Nils Baker Case Study

Chaman Preet Kaur and Chris Atterbury

October 15, 2015

1 Executive Summary

This section is to be written last and contains a short (approx 250 words) summary of the case study as a whole.

2 Background

The financial success of a bank depends on its total holdings, which in turn depends on the size of its customer base. The case study revolves around Nils Baker, vice president of a regional retail bank in US. Based on the feedback received from a promising customer, he develops a hypothesis that having more physical branches motivates people to open a checking account with the bank. In order to make a strategic business plan based on this hypothesis, it needs to proven statistically. The data available for analysis includes: total households in a particular area, total number of checking accounts with the bank in that area and the presence of physical branch or an ATM in that area. Setting up a physical branch involves a lot of investment and maintenance fee but if it can be proven to help gain more checking accounts in turn expanding the customer base and adding to the the total holdings, it can be a profitable business move.

3 Models

3.1 Initial Considerations

The data that we have received to pursue Nils Baker's hypothesis includes four columns. Initially, they are *ID*, *Total.Households.in.Area*, *Households.with.Account*, and *Inside.Outside.Footprint*. First, we drop the last two rows of the data set because they are empty and change the column names to *ID*, *Total.Households*, *Accounts*, and *Footprint*, respectively. Next, further preprocessing of the data is performed so that we have numeric data (see **Additional Codeblock 1** in the *Appendix* for the code used to do this).

The first column, *ID*, is identical to the row number and, thus, will not be considered in this case study. Each row itself is a separate Metropolitan Statistical Area (MSA). For our purposes, we need only understand that each row contains a diferent geographic region. The next two columns, *Total.Households* and *Accounts*, contain the number of total households and the number of those households that have a checking account with the bank. The last column, *Footprint*, was originally coded with either "Inside" or "Outside." A region was considered "Inside" if there was both a physical bank location and an ATM in the region. This was recoded as 1 for easier analysis of the data. A region was considered "Outside" if there was not a physical bank location and only an ATM in the region. This was recoded as 0.

Finally, we will add a column called *Accts.Hsehld* that takes the *Accounts* column and divides it by the *Total.Households* column. We may want to choose this as the response variable since this allows us to get closer to evaluating Nils Baker's hypothesis without as much interpretation of the model.

With the data in a usable form, we can now view the correlation matrix to get an initial feel for the relationships amongst the variables.

Figure 1

```
##                           ID Total.Households    Accounts  Footprint
## ID                1.0000000      -0.67957529 -0.66181262  0.3030236
## Total.Households -0.6795753       1.00000000  0.91121152 -0.3004534
## Accounts         -0.6618126       0.91121152  1.00000000 -0.2171381
## Footprint         0.3030236      -0.30045335 -0.21713807  1.0000000
## Accts.Hsehld      0.1732239      -0.08685308  0.07169036  0.1523331
##                  Accts.Hsehld
## ID                 0.17322388
## Total.Households  -0.08685308
## Accounts           0.07169036
## Footprint          0.15233308
## Accts.Hsehld       1.00000000
```

We can ignore the values involving ID since this is basically just a running counter of which number region we see. Its correlation with Total.Households is based on the fact that the data set's observations are organized by the number of households in the region with the largest first. We see that Accounts has a strong correlation with Total.Households, which is reasonable since we expect that regions with more households will have more accounts, just based on volume. Also, this explains the negative correlation between Accounts and ID. The numbers indicate what we would expect for Footprint; the correlations with the households and accounts do not show anything. That is, the locations of physical banks were not chosen based on the number of households or accounts in a region. This correlation matrix ends up showing us that trying to predict Accts.Hsehld may be difficult, but hopefully combining a few features will help. Please see **Additional Figure 1** in the *Appendix* for the pairs plot that shows these relationships graphically.

3.2 Procedure

The goal in this case study, as stated by Nils Baker, is to ascertain whether or not the presence of a physical bank in a region increases the likelihood of a given household possessing a checking account. We will start with Simple Linear Regression (SLR) models, before considering more complex Multiple Linear Regression (MLR) models.

