Brand Equity Case Analysis

Alicja Wilk

3/4/2019

Measurement of Brand Equity

Brand Equity per definition is value added to the company and its brand due to its name. The higher the brand equity the better the competitive advantage of a given brand. Ariel measured brand equity with five factors: Familiarity, popularity, relevancy, loyalty & uniqueness.Brand Equity is not an easy thing to measrue due to its complexity. Brand Equity requires us to often try and capture the intangible assets. The best way to discover the equity of the brand is to convey a survey(form of a research design) of the valued customers.

Ariel measured brand equity through mulitdimensional factors. The brand equity score is the measurement of the mean score of all 5 variables.

Besides descriptive statistics and visualizations, statistical measures were used in the report:

ANOVA, MANOVA - good statistical tool when comparing multiple groups of independent categorical variables,

CROSSTAB - statistical tool to measure the frequency of the variable in the data and to break down two categorical varaibles and the expected count between each level. Gives us instant quick insights into the variable that we are interested in,

CHI-SQRT TEST - since we are dealing with categorical data, the test of independance between two variables was used to determine wheter there is a significant relationship between them. This test is suitable to discover whether or not the variables are related since we are trying to discover which demographics influence loyalty,

MULTIPLE COMPARISIONS - interest for conducting multiple comparisons test is to see how the variable of interest(in our case loyalty) differs among age groups or income groups etc,

REGRESSION ANALYSIS - used in the report to discover the relationship and significance between variables and whether targeted customers characeristics would improve the relationship or not.

Descriptive Statistics

First, descriptive statistics were conduted on the data set. The missign values were removed and disregarded as it was assumed the sample size allowed that. There are 1459 observations in the data with 18 variables. From the case study, the description on the variables:

Demographic Variables Region 1=Maritimes; 2=Quebec; 3=Ontario; 4=West Gender 1=Female; 2=Male Age in years Children (1=Yes; 2=No) Income 1= < $30k; 2= $30-$49.9k; 3= $50-$74.9k; 4=$75k+

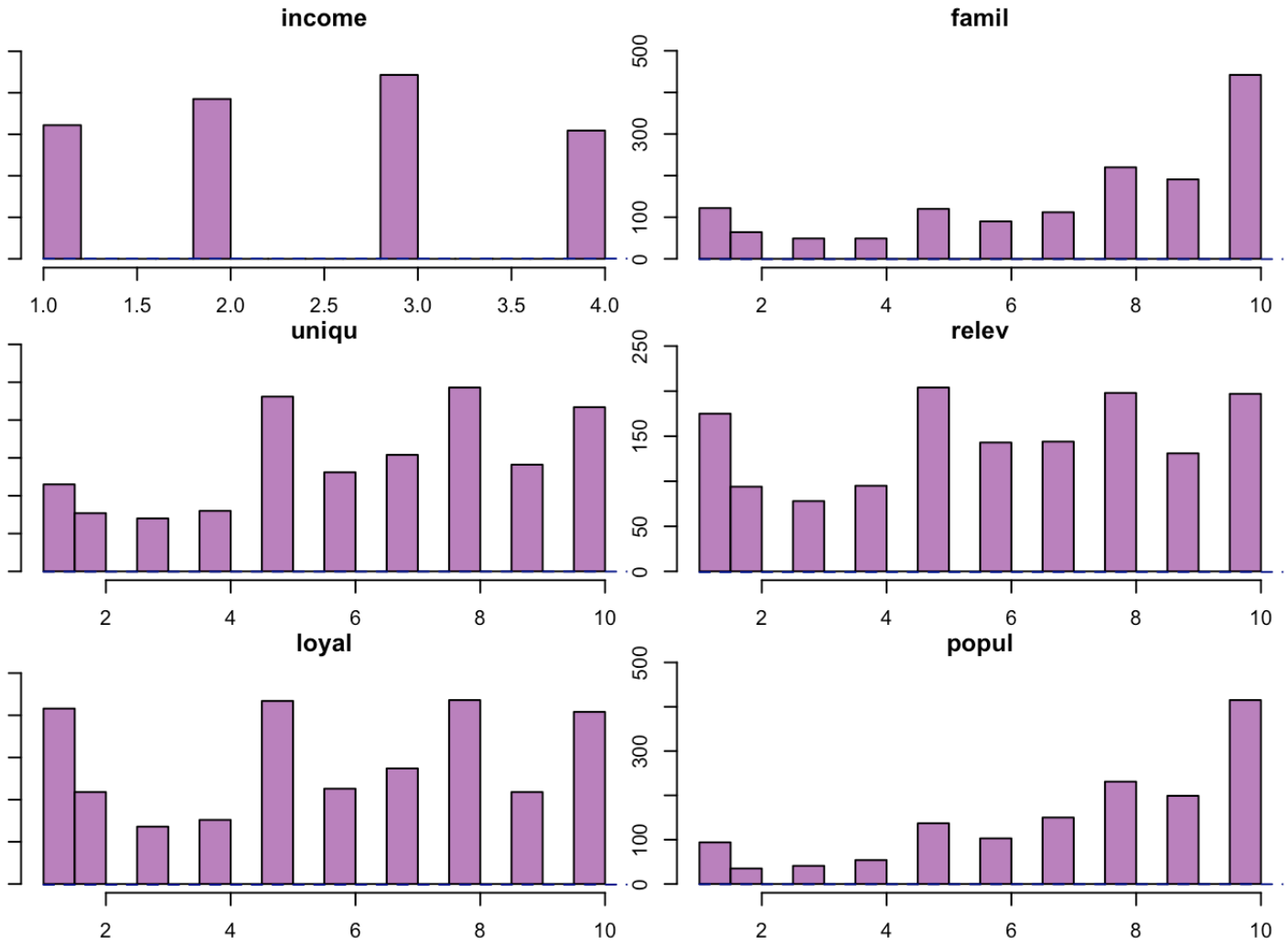
Famil I am familiar and understand what this brand is about. Uniqu This brand has unique or different features or a distinct image other brands in this category don't have. Relev This brand is appropriate and fits my lifestyle and needs.

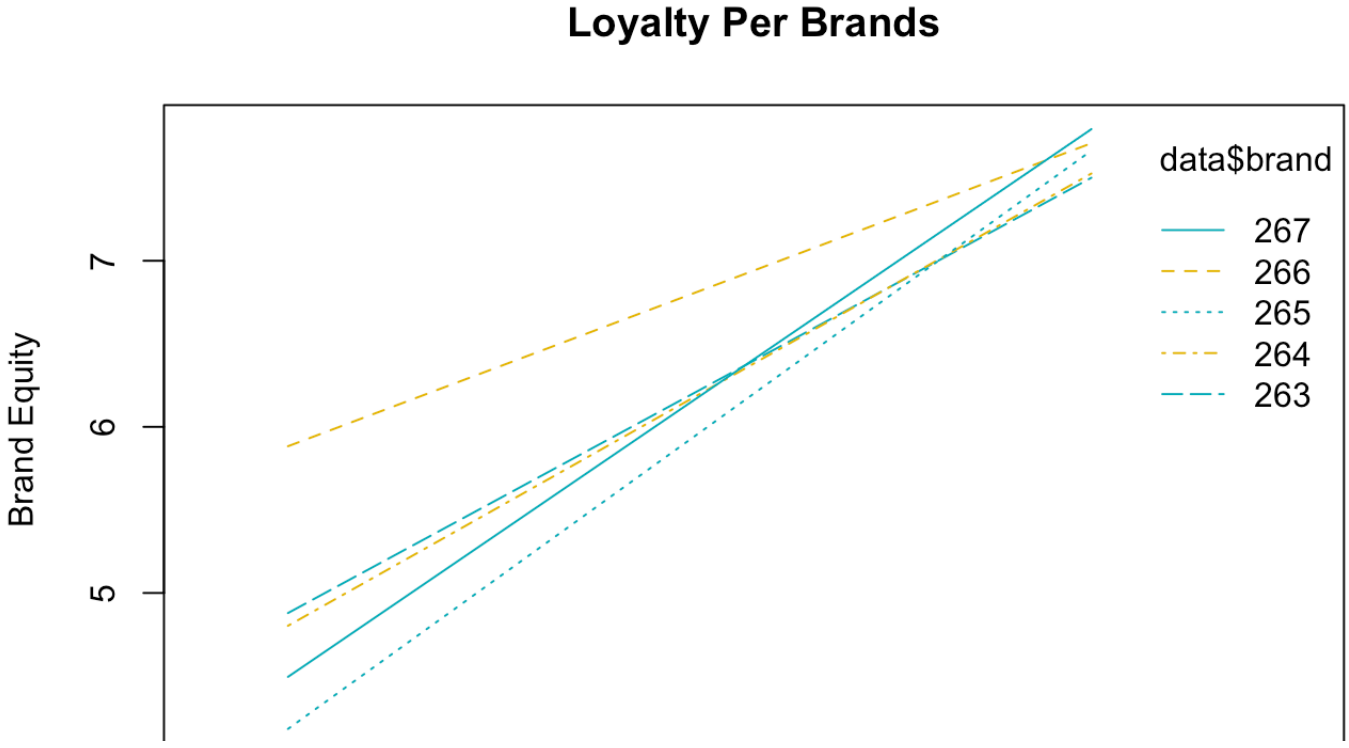
Familbin 0 = not loyal (responses of 1 to 7); 1 = loyal (responses of 8-10) Uniqubin 0 = not loyal (responses of 1 to 7); 1 = loyal (responses of 8-10) Relevbin 0 = not loyal (responses of 1 to 7); 1 = loyal (responses of 8-10) Loyalbin 0 = not loyal (responses of 1 to 7); 1 = loyal (responses of 8-10) Populbin 0 = not loyal (responses of 1 to 7); 1 = loyal (responses of 8-10)

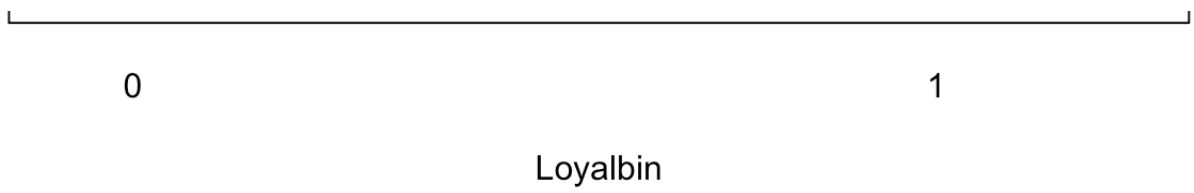
Comparing loyalty, relevance, familiarity, uniqueness and popularity for its brands.

