

---

# Visualizing Multi-Way Contingency Tables

David Meyer<sup>1</sup>, Achim Zeileis<sup>2</sup>, and Kurt Hornik<sup>2</sup>

<sup>1</sup> Department of Information Systems and Process Management

<sup>2</sup> Department of Statistics and Mathematics

Wirtschaftsuniversität Wien

Augasse 2-6, 1090 Vienna, Austria

{David.Meyer,Achim.Zeileis,Kurt.Hornik}@wu-wien.ac.at

## 1 Introduction

Categorical data analysis is typically based on two- or higher-dimensional contingency tables, cross-tabulating the co-occurrences of levels of nominal and/or ordinal data. In order to explain these, statisticians typically look for (conditional) independence structures using common statistical methods such as independence tests and log-linear models. The analysis quickly becomes complex with increasing dimensionality. One idea of visualization techniques is to use the human visual system to detect structures in the data that possibly are not obvious from solely numeric output (e.g., test statistics). Whether the task is purely exploratory or model based, techniques such as mosaic, sieve, and association plots offer good support for visualization. All three have been extended over the last two decades, and implementations exist in many statistical environments.

All three graphical methods visualize aspects of (possibly higher-dimensional) contingency tables. A *mosaic plot* [13] is basically an area-proportional visualization of (typically observed) frequencies, composed of tiles (corresponding to the cells) created by recursive vertical and horizontal splits of a square. Thus, the area of each tile is proportional to the corresponding cell entry *given* the dimensions of previous splits. *Sieve plot* are similar to mosaic plots, but the area of each tile is proportional to the *expected* cell entry, and each tile is filled with a number of rectangles corresponding to the observed value. An *association plot* [4] visualizes the standardized deviations of observed frequencies from those expected under a certain independence hypothesis. Each cell is represented by a rectangle that has (signed) height proportional to the residual and width proportional to the square root of the expected counts, so that the area of the box is proportional to the difference in observed and expected frequencies.

Over the years, extensions to these techniques were mainly focused on five aspects:

1. Varying the shape of bar plots and mosaic displays to yield, e.g., double-decker plots [14], spine plots [31], or spinograms.
2. Using residual-based shadings to visualize log-linear models [6, 8] and significance of statistical tests [22].
3. Using pairs plots and trellis-like layouts for marginal, conditional and partial views [7].
4. Adding direct user interaction, allowing quick exploration and modification of the visualized models [34, 32].
5. Providing a modular and flexible implementation to easily allow user extensions [24].

Current implementations of mosaic displays can be found, e.g., for SAS [29], ViSta [35], MANET [34], Mondrian [32], R [30], and S-PLUS [17]. Table 1 gives an overview of the available functionality in these systems. The figures in this chapter have all been produced using the R system, using the extension packages `vcd` [23] and `scatterplot3d` [20] (Fig. 2 only), all freely available from the Comprehensive R Archive Network (<http://cran.r-project.org/>).

	SAS	S-PLUS	R	ViSta	MANET/Mondrian
Basic functionality	×	×	×	×	×
Shape			×		×
Residual-based shadings	×		×	×	(×)
Conditional Views	×		×		×
Interaction				×	×
Extensible Design			×		

**Table 1.** Comparison of current software environments.

Our contribution will give an overview of the state of the art of mosaic, sieve, and association plots, both for exploratory visualization and model-based analysis. Exploratory techniques will include specialized displays for the bivariate case, as well as pairs plot-like displays for higher-dimensional tables. As for the model-based tools, particular emphasis will be given to methods suitable for the visualization of conditional independence tests (including permutation tests), as well as for the visualization of particular GLMs (such as log-linear models). In Sect. 2, we start with the simple bivariate case. Section 3 explains how the use of color in residual-based shadings can support data exploration, and even promotes the methods to diagnostic and model-based tools by visualizing test statistics and residuals of independence models. In Sect. 4, we show how the basically bivariate methods straightforwardly extend to the multivariate case by using ‘flat’ representations of the multi-way tables. In this section, we also introduce specialized displays for conditional independence structures. Sect. 5 concludes the chapter.

## 2 Two-way Tables

Throughout these section, our examples will be based on the hospital data [10] (see Tab. 2).

	Length of stay			TOTAL
	2–9 days	10–19 days	more than 20 days	
Visit frequency				
Regular	43	16	3	62
Less than monthly	6	11	10	27
Never	9	18	16	43
TOTAL	58	45	29	132

**Table 2.** The hospital data.

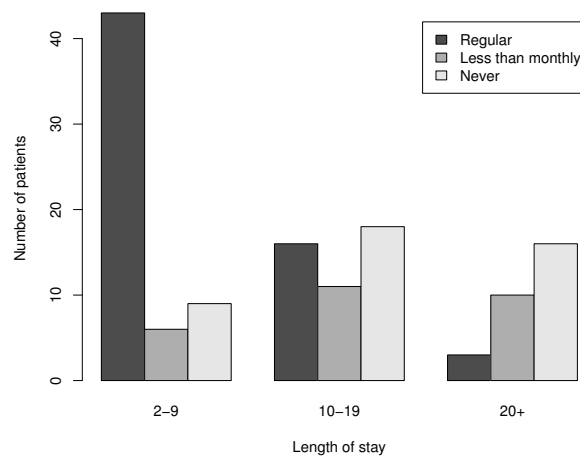
The table relates the length of stay (in days) in hospital and the visit frequency for mental patients, giving evidence to the fact that the longer the length of stay in hospital, the less frequent the visits.

Although far from optimal, contingency tables are frequently visualized using grouped bar charts (see Fig. 1) or even by means of 3D-bar charts (see Fig. 2). It seems hard to detect the aforementioned pattern in these, especially in the 3D plot where the perspective view tends to distort the true proportions of the bars. In the following, we will introduce three graphical methods that are better suited for contingency tables.

