Milwaukee Eviction Court Outcomes.

Part 2: Exploration and Visualization

DEBORAH J SCHLEIF

COSC 5610: Data Mining, Marquette University

1 INTRODUCTION

The Milwaukee Eviction Court Study (MECS) [2] is from in-person surveys of people in Milwaukee County Small Claims Court for an eviction hearing. This data set focuses on analysis on the question: Of tenants appearing in court, what characteristics make one likely to receive an eviction judgement?
Relative sample sizes in categories for some of the variables prompts some questions about whether they are predictors of whether a renter actually gets an eviction notice, but that analysis would be better considered with Desmond's other data set, the Milwaukee Area Renters Study (MARS), which surveyed renters about rental issues and included both renters who have and have not had an eviction.

1.1 Eviction Judgement Outcomes

For the full data set, 28.7% of renters received an eviction judgement, and 71.3% had a different outcome, working out an agreement with the landlord, case dismissed, or needing to come back at a later date. Categories that are not close to a 28.7% eviction judgement proportion might be variables of interest. (See Table 1 and Figure 1).

Table 1: Eviction Outcome by Reason on Notice

Reason	Judgement: 0	Judgement: 1
Default: Nonpayment of Rent/Other Charges	151	65
Other Dispute	2	0
Fourteen Day Notice	1	0
Lease Violation	2	2
Multiple	17	1
Nuisance: Landlord received nuisance property citation	1	0
Twenty-Eight Day Notice Without Cause	3	3
Total	177	71
	71.3%	28.7%

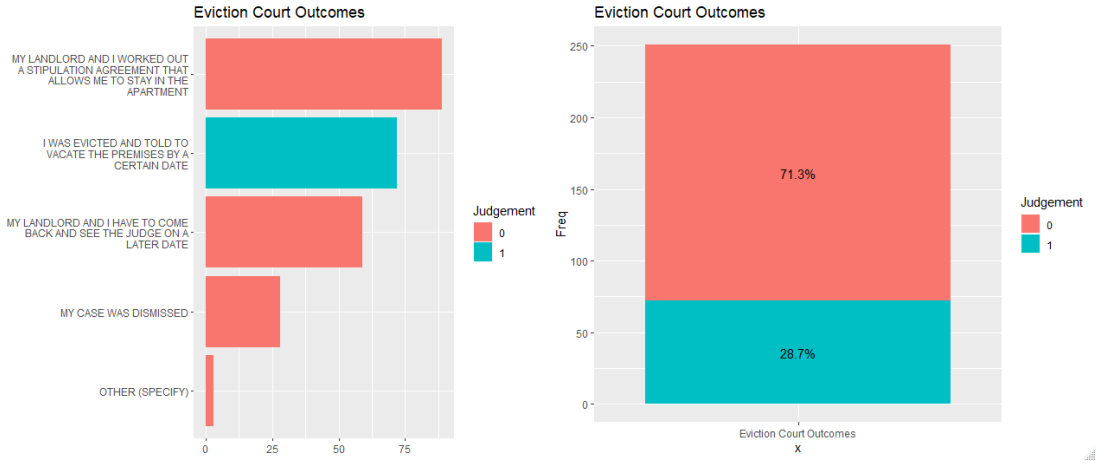

Figure 1: Eviction Court Judgements, Total Proportions

1.2 Reasons for Eviction

The reason listed on the eviction notice for most of the eviction cases in this data set is for defaulting on rent owed; even in the category where multiple reasons are applicable, in all cases except one, defaulting is one of the reasons (See Table 1 and Figure 2). This is consistent with some of the narratives in Desmond's book, where eviction court is sometimes used more as a process to get payment than to get a tenant to leave [1]. If the reason for eviction is nonpayment, variables that indicate a tenant's ability to pay would be expected to be significant in predicting the court outcome. If the outcome is not dependent on financial variables, it could indicate that even when nonpayment is the official reason for eviction, there could be other underlying reasons.

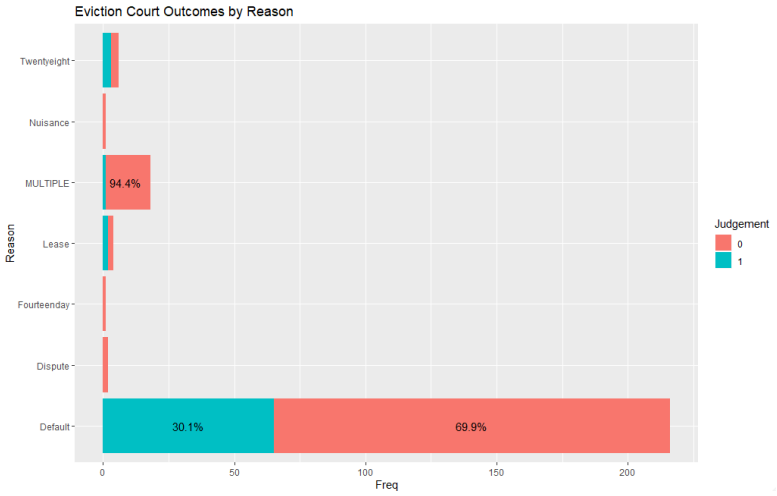

Figure 2: Eviction Court Judgements, Proportions by Reason

2 FINANCIAL METRICS

Data fields that might indicate a tenant's ability to pay are monthly rent, monthly household income, amount owed, and whether the tenant is receiving assistance in paying rent through a government program.

