# Intro

In this latest of posts, we examine the question how does marriage affect lifetime earnings?

Spurred by the question of how marriage decisions affect income and household earnings, I set out to examine a range of variables, such as age, race, work status, but contented myself with the ones I could find in my dataset.

# The Data:

The data comes from the ACS’s 1-year surveys, then aggregated down for each year and state, to the FES (Family Employment Status) components. There are 8 such components of this ordinal variable.

# The Hypothesis

My hypothesis was that the FES of 1, with a married couple, both male and female in the labor force would be strongly predictive of household earnings.

# Results

I was surprised to find that while it is significantly related at the 10% confidence level, it has a rather weak effect compared with other variables such as family size, or number of related children in the household.

# EDA

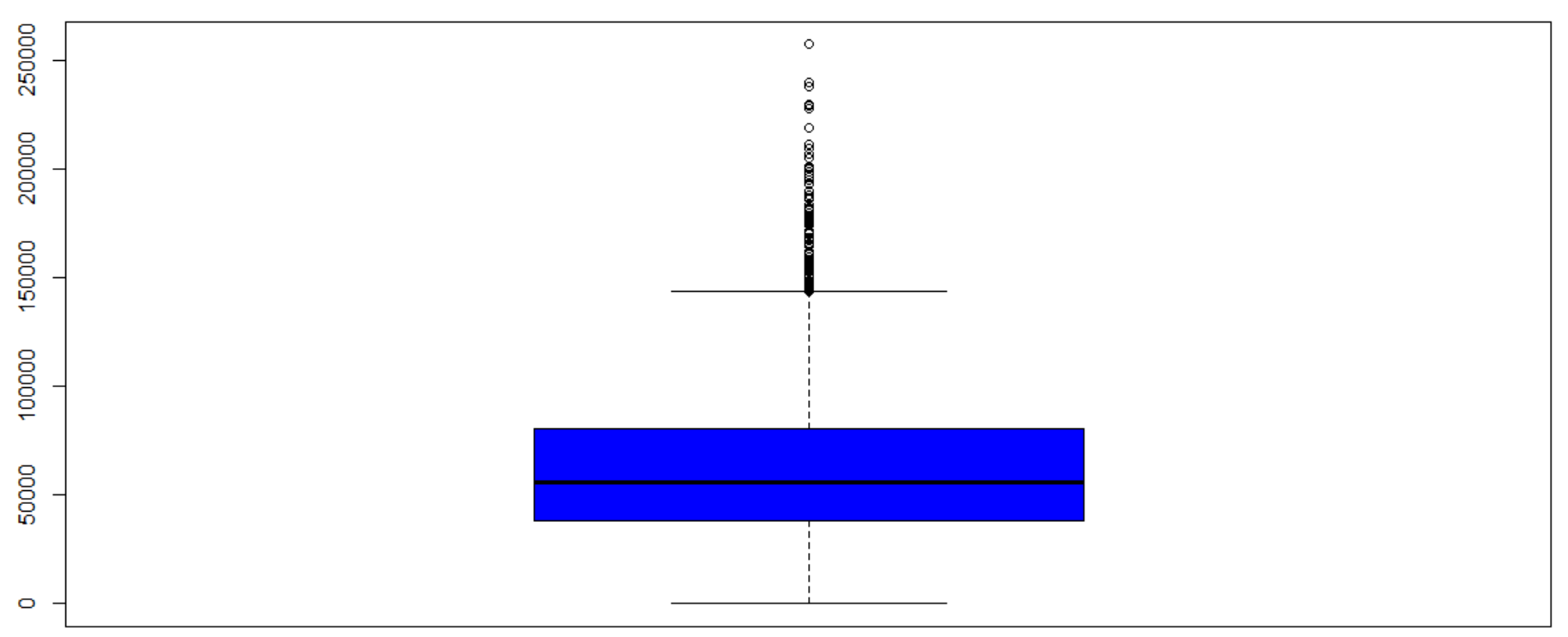
I believe some very interesting findings are in order. No doubt, some invidious ones too (as in Veblen’s usage of the term). But all jokes aside, having points of reference is key to surviving and thriving in today’s job market, what ever your field or background.

At the outset, some indulgent graphs. Some priceless invaluable boxplots of income and number of children.

Income Variables

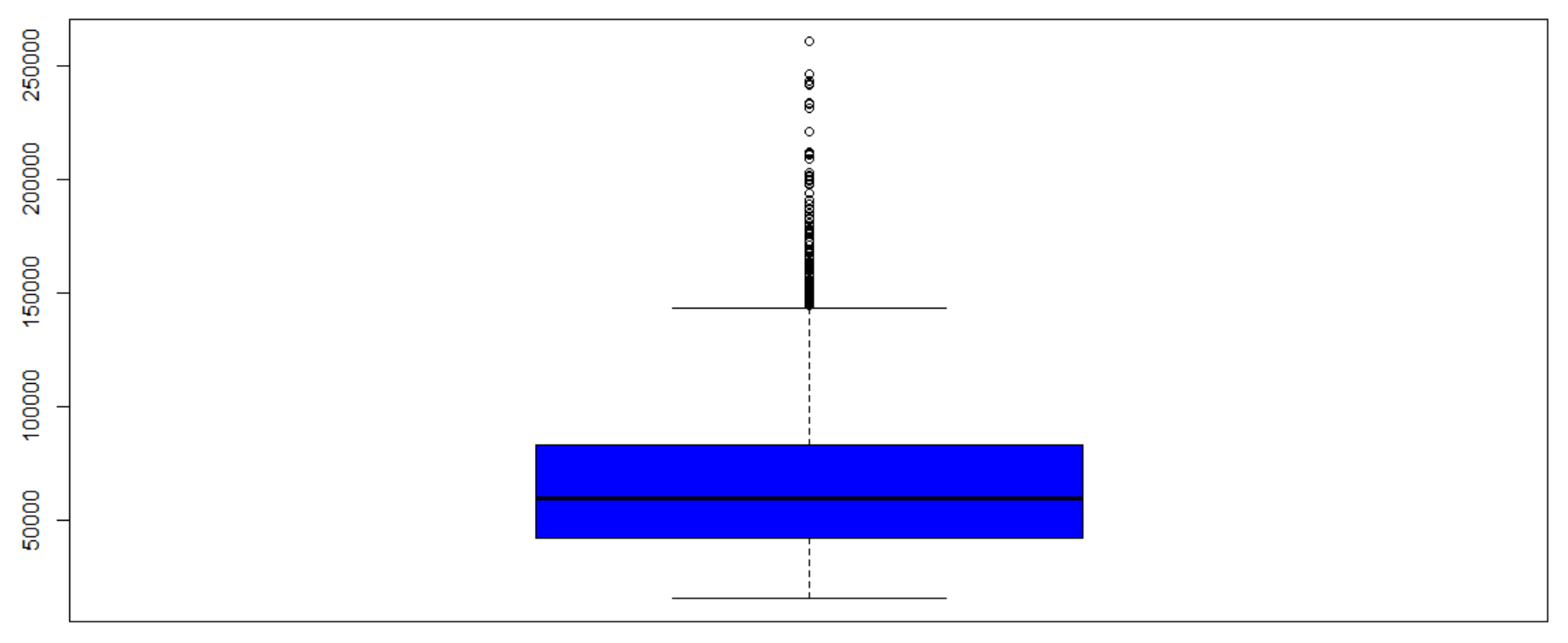
Fincp:

Our first graph is of family income, with a median set at around a little more than 50k.



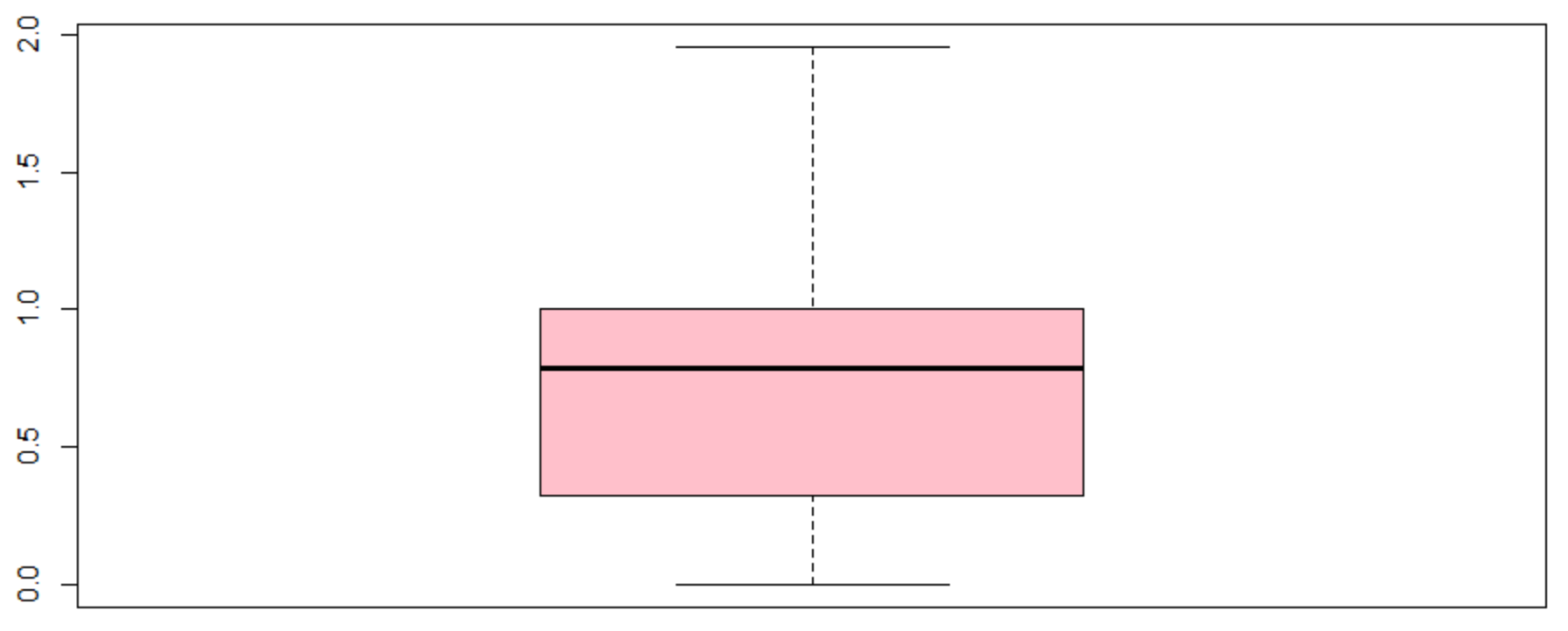
Hincp:

Our next boxplot, of household income, also shows an overall median slight above 50 k, this time higher, with longer right tails.