Descriptive Statistics On Entire Dataset:

```
##              vars    n   mean    sd median trimmed   mad min max range
## gender          1 1459   1.41  0.49   1.00    1.38  0.00   1   2     1
## age             2 1459  40.58 10.67  39.00   40.14 11.86  20  64    44
## region          3 1459   2.92  0.93   3.00    3.01  1.48   1   4     3
## children        4 1459   1.51  0.50   2.00    1.51  0.00   1   2     1
## income          5 1459   2.51  1.06   3.00    2.51  1.48   1   4     3
## famil           6 1459   7.14  2.95   8.00    7.53  2.97   1  10     9
## uniqu           7 1459   6.31  2.75   7.00    6.48  2.97   1  10     9
## relev           8 1459   5.89  2.92   6.00    5.99  2.97   1  10     9
## loyal           9 1459   5.77  3.02   6.00    5.84  2.97   1  10     9
## popul          10 1459   7.30  2.71   8.00    7.68  2.97   1  10     9
## category       11 1459  12.00  0.00  12.00   12.00  0.00  12  12     0
## brand          12 1459 265.04  1.40 265.00  265.05  1.48 263 267     4
## familbin       13 1459   0.58  0.49   1.00    0.61  0.00   0   1     1
## uniqubin       14 1459   0.41  0.49   0.00    0.39  0.00   0   1     1
## relevbin       15 1459   0.36  0.48   0.00    0.33  0.00   0   1     1
## loyalbin       16 1459   0.36  0.48   0.00    0.33  0.00   0   1     1
## populbin       17 1459   0.58  0.49   1.00    0.60  0.00   0   1     1
## brand_equity   18 1459   5.82  2.02   6.17    5.97  1.99   1   9     8
##                skew kurtosis   se
## gender         0.38    -1.86 0.01
## age            0.33    -0.78 0.28
## region        -0.44    -0.75 0.02
## children      -0.03    -2.00 0.01
## income        -0.04    -1.21 0.03
## famil         -0.84    -0.55 0.08
## uniqu         -0.41    -0.85 0.07
## relev         -0.25    -1.10 0.08
## loyal         -0.22    -1.20 0.08
## popul         -0.91    -0.15 0.07
## category        NaN      NaN 0.00
## brand         -0.04    -1.27 0.04
## familbin      -0.34    -1.88 0.01
## uniqubin       0.36    -1.87 0.01
## relevbin       0.58    -1.66 0.01
## loyalbin       0.56    -1.68 0.01
## populbin      -0.32    -1.90 0.01
## brand_equity -0.57    -0.47 0.05
```


income
famil
uniqu
relev
loyal
popul


Loyalty Per Brands
Brand Equity
data$brand
267
266
265
264
263



Quick insights from the descriptive statistics: - the most frequent region is Ontario - the most frequent customer age between 30 and 40 years old - there are more females than males in the sample - the mean for brand equity is 6(rounded) - se is small for most variables - the mean for 'brand' is Harvey's

From the graphs we can infer that 10 was the most frequent score for 'popul' and 'famil' which means that most of the brands are well known. The most frequent income group is between 50K and 74.9K. 'Uniqu','relev' and 'loyal' and more normally distributed the other variables. From the graph we can indicate that McDonalds has higher scores than other brands.

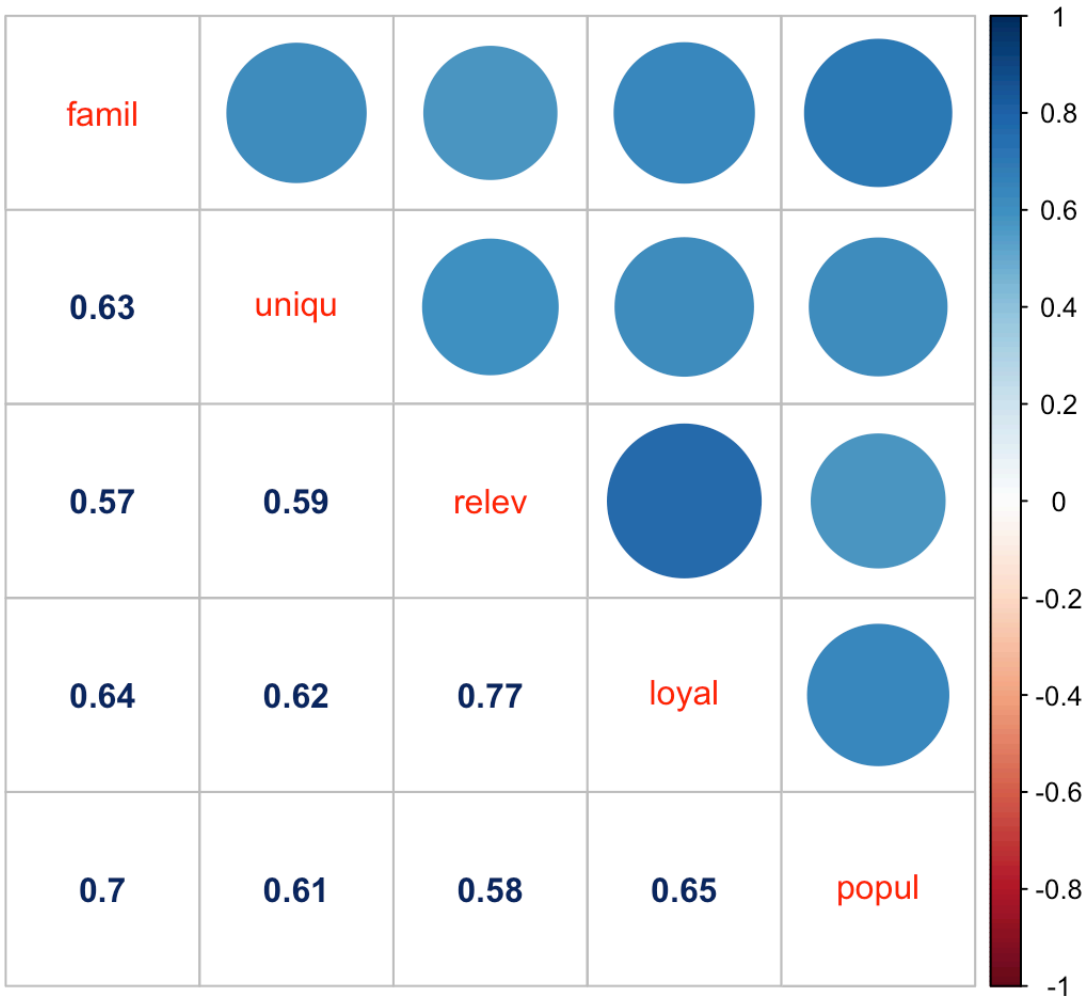
MANOVA analysis

Since there are multiple dependent variables MANOVA was used to analyse and compare loyalty, relevance, familiarity, uniqueness and popularity against the brands. We assume that observations are independent from each other, absense of multicollinearity and multivariate normality. Same sizes are equal among variables, we assume robustness of Anova and use it in our analysis.

Hypothesis Test:

The null hypothesis is that there the vectors are of equal mean and the alternative hypothesis is that at least one means is different.

$H0{:}\mu 1{=}\mu 2{=}\ldots{=}\mu k.\ H1$ μi⁄=μi′ forsome*i*=i




```
##             Df   Pillai approx F num Df den Df   Pr(>F)
## brand_var    1 0.012433   3.6586      5   1453 0.002709 **
## Residuals 1457
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##             Df   Wilks approx F num Df den Df   Pr(>F)
## brand_var    1 0.98757   3.6586      5   1453 0.002709 **
## Residuals 1457
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##             Df Hotelling-Lawley approx F num Df den Df   Pr(>F)
## brand_var    1          0.01259   3.6586      5   1453 0.002709 **
## Residuals 1457
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

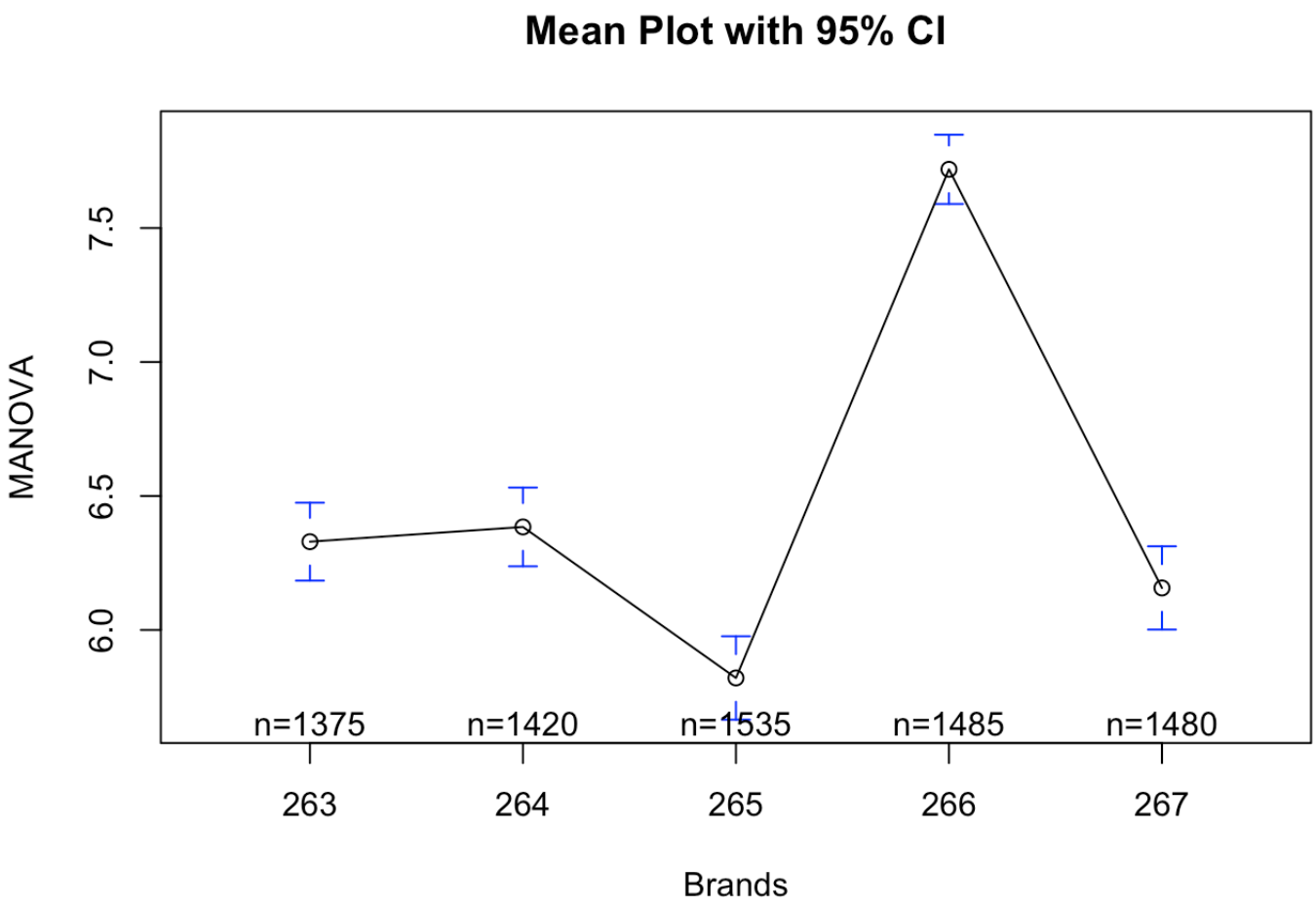
```
##             Df     Roy approx F num Df den Df   Pr(>F)
## brand_var    1 0.01259   3.6586      5   1453 0.002709 **
## Residuals 1457
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