To explore whether some of these variables matter more relative to each other than as individual variables, several ratios are calculated: Months Behind=Income/Rent, Amount Owed/Income, Rent/Income, Months Behind/Income, and Income/Residents.

As expected, owing more and having less income appear to correlate with a higher probability of eviction, but it is difficult to know from these plots if the difference is statistically significant (see Figure 3). The outlier renter who has a $25,000 monthly income also influences the scales of the plots. When separated by the categories Gender, Age, and Having Children, the amount owed/income appears to make more of a difference for a renter without children.

Figure 9 lists the p-value for the Chi-Squared Test for Independence for various variables paired with the Judgement Outcome. None of these financial metrics are statistically significant with a level of significance of 0.05, the closest being Income with a p-value of 0.15. The ratios Owed/Income, Rent/Income, and Income/Residents do have lower p-values, 0.10, 0.13, and 0.13.

As shown in Figure 4, few renters who are receiving rent assistance appeared in eviction court than those without. This might indicate that rent assistance could help prevent evictions, but this dataset does not provide enough information to make that conclusion because only people in eviction court are covered, not renters who were not evicted.

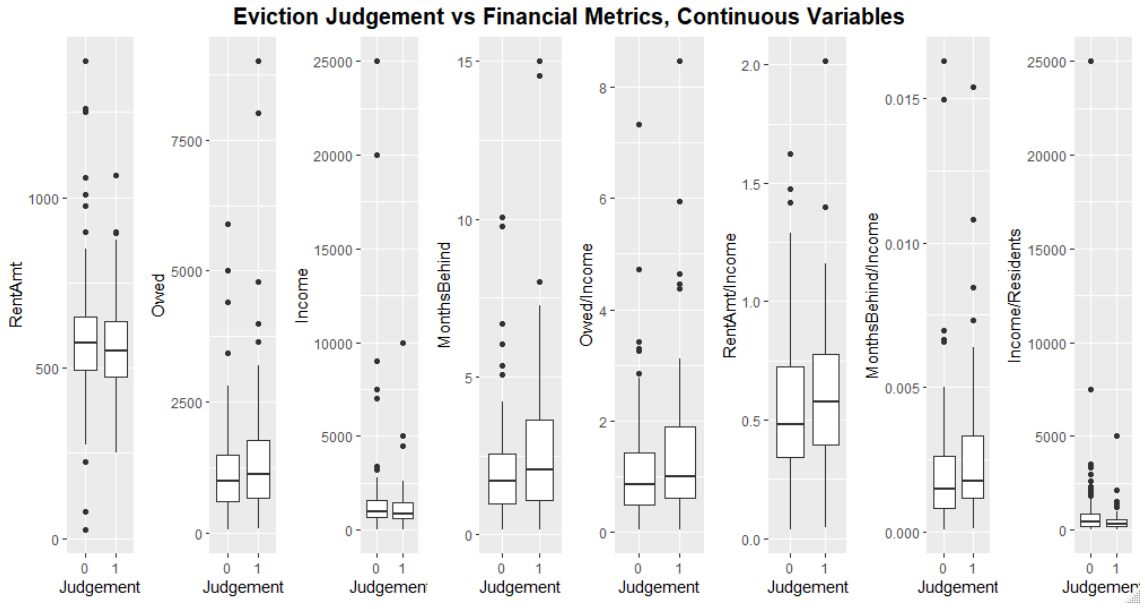

Figure 3: Box Plots, Judgement and Financial Metrics

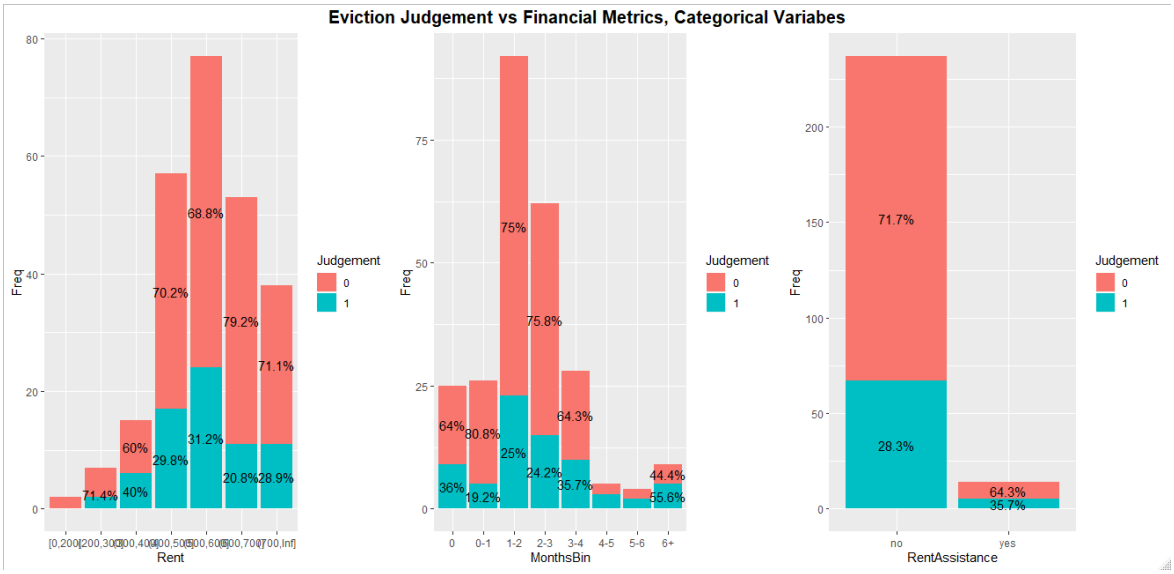

Figure 4: Bar Plots, Categorical Financial Metrics

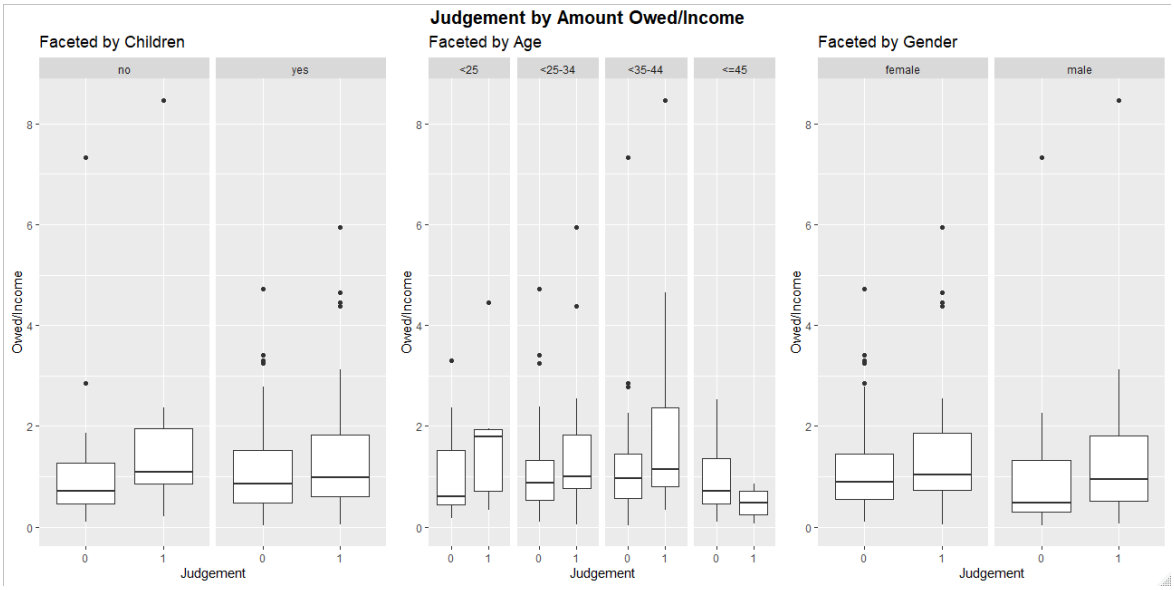

Figure 5: Owed/Income Ratio, Faceted by Categories

3 DEMOGRAPHIC METRICS

3.1 Race

Table 2 and Figure 6 illustrate the outcomes by race category. Black or African American renters are the largest category included in this data set, but from Figure 6 it appears that Hispanic and White renters might be more likely to get an eviction judgement. Race is not very statistically significant (p-value=0.16, Figure 9) in the dependence test with judgement. There is some statistically significant dependence in contingency tables with more significant variables, like Gender (Figure 7). Race also correlates with Amount Owed and whether living with children (see Figure 10).

Table 2: Eviction Outcome by Race and Gender

	Female	Male
Judgement:0		
American Indian or Alaskan Native	1	1
Black or African American	112	24
Hispanic	4	1
Mixed	5	4
Native Hawaiian or Pacific Islander	3	0
Refused	1	0
Something Else	1	0
White	12	10
Judgement:1		
American Indian or Alaskan Native	0	0
Black or African American	35	13
Hispanic	3	4
Mixed	1	2
Native Hawaiian or Pacific Islander	0	0
Refused	0	0
Something Else	0	0
White	3	11