3.3 SLR Models

3.3.1 Predicting Number of Accounts

We will start by evaluating SLR models for Accounts. Even though these models may be more difficult to interpret in terms of Nils Baker's hypothesis, if we find an especially good model, then we may make an exception. The summaries of these three models are below.

Figure 2

```
##
## Call:
## lm(formula = Accounts ~ Total.Households, data = d)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4645.8  -358.6  -223.9    77.7  7823.8
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.260e+02  1.447e+02   1.562    0.121
## Total.Households 1.086e-02  4.521e-04  24.029   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1366 on 118 degrees of freedom
## Multiple R-squared:  0.8303, Adjusted R-squared:  0.8289
## F-statistic: 577.4 on 1 and 118 DF,  p-value: < 2.2e-16
```

```
##
## Call:
## lm(formula = Accounts ~ Footprint, data = d)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2540.2 -1778.4  -819.1   -40.1 16943.4
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2627.2      395.4   6.645 9.86e-10 ***
## Footprint    -1437.6      594.9  -2.416   0.0172 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3236 on 118 degrees of freedom
## Multiple R-squared:  0.04715,    Adjusted R-squared:  0.03907
## F-statistic: 5.839 on 1 and 118 DF,  p-value: 0.01721
```

```
##
## Call:
## lm(formula = Accounts ~ Accts.Hsehld, data = d)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2671.2 -1594.0 -1433.4   143.6 16047.8
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1742.8      439.6   3.965 0.000126 ***
## Accts.Hsehld  17527.1    22448.6   0.781 0.436504
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3307 on 118 degrees of freedom
## Multiple R-squared:  0.00514,    Adjusted R-squared:  -0.003292
## F-statistic: 0.6096 on 1 and 118 DF,  p-value: 0.4365
```

Considering the results of the correlation matrix above, none of these results are particularly surprising. The total number of households in a region proves to be a good predictor of the number of accounts in the region. The coefficient of determination, denoted R^2, is high, so we know that a significant amount of the variation in `Accounts` is explained by `Total.Households`. Additionally, the p-value for the coefficient and the model are significant. However, this model only serves to highlight the drawbacks of choosing `Accounts` as the response variable since we can only determine that regions with more households have more accounts. We can say nothing about how likely a given household is to have an account.

The other two models, which use `Footprint` and `Accts.Hsehld`, are exceptionally poor. `Footprint` is the feature that most interests us since it tells us if a physical bank is in the region. While we would have been surprised to see the presence of a physical bank be a good predictor of the gross number of accounts, seeing that it is not fits in with what we would expect. Now we will look at SLR models predicting the number of accounts per household in a region.

3.3.2 Predicting Accounts per Household

Now we look at how each of the individual potential predictors does with `Accts.Hsehld` as the response variable. By using this response variable, we can easily understand which potential predictors are involved in determining the likelihood of a given household having an account with the bank. The summaries of these three models are below.

Figure 3

```
##
## Call:
## lm(formula = Accts.Hsehld ~ Total.Households, data = d)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.010135 -0.006645 -0.003025  0.001904  0.121737
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)       1.492e-02  1.432e-03  10.423   <2e-16 ***
## Total.Households -4.235e-09  4.472e-09  -0.947    0.346
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01351 on 118 degrees of freedom
## Multiple R-squared:  0.007543,   Adjusted R-squared:  -0.0008672
## F-statistic: 0.8969 on 1 and 118 DF,  p-value: 0.3456
```

```
##
## Call:
## lm(formula = Accts.Hsehld ~ Accounts, data = d)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.009282 -0.006456 -0.003580  0.002018  0.122535
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.365e-02  1.444e-03   9.454 3.94e-16 ***
## Accounts    2.932e-07  3.756e-07   0.781    0.437
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01353 on 118 degrees of freedom
## Multiple R-squared:  0.00514,    Adjusted R-squared:  -0.003292
## F-statistic: 0.6096 on 1 and 118 DF,  p-value: 0.4365
```

```
##
## Call:
## lm(formula = Accts.Hsehld ~ Footprint, data = d)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.011787 -0.006085 -0.003160  0.002384  0.120077
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.012411   0.001637   7.580 8.62e-12 ***
## Footprint   0.004125   0.002464   1.674   0.0967 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0134 on 118 degrees of freedom
## Multiple R-squared:  0.02321,    Adjusted R-squared:  0.01493
## F-statistic: 2.803 on 1 and 118 DF,  p-value: 0.09672
```