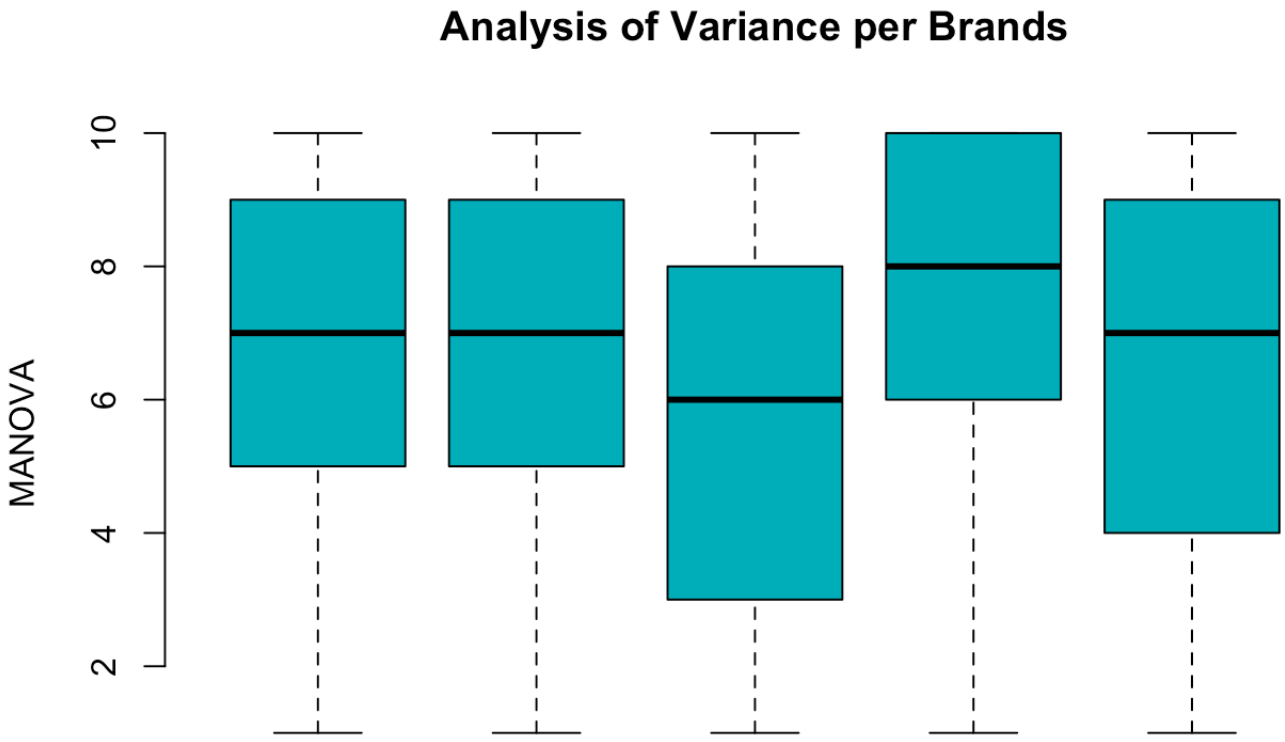
Results

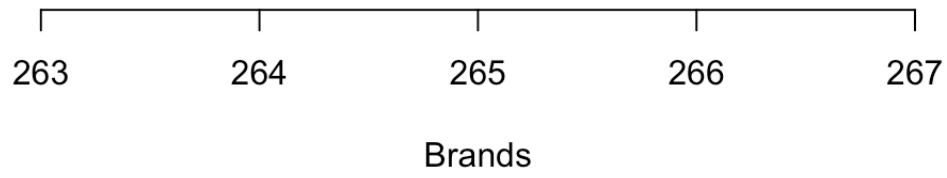
When conducting MANOVA we look at the within groups and between groups variation and significant differences. The most popular tests to test the for significance are:“Pillai”,“Wilks”(likelihood, ratio of determinants),“Hotelling-Lawley” and “Roy”.

With $p < .05$ we can reject the null hypothesis and conduct that at least one pair of means is different.

We use mutiple comparisons to further determine which specific groups have different mean vectors. The method chosen for this is The Tukey Honest Significant Difference.The method is chosen over the Bonferroni Adjustment or Holm Adjustment as those method might seem conservative. We are given adjusted p-values for the pair wise tests. Looking at statistics it shows that brand has statistical inluence on dependent variables ‘uniqu’,‘relev’ and ‘popul’ are significant for the chosen brands so we can conclude that there is statistically significant association. It is possible that the values and significance would change if we played with the brands that we are taking into account especially knowing that McDonalds is one of the brands that ranks higher and gets better responses than rest.


Mean Plot with 95% CI
MANOVA
6.0
6.5
7.0
7.5
n=1375
n=1420
n=1535
n=1485
n=1480
263
264
265
266
267
Brands