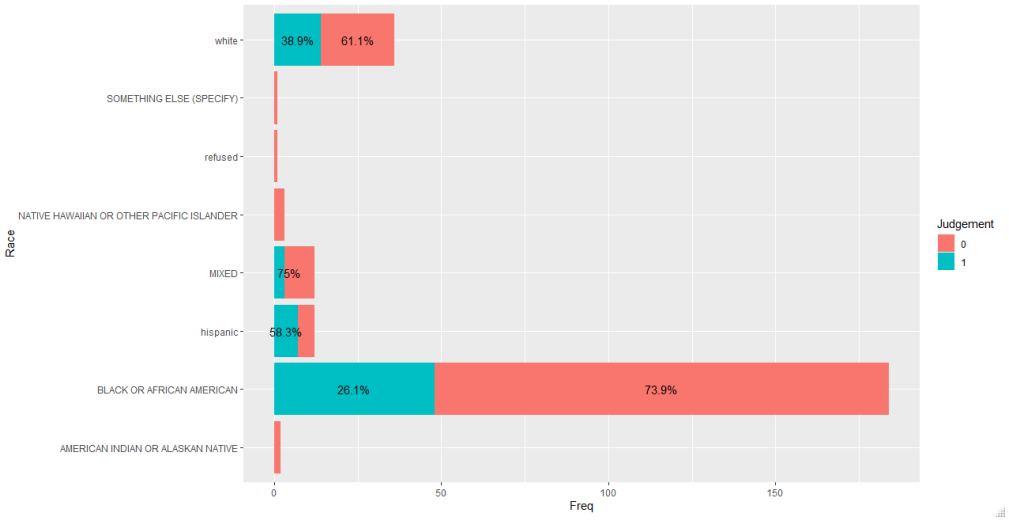

Figure 6: Eviction Judgement by Race

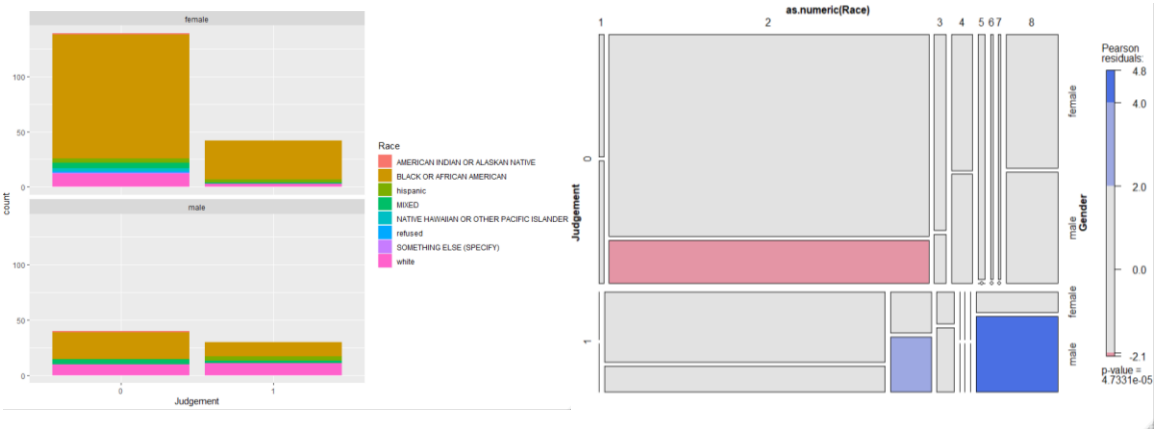

Figure 7: Eviction Judgement by Race and Gender

3.2 Children, Age, and Gender

Desmond reported that “The effect of living with children on receiving an eviction judgement was equivalent to falling four months behind in rent” [1]. From Figure 8, it does appear that renters without children are less likely to be evicted. However, with a p-value of 0.24 (Figure 9) compared to a p-value of for months behind at 0.18, it seems this might not support Desmond’s statement. The renter’s age and gender are the statistically significant variables, with p-values of 0.06 and 0.003.

Considering variables that correlate with each other might reveal the influence of children as a predictor. For example, females are less likely to get an eviction judgement and renters with children are more likely, but children are also more likely to be living with females than males (Figure 13). The age group 25-34 is more

likely than other age groups to get an eviction outcome (Figure 8) and also more likely to have children (Figure 13).