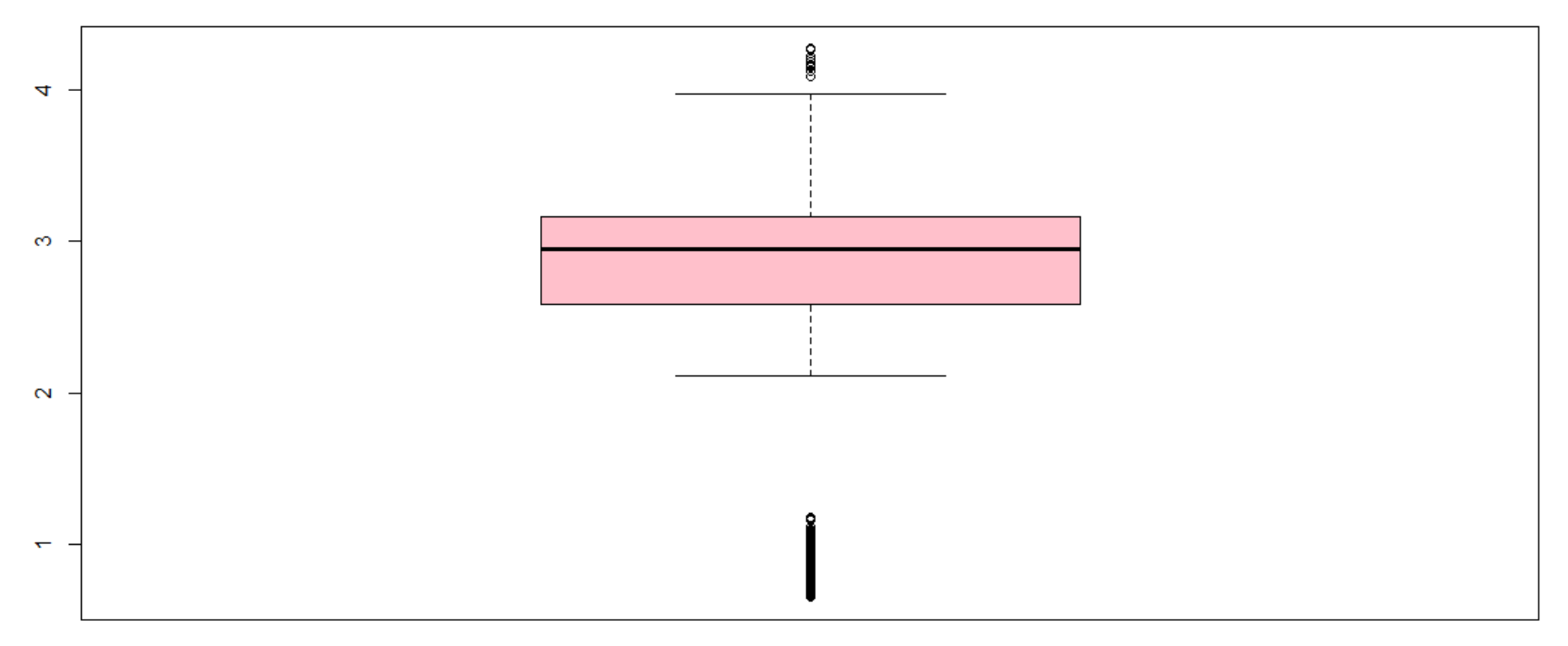


Children and Head Count Variables

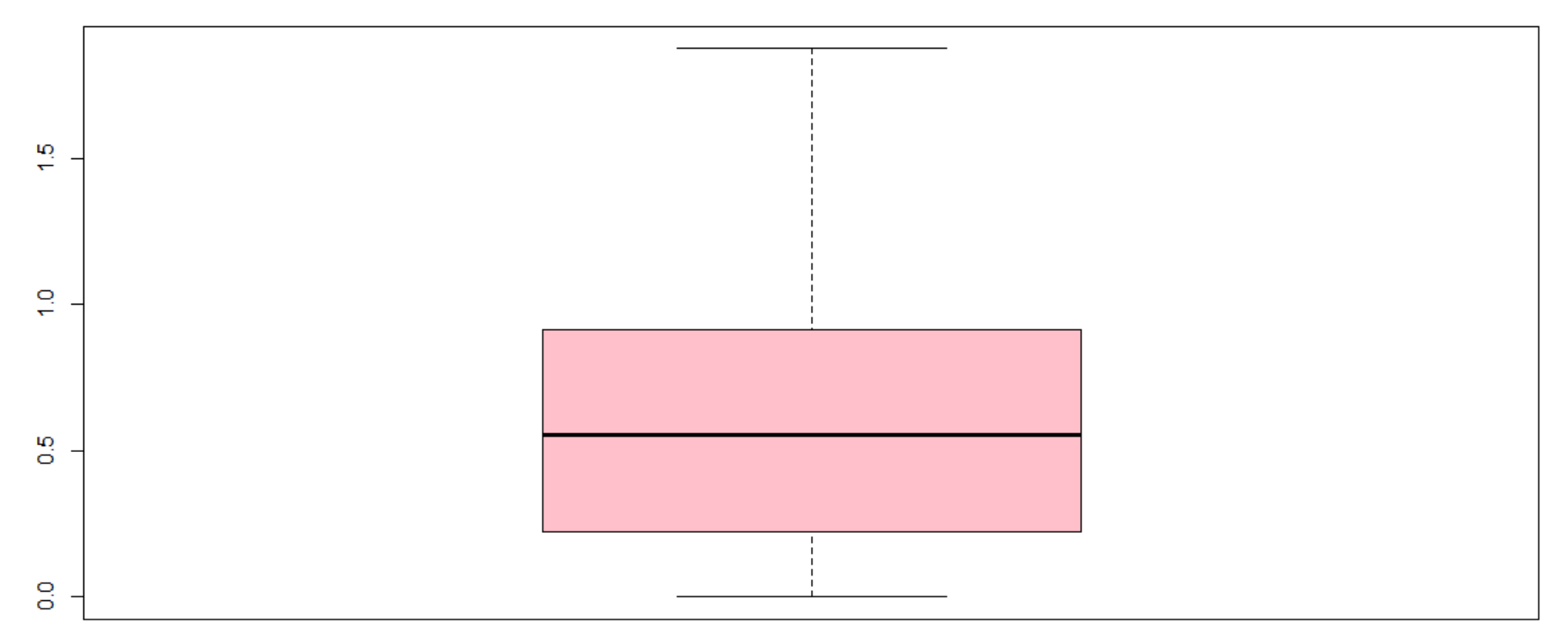
NRC, or the number of related children, is inclusive of children that survey respondents are responsible for. This metric doubtless places a burden on caretakers, and as we shall see, is negatively related with their household’s income on average. Of course, this is not a common metric reported, so we find a median below 1.



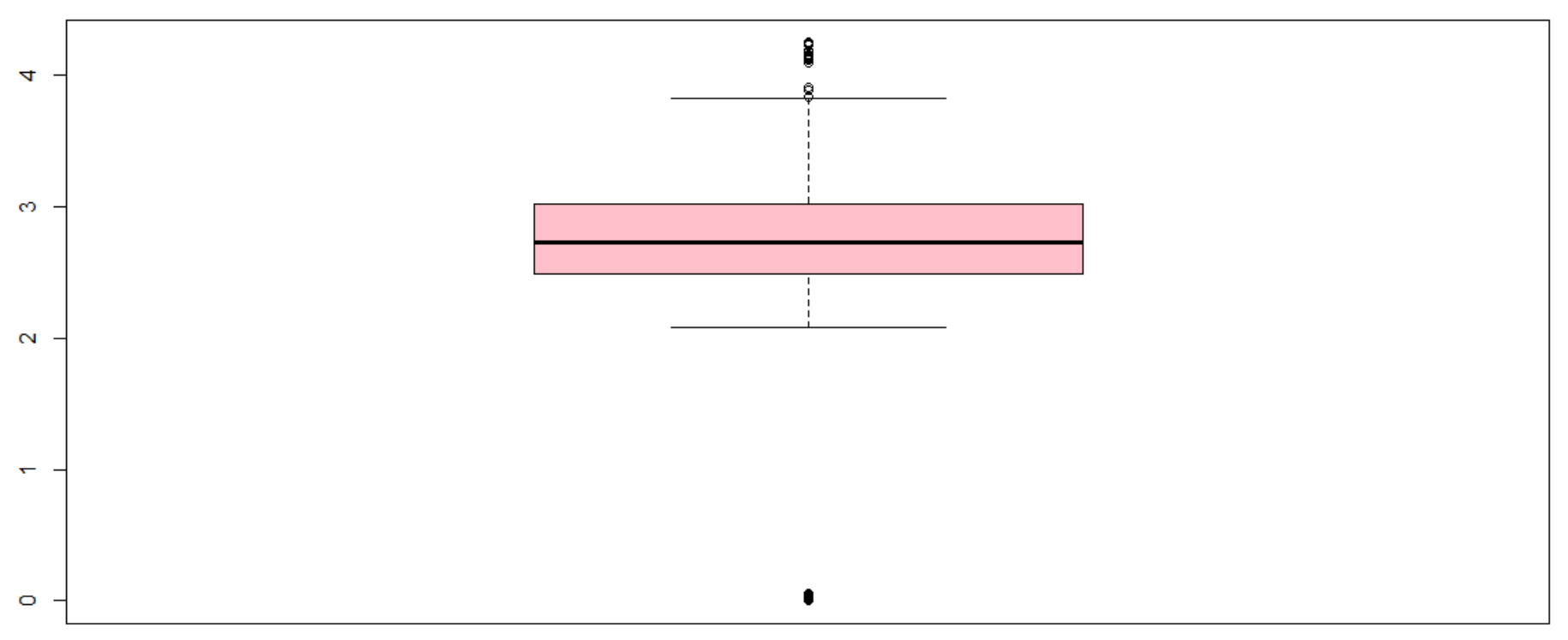
NP, number of persons pertaining to household measures, is roughly above 3 on average:



NOC, number of own children, much like number of related children, is less than 1. However, NOC does not have the same effect as NRC when it comes to household income:

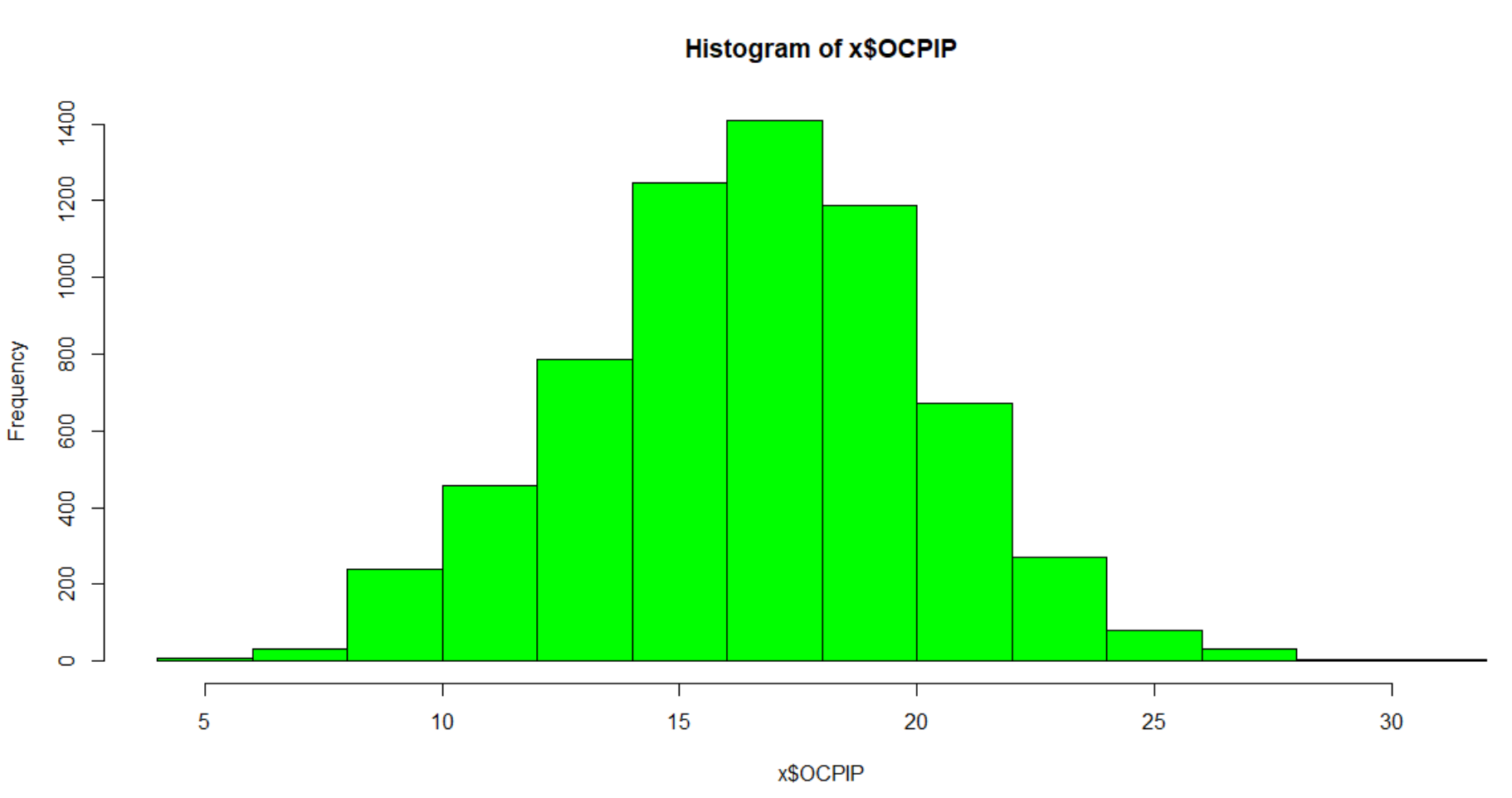


NPF, number of persons in family much like NP, is around 3 on average:

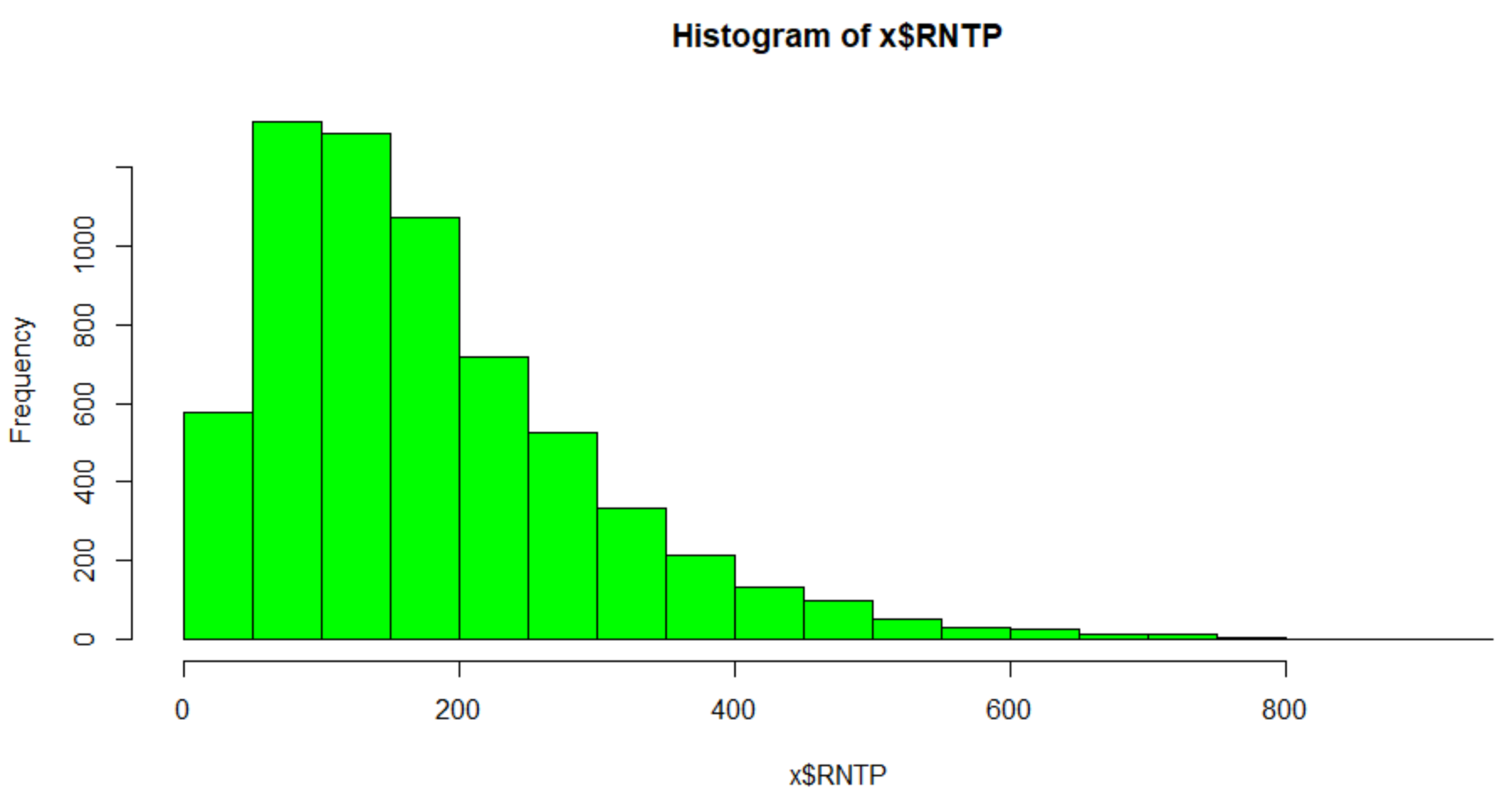


Charts of rents and mortgage payments

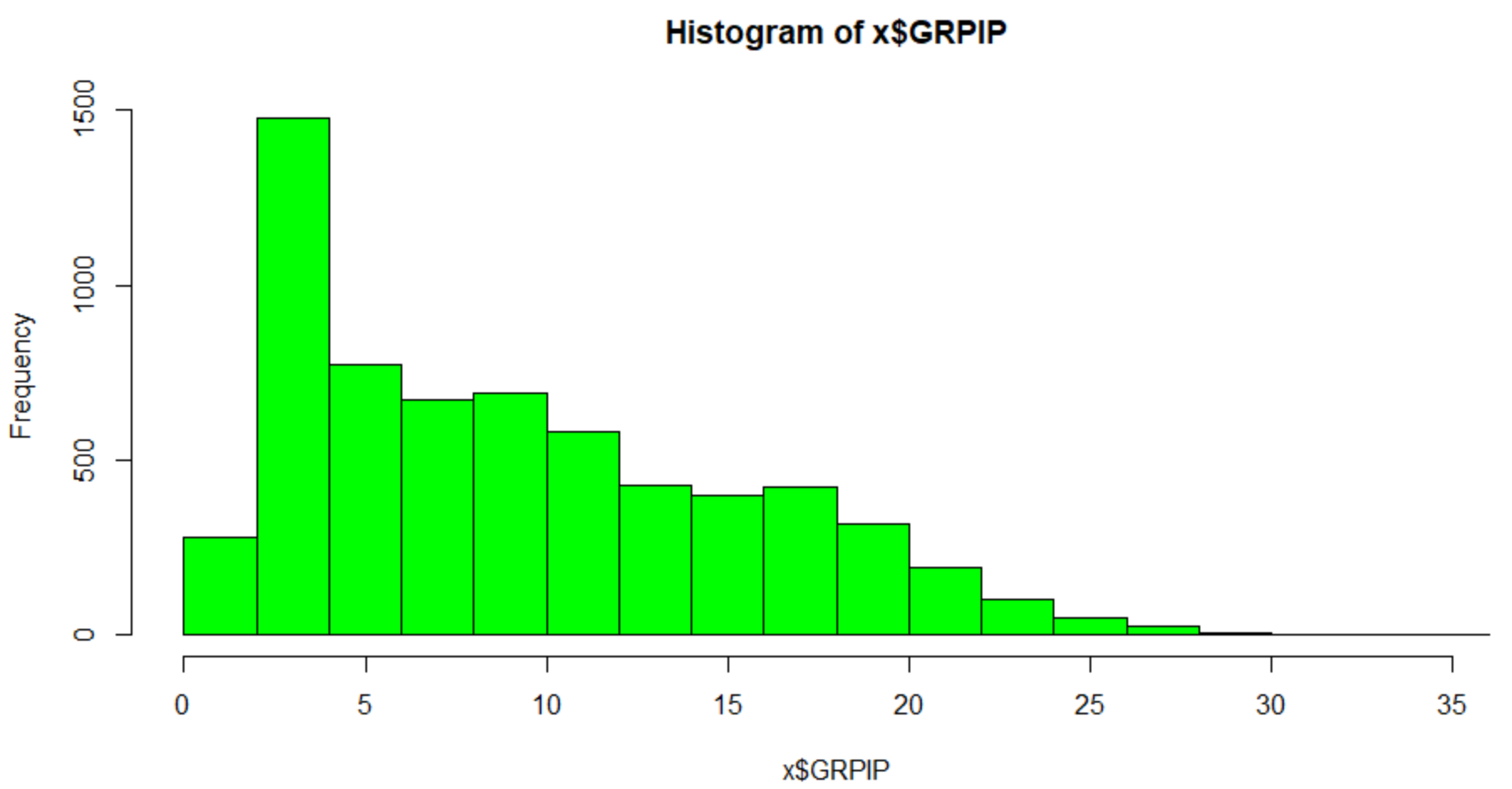
OCPIP: The Ownership costs as a percentage of monthly incomes are rather evenly distributed around 15%.



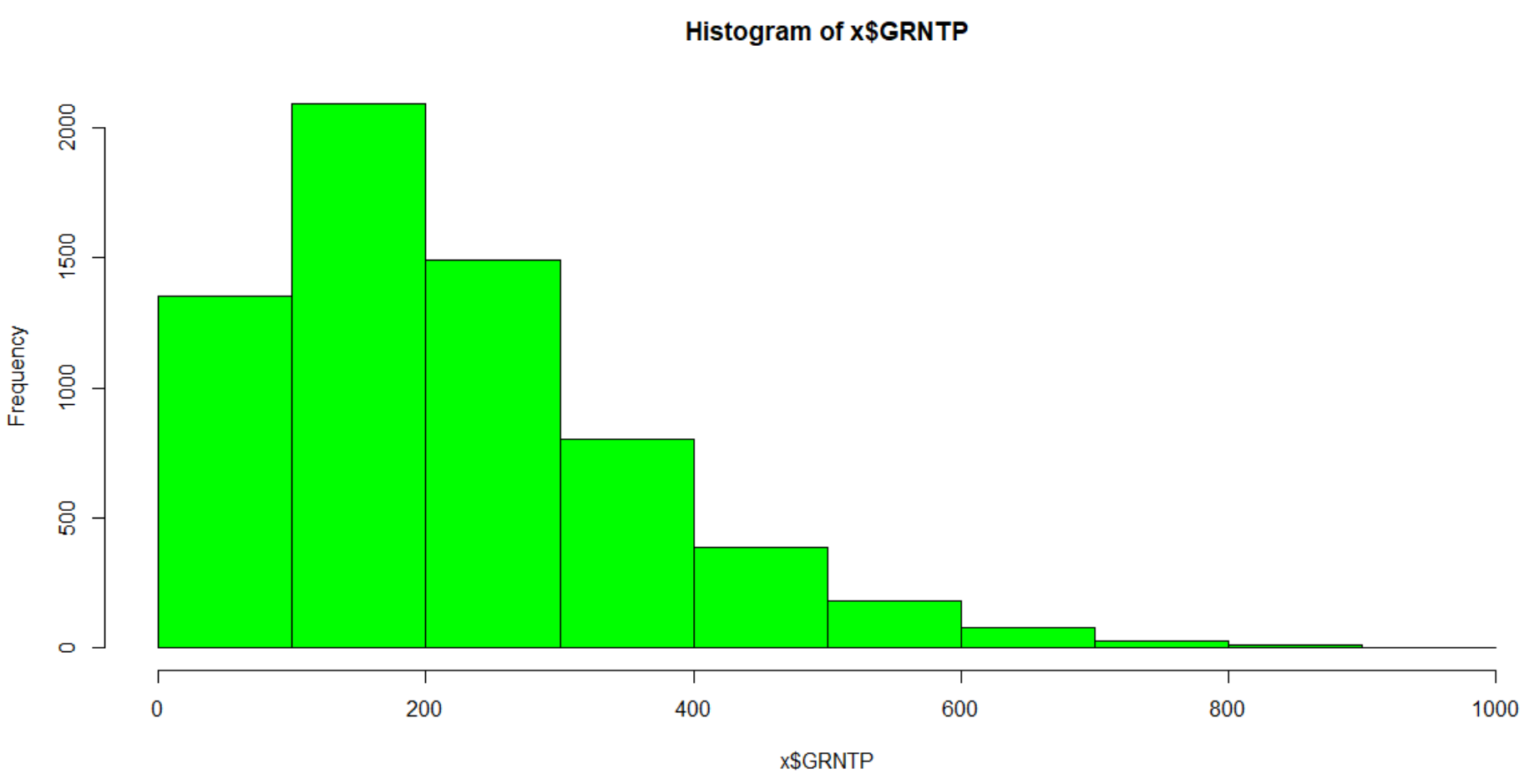
RNTP, the rent per month is obviously skewed right. Since these are already frequency counts of averages, they are not very reflective of the details. However, the right skew is too be expected, given that for many, ownership is just not at the top of their financial priorities. Such is the case in countries like Germany as well. The results and outcomes are debatable. But the highest frequencies occur among the lower ends of the spectrum.



GRPIP, gross rent as a % of monthly income shows a very interesting spike below 5%. This must be for people who receive public assistance for housing. The rest of the distribution is somewhat evenly spread, while still being rightly skewed.



GRNTP, gross monthly rent, probably includes features such as utilities, and gives us a fuller picture of average monthly distributions of rent around the country.