Analysis of Variance per Brands
MANOVA
2
4
6
8
10

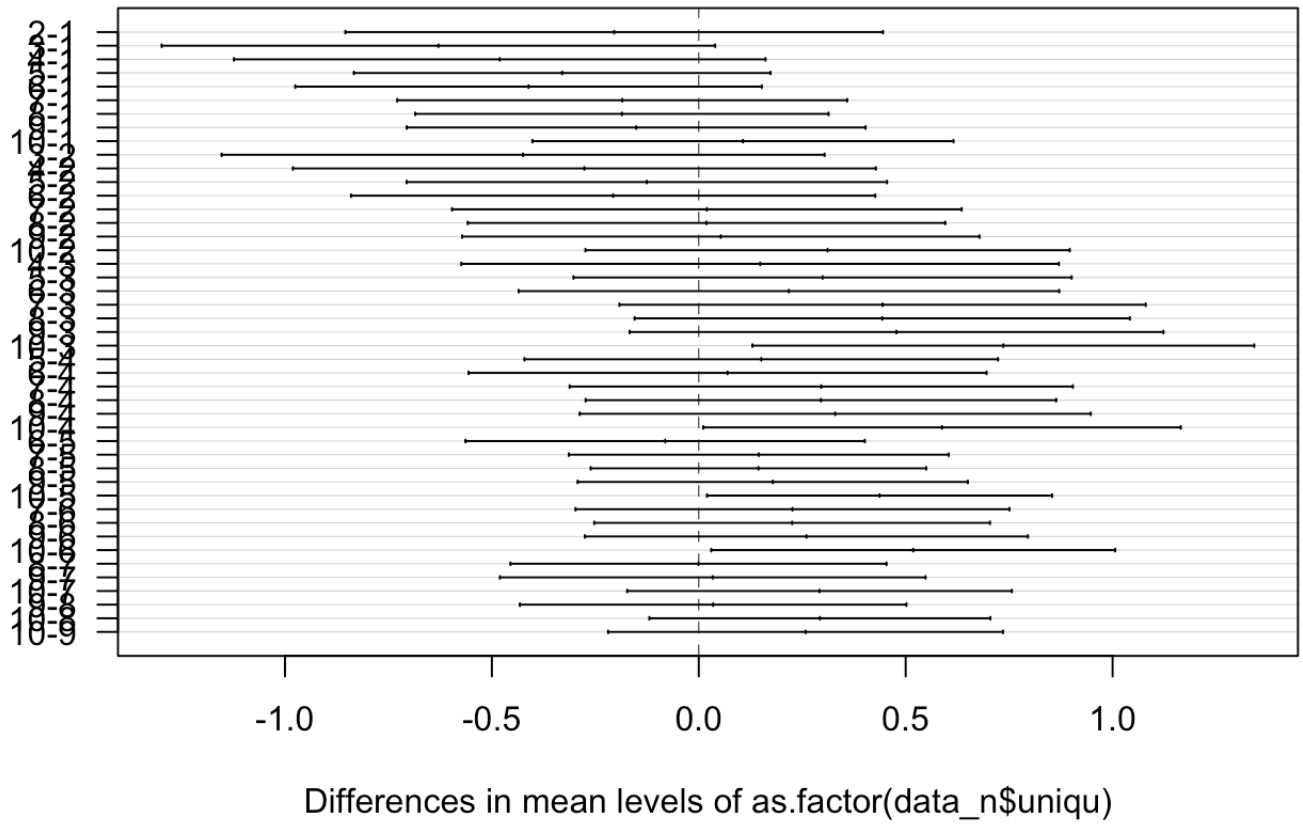



```
##               Df Sum Sq Mean Sq F value  Pr(>F)
## dep_vars       5   35.6   7.119   3.659 0.00271 **
## Residuals   1453 2827.3   1.946
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## Analysis of Variance Table
##
##               Df  Pillai approx F num Df den Df    Pr(>F)
## (Intercept)    1 0.89355  2439.21      5   1453 < 2.2e-16 ***
## brand_var      1 0.01243     3.66      5   1453  0.002709 **
## Residuals   1457
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

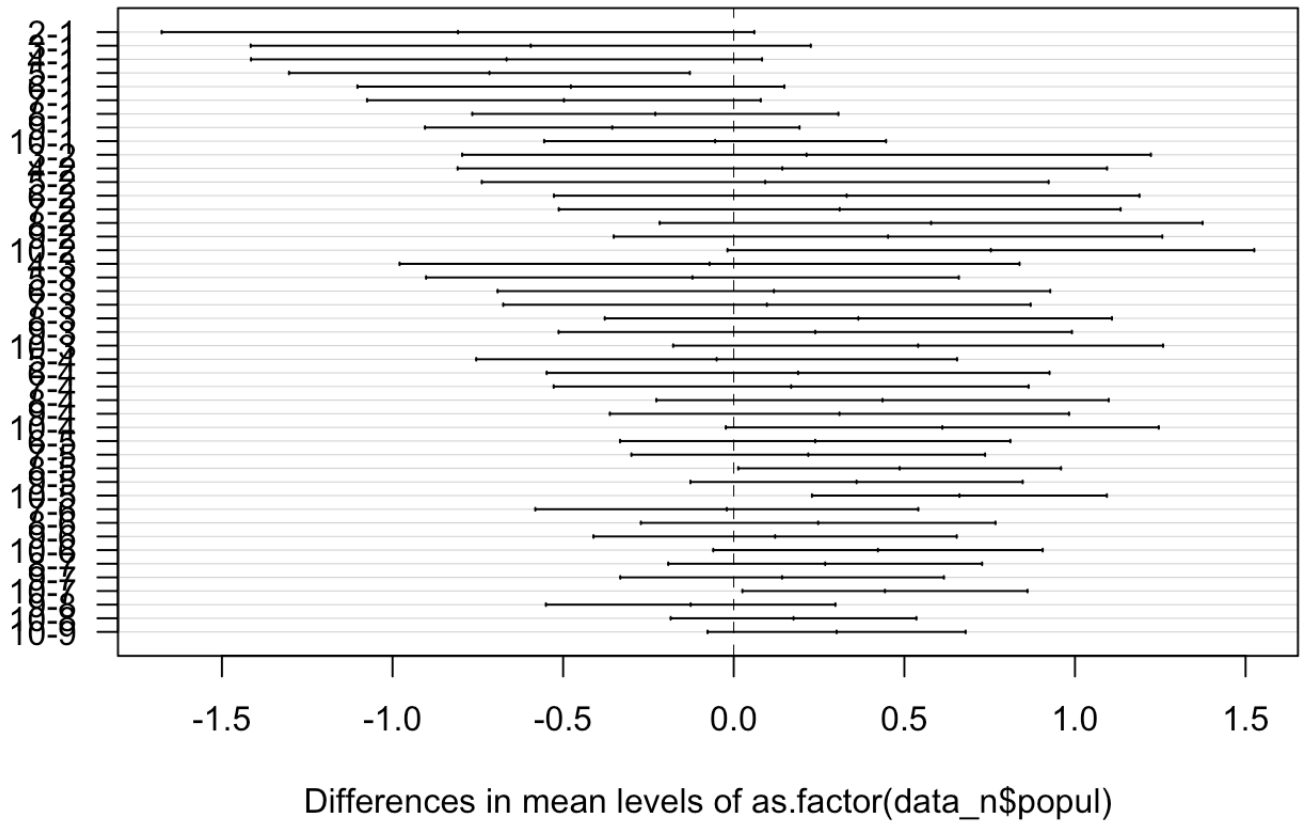
Pair Wise Comparison with Adjusted P Value

Significant difference for brands with ‘popul’ score 5 on one side and 1 on another -> ## 5-1 , 0.0045858 and also for brands with score 10 on one side and 5 on another -> 10-5 , 0.0000598. For variable ‘uniqu’ 10-3 , 0.0049017.




```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = data_n$brand ~ as.factor(data_n$uniqu))
##
## $`as.factor(data_n$uniqu)`
##                diff         lwr        upr     p adj
## 2-1  -0.2045172219 -0.85410591 0.44507146 0.9924219
## 3-1  -0.6291925466 -1.29795217 0.03956707 0.0856020
## 4-1  -0.4809782609 -1.12323033 0.16127381 0.3429300
## 5-1  -0.3300583474 -0.83351868 0.17340199 0.5439593
## 6-1  -0.4114171922 -0.97513935 0.15230496 0.3812965
## 7-1  -0.1850367024 -0.72872374 0.35865033 0.9867386
## 8-1  -0.1858650921 -0.68517660 0.31344642 0.9754244
## 9-1  -0.1512796793 -0.70557808 0.40301872 0.9974128
## 10-1  0.1067521539 -0.40207745 0.61558176 0.9996835
## 3-2  -0.4246753247 -1.15320359 0.30385294 0.7048019
## 4-2  -0.2764610390 -0.98073493 0.42781285 0.9651178
## 5-2  -0.1255411255 -0.70604634 0.45496409 0.9995941
```

```
## 6-2  -0.2068999703 -0.84037978 0.42657984 0.9901102
## 7-2   0.0194805195 -0.59623824 0.63519928 1.0000000
## 8-2   0.0186521298 -0.55825859 0.59556285 1.0000000
## 9-2   0.0532375426 -0.57187102 0.67834610 0.9999999
## 10-2  0.3112693758 -0.27389861 0.89643736 0.8038036
## 4-3   0.1482142857 -0.57377996 0.87020853 0.9997354
## 5-3   0.2991341991 -0.30274642 0.90101482 0.8603020
## 6-3   0.2177753544 -0.43534840 0.87089911 0.9885310
## 7-3   0.4441558442 -0.19175576 1.08006745 0.4480515
## 8-3   0.4433274544 -0.15508708 1.04174199 0.3589056
## 9-3   0.4779128673 -0.16709464 1.12292038 0.3586955
## 10-3  0.7359447005  0.12956566 1.34232374 0.0049017
## 5-4   0.1509199134 -0.42136373 0.72320355 0.9980171
## 6-4   0.0695610687 -0.55639334 0.69551548 0.9999986
## 7-4   0.2959415584 -0.31203200 0.90391512 0.8748348
## 8-4   0.2951131687 -0.27352401 0.86375035 0.8259483
## 9-4   0.3296985816 -0.28778257 0.94717973 0.8002464
## 10-4  0.5877304147  0.01071757 1.16474326 0.0418171
## 6-5  -0.0813588447 -0.56385601 0.40113832 0.9999491
## 7-5   0.1450216450 -0.31390802 0.60395131 0.9922157
## 8-5   0.1441932553 -0.26118660 0.54957311 0.9819587
## 9-5   0.1787786681 -0.29267394 0.65023128 0.9721904
## 10-5  0.4368105013  0.01976324 0.85385776 0.0314047
## 7-6   0.2263804897 -0.29795414 0.75071512 0.9367445
## 8-6   0.2255521000 -0.25261439 0.70371859 0.8945045
## 9-6   0.2601375129 -0.27519222 0.79546725 0.8759699
## 10-6  0.5181693460  0.03007225 1.00626644 0.0271560
## 8-7  -0.0008283897 -0.45520281 0.45354603 1.0000000
## 9-7   0.0337570231 -0.48043248 0.54794653 1.0000000
## 10-7  0.2917888563 -0.17302477 0.75660248 0.6068854
## 9-8   0.0345854128 -0.43243411 0.50160494 1.0000000
## 10-8  0.2926172460 -0.11941199 0.70464648 0.4226276
## 10-9  0.2580318332 -0.21915034 0.73521400 0.7879782
```




```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = data_n$brand ~ as.factor(data_n$popul))
##
## $`as.factor(data_n$popul)`
##              diff         lwr         upr     p adj
## 2-1  -0.80820669 -1.67655694  0.06014357 0.0933680
## 3-1  -0.59496627 -1.41571281  0.22578027 0.3914785
## 4-1  -0.66587864 -1.41468252  0.08292523 0.1313404
## 5-1  -0.71602733 -1.30335412 -0.12870054 0.0045858
## 6-1  -0.47727742 -1.10280809  0.14825324 0.3158480
## 7-1  -0.49773050 -1.07460771  0.07914672 0.1611115
## 8-1  -0.22985171 -0.76635179  0.30664838 0.9396250
## 9-1  -0.35608896 -0.90492342  0.19274551 0.5593049
## 10-1 -0.05467829 -0.55559895  0.44624238 0.9999988
## 3-2   0.21324042 -0.79596338  1.22244422 0.9996645
## 4-2   0.14232804 -0.80928883  1.09394491 0.9999818
## 5-2   0.09217935 -0.73837479  0.92273350 0.9999986
```