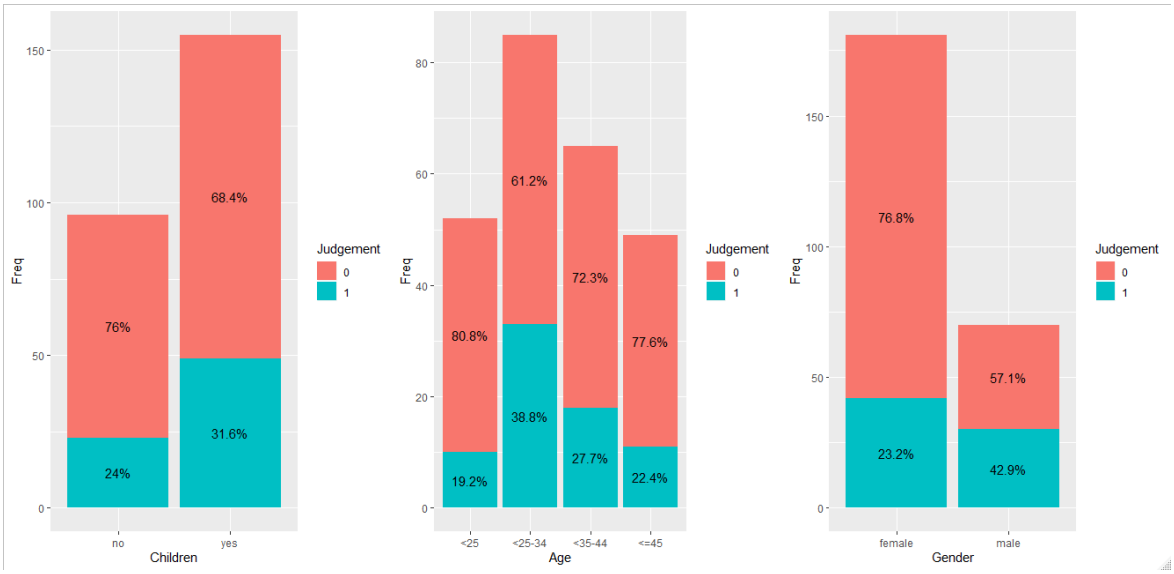

Figure 8: Bar Plots, Eviction Judgement by Children, Age, and Gender Categories

The data set also includes some more details about the household's makeup not included in Figure 9 that are not statistically significant with judgement outcome: the age of children living with the renter (p-value 0.44), the relationship with other adults living with the renter (spouse, significant other, family member, etc., p-value 0.53), and whether another adult is included on the eviction notice (p-value 0.24).

4 CORRELATION BETWEEN PREDICTOR VARIABLES

4.1 Chi-Squared Test for Independence

Figure 9 lists the p-values for the Chi-Squared test for Independence for each variable paired with the outcome judgement, and with each of the other predictor variables. Figure 10 presents similar information, for dummy variables with p-values less than 0.3 in the Chi-Squared test.