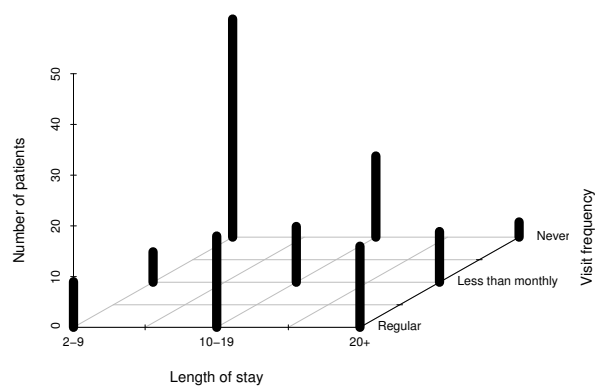
### 2.1 Mosaic Displays

Mosaic displays have been introduced by [12, 13] and extended, e.g., by [6, 7, 8]. They visualize the observed values of a contingency table by area-proportional tiles, arranged in a squared mosaic. The tiles are obtained by recursive partitioning splits of a square. Since chapter XXX in this book (**Reference to the contribution of Heike Hofmann**) is entirely devoted to mosaic plots, in the following we only describe the main idea. Consider our example of the hospital data from above. Step 1 consists of splitting a square according to the marginals of one of the variables. To be consistent with the textual representation, we choose ‘Length of stay’ with vertical splits (see Fig. 3).

The result is similar to a bar plot where not the height, but the width is adapted to visualize the counts for each level, which some authors [15] call a *spine plot*. From this plot, we see that the number of patients decreases with the length of stay. Step 2 now is to add further splits in the other direction, i.e., horizontal splits, for the second variable. This means that each vertical bar is split according to the marginals of the second variable, *given* the first variable (see Fig. 4). The resulting plot visualizes the contingency table where



**Fig. 1.** Barplot for the hospital data.



**Fig. 2.** 3D-barplot for the hospital data.

each cell has a size proportional to the corresponding table entry. We can still depict the marginal distribution of ‘Length of stay’, and additionally, the visit frequency *given* each category of ‘Length of stay’. As we will see in the next section, the grid would be regular if the two variables were independent. Clearly, compared to a length of stay of 10–19 days, more patients get regular visits for stays from 2–9 days, and conversely, less patients get regular visits for stays for more than 20 days. For patients that get no visits, the pattern is inversed.

Since mosaic plots are asymmetric by construction, the choice of the variable order matters, as the first splitting variable dominates the plot. In our example, if we use ‘Visit frequency’ as the first splitting variable, the impression is very different compared to the previous mosaic (see Fig. 5). In this alternative display, we see the marginal distribution of ‘Visit frequency’ in the rows: about half of the patients get visited regularly. This group is dominated by patients staying between 2 and 9 days. It seems apparent that the distribution of ‘Length of stay’ is similar for monthly and never visited patients, so this two categories actually represent one homogenous group (patients visited only casually).

## 2.2 Sieve Plots

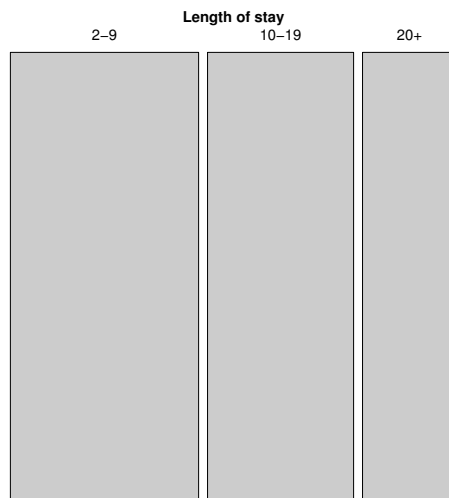
When we try to explain data, we suppose the validity of a certain model for the generating process. In the case of two-way contingency tables, the two most common hypotheses are

1. independence of the two variables, and
2. homogeneity of one variable among the strata defined by the second.

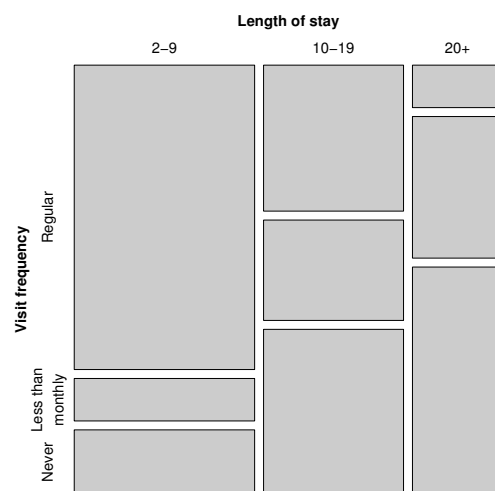
It is easy to compute the *expected* table under either of these hypotheses. Consider a 2-way contingency table with  $I$  rows and  $J$  columns, cell frequencies  $\{n_{ij}\}$  for  $i = 1, \dots, I$  and  $j = 1, \dots, J$ , and row and column sums  $n_{i+} = \sum_j n_{ij}$  and  $n_{+j} = \sum_i n_{ij}$ , respectively. For convenience, the number of observations is denoted  $n = n_{++}$ . Given an underlying distribution with theoretical cell probabilities  $\pi_{ij}$ , the null hypothesis of independence of the two categorical variables can be formulated as

$$H_0 : \pi_{ij} = \pi_{i+}\pi_{+j}. \quad (1)$$

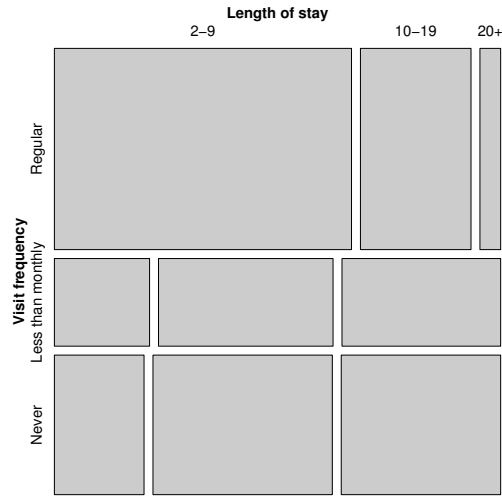
Now, the expected cell frequencies in this model are simply  $\hat{n}_{ij} = n_{i+}n_{+j}/n$ . The expected table for our sample data is given in Tab. 3. It could again be visualized using a mosaic plot, this time applied to the table of expected frequencies. In Fig. 6, we can see such a plot in which, additionally, a number of squares equal to the corresponding cell entry has been added to each tile. We see that the grid formed by the tiles is completely regular. The added rectangles in the tiles become more useful when they are not based on the expected



**Fig. 3.** Construction of a mosaic plot for a two-way table, step 1.



**Fig. 4.** Construction of a mosaic plot for a two-way table, step 2.



**Fig. 5.** Mosaic plot for the hospital data, using ‘Visit frequency’ as first splitting variable.

	Length of stay			TOTAL
	2-9	10-19	20+	
Visit frequency				
Regular	27.24	21.14	13.62	62
Less than monthly	11.86	9.20	5.93	27
Never	18.89	14.66	9.45	43
TOTAL	58.00	45.00	29.00	132

**Table 3.** The hospital data—expected values.

but the observed values instead (see Fig. 7). Such a *sieve plot* implicitly compares expected and observed values since the density of the grid will increase with the deviation of the observed from the expected values. This allows the detection of general association patterns (for nominal variables) and of linear association (for ordinal variables), the latter producing tiles of either very high or very low density along one of the diagonals. In the case of our data, the plot reveals indeed a negative association of the two variables since the density of the rectangles is marked along the secondary diagonal, giving evidence to the fact that for these patients, visit frequency decreases with the length of stay.