```
## 6-2   0.33092926 -0.52706610  1.18892463 0.9688971
## 7-2   0.31047619 -0.51272169  1.13367407 0.9731734
## 8-2   0.57835498 -0.21706929  1.37377925 0.3869187
## 9-2   0.45211773 -0.35167745  1.25591291 0.7465163
## 10-2  0.75352840 -0.01834525  1.52540205 0.0624185
## 4-3  -0.07091238 -0.97929962  0.83747486 0.9999999
## 5-3  -0.12106106 -0.90171009  0.65958796 0.9999751
## 6-3   0.11768885 -0.69209430  0.92747200 0.9999857
## 7-3   0.09723577 -0.67558210  0.87005364 0.9999959
## 8-3   0.36511456 -0.37804927  1.10827839 0.8685165
## 9-3   0.23887731 -0.51323931  0.99099393 0.9919222
## 10-3  0.54028798 -0.17761288  1.25818884 0.3357118
## 5-4  -0.05014869 -0.75477288  0.65447550 1.0000000
## 6-4   0.18860122 -0.54816951  0.92537195 0.9984291
## 7-4   0.16814815 -0.52778994  0.86408623 0.9990059
## 8-4   0.43602694 -0.22682662  1.09888049 0.5389227
## 9-4   0.30978969 -0.36308605  0.98266543 0.9079074
## 10-4  0.61120036 -0.02320015  1.24560087 0.0701146
## 6-5   0.23874991 -0.33315622  0.81065605 0.9484873
## 7-5   0.21829684 -0.29994571  0.73653938 0.9455579
## 8-5   0.48617562  0.01329049  0.95906076 0.0380845
## 9-5   0.35993838 -0.12689556  0.84677232 0.3619013
## 10-5  0.66134905  0.22925007  1.09344802 0.0000598
## 7-6  -0.02045307 -0.58162256  0.54071641 1.0000000
## 8-6   0.24742571 -0.27214739  0.76699882 0.8888360
## 9-6   0.12118847 -0.41111139  0.65348833 0.9993834
## 10-6  0.42259913 -0.06014861  0.90534687 0.1463626
## 8-7   0.26787879 -0.19196349  0.72772107 0.7055924
## 9-7   0.14164154 -0.33253338  0.61581646 0.9948634
## 10-7  0.44305221  0.02526747  0.86083695 0.0274745
## 9-8  -0.12623725 -0.55036797  0.29789347 0.9949999
## 10-8  0.17517342 -0.18481145  0.53515829 0.8750462
## 10-9  0.30141067 -0.07671106  0.67953239 0.2553958
```

Business Insignts -

As a brand manager it is really important to be able to recognize which factors influence the brand attractivness, due to our analysis we can conclude that brand having significant effect on popularity among customers of the brand. This can be expanded and/ or achieved through word of mouth marketing. The uniqueness of the brand, something that is very much its own is also importatnt. This information gives us insight on how important it is to have a differentiation strategy from the very begining and how it is crucial to pick the right colors and logos. Something that the customer can be drawn to. Besdies those two, relevance is important to mention. As a marketing team knowing that brand fits the lifestyle of customers matters confirms strtegy “How would my loyal customer base feel the product was gone?”, “How easy would it be to substitue me?”, “Could my customer base just move on right away to a different brand?”.

Relationships between loyalty and the respondent profiles for McDonalds

Descriptive statistic conducted earlied showed that McDonalds has the highest brand equity of all of the brands(6.85), McDonalds also has highest popularity. We will further look into McDonalds data and compare the brand against customer demographics. For each of the demographics variables that were analyized : age, region, gender, children and income, there are examinations of :

1. CrossTables where we inspect : are the groups within data the same or statistically different?
2. Chi Sq Test of Independance : groups variables to understand the significance of the relationship between the variables
3. Visuals of the frequency tables : visual representation between the variable loyalbin(1 = loyal, 0 = not loyal) and the demographics.

Age Analysis

The variable age is binned into 5 bins based on the frequency of occurences per each age year:

Bins :

1: 20 - 31 2: 31 - 37 3: 37 - 43 4: 43 - 51 5: 51 - 64

Cross Tabulation and Chi^2 Test for Independence:

Through the cross tabulation we test whether or not the variable is independent.

Null hypothesis: wheter the customer is loyal is not affected by the age. H0 = There is no relationship between loyalty and age (independent) H1 = There is a relationship between loyalty and age (dependent)

```
##
##    Cell Contents
## |-------------------------|
## |                   Count |
## |         Expected Values |
## |             Row Percent |
## |          Column Percent |
## |           Total Percent |
## |-------------------------|
##
## Total Observations in Table:  297
##
##                  | data_mcd$loyalbin
## data_mcd$age_bin |         0  |         1  | Row Total |
## -----------------|------------|------------|-----------|
##                1 |        18  |        42  |        60  |
##                  |    28.081  |    31.919  |           |
##                  |    30.000% |    70.000% |   20.202% |
##                  |    12.950% |    26.582% |           |
```

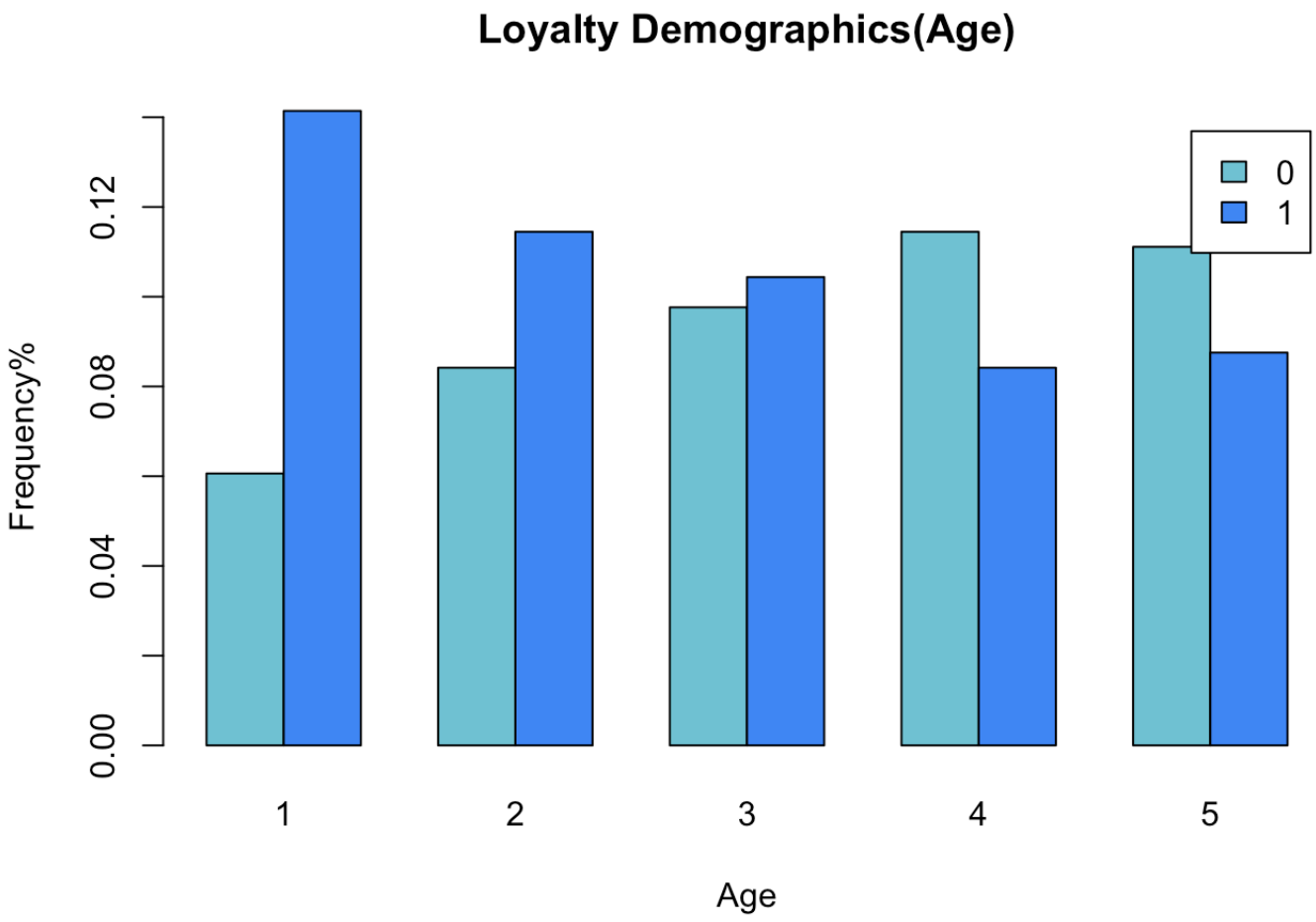
```
##                  |    6.061% |   14.141% |           |
## -----------------|-----------|-----------|-----------|
##                2 |        25 |        34 |        59 |
##                  |   27.613  |   31.387  |           |
##                  |   42.373% |   57.627% |   19.865% |
##                  |   17.986% |   21.519% |           |
##                  |    8.418% |   11.448% |           |
## -----------------|-----------|-----------|-----------|
##                3 |        29 |        31 |        60 |
##                  |   28.081  |   31.919  |           |
##                  |   48.333% |   51.667% |   20.202% |
##                  |   20.863% |   19.620% |           |
##                  |    9.764% |   10.438% |           |
## -----------------|-----------|-----------|-----------|
##                4 |        34 |        25 |        59 |
##                  |   27.613  |   31.387  |           |
##                  |   57.627% |   42.373% |   19.865% |
##                  |   24.460% |   15.823% |           |
##                  |   11.448% |    8.418% |           |
## -----------------|-----------|-----------|-----------|
##                5 |        33 |        26 |        59 |
##                  |   27.613  |   31.387  |           |
##                  |   55.932% |   44.068% |   19.865% |
##                  |   23.741% |   16.456% |           |
##                  |   11.111% |    8.754% |           |
## -----------------|-----------|-----------|-----------|
##     Column Total |       139 |       158 |       297 |
##                  |   46.801% |   53.199% |           |
## -----------------|-----------|-----------|-----------|
##
##
## Statistics for All Table Factors
##
##
## Pearson's Chi-squared test
## ------------------------------------------------------------
## Chi^2 =  12.07687     d.f. =  4     p =  0.01678867
##
##
##
##        Minimum expected frequency: 27.61279
```

The observed values are different from the expected values. Group 1(ages 20-31) has the highest observed value for brand loyalty = 1, 10 more than expected. For group 4 and 5 the observed value was lower than expected. Age and Loyalty are related for the brand McDonalds. We can infer that there is a relationship between age and loyalty since all variables have much differet observed and expected counts.

Pearson's Chi-squared test: Chi^2 = 12.07687 d.f. = 4 p = 0.01678867

p < alpha level of significance

We reject the null hypothesis at 5 % level of significance, and conclude that there is assosiation between the two.



Region Analysis

Region:

1=Maritimes; 2=Quebec; 3=Ontario; 4=West

Cross Tabulation and Chi^2 Test for Independence:

Through the cross tabulation we test whether or not the variable is independent.

Null hypothesis: wheter the customer is loyal is not affected by the region. H0 = There is no relationship between region and loyal (independent) H1 = There is a relationship between region and loyal (dependent)