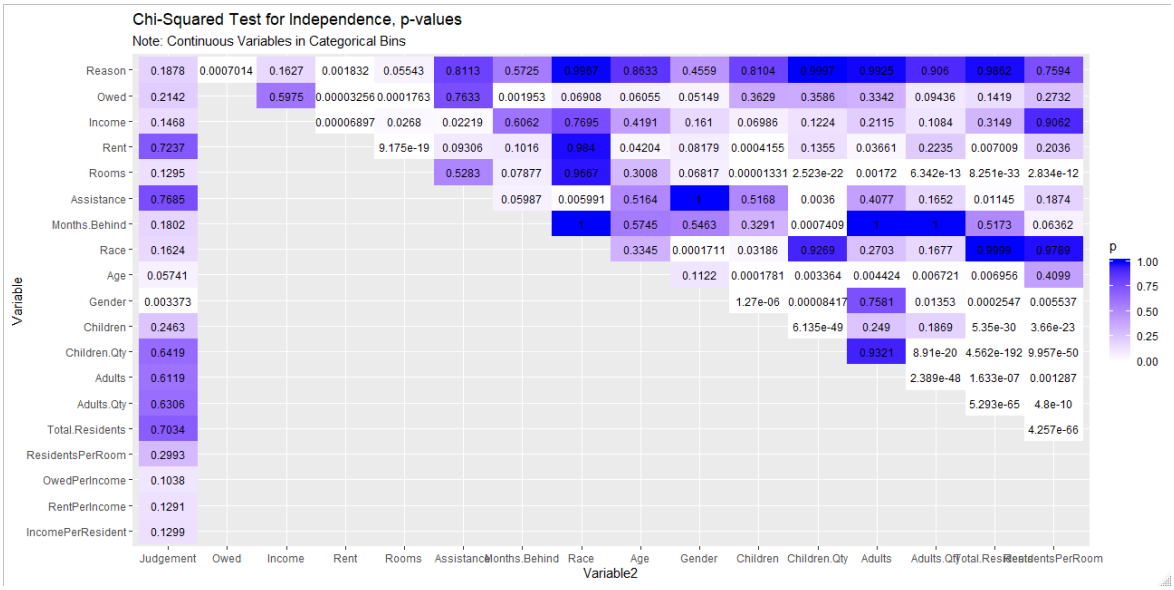

Figure 9: Tests for Independence Between Variables

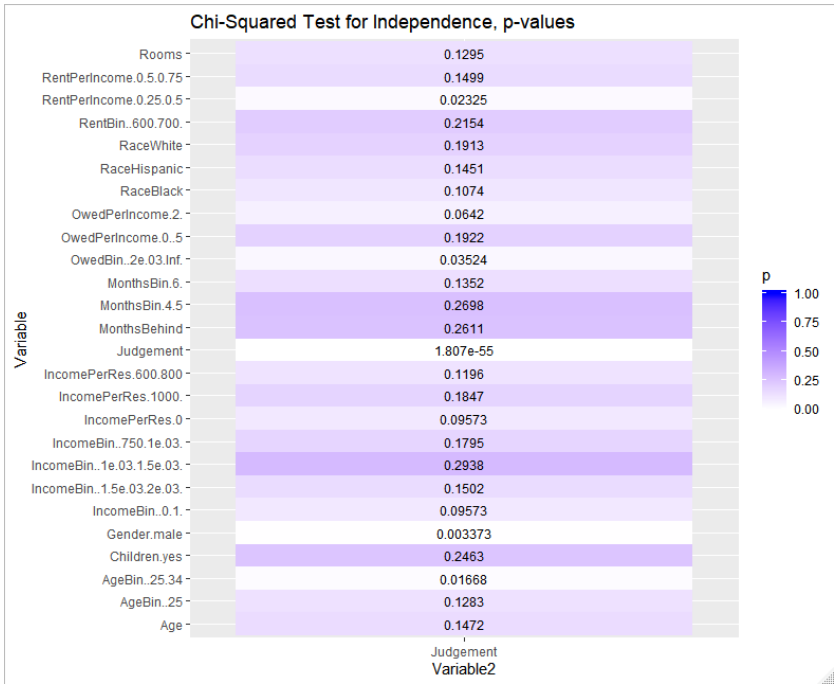

Figure 10: Tests for Independence Between Variables

4.2 Dependence

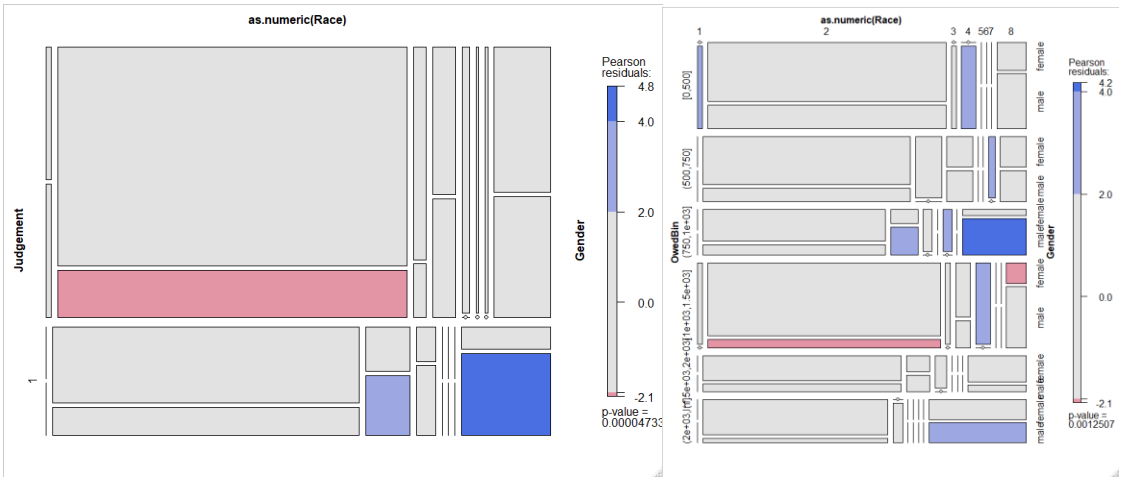

Figure 11: Mosaic Plots, Correlation between Race & Gender, Amount Owed, and Children

In Figure 11, race is listed as a numeric factor to avoid clutter (see Figure 7 for descriptions of race category).

The bar plots in Figures 12 and 13 show the categorical variables by whether the renter has children. Similar information for the variables that have a statistically significant dependence with children are shown less colorfully in the mosaic plots in Figure 14.