### 2.3 Association Plots

In the last section, we have seen how to compare observed and expected values of a contingency table using sieve plots. We can do this more straightforwardly

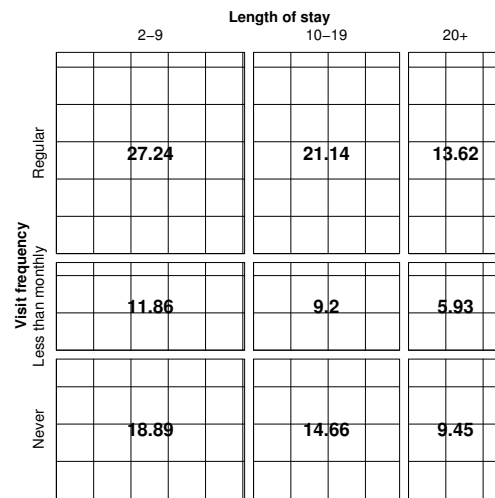


Fig. 6. Expected values for the hospital data.

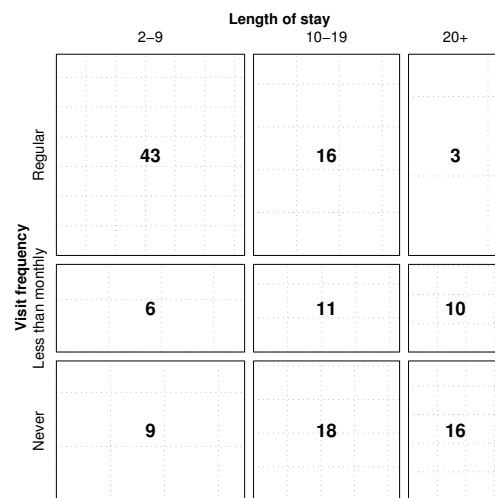


Fig. 7. Sieve plot for the hospital data.



by using a plot that directly visualizes the residuals. The most widely known residuals are the Pearson residuals

$$r_{ij} = \frac{n_{ij} - \hat{n}_{ij}}{\sqrt{\hat{n}_{ij}}}. \quad (2)$$

that are standardized raw residuals. They are best visualized by association plots [4]: each cell is represented by a rectangle that has (signed) height proportional to the corresponding Pearson residual  $r_{ij}$  and width proportional to the square root of the expected counts  $\sqrt{\hat{n}_{ij}}$ . Thus, the area is proportional to the raw residuals  $n_{ij} - \hat{n}_{ij}$ . In available implementations, the sign of a residual is often redundantly coded, e.g., by the rectangle's color (see Sect. 3.1) and its position relative to the baseline. Figure 8 shows the association plot for the hospital data. Consistent with the corresponding mosaic and sieve plots, we clearly see that too many (too less) patients that stay from 2 to 9 (more than 20) days get visited regularly than would be expected under the null of independence, and that this pattern is reversed for patients never visited.

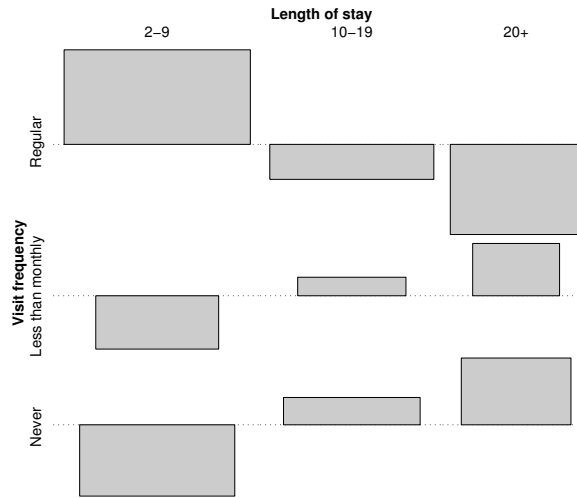


Fig. 8. Association plot for the hospital data.

### 3 Residual-based Shadings

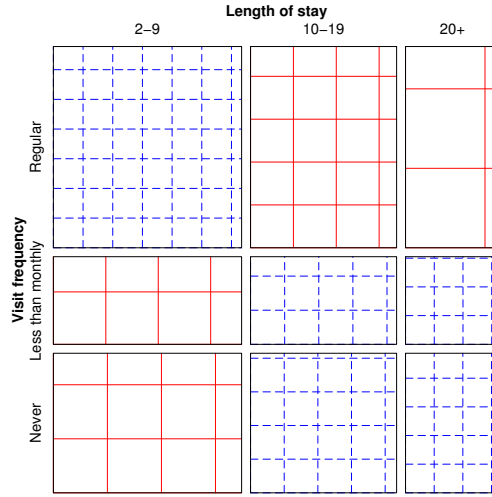
As introduced in the previous section for association plots, the investigation of residuals from a posited independence model is of major interest in analyzing

contingency tables. In the following, we will demonstrate how the use of colors can greatly facilitate the detection of interesting patterns.

### 3.1 Using Colors for Residuals

All plots introduced in the previous section can be improved using colors, either by adding additional information, or by redundantly coding information already visualized by the ‘raw’ plot to support our perceptual system.

First, we consider the sieve plots. The density of the grid in the raw version implicitly gives us an idea of the residuals’ size, but since the plot does not include the density corresponding to zero residuals (the null model) for comparison, we cannot easily assess whether there are more, or less counts in a cell than expected under the null hypothesis. It would help if we knew the sign of the residuals. Using color, we can add this information, for example using blue for positive, red for negative, and gray for zero residuals. Figure 9 demonstrates the effect of applying such a color shading to a sieve plot for the hospital data. Color clearly emphasizes the linear association pattern in this data.