```
##
##    Cell Contents
## |-------------------------|
```

```
## |                   Count |
## |         Expected Values |
## |             Row Percent |
## |          Column Percent |
## |           Total Percent |
## |-------------------------|
##
## Total Observations in Table:  297
##
##                  | data_mcd$loyalbin
## data_mcd$region |        0  |        1  | Row Total |
## ----------------|-----------|-----------|-----------|
##               1 |       13  |       16  |       29  |
##                 |   13.572  |   15.428  |           |
##                 |   44.828% |   55.172% |    9.764% |
##                 |    9.353% |   10.127% |           |
##                 |    4.377% |    5.387% |           |
## ----------------|-----------|-----------|-----------|
##               2 |       23  |       41  |       64  |
##                 |   29.953  |   34.047  |           |
##                 |   35.938% |   64.062% |   21.549% |
##                 |   16.547% |   25.949% |           |
##                 |    7.744% |   13.805% |           |
## ----------------|-----------|-----------|-----------|
##               3 |       57  |       51  |      108  |
##                 |   50.545  |   57.455  |           |
##                 |   52.778% |   47.222% |   36.364% |
##                 |   41.007% |   32.278% |           |
##                 |   19.192% |   17.172% |           |
## ----------------|-----------|-----------|-----------|
##               4 |       46  |       50  |       96  |
##                 |   44.929  |   51.071  |           |
##                 |   47.917% |   52.083% |   32.323% |
##                 |   33.094% |   31.646% |           |
##                 |   15.488% |   16.835% |           |
## ----------------|-----------|-----------|-----------|
##    Column Total |      139  |      158  |      297  |
##                 |   46.801% |   53.199% |           |
## ----------------|-----------|-----------|-----------|
##
##
## Statistics for All Table Factors
##
##
## Pearson's Chi-squared test
## ------------------------------------------------------------
## Chi^2 =  4.676495     d.f. =  3     p =  0.1970773
```

```
##
##
##
##        Minimum expected frequency: 13.57239
```

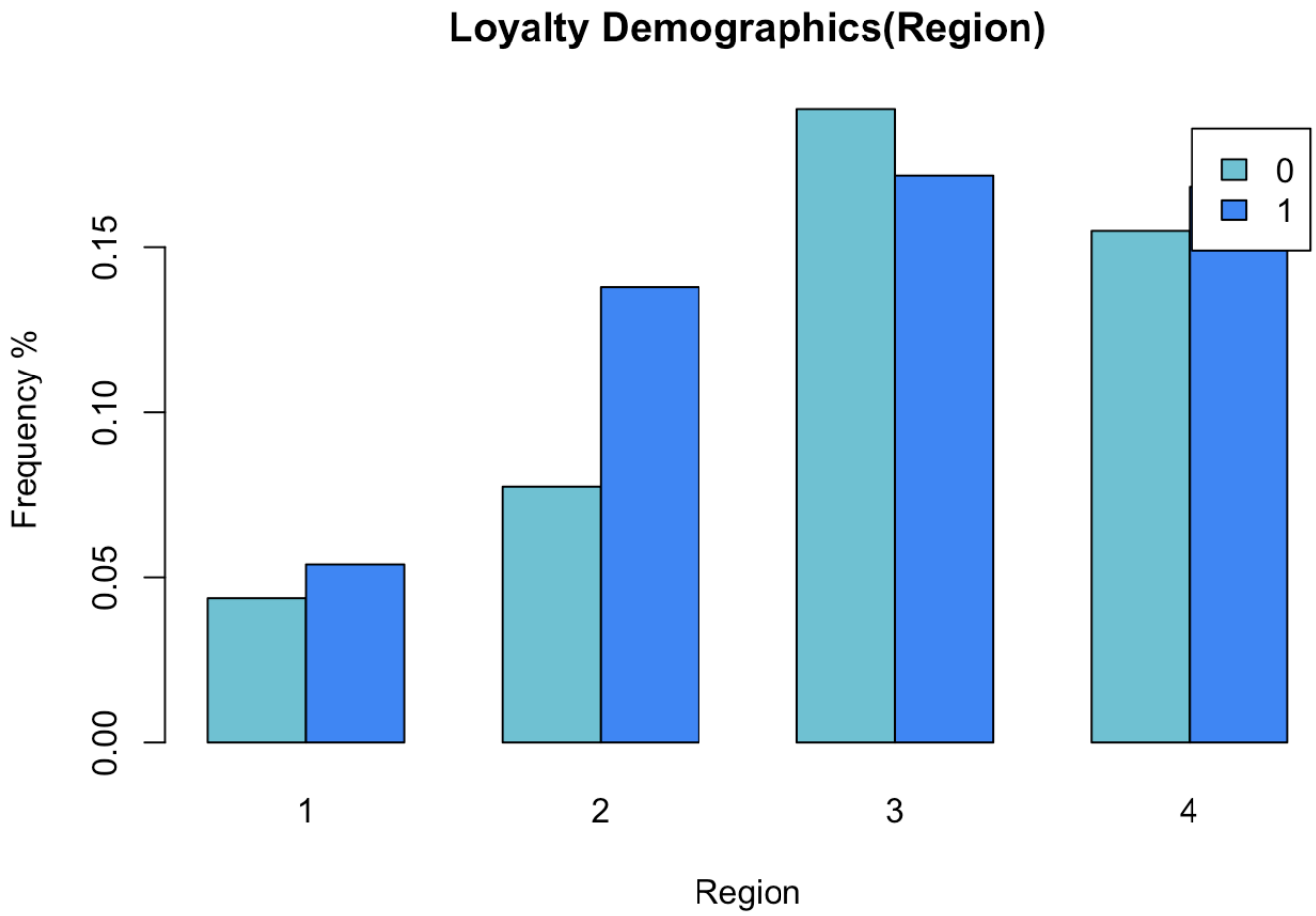
The observed values are different from the expected values but the differences are not big. Region and loyalty might not be associated with each other for the brand McDonalds.

Pearson's Chi-squared test

Chi^2 = 4.676495 d.f. = 3 p = 0.1970773

p > alpha level of significance

We fail to reject the null hypothesis at 5 % level of significance, and conclude that there is no assosiation between the two variables.



Gender Analysis Gender:

1 = Female and 2 = Male

Cross Tabulation and Chi^2 Test for Independence:

Through the cross tabulation we test whether or not the variable is independent.