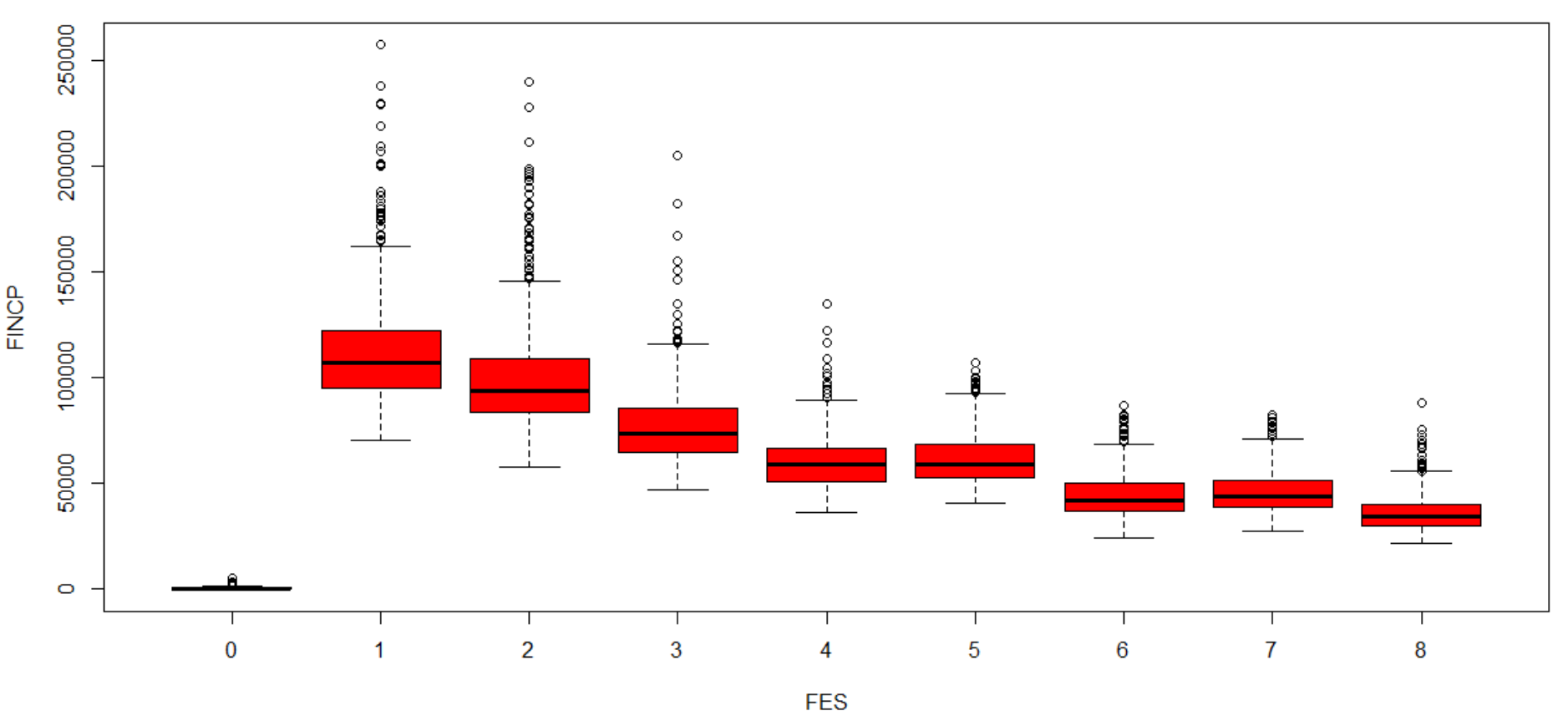


# Hypothesis Testing

2 way boxplot: FINCP vs. FES

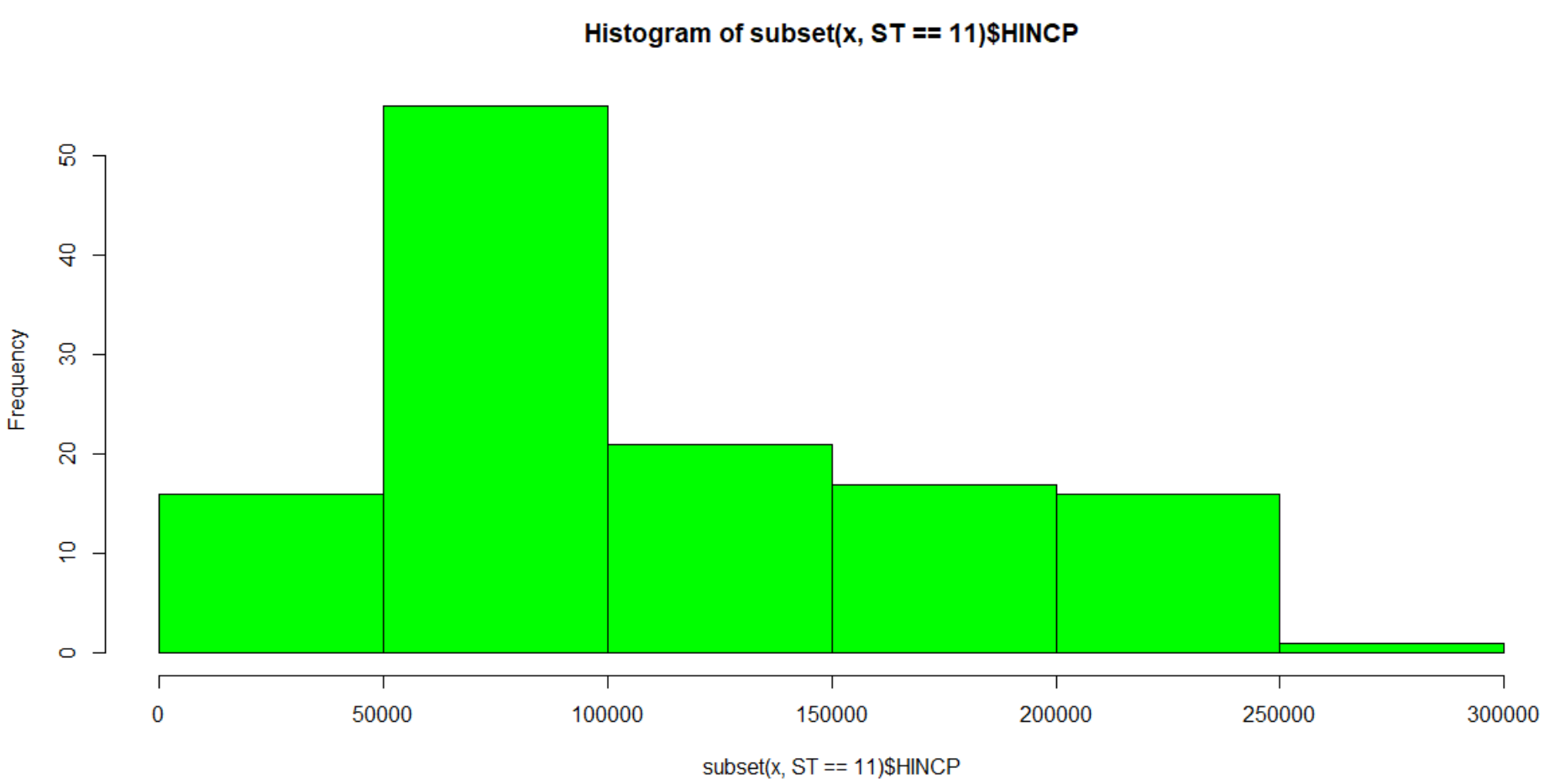
Now this! Is the reason why I formulated this hypothesis in the first place. Doubtless, if I plot other variables on a scatter plot, I’m bound to see a similar relation between say food stamps and family income. But FES, basically akin to SES (with 1 corresponding to a higher SES, and 8 to a lower one), is negatively related with family income.

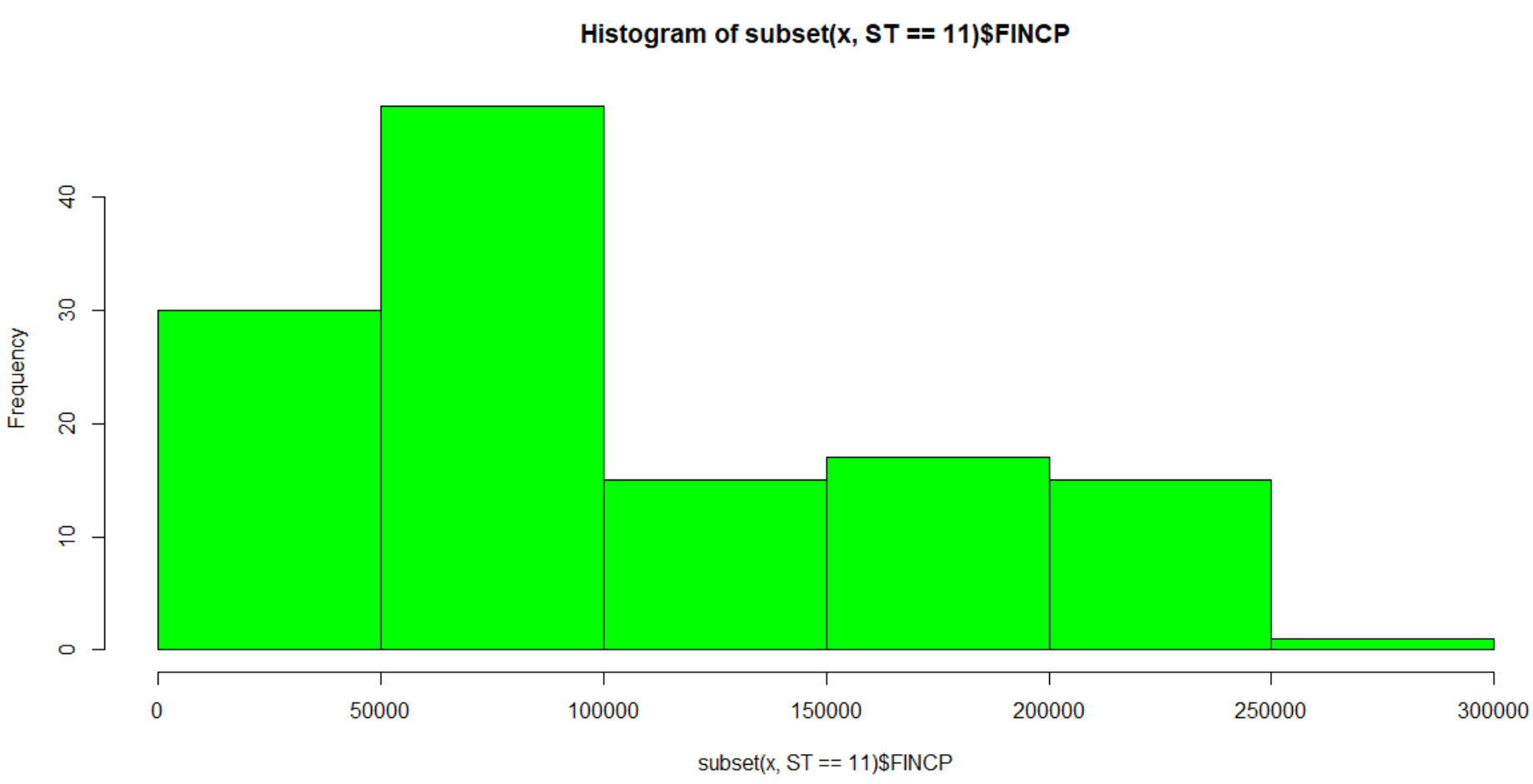
1: where a married couple has both partners in the workforce, sees a median income of above 100k. 2: where only the male is in the workforce, also sees a median above a 100k, albeit slightly lesser. And at the lower end of the spectrum, 8: single female, not in workforce sees a median income slightly less than 40k. These are the facts. To make a value judgement would be besides the point. But it is informative.