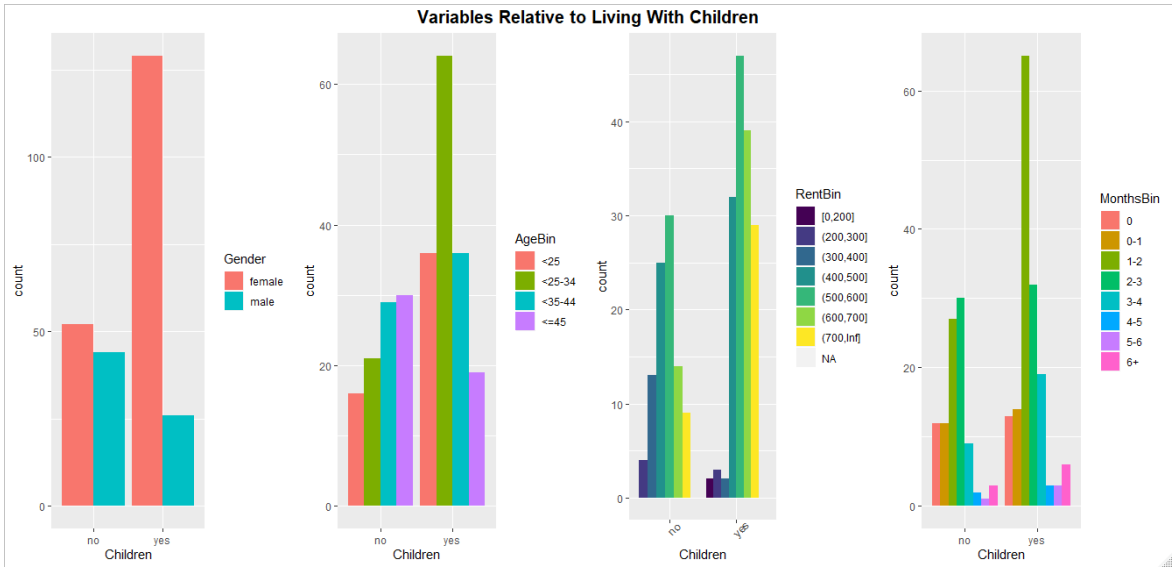

Figure 12: Bar Plots, Categorical Variables by Children

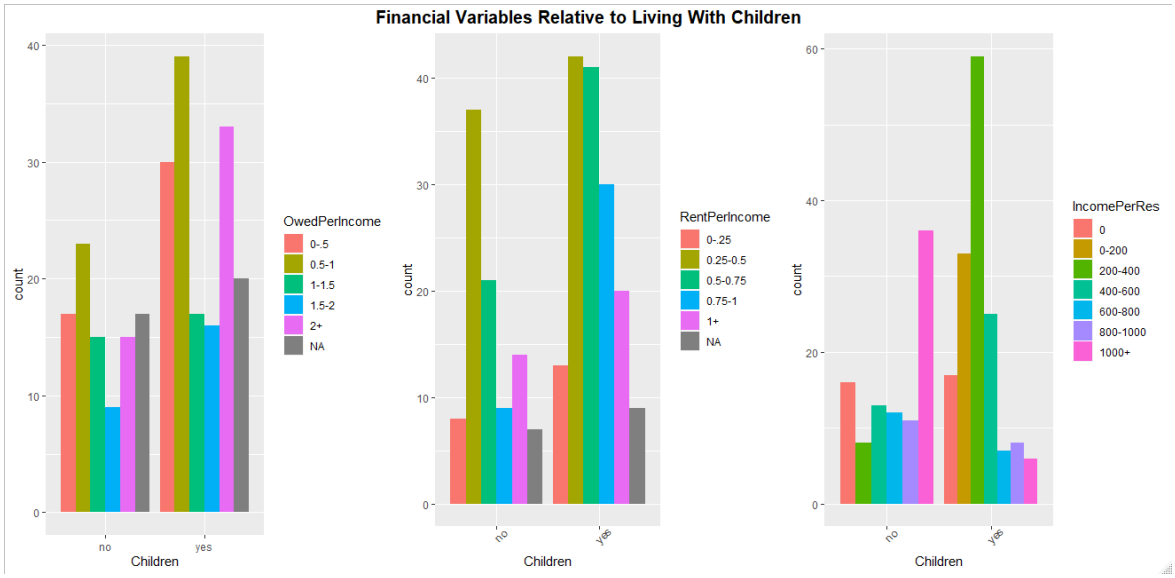

Figure 13: Bar Plots, Financial Ratio Variables by Children

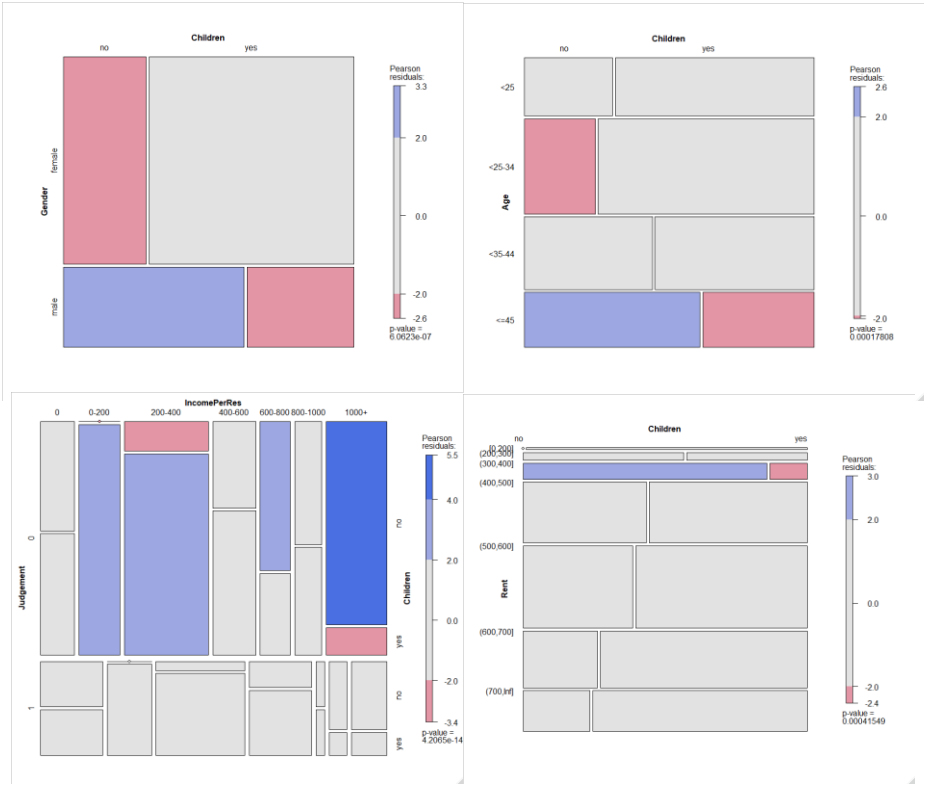

Figure 14: Mosaic Plots, Significant Variable Dependence on Children

5 SUMMARY

The data set includes potential predictor variables describing the reason on the eviction notice, financial metrics, and demographic information. In a chi-squared test for independence between each variable and the outcome judgement, if a level of significance of 0.05 is used, only age gender would show a statistically significant dependence. With a looser level of significance of 0.15, Amount Owed/Income, Rent/Income, number of bedrooms, Income/Resident, and income could be considered significant.