```
H0 = There is no relationship between loyalty and gender   (independent)
H1 = There is a relationship between loyalty and gender    (dependent)
```

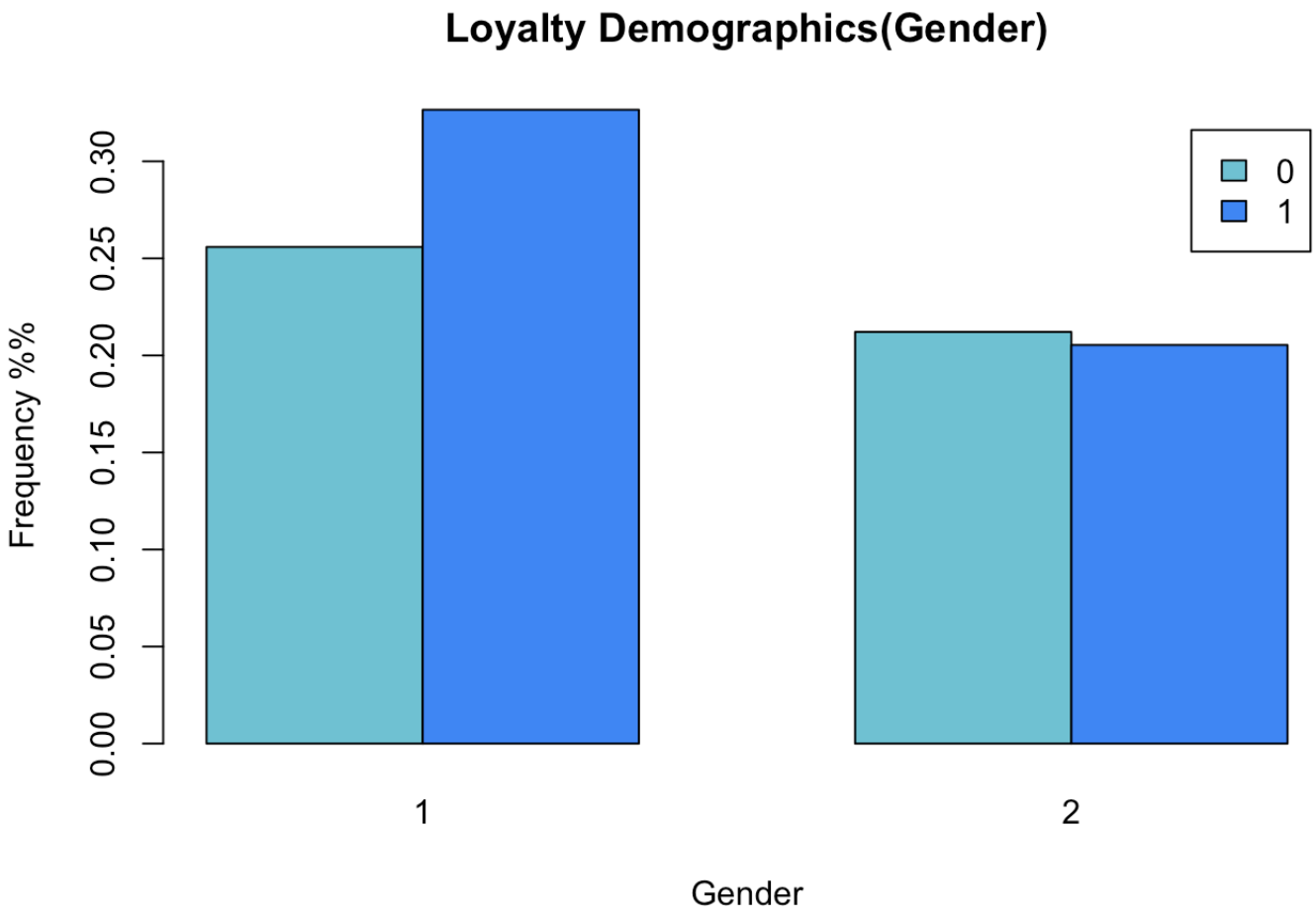
```
##
##    Cell Contents
## |-------------------------|
## |                   Count |
## |         Expected Values |
## |             Row Percent |
## |          Column Percent |
## |           Total Percent |
## |-------------------------|
##
## Total Observations in Table:  297
##
##                  | data_mcd$loyalbin
## data_mcd$gender |         0  |         1  | Row Total |
## ----------------|------------|------------|-----------|
##               1 |        76  |        97  |       173  |
##                 |   80.966  |   92.034  |            |
##                 |   43.931% |   56.069% |   58.249% |
##                 |   54.676% |   61.392% |            |
##                 |   25.589% |   32.660% |            |
## ----------------|------------|------------|-----------|
##               2 |        63  |        61  |       124  |
##                 |   58.034  |   65.966  |            |
##                 |   50.806% |   49.194% |   41.751% |
##                 |   45.324% |   38.608% |            |
##                 |   21.212% |   20.539% |            |
## ----------------|------------|------------|-----------|
##    Column Total |       139  |       158  |       297  |
##                 |   46.801% |   53.199% |            |
## ----------------|------------|------------|-----------|
##
##
## Statistics for All Table Factors
##
##
## Pearson's Chi-squared test
## ------------------------------------------------------------
## Chi^2 =  1.371516     d.f. =  1     p =  0.2415516
##
## Pearson's Chi-squared test with Yates' continuity correction
## ------------------------------------------------------------
## Chi^2 =  1.109255     d.f. =  1     p =  0.292244
##
##
##        Minimum expected frequency: 58.03367
```

Again, the observed values are different from the expected values but the differences are not big. Gender and loyalty might not be associated with each other for the brand McDonalds.

Pearson's Chi-squared test Chi^2 = 1.371516 d.f. = 1 p = 0.2415516

p > alpha level of significance

We fail to reject the null hypothesis at 5 % level of significance, and conclude that there is no assosiation between the two variables.



Analysis of Children

Children: Binary variable, whether or not there are children present at home. H0 = There is no relationship between loyalty and children (independent) H1 = There is a relationship between loyalty and children (dependent)

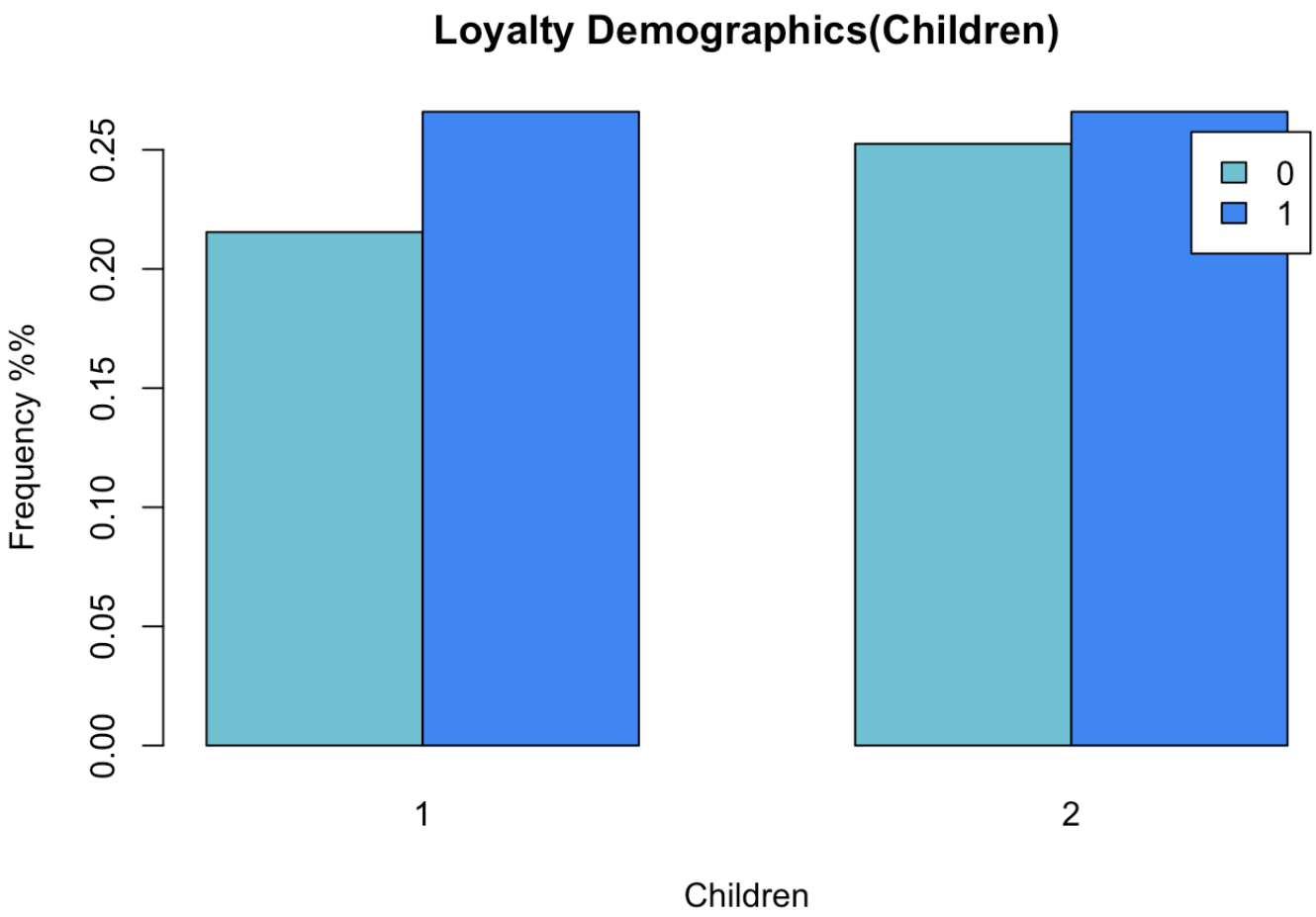
```
##
##    Cell Contents
## |-------------------------|
## |                   Count |
## |         Expected Values |
## |             Row Percent |
## |          Column Percent |
## |           Total Percent |
## |-------------------------|
##
## Total Observations in Table:  297
##
##                   | data_mcd$loyalbin
## data_mcd$children |        0  |        1  | Row Total |
## ------------------|-----------|-----------|-----------|
##                 1 |       64  |       79  |      143  |
##                   |   66.926  |   76.074  |           |
##                   |   44.755% |   55.245% |   48.148% |
##                   |   46.043% |   50.000% |           |
##                   |   21.549% |   26.599% |           |
## ------------------|-----------|-----------|-----------|
##                 2 |       75  |       79  |      154  |
##                   |   72.074  |   81.926  |           |
##                   |   48.701% |   51.299% |   51.852% |
##                   |   53.957% |   50.000% |           |
##                   |   25.253% |   26.599% |           |
## ------------------|-----------|-----------|-----------|
##      Column Total |      139  |      158  |      297  |
##                   |   46.801% |   53.199% |           |
## ------------------|-----------|-----------|-----------|
##
##
## Statistics for All Table Factors
##
##
## Pearson's Chi-squared test
## ------------------------------------------------------------
## Chi^2 =  0.4637323     d.f. =  1     p =  0.4958848
##
## Pearson's Chi-squared test with Yates' continuity correction
## ------------------------------------------------------------
## Chi^2 =  0.3187835     d.f. =  1     p =  0.5723397
##
##
##        Minimum expected frequency: 66.92593
```

Again, the observed values are different from the expected values but the differences are not big. Children and loyalty might not be associated with each other for the brand McDonalds.

Pearson's Chi-squared test Chi^2 = 0.4637323 d.f. = 1 p = 0.4958848

p > alpha level of significance

We fail to reject the null hypothesis at 5 % level of significance, and conclude that there is no assosiation between the two variables.



Analysis of Income

Income : 1= < $30k; 2= $30-$49.9k; 3= $50-$74.9k; 4=$75k+

```
    H0 = There is no relationship between loyalty and income   (independent)
    H1 = There is a relationship between loyalty and income    (dependent)
```

```
##
##    Cell Contents
## |-------------------------|
## |                   Count |
```