Once again, these results are aligned with those from the correlation matrix. All three SLR models are quite poor. Their respective R^2 values are very low, and their respective p-values are high. With a higher p-value we cannot reject the null hypothesis these models are just as good as the average at predicting `Accts.Hsehld`. Interestingly, the R^2 and p-values are best, albeit still terrible, for the SLR that used `Footprint`. Depending on what we find when looking at MLR models, Nils Baker's hypothesis still has a chance at being correct. Now we will turn our attention to MLR models to try to find one that does a decent job at predicting `Accts.Hsehld`.

3.4 MLR Models

3.4.1 Initial Considerations

MLR models can get quite complicated. We can look at how different terms interact with each other, quadratic terms, etc. It is imperative that we keep Nils Barker's hypothesis in mind as we look at MLR models. We want to see how well `Footprint` predicts `Accts.Hsehld`. Thus, we will focus on generating models that will help us answer that question.

3.4.2 Ramsey RESET Test

The Ramsey RESET Test attempts to determine if quadratic terms add value to a model. It does not take into consideration the interaction between quadratic terms and other variables, so this will only help to give us an intuition. The results of the test are below.

Figure 4

```
##
##  RESET test
##
## data:  lm(Accts.Hsehld ~ Accounts + Total.Households, data = d)
## RESET = 0, df1 = 1, df2 = 116, p-value = 0.9965
```

We see that the p-value is about as bad as is possible at 0.9965. This test was run for just possible quadratic terms. We will not evaluate any higher order terms with such a poor value. We did not run the test with

Footprint in the model because a qualitative variable cannot be squared. Thus, we expect that quadratic terms will be most valuable for their interactions, if at all. Now we turn our attention to actual MLR models.

3.4.3 Footprint with one other term

In the correlation matrix (**Figure 1** above) we saw that Accounts and Total.Households have a very high correlation (0.911). Thus, we will explore models with only one or the other.
The code and summaries for each of the models explored can be found in **Additional Codeblock 2** in the *Appendix*. However, they will not be displayed here because none offered promising results. R^2 only goes up to 0.02741. When R^2 is adjusted to account for the number of predictors, R_a^2 only gets up to a miniscule 0.008395. With these dispiriting results we turn our attention to MLR models with both Accounts and Total.Households.

3.4.4 Using all potential predictors

We know that error cannot be unexplained. Thus, we can look at the R^2 value with all possible quadratic and interaction terms to get a sense of just how high we can hope to get that value when building our model. Additionally, we can compare R_a^2 to make sure that our chosen model improves on this. See below.