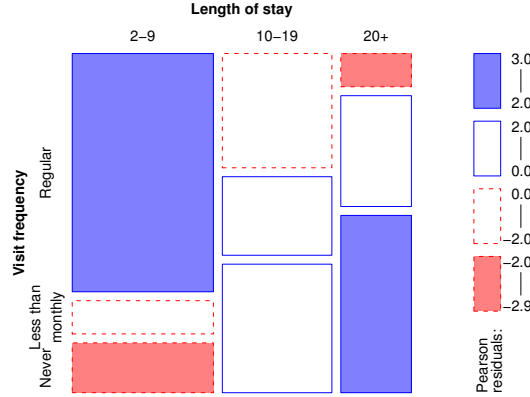


**Fig. 9.** Sieve plot with color coding of the residuals.

Mosaic plots in their initial version are monochrome displays. Friendly [6] introduced a residual-based shading of the tiles to additionally visualize the residuals from a given independence model fitted to the table. The idea is to use a color coding for the mosaic tiles that visualizes the sign and absolute size of each residual  $r_{ij}$ : Cells corresponding to small residuals ( $|r_{ij}| < 2$ ) are

shaded white. Cells with medium sized residuals ( $2 \leq |r_{ij}| < 4$ ) are shaded light blue and light red for positive and negative residuals, respectively. Cells with large residuals ( $|r_{ij}| \geq 4$ ) are shaded with a fully saturated blue and red, respectively. The heuristic for choosing the cut offs 2 and 4 is that the Pearson residuals are asymptotically standard normal which implies that the highlighted cells are those with residuals *individually* significant at approximately the  $\alpha = 0.05$  and  $\alpha = 0.0001$  levels. However, the main purpose of the Friendly shading is not to visualize significance but the *pattern* of deviation from independence [8]. In addition to the shading of the rectangles themselves, the Friendly shading also encompasses a choice of line type and line color of the rectangle borders with similar ideas as described above.

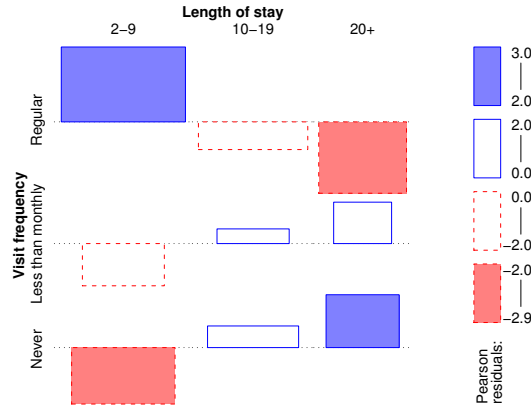
In Fig. 10, we again show the mosaic for the hospital data, this time using the Friendly color coding.



**Fig. 10.** Mosaic display with Friendly color coding of the residuals.

Clearly, the asymmetry for regular and never visited students, and the pattern inversion for lengths of stay of 2–9 and more than 20 days are emphasized using the color shading.

For association plots, residual-based shadings are redundant since all relevant information is already contained in the plot by construction. But using one of the shadings discussed above will nevertheless support the analysis process and is therefore recommended. For example, applying the Friendly shading in Fig. 11 facilitates the discrimination between positive and negative residuals, and therefore to get a quick overview on the discussed deviation pattern.

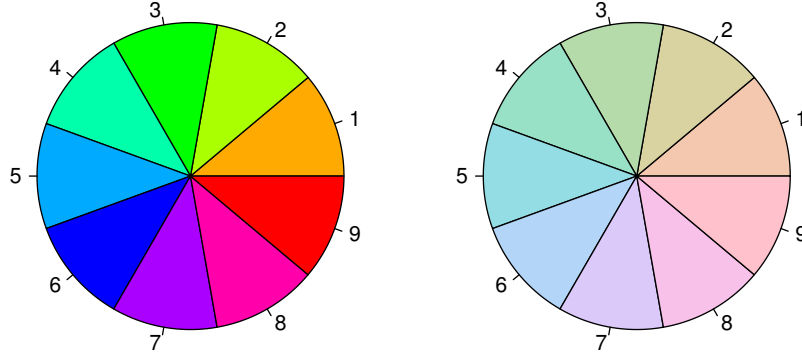


**Fig. 11.** Association plot with Friendly color coding of the residuals.

### 3.2 A Note on Color Palettes

The color shading introduced in the previous section but can still be improved. One of the issues is the suboptimal color palette used in the Friendly shading. Implementations are often based on so-called Hue-Saturation-Value (HSV) colors (or the comparable Hue-Luminance-Saturation color scheme). Both spaces are rather similar transformations of RGB (Red-Green-Blue) space [2, 28] and are very common implementations of colors in many computer packages [25] making the generation of the Friendly shading very simple. Each color in this space is represented by three dimensions: the hue, the saturation (‘colorfulness’), and the value (‘lightness’, amount of gray). But HSV colors have several disadvantages. Most importantly, HSV colors are not perceptually uniform because the three HSV dimensions map only poorly to the three perceptual dimensions of the human visual system [2, 16]. Consequently, the HSV dimensions are confounded, e.g., saturation is not uniform across different hues. For example, see Fig. 12, left-hand side, showing a qualitative color palette in the HSV space: although saturation and value are fixed, the fully saturated blue is perceived much darker than the fully saturated red or green. This makes it more difficult for the human eye to judge the size of shaded areas and can therefore lead to color-caused optical illusions when used in statistical graphs [3]. Furthermore, flashy fully saturated HSV colors are good for drawing attention to a plot, but hard to look at for a longer time [16] which makes HSV-shaded graphics harder to interpret.

The use of colors that are more ‘in harmony’ go back to Munsell [26] who introduced a color notation for balanced colors. Based on those, tools



**Fig. 12.** Qualitative color palette for the HSV (left hand side) and HCL (right hand side) spaces.

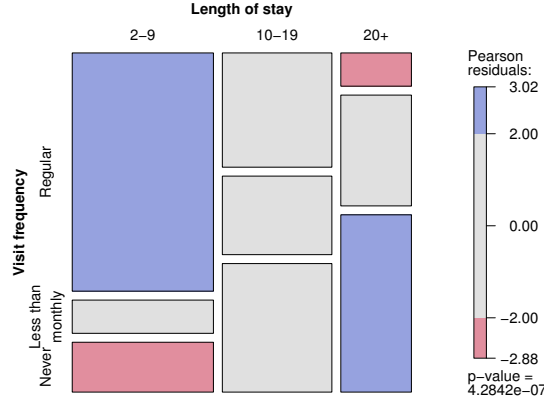
producing better palettes for specific tasks have been developed [11]. Other perceptually-based color spaces, especially suited for displays, are the CIELAB and CIELUV spaces [5] from which qualitative palettes for statistical graphics have been derived [16]. A transformation of the CIELUV space leads to the HCL (Hue-Chroma-Luminance) space. HCL colors with fixed chroma and luminance are always balanced towards the same grey and thus do not have the problem of varying saturations as the HSV colors (see Fig. 12, right hand side). Similarly, diverging color palettes can be derived that are suited for visualizing residuals in a similar way than the Friendly shading [36]. Figure 13 shows again the mosaic plot for the hospital data, this time using HCL colors and an alternative legend, showing the whole range of residuals.