#Income plots by state: here are some fun charts

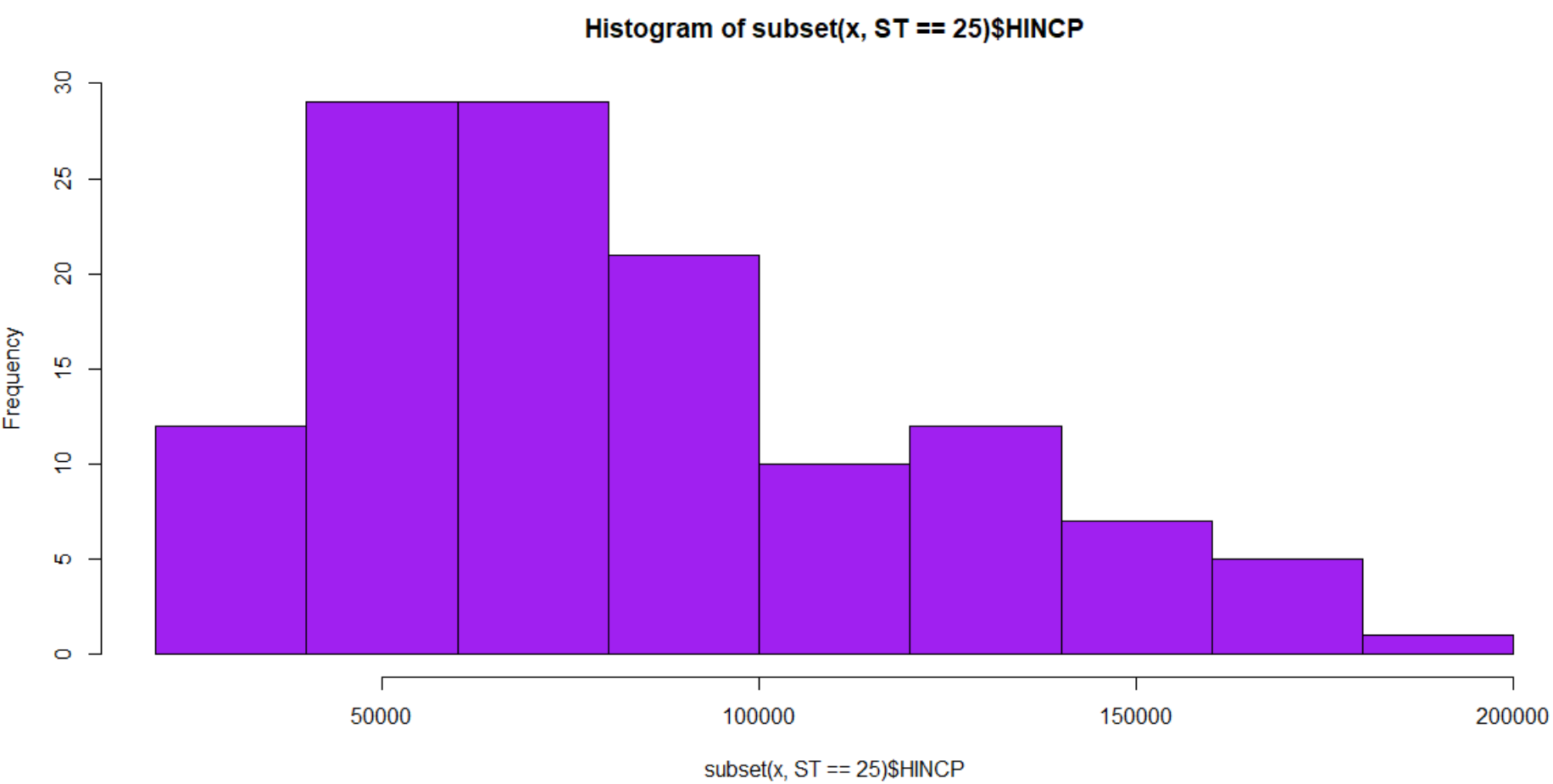
DC, with its limited sample size:

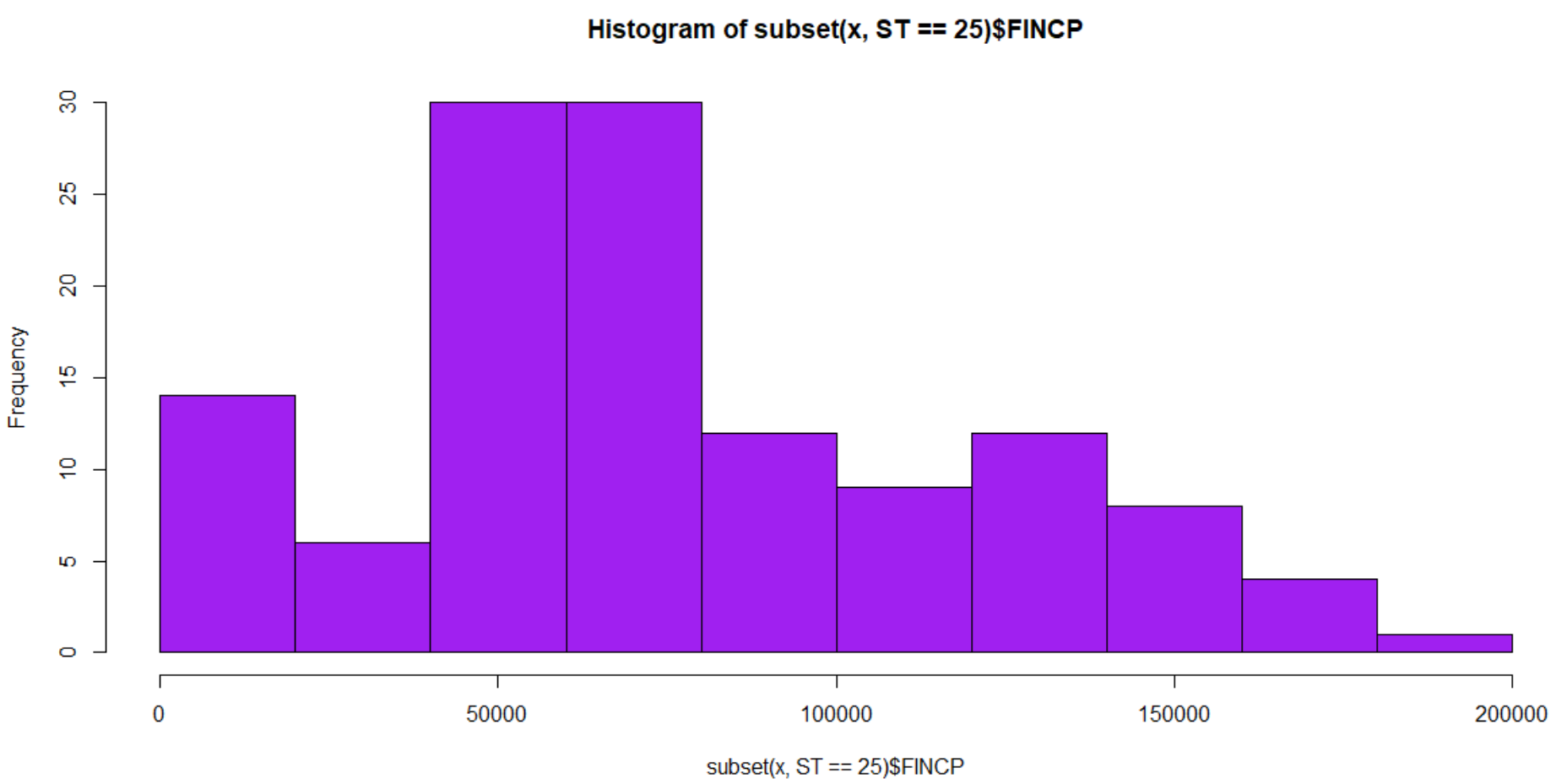




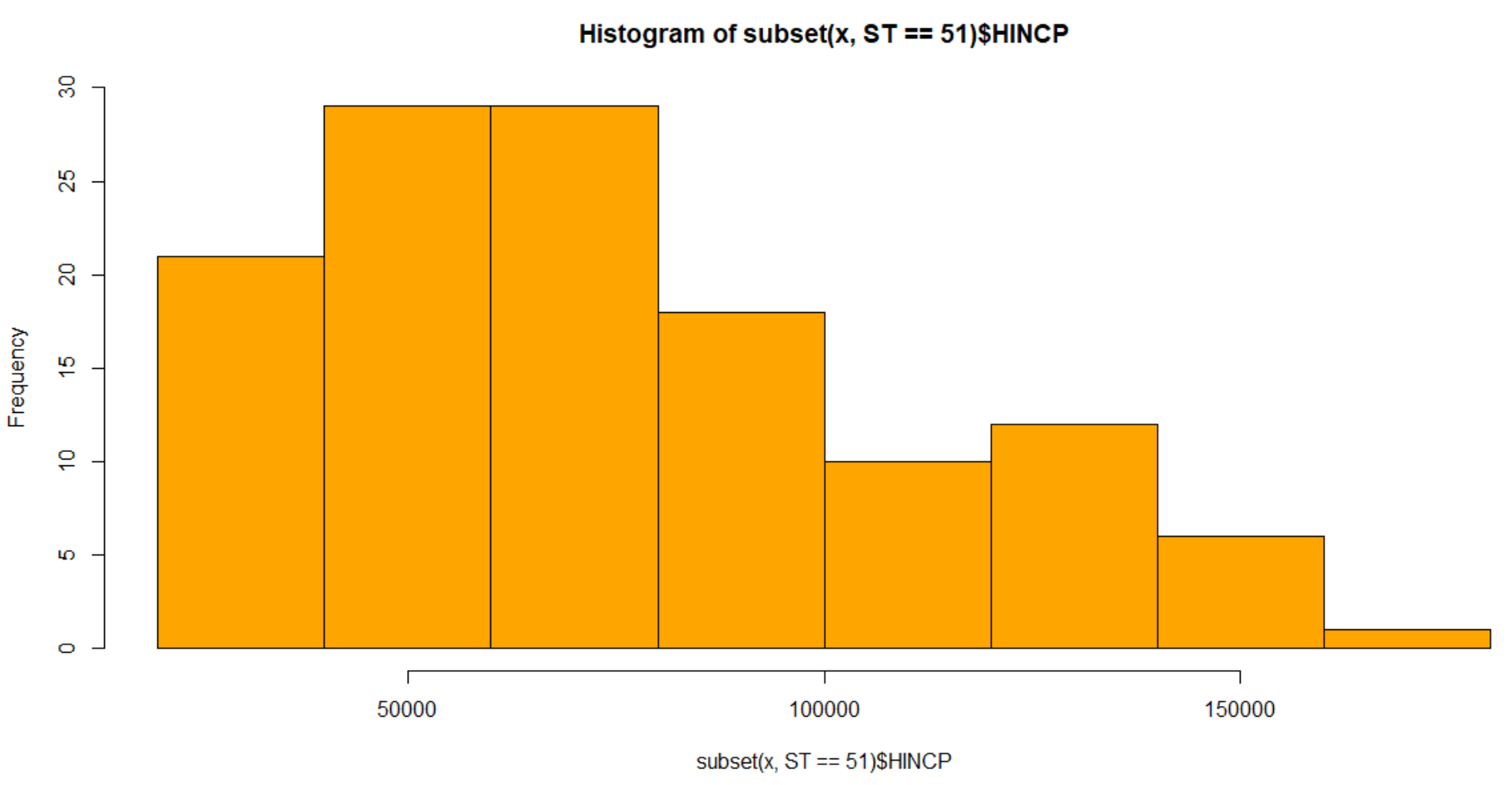
MD:

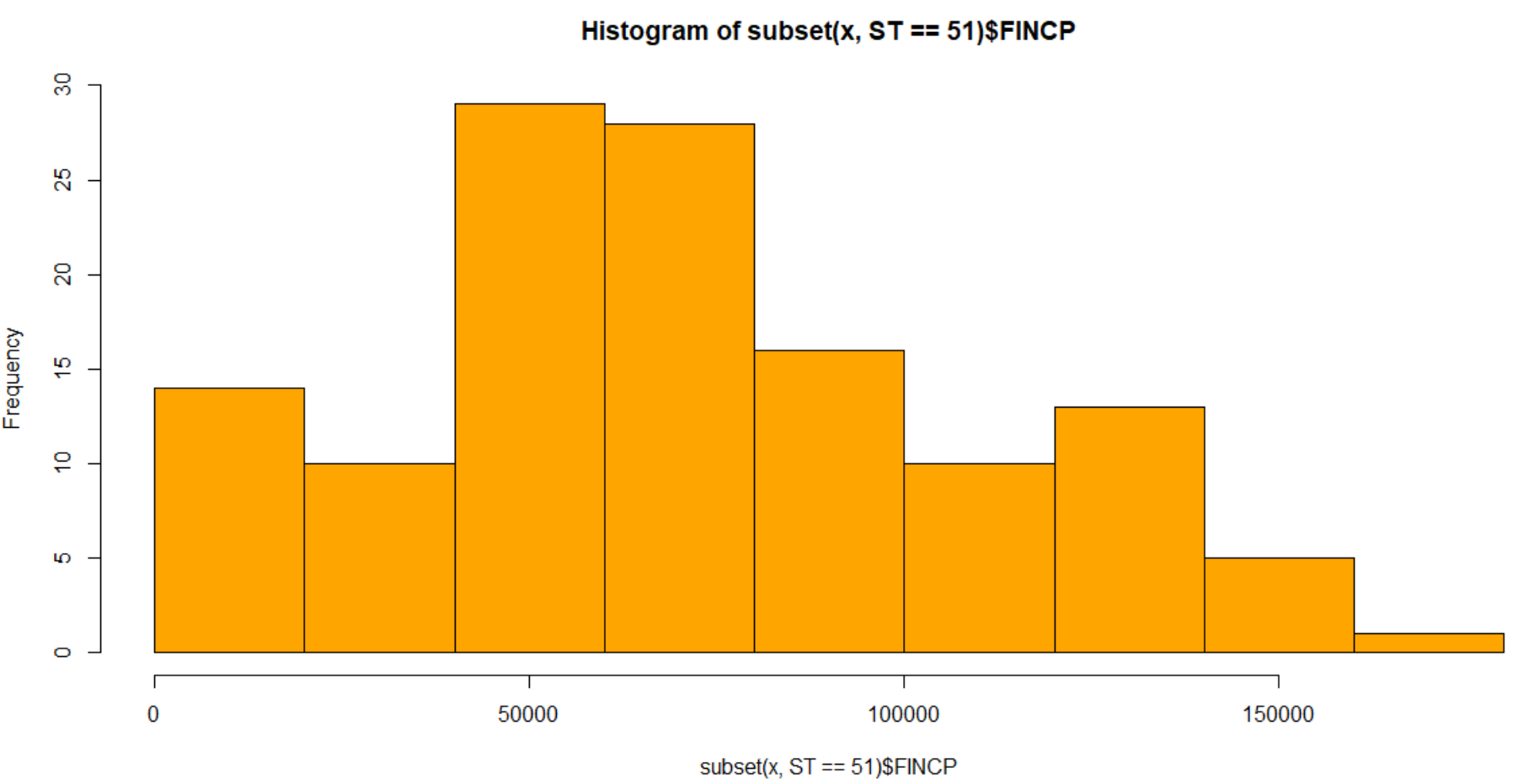
For Maryland and Virginia, the household incomes mean is centered slightly more to the right than that of the family income. I don’t know what the reason for this is!





VA:



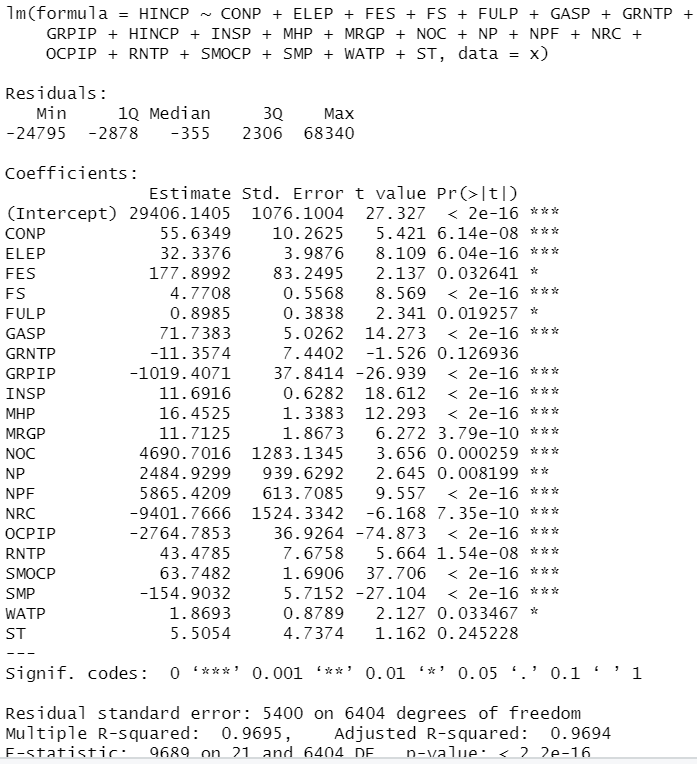


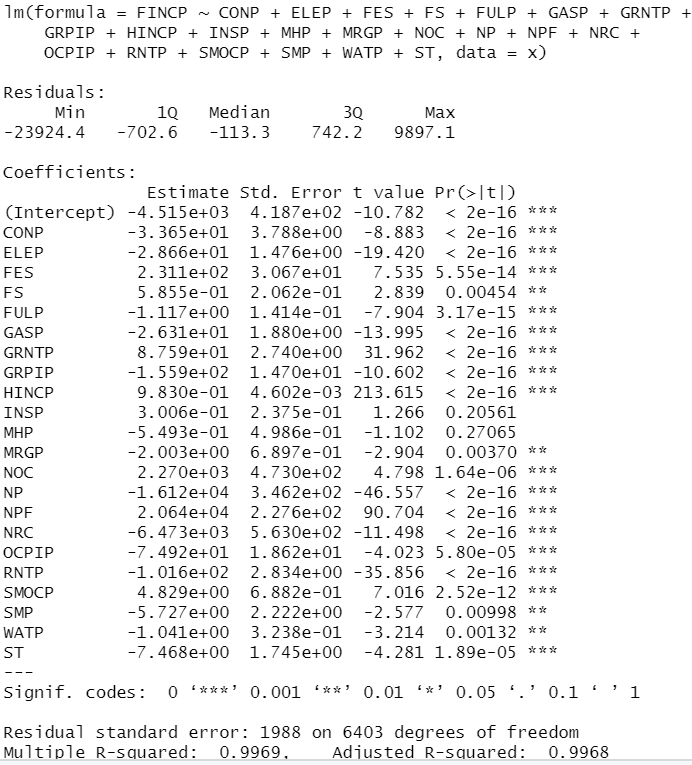
# And Finally, the Linear Regression Model:

For predicting household income, and family income respectively. I’ve shown both tables here. But because the family income model had a negative intercept, I’ve decided instead to focus in on the household income model. Besides, the family income model had coefficients having multiple powers of 10, making it more difficult to make sense of.