Figure 5

```
##
## Call:
## lm(formula = Accts.Hsehld ~ (. - ID + I(Accounts^2) + I(Total.Households^2))^2,
##     data = d)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.023668 -0.002990 -0.000442  0.002224  0.067033
##
## Coefficients:
##                                        Estimate Std. Error t value
## (Intercept)                           1.421e-02  2.309e-03   6.155
## Total.Households                     -2.303e-07  6.991e-08  -3.294
## Accounts                              1.649e-05  5.187e-06   3.179
## Footprint                             4.773e-03  3.204e-03   1.490
## I(Accounts^2)                        -1.645e-09  1.657e-09  -0.992
## I(Total.Households^2)                 4.947e-13  1.847e-13   2.678
## Total.Households:Accounts            -8.893e-12  2.363e-11  -0.376
## Total.Households:Footprint           -5.193e-07  1.193e-07  -4.353
## Total.Households:I(Accounts^2)       -8.068e-15  5.562e-15  -1.451
## Total.Households:I(Total.Households^2) -5.887e-19  2.365e-19  -2.490
## Accounts:Footprint                    3.012e-05  5.980e-06   5.037
## Accounts:I(Accounts^2)                2.793e-13  1.975e-13   1.415
## Accounts:I(Total.Households^2)        8.509e-17  5.784e-17   1.471
## Footprint:I(Accounts^2)              -3.212e-09  7.623e-10  -4.213
## Footprint:I(Total.Households^2)       1.009e-12  3.746e-13   2.693
## I(Accounts^2):I(Total.Households^2)   8.509e-22  6.355e-22   1.339
##                                      Pr(>|t|)
## (Intercept)                          1.43e-08 ***
## Total.Households                      0.00135 **
## Accounts                              0.00195 **
## Footprint                               0.13937    
## I(Accounts^2)                           0.32332    
## I(Total.Households^2)                   0.00860 ** 
## Total.Households:Accounts               0.70747    
## Total.Households:Footprint             3.15e-05 ***
## Total.Households:I(Accounts^2)          0.14990    
## Total.Households:I(Total.Households^2)  0.01438 *  
## Accounts:Footprint                     2.00e-06 ***
## Accounts:I(Accounts^2)                  0.16018    
## Accounts:I(Total.Households^2)          0.14428    
## Footprint:I(Accounts^2)                5.38e-05 ***
## Footprint:I(Total.Households^2)         0.00825 ** 
## I(Accounts^2):I(Total.Households^2)     0.18352    
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 0.008973 on 104 degrees of freedom
## Multiple R-squared:  0.6141, Adjusted R-squared:  0.5585 
## F-statistic: 11.03 on 15 and 104 DF,  p-value: 1.699e-15
```

We find that $R^2 = 0.6141$ and $R^2_a = 0.5585$. Now we can evaluate how less bloated models compare.

```
## 
## Call:
## lm(formula = Accts.Hsehld ~ . - ID + Footprint:Accounts + Total.Households:Footprint + 
##     I(Accounts^2) + I(Total.Households^2) + Footprint:I(Accounts^2) + 
##     Footprint:I(Total.Households^2), data = d)
## 
## Residuals:
##       Min        1Q    Median        3Q       Max 
## -0.024939 -0.003457 -0.001472  0.002135  0.068849 
## 
## Coefficients:
##                                   Estimate Std. Error t value Pr(>|t|)    
## (Intercept)                      1.182e-02  1.653e-03   7.150 1.01e-10 ***
## Total.Households                -9.749e-08  2.366e-08  -4.121 7.32e-05 ***
## Accounts                         9.509e-06  2.362e-06   4.026 0.000104 ***
## Footprint                        7.859e-03  2.680e-03   2.932 0.004096 ** 
## I(Accounts^2)                   -5.234e-10  1.846e-10  -2.836 0.005439 ** 
## I(Total.Households^2)            5.402e-14  1.860e-14   2.904 0.004450 ** 
## Accounts:Footprint               3.403e-05  4.873e-06   6.984 2.30e-10 ***
## Total.Households:Footprint      -6.379e-07  7.951e-08  -8.023 1.22e-12 ***
## Footprint:I(Accounts^2)         -3.203e-09  5.120e-10  -6.255 7.78e-09 ***
## Footprint:I(Total.Households^2)  1.244e-12  1.672e-13   7.443 2.33e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 0.009204 on 110 degrees of freedom
## Multiple R-squared:  0.5706, Adjusted R-squared:  0.5354 
## F-statistic: 16.24 on 9 and 110 DF,  p-value: < 2.2e-16
```

Here we find all coefficients are significant at the $\alpha = 0.05$ level. That is, we can be at least 95% confident in the significance of these coefficients. $R^2 = 0.5706$ and $R^2_a = 0.5354$. These values are close to the obese one above. Also, the interaction terms with `Footprint` are all the most significant ones.

Details how we arrived at the models that we tried and how they worked. Include details about both the processes and the model to which they led. Our thought processes should be detailed so that there is no question how we got to our models. This needs to do all of the leg work so that the Conclusions section can focus on the actual meaning of the model. *This section should include subsections for each model with two hashes and then further subsectioniong using three hashes for each part of the model discussion.*

4 Conclusions

The goal is that everything is built up to this point so that little we can just plow right into the meaning of the model. Other general conclusions can be included.