### 3.3 Visualizing Test Statistics

Figure 13 includes the  $p$  value of a  $\chi^2$  test of independence, frequently used to assess the significance of the hypothesis of independence (or homogeneity for stratified data) in two-way tables. The test statistic is just the sum of the squared Pearson residuals

$$X^2 = \sum_{i,j} r_{ij}^2, \quad (3)$$

known to have a limiting  $\chi^2$  distribution with  $(I - 1)(J - 1)$  degrees of freedom under the null hypothesis. Since the HCL space is three-dimensional and we only used two dimensions so far for coding information (hue for the sign and colorfulness for the size of the residuals), we can use the third dimension (lightness) for visualizing the significance of some specified test statistic, for example the  $\chi^2$  test statistic, using darker (‘uninteresting’) colors for non-significant results. The following example using the Bundesliga data [18] shows



**Fig. 13.** Mosaic display with HCL shading of the residuals.

the relationship of home goals and away goals of Germany’s premier soccer league in 1995: although there are two “larger” residuals (one greater than 2, one less than  $-2$ ), the  $\chi^2$  test does not reject the null hypothesis of independence on a significance level of 95% since the  $p$  value is 0.12133. Consequently, the employed shading uses a darker overall color scheme (see Fig. 14).

The previous example shows that the heuristic for choosing the cut-off points in the Friendly shading may lead to wrong conclusions: the test of independence is not significant even though some of the residuals are “large”. The reason here is that the cut-off points are really data-dependent. Another example where the Friendly-shading may trap the rush statistician is the case of the arthritis data [19], resulting from a double-blind clinical trial investigating a new treatment for rheumatoid arthritis, stratified by gender. Figure 16 visualizes the results for the female patients, again by means of a mosaic display. Clearly, the hypothesis of independence is rejected by the  $\chi^2$  test even on a 99% level ( $p = 0.0035$ ), but since all residuals are in the  $[-1.72, 1.87]$  interval, the tiles remain uncolored. A solution to this issue is to use a different test statistic, for example the maximum of the absolute values of the Pearson residuals [22] instead of the sum of squares:

$$M = \max_{i,j} |r_{ij}|. \quad (4)$$

Given a critical value  $c_\alpha$  for this test statistic, all residuals whose absolute value exceeds  $c_\alpha$  violate the hypothesis of independence at level  $\alpha$  [21, ch. 7]. Thus, the interesting cells giving evidence for the rejection of the independence

hypothesis can easily be identified. The distribution of this test statistic under the null can be obtained by simulation, sampling tables with the same row and column sums  $n_{i+}$  and  $n_{+j}$  using the Patefield algorithm [27] and computing the maximum statistic for each of these tables. In Figs. 15 and 17, we again visualize the Bundesliga and the arthritis data, this time using the maximum test statistic. We see that for the Bundesliga data no cell is colored, and we now can indeed conclude from the visualization that neither of the home and away teams seems advantaged. On the other hand, for the arthritis data, the shading of the tiles clearly shows that the treatment is effective: too many patients in the treatment group exhibit marked improvement than would be expected under independence.

## 4 Multi-way Tables

In Sect. 2, we have presented basic visualization methods for two-way tables. In this section, we will show how these methods extend to multi-way tables by applying them to ‘flat’ representations, and we will treat specialized displays for conditional independence models.

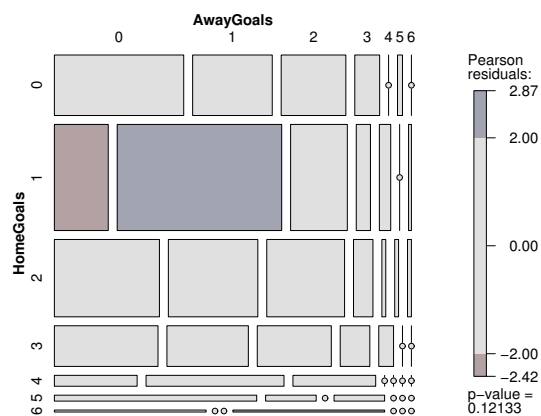
### 4.1 Visualize Flat-Tables

The main idea of the mosaic, sieve, and association plots presented in the previous sections is to visualize information on the tables’ cells, arranged in rectangular form. For multi-way tables, mosaic plots can directly be used by simply adding further splits for each additional variable. For sieve and association plots, we apply the basic idea of mosaic plots to the table itself, i.e., simply nest the variables into rows and columns using recursive conditional splits, given the margins. The result is a ‘flat’ representation of the multi-way table that can be visualized in ways similar to a two-dimensional table. As an example, consider the HairEyeColor data containing two polytomous variables (hair and eye color), as well as one (artificial) dichotomous variable (gender). A ‘flattened’ contingency table, putting eye color in the columns and hair color—nested in gender—in the rows, is given by Tab. 4.

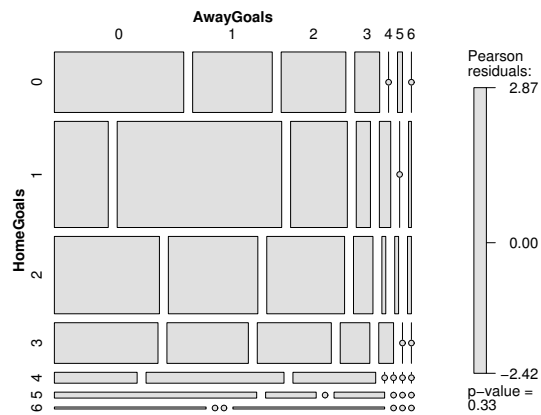
The corresponding mosaic, sieve, and association plots are shown in Figs. 18–20. Note that with three variables, we have the choice of four kinds of independence models that can be fitted to the table: partial, joint, and conditional independence, as well as the model of no three-way interaction. For demonstration purposes, the shadings in these plots visualize the model of hair and eye color being *jointly* independent from gender. The model fits better than others, but still is significant at the 95% level.

