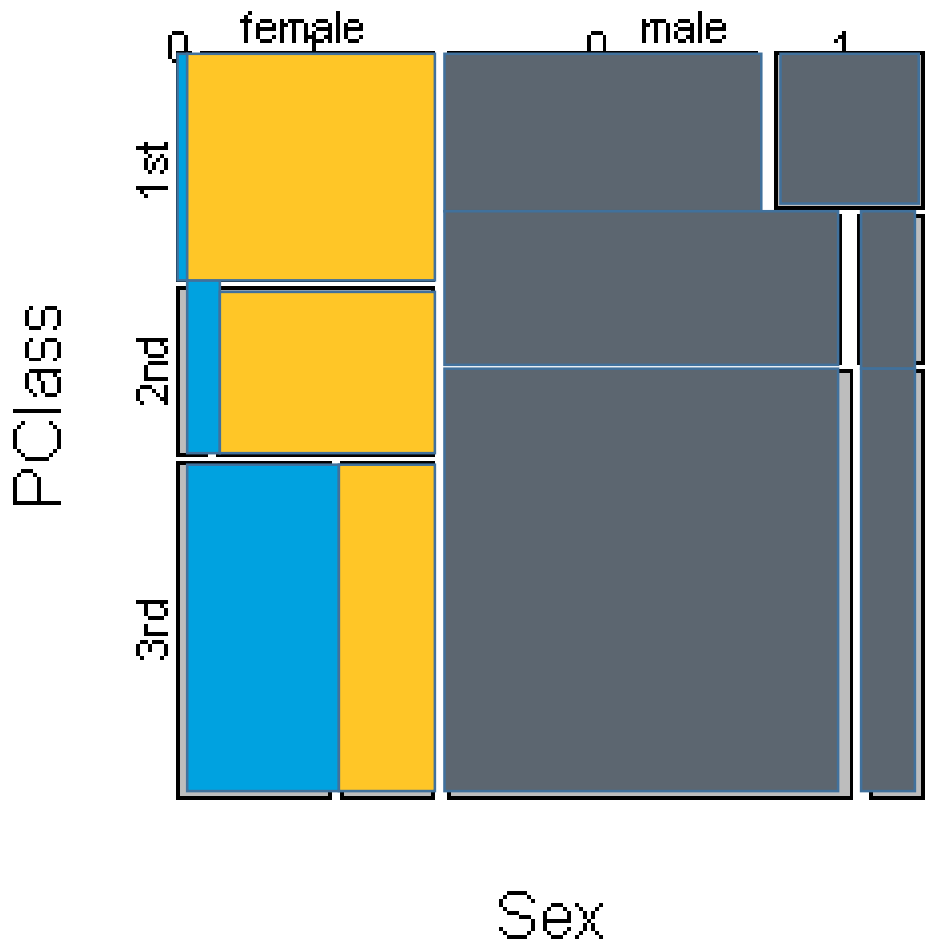
Adults	*Survivors*		*Non-Survivors*	
	Male	*Female*	*Male*	*Female*
1st Class	57	140	118	4
2nd Class	14	80	154	13
3rd Class	75	76	387	89
Crew	192	20	670	3

Children	*Survivors*		*Non-Survivors*	
	Male	*Female*	*Male*	*Female*
1st Class	5	1	0	0
2nd Class	11	13	0	0
3rd Class	13	14	35	17
Crew	0	0	0	0



Examples


1st
2nd
3rd
0
1
Survived
PClass


female
male
0
1
1st
2nd
3rd
PClass
Sex

Mosaic Plots

- It is tempting to dismiss mosaic plots because they represent counts as rectangular areas and so provide a distorted perceptual encoding
- In fact, the important encoding is the length
- At each stage, the comparison of interest is of the length of the sides