5 Appendix

5.1 Additional Codeblocks

Additional Codeblock 1

```
d <- read.csv("41330723.csv", header = TRUE, stringsAsFactors = FALSE)
d <- d[1:120, ] # last two rows contain no data
names(d) <- c("ID", "Total.Households", "Accounts", "Footprint")

for (i in 2:3) {
  d[[i]] <- gsub(",", "", d[[i]])
  d[[i]] <- as.numeric(d[[i]])
}
d[["Footprint"]][d[["Footprint"]] == "Outside"] <- 0
d[["Footprint"]][d[["Footprint"]] == "Inside"] <- 1
d[["Footprint"]] <- as.numeric(d[["Footprint"]])
d[["ID"]] <- as.numeric(d[["ID"]])
d[["Accts.Hsehld"]] <- d[["Accounts"]] / d[["Total.Households"]]
```

```
head(d)
```

```
##   ID Total.Households Accounts Footprint Accts.Hsehld
## 1  1          1772960    17563         0  0.009906033
## 2  2          1345209    14547         0  0.010813933
## 3  3           960434    10847         0  0.011293853
## 4  4           928274    18133         1  0.019534103
## 5  5           893995     5291         0  0.005918378
## 6  6           812137     6297         0  0.007753618
```

Additional Codeblock 2

```
summary(lm(Accts.Hsehld ~ Footprint + Total.Households, data = d))
```

```
##
## Call:
## lm(formula = Accts.Hsehld ~ Footprint + Total.Households, data = d)
##
```

```
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.011920 -0.006122 -0.003293  0.002291  0.119948
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)       1.293e-02  1.979e-03   6.536 1.72e-09 ***
## Footprint         3.758e-03  2.592e-03   1.450    0.150
## Total.Households -2.202e-09  4.667e-09  -0.472    0.638
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01345 on 117 degrees of freedom
## Multiple R-squared:  0.02506,    Adjusted R-squared:  0.008395
## F-statistic: 1.504 on 2 and 117 DF,  p-value: 0.2266
```

```
summary(lm(Accts.Hsehld ~ Footprint + Total.Households +
             Footprint:Total.Households, data = d))
```

```
##
## Call:
## lm(formula = Accts.Hsehld ~ Footprint + Total.Households + Footprint:Total.Households,
##     data = d)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.011634 -0.006183 -0.003389  0.002127  0.120225
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 1.306e-02  2.022e-03   6.461 2.54e-09 ***
## Footprint                   3.299e-03  2.922e-03   1.129    0.261
## Total.Households           -2.754e-09  4.950e-09  -0.556    0.579
## Footprint:Total.Households  5.284e-09  1.532e-08   0.345    0.731
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0135 on 116 degrees of freedom
## Multiple R-squared:  0.02606,    Adjusted R-squared:  0.0008719
## F-statistic: 1.035 on 3 and 116 DF,  p-value: 0.38
```

```
summary(lm(Accts.Hsehld ~ Footprint + Total.Households +
             Footprint:Total.Households + I(Total.Households^2),
             data = d))
```

```
##
## Call:
## lm(formula = Accts.Hsehld ~ Footprint + Total.Households + Footprint:Total.Households +
##     I(Total.Households^2), data = d)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.011703 -0.006198 -0.003313  0.002467  0.120160
```