### 4.2 Displays for Conditional Independence

Apart from the generic approach to the visualization of multi-way tables outlined in the previous section, there are modifications and specialized displays

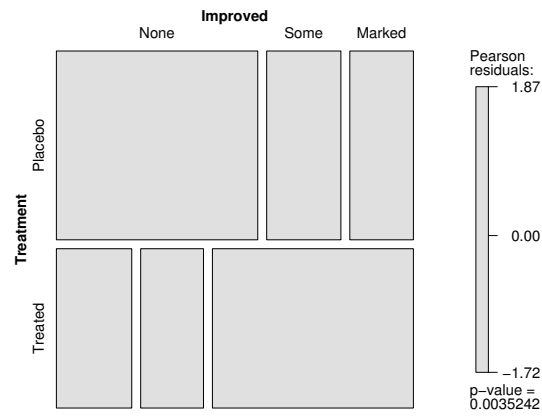


**Fig. 14.** Mosaic plot for part of the Bundesliga, using a  $\chi^2$  test and fixed cut-off points for the residuals.

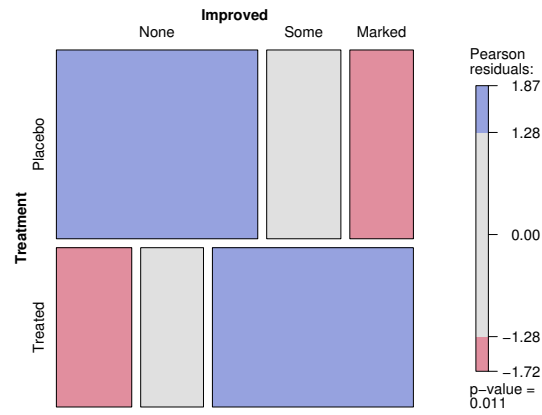


**Fig. 15.** Mosaic plot for part of the Bundesliga, using the maximum statistic and data-driven cut-off points for the residuals.





**Fig. 16.** Mosaic plot for the Arthritis data, using the  $\chi^2$  test and fixed cut-off points for the shading.



**Fig. 17.** Mosaic plot for the Arthritis data, using the maximum test and data-driven cut-off points for the residuals.

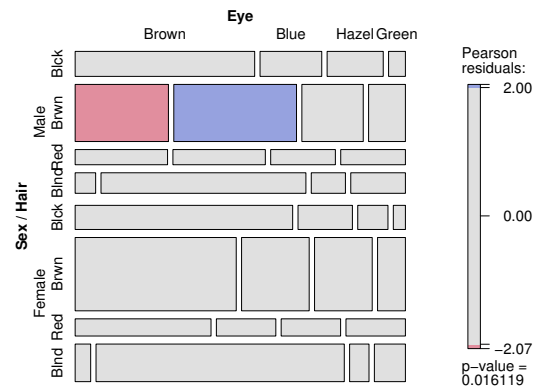


Fig. 18. Mosaic plot for the HairEyeColor data.

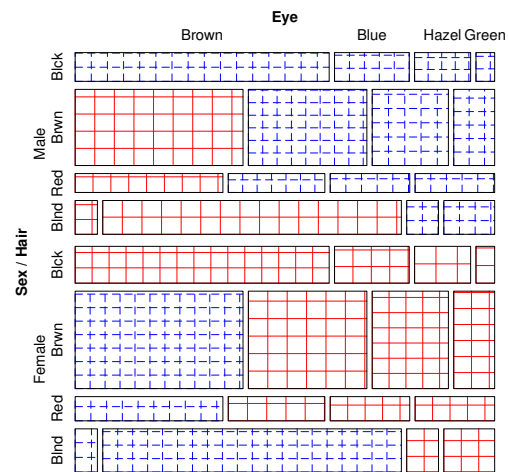


Fig. 19. Sieve plot for the HairEyeColor data.

		EYE			
		Brown	Blue	Hazel	Green
GENDER HAIR					
Male	Black	32	11	10	3
	Brown	38	50	25	15
	Red	10	10	7	7
	Blond	3	30	5	8
Female	Black	36	9	5	2
	Brown	81	34	29	14
	Red	16	7	7	7
	Blond	4	64	5	8

Table 4. The HairEyeColor data, in flat representation.

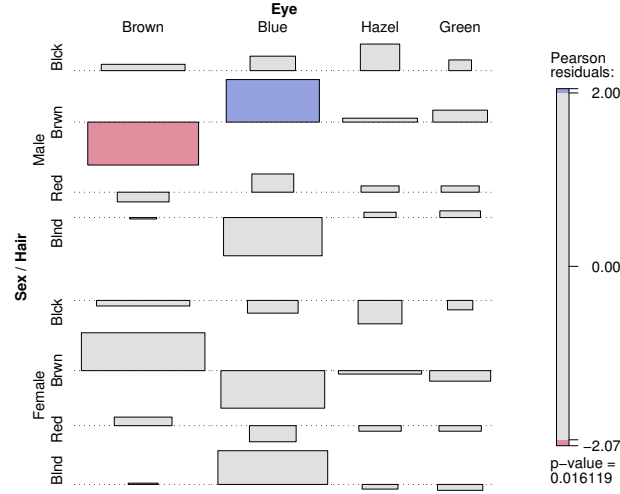
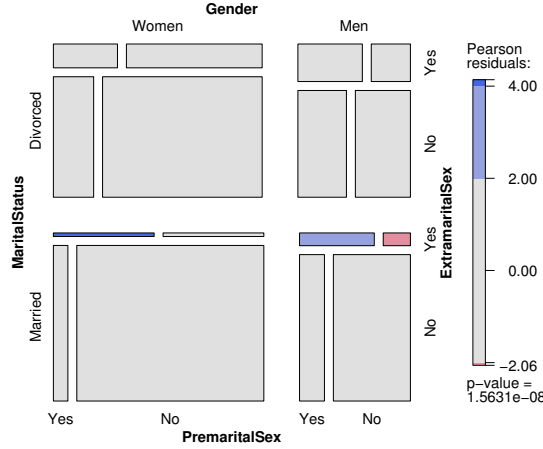


Fig. 20. Association plot for the HairEyeColor data.