Household income regression model:

In the model below, somewhat disappointingly, FES turned out to be a rather weak predictor of Family income. Meanwhile, the strongest and largest predictors were NPF, or Number of Persons per family, and NRC, or Number of Related Children.

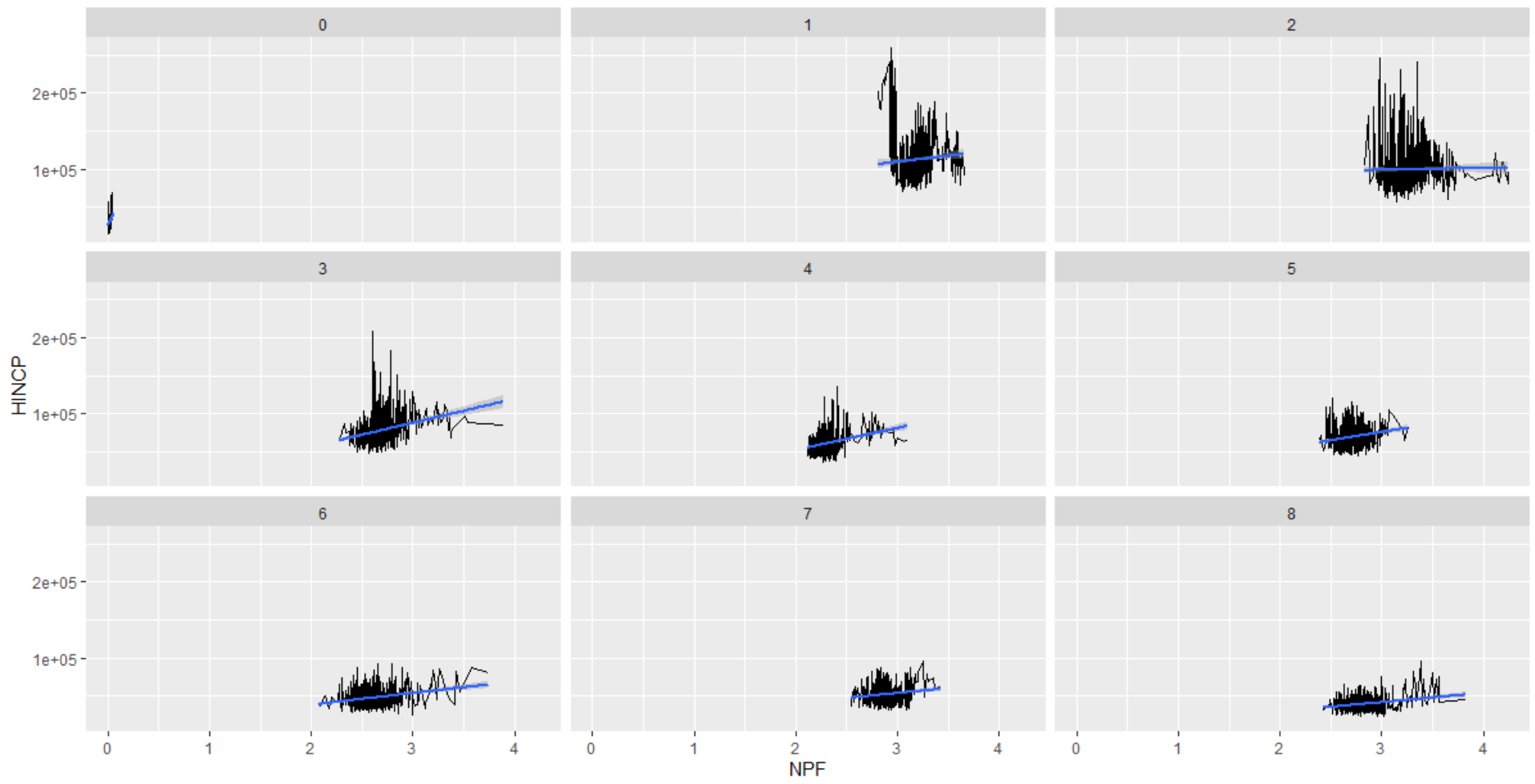




Conclusion:

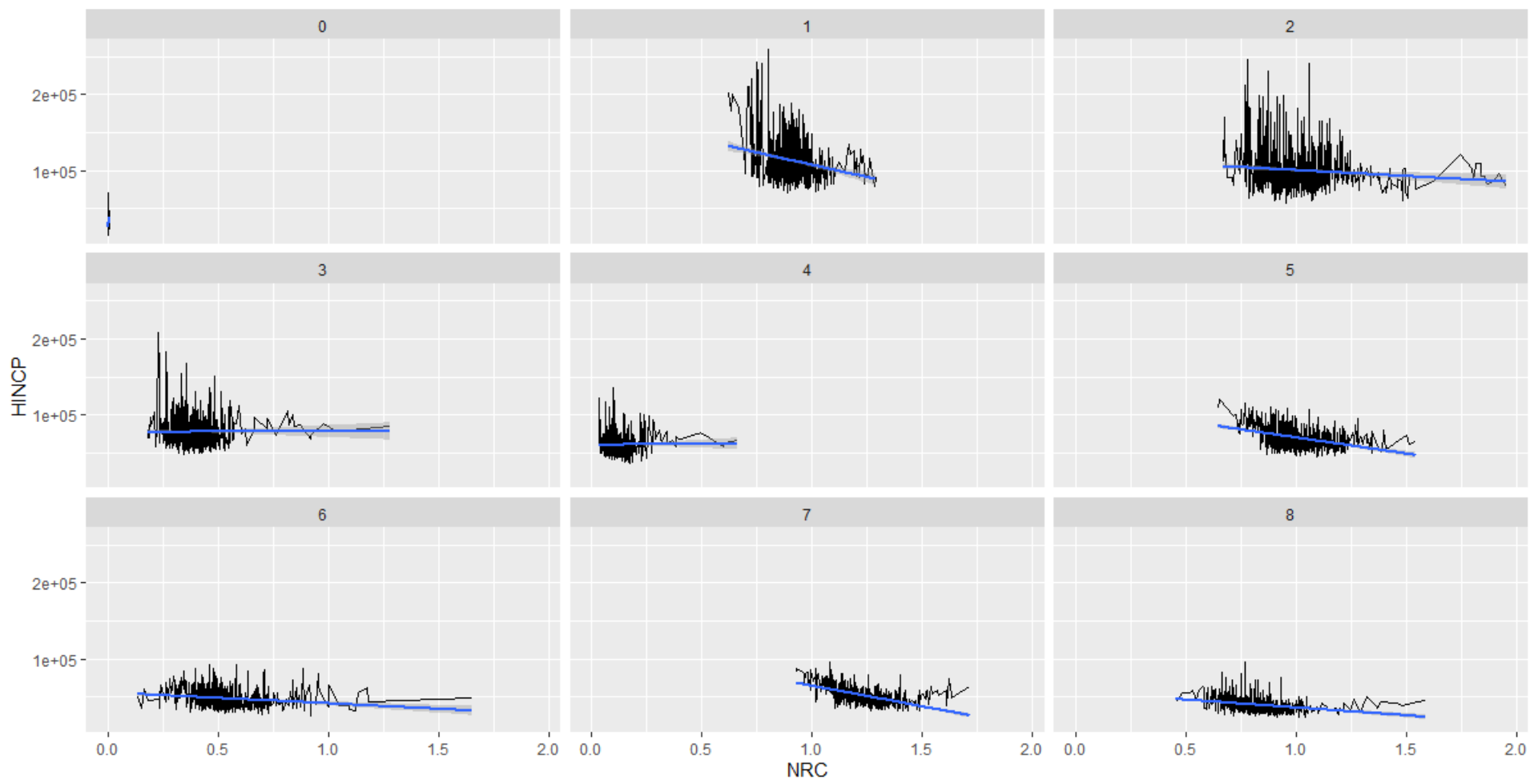
Family Income vs. Number of persons in family

In plotting the income data against the headcount data, and cutting it by the FES, we find that at every level of FES, the number of persons in family, at any given time, correlates with a higher family income. This may be due to more helping hands. But also due to the strength of reinforcement, and teamwork.



Family Income vs. Number of Related Children:

Likewise, at every level of FES, the number of related children correlates negatively with the overall family income, and likely places a burden on the adults in the workforce. This becomes an added burden for lower income households, with lower FES, where one adult is responsible for another’s child. For higher FES households too, the trend is negative.



This may be a call to action for more apprenticeships, more worker training programs, or more affordable education as intermediary steps for willing workers to join the workforce.

As the saying goes, people need a hand up, not a hand out!