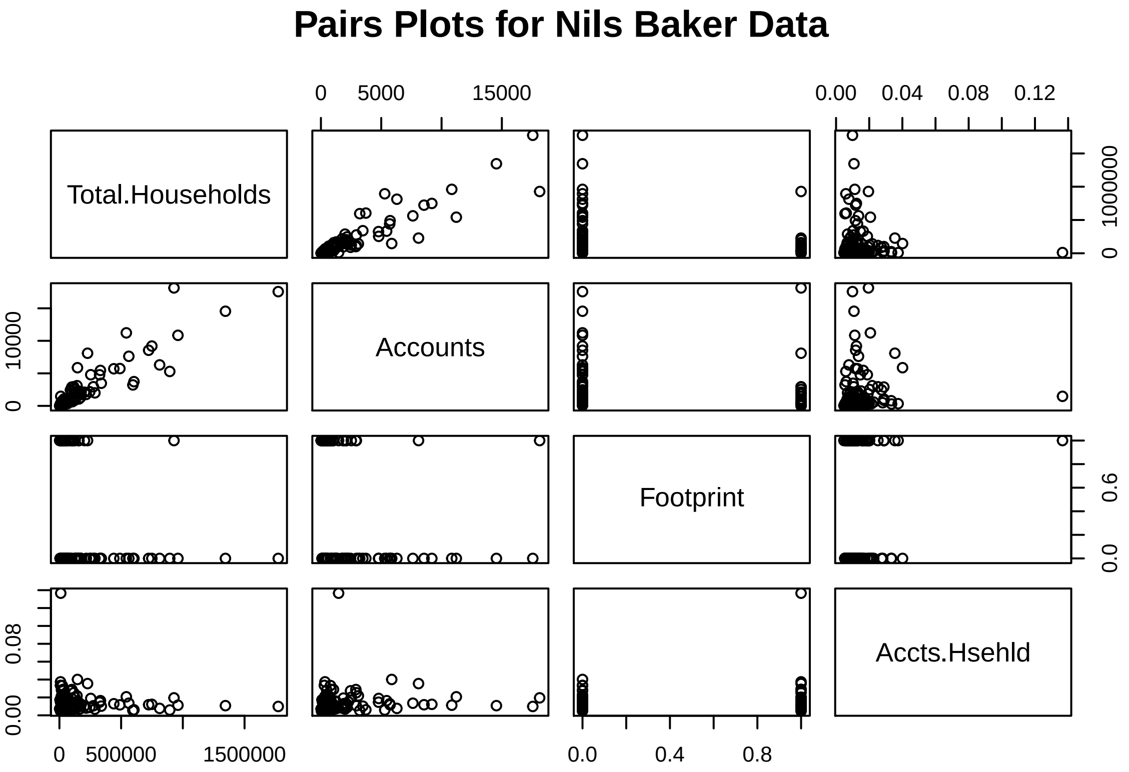
```
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 1.336e-02  2.363e-03   5.652 1.17e-07 ***
## Footprint                   3.091e-03  3.056e-03   1.011    0.314
## Total.Households           -5.533e-09  1.246e-08  -0.444    0.658
## I(Total.Households^2)       2.172e-15  8.934e-15   0.243    0.808
## Footprint:Total.Households  6.161e-09  1.580e-08   0.390    0.697
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01355 on 115 degrees of freedom
## Multiple R-squared:  0.02656,    Adjusted R-squared:  -0.007298
## F-statistic: 0.7844 on 4 and 115 DF,  p-value: 0.5375
```

```
summary(lm(Accts.Hsehld ~ Footprint + Total.Households +
             Footprint:Total.Households + I(Total.Households^2) +
             Footprint:I(Total.Households^2),
           data = d))
```

```
##
## Call:
## lm(formula = Accts.Hsehld ~ Footprint + Total.Households + Footprint:Total.Households +
##     I(Total.Households^2) + Footprint:I(Total.Households^2),
##     data = d)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.012253 -0.006095 -0.003269  0.002092  0.119637
##
## Coefficients:
##                                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)                      1.329e-02  2.381e-03   5.583 1.62e-07 ***
## Footprint                        3.836e-03  3.870e-03   0.991    0.324
## Total.Households                -4.936e-09  1.265e-08  -0.390    0.697
## I(Total.Households^2)            1.706e-15  9.090e-15   0.188    0.851
## Footprint:Total.Households      -9.514e-09  5.207e-08  -0.183    0.855
## Footprint:I(Total.Households^2)  1.769e-14  5.597e-14   0.316    0.753
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01361 on 114 degrees of freedom
## Multiple R-squared:  0.02741,    Adjusted R-squared:  -0.01524
## F-statistic: 0.6426 on 5 and 114 DF,  p-value: 0.6676
```

5.2 Additional Figures

Additional Figure 1


Pairs Plots for Nils Baker Data
0
5000
15000
0.00
0.04
0.08
0.12
Total.Households
0
1000000
Accounts
0
10000
Footprint
0.0
0.6
Accts.Hsehld
0.00
0.08
0
500000
1500000
0.0
0.4
0.8