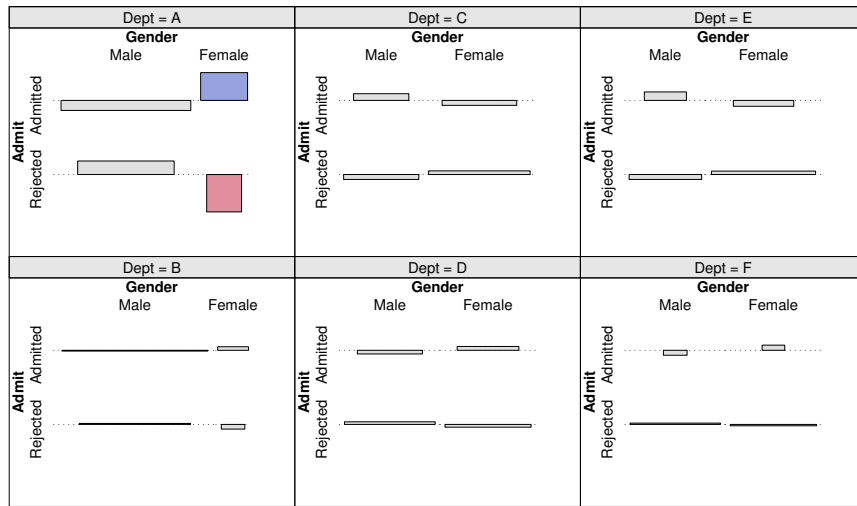
designed for the visualization of conditional independence structures. A first approach is to modify the spacing of the standard mosaic display—which already is a conditional plot by construction—to better distinguish conditional from conditioned variables. Figure 21 illustrates the conditional independence of premarital and extramarital sex, given gender and marital status, in the PreSex data [33, 9]. The  $\chi^2$  test of independence rejects the null hypothesis: possibly, because too many men who had premarital sex also had extramarital sex? For the conditioned variables premarital and extramarital sex, we use a fixed spacing, whereas for the conditioning variables gender and marital status, the spacing is increasing.



**Fig. 21.** Mosaic plot for the pre- and extramarital sex data.

Another possibility is to use trellis-like layouts for visualizing partial tables for given strata, defined by the conditioning variables. Figure 22 visualizes the well-known UCB admissions data [1] by the means of a conditional association plot. The panels show the residuals from a conditional independence model (independence of gender and admission, given department), stratified by department. Clearly, the situation in department A (more women/less men accepted than would be expected under the null) causes the rejection of the hypothesis of conditional independence, indicating that female students are more successful in that department than their male colleagues.

Finally, it is possible for all basic plots to create pairwise displays, arranged in a matrix similar to scatterplots in a pairs plot. The diagonal cells contain the variable names, optionally with univariate statistics, whereas the off-diagonal cells contain plots whose variables are implicitly specified by the cells' position in the matrix. More formally, each cell  $a_{ij}$  in such a matrix defines two variables  $i$  and  $j$  that can be used to specify the model visualized in that cell. Typical hypotheses are: Variables  $i$  and  $j$  are marginally independent; variables  $i$  and  $j$  are conditionally independent, given all others; variables  $i$  and  $j$  are jointly independent from all others, etc. Figure 23 shows a pairs display with mosaic plots visualizing mutual independence in the upper triangle, sieve plots for the same in the lower triangle, and bar charts in the diagonal. Immediately, we can detect the independence of hair and gender and eye and gender from the non-colored cells in the corresponding mosaic plots. Thus, the associational pattern present in the sieve plots are not significant. On the other hand, hair and eye color clearly are not independent.



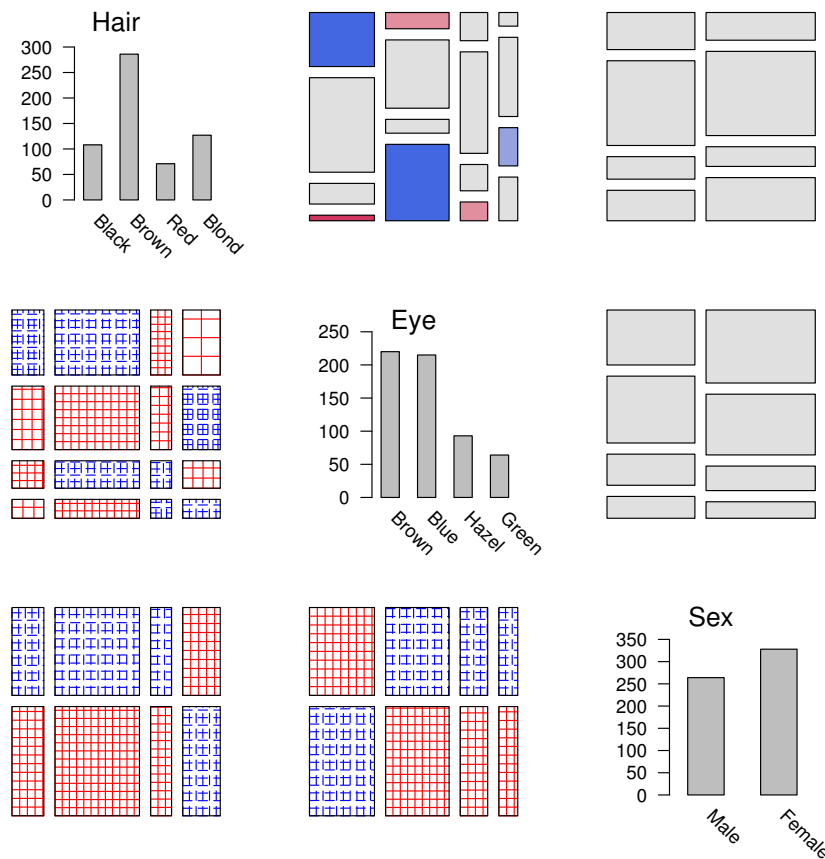
**Fig. 22.** Conditional association plot for the UCB admissions data.

## 5 Conclusion

This chapter reviews several alternatives for the visualization of multi-way contingency tables. For two-way tables, mosaic, sieve, and association plots are suitable for the visualization of observed and expected values and the Pearson residuals, respectively. These basic methods are enhanced by using residual-based shadings, preferably based on perceptual color palettes such as those derived from the HCL space. Residual-based shadings can be used to visualize sign and size of the residuals, as well as the significance of test statistics such as the  $\chi^2$  or the maximum test statistic. The latter has the advantage of detecting residuals causing the rejection of the hypothesis of independence. The methods directly extend to the multi-way case by using ‘flat’ representations of the multi-way tables, and specialized displays for conditional independence such as trellis layouts of partial tables and pairs plots.