Reason for eviction as a predictor of eviction outcome would add six dummy variables. The reason for eviction is defaulting on rent payment for 92% of records in this dataset, and reasons are not statistically significant in testing for dependence with the judgement outcome; it might not be useful as a predictor variable.

The data includes variables for rent amount, income, amount owed, and whether the renter is receiving government assistance. None of these variables are statistically significant in the test for independence with judgement, with p-values 0.15 or higher. Desmond states one of his conclusions about the effect of Months Behind, which is a ratio of Rent Amount/Month. This has a p-value of 0.18, which might not be statistically significant. The other financial metrics might be more important relative to each other than as individual predictors: Owed/Income, Rent/Income, and Income/Residents have p-values, 0.10, 0.13, and 0.13. If all of these predictors were included, there would be some redundancy, though.

Race would add seven dummy variables if used as predictors. The racial categories that have significance with judgement, Black males, White males, and Hispanic males also correlate with having children, Black males less likely to be living with children and also less likely to get an eviction judgement, while White and Hispanic males are more likely to be living with children and more likely to get an eviction judgement. Black is also the largest racial category represented in the sample (73%). Race appears to be less important than other variables in predicting the judgement.

Having children and the number of months behind, which Desmond references in his conclusions, do not show up as having statistically significant dependence with the outcome judgement. Because having children does correlate with age and gender, which are significant, it could be an indirect indication that children is a significant predictor. Having children also correlates with income per resident, race, rent, and income, which are potentially significant predictors of the outcome. These correlations also point to a potential relationship between having children and the reason for eviction. If having children indicates that a renter would be unable to pay back owed rent, there would be a reason that children could be a predictor of outcome for an eviction notice listed because of rent nonpayment. This would be an indirect and potentially erroneous indicator, so such a conclusion is beyond the scope of this analysis.

Over-represented categories in this data set are people of Black or African American race, and people not receiving government assistance for rent might suggest that being Black would increase one's probability of getting evicted and that receiving assistance would decrease the probability. Without also considering the same probabilities for renters who did not get evicted, this is also beyond the scope of the analysis. However, it is consistent with descriptions of landlord-renter relationships that involve eviction court more as a process to get payment from renters than as a process to get them to vacate the property [1].

REFERENCES

[1] Matthew Desmond. *Evicted*. Crown Publishing Group, New York, 2016.

[2] Matthew Desmond., 2016, "Milwaukee Eviction Court Study (MECS)", https://doi.org/10.7910/DVN/MBDI45, Harvard Dataverse, V1, UNF:6:ZJX/wZG0nO4yU+Bv2qgJpA== [fileUNF]

[3] Matthew Desmond., 2016, " Milwaukee Area Renters Study (MARS)", https://doi.org/10.7910/DVN/BLUU3U, Harvard Dataverse, V3, UNF:6:i0T2EeAP8Q73PYrYkuLh3Q== [fileUNF]

[4] Achim Zeileis, David Meyer, and Kurt Hornik (2007). Residual-based Shadings for Visualizing (Conditional) Independence. Journal of Computational and Graphical Statistics, 16(3), 507-525.

[5] Baptiste Auguie (2017). gridExtra: Miscellaneous Functions for "Grid" Graphics. R package version 2.3. https://CRAN.R-project.org/package=gridExtra

[6] David Meyer, Achim Zeileis, and Kurt Hornik (2006). The Strucplot Framework: Visualizing Multi-Way Contingency Tables with vcd. Journal of Statistical Software, 17(3), 1-48. URL http://www.jstatsoft.org/v17/i03/

[7] David Meyer, Achim Zeileis, and Kurt Hornik (2020). vcd: Visualizing Categorical Data. R package version 1.4-7.

[8] Emmanuel-Lin Toulemonde (2020). dataPreparation: Automated Data Preparation. R package version 0.4.3. https://CRAN.R-project.org/package=dataPreparation

[9] H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

[10] Hadley Wickham (2017). tidyverse: Easily Install and Load the 'Tidyverse'. R package version 1.2.1. https://CRAN.R-project.org/package=tidyverse

[11] Hadley Wickham (2019). forcats: Tools for Working with Categorical Variables (Factors). R package version 0.4.0. https://CRAN.R-project.org/package=forcats

[12] Hadley Wickham (2019). stringr: Simple, Consistent Wrappers for Common String Operations. R package version 1.4.0. https://CRAN.R-project.org/package=stringr

[13] Hadley Wickham and Jennifer Bryan (2019). readxl: Read Excel Files. R package version 1.3.1. https://CRAN.R-project.org/package=readxl