## References

1. P.J. Bickel, E.A. Hammel, and J.W. O’Connell. Sex bias in graduate admissions: Data from Berkeley. *Science*, 187(398–403), 1975.
2. Cynthia A. Brewer. Color use guidelines for data representation. In *Proceedings of the Section on Statistical Graphics, American Statistical Association*, pages 55–60, Alexandria, VA, 1999.
3. William S. Cleveland and Robert McGill. A color-caused optical illusion on a statistical graph. *The American Statistician*, 37:101–105, 1983.



**Fig. 23.** Pairs plot for the HairEyeColor data.

4. A. Cohen. On the graphical display of the significant components in a two-way contingency table. *Communications in Statistics—Theory and Methods*, A9:1025–1041, 1980.
5. Commission Internationale de l’Éclairage. *Colorimetry*. Publication CIE 15:2004, Vienna, Austria, 3rd edition, 2004. ISBN 3-901-90633-9.
6. Michael Friendly. Mosaic displays for multi-way contingency tables. *Journal of the American Statistical Association*, 89:190–200, 1994.
7. Michael Friendly. Extending mosaic displays: Marginal, conditional, and partial views of categorical data. *Journal of Computational and Graphical Statistics*, 8(3):373–395, 1999.
8. Michael Friendly. *Visualizing Categorical Data*. SAS Institute, Cary, NC, 2000.
9. G.N. Gilbert. *Modelling Society: An Introduction to Loglinear Analysis for Social Researchers*. Allen and Unwin, London, 1981.
10. S.J. Haberman. Log-linear models for frequency tables with ordered classifications. *Biometrics*, 30:689–600, 1974.

11. Mark A. Harrower and Cynthia A. Brewer. **ColorBrewer.org**: An online tool for selecting color schemes for maps. *The Cartographic Journal*, 40:27–37, 2003.
12. J. A. Hartigan and B. Kleiner. Mosaics for contingency tables. In W.F. Eddy, editor, *Computer Science and Statistics: Proceedings of the 13th Symposium on the Interface*, pages 268–273, New York, 1981. Springer.
13. J. A. Hartigan and B. Kleiner. A mosaic of television ratings. *The American Statistician*, 38:32–35, 1984.
14. Heike Hofmann. Generalized odds ratios for visual modelling. *Journal of Computational and Graphical Statistics*, 10:1–13, 2001.
15. J. Hummel. Linked bar charts: Analysing categorical data graphically. *Computational Statistics*, 11:23–33, 1996.
16. Ross Ihaka. Colour for presentation graphics. In Kurt Hornik, Friedrich Leisch, and Achim Zeileis, editors, *Proceedings of the 3rd International Workshop on Distributed Statistical Computing, Vienna, Austria*, 2003. ISSN 1609-395X.
17. Insightful Inc. *S-PLUS 7*. Seattle, WA, 2005.
18. Leonhard Knorr-Held. Dynamic rating of sports teams. Discussion Paper 98, SFB 386 “Statistical Analysis of Discrete Structures”, 1999.
19. G. Koch and S. Edwards. Clinical efficiency trials with categorical data. In K. E. Peace, editor, *Biopharmaceutical Statistics for Drug Development*, pages 403–451. Marcel Dekker, New York, 1988.
20. Uwe Ligges and Martin Mächler. Scatterplot3d - an r package for visualizing multivariate data. *Journal of Statistical Software*, 8(11):1–20, 2003.
21. Josef A. Mazanec and Helmut Strasser. *A Nonparametric Approach to Perceptions-based Market Segmentation: Foundations*. Springer, Berlin, 2000.
22. David Meyer, Achim Zeileis, and Kurt Hornik. Visualizing independence using extended association plots. In Kurt Hornik, Friedrich Leisch, and Achim Zeileis, editors, *Proceedings of the 3rd International Workshop on Distributed Statistical Computing, Vienna, Austria*, 2003. ISSN 1609-395X.
23. David Meyer, Achim Zeileis, and Kurt Hornik. *vcd: Visualizing Categorical Data*, 2005. R package version 0.9-6.
24. David Meyer, Achim Zeileis, and Kurt Hornik. The strucplot framework: Visualizing multi-way contingency tables with **vcd**. Report 22, Department of Statistics and Mathematics, Wirtschaftsuniversität Wien, Research Report Series, November 2005.
25. Giovanni Moretti and Paul Lyons. Tools for the selection of colour palettes. In *Proceedings of the New Zealand Symposium On Computer-Human Interaction (SIGCHI 2002)*, University of Waikato, New Zealand, July 2002.
26. Albert H. Munsell. *A Color Notation*. Munsell Color Company, Boston, Massachusetts, 1905.
27. W. M. Patefield. An efficient method of generating  $r \times c$  tables with given row and column totals. *Applied Statistics*, 30:91–97, 1981. Algorithm AS 159.
28. Charles Poynton. Frequently-asked questions about color. URL <http://www.poynton.com/ColorFAQ.html>, 2000.
29. SAS Institute Inc. *SAS/STAT Version 9*. Cary, NC, 2005.
30. R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2005. ISBN 3-900051-00-3.
31. H. Riedwyl and M. Schüpbach. Parquet diagram to plot contingency tables. In F. Faulbaum, editor, *Softstat '93: Advances in Statistical Software*, pages 293–299, New York, 1994. Gustav Fischer.

32. Martin Theus. Interactive data visualization using Mondrian. *Journal of Statistical Software*, 7(11):1–9, 2003.
33. B. Thornes and J. Collard. *Who Divorces?* Routledge & Kegan, London, 1979.
34. Antony R. Unwin, G. Hawkins, Heike Hofmann, and B. Siegl. Interactive graphics for data sets with missing values - MANET. *Journal of Computational and Graphical Statistics*, 4(6):113–122, 1996.
35. Forrest W. Young. ViSta: The visual statistics system. Technical Report 94–1(c), UNC L.L. Thurstone Psychometric Laboratory Research Memorandum, 1996.
36. Achim Zeileis, David Meyer, and Kurt Hornik. Residual-based shadings for visualizing (conditional) independence. Report 20, Department of Statistics and Mathematics, Wirtschaftsuniversität Wien, Research Report Series, August 2005.