```
## |         Expected Values |
## |             Row Percent |
## |          Column Percent |
## |           Total Percent |
## |-------------------------|
##
## Total Observations in Table:  297
##
##                  | data_mcd$loyalbin
## data_mcd$income |         0  |         1  | Row Total |
## ----------------|------------|------------|-----------|
##               1 |        26  |        40  |        66  |
##                 |    30.889  |    35.111  |           |
##                 |    39.394% |    60.606% |   22.222% |
##                 |    18.705% |    25.316% |           |
##                 |     8.754% |    13.468% |           |
## ----------------|------------|------------|-----------|
##               2 |        25  |        47  |        72  |
##                 |    33.697  |    38.303  |           |
##                 |    34.722% |    65.278% |   24.242% |
##                 |    17.986% |    29.747% |           |
##                 |     8.418% |    15.825% |           |
## ----------------|------------|------------|-----------|
##               3 |        48  |        47  |        95  |
##                 |    44.461  |    50.539  |           |
##                 |    50.526% |    49.474% |   31.987% |
##                 |    34.532% |    29.747% |           |
##                 |    16.162% |    15.825% |           |
## ----------------|------------|------------|-----------|
##               4 |        40  |        24  |        64  |
##                 |    29.953  |    34.047  |           |
##                 |    62.500% |    37.500% |   21.549% |
##                 |    28.777% |    15.190% |           |
##                 |    13.468% |     8.081% |           |
## ----------------|------------|------------|-----------|
##    Column Total |       139  |       158  |       297  |
##                 |    46.801% |    53.199% |           |
## ----------------|------------|------------|-----------|
##
##
## Statistics for All Table Factors
##
##
## Pearson's Chi-squared test
## ------------------------------------------------------------
## Chi^2 =  12.53827     d.f. =  3     p =  0.005749369
##
```

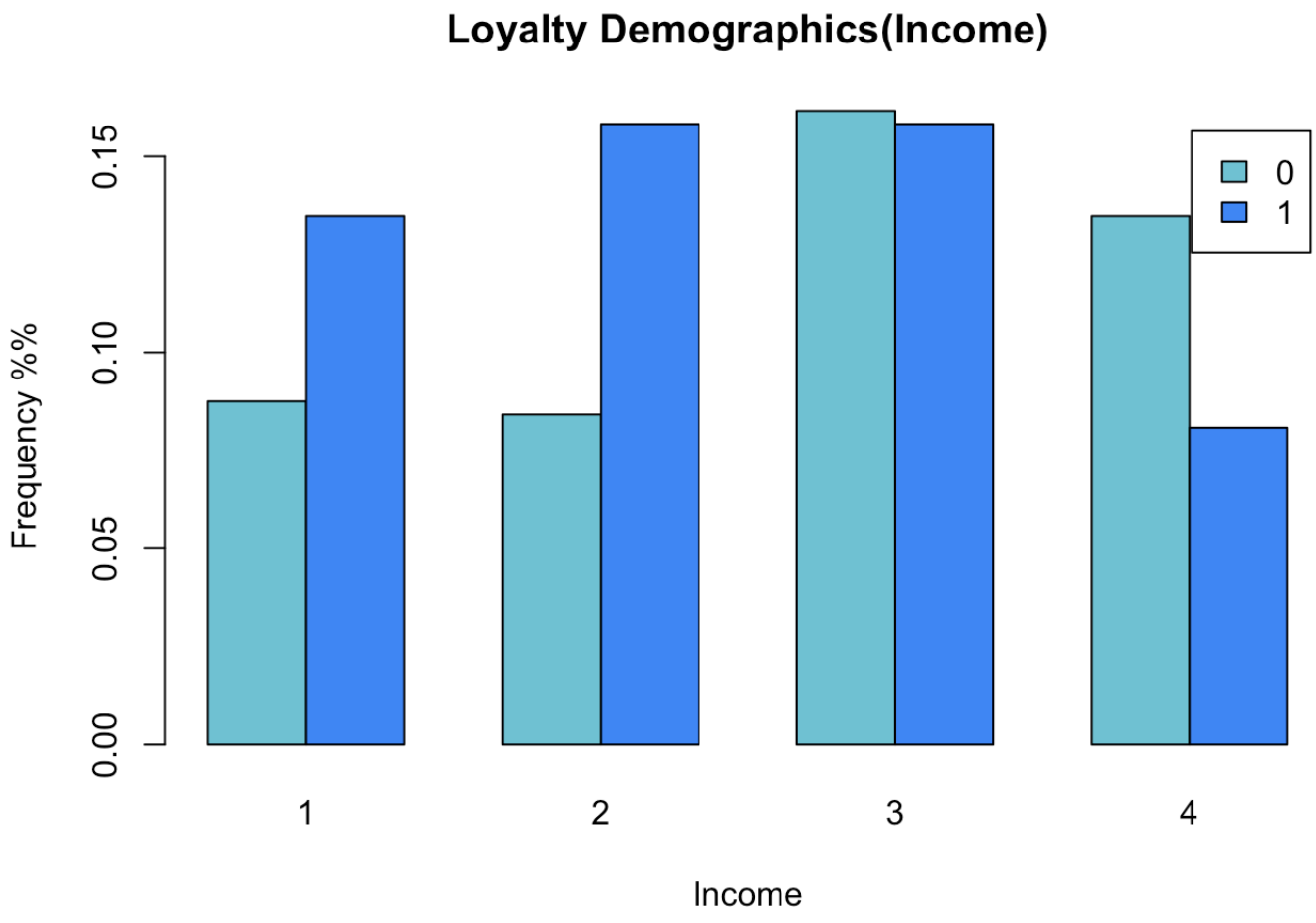
```
##
##
##        Minimum expected frequency: 29.95286
```

The observed values are different from the expected values. Group 2 has the highest observed value, for brand loyalty = 1, 9 more than expected. For group 3 and 4 the observed value was lower than expected. Income and Loyalty are related for the brand McDonalds. We can infer that there is a relationship between income and loyalty since all variables have much differet observed and expected counts.

Pearson's Chi-squared test Chi^2 = 12.53827 d.f. = 3 p = 0.005749369

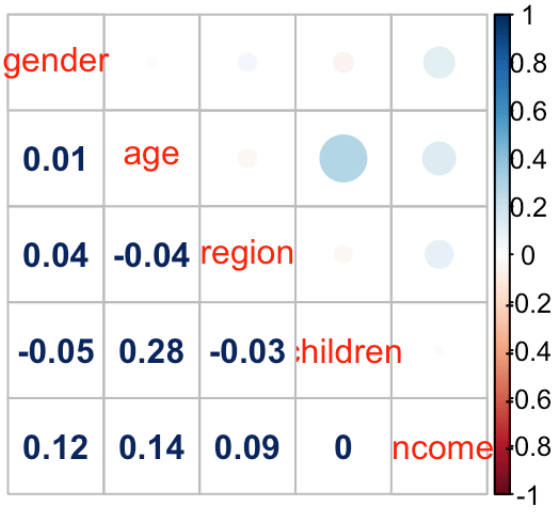
p < alpha level of significance

We reject the null hypothesis at 5 % level of significance, and conclude that there is assosiation between the two.



Regression Analysis

To further examine the variable loyal and age, we used a regression analysis.



Correlation test shows that the variables are not correlated with each other for the brand McDonals.

Regression analysis is perfermed on the data. With brand variable we conclud the test that examine the relationship between deppendent variable loyal and independent demographic variables.

The statistical results show, once again income and age being significant variables.

Null hypothesis: the slope is equal to 0

The regression equation takeaways, each increase in the independent variable will cause the dependent variable to decrease(besides gender). Demographics and brand loyalty have negative relationship.

The slope is significant and not 0. We can reject the null.

Pairwise Test - Age Groups

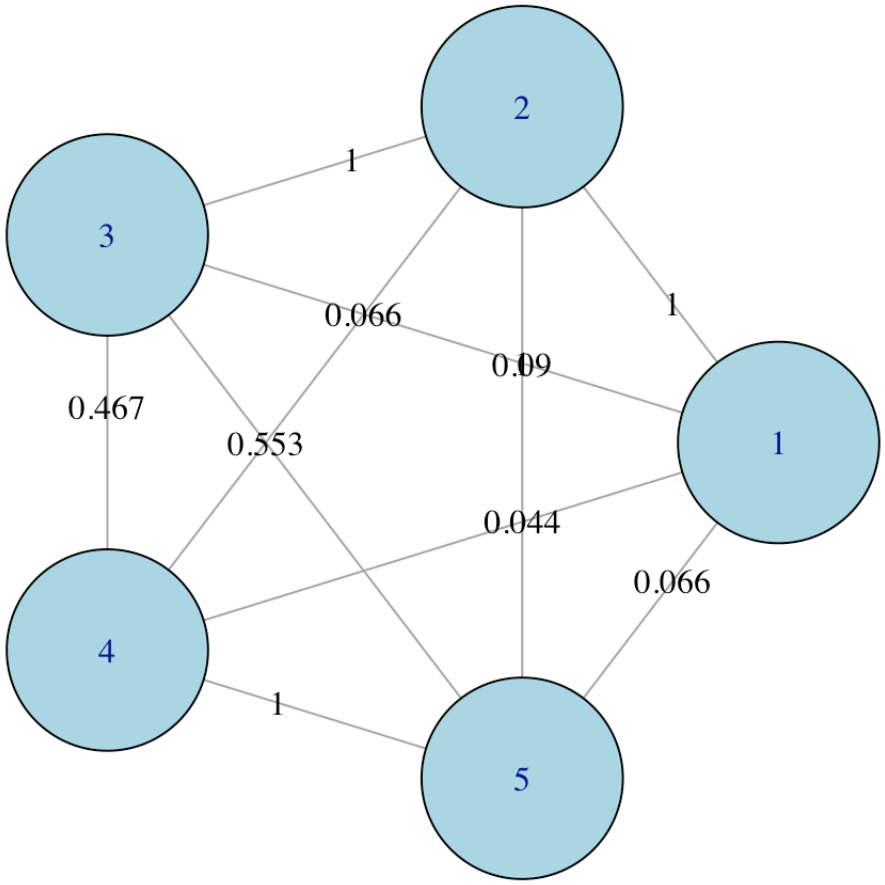
Multiple comparisons and pairwise test is performed to take a look at the significant variable age_bin. We concluded that there is a relationship with the variable and we also saw the significance in the regression analysis. The particular subgrups of variable age is analyzed. We first run a t test to compare each pair.

H0 = All means are equal within the group AGE. H1 = At least one pair has different means

Through the Symmetric Matrix we can observe that for age group 4 and age group 1 the corresponding p - value is .043 which rejects are null that all means are equal within the group AGE. Below there is a visual representation of the graph:

1 indicates uninteresting relationship.

```
##             1           2         3           4           5
## 1 0.00000000 1.00000000 1.0000000 0.04372790 0.06556481
## 2 1.00000000 0.00000000 1.0000000 0.06556481 0.08991290
## 3 1.00000000 1.00000000 0.0000000 0.46666763 0.55329447
## 4 0.04372790 0.06556481 0.4666676 0.00000000 1.00000000
## 5 0.06556481 0.08991290 0.5532945 1.00000000 0.00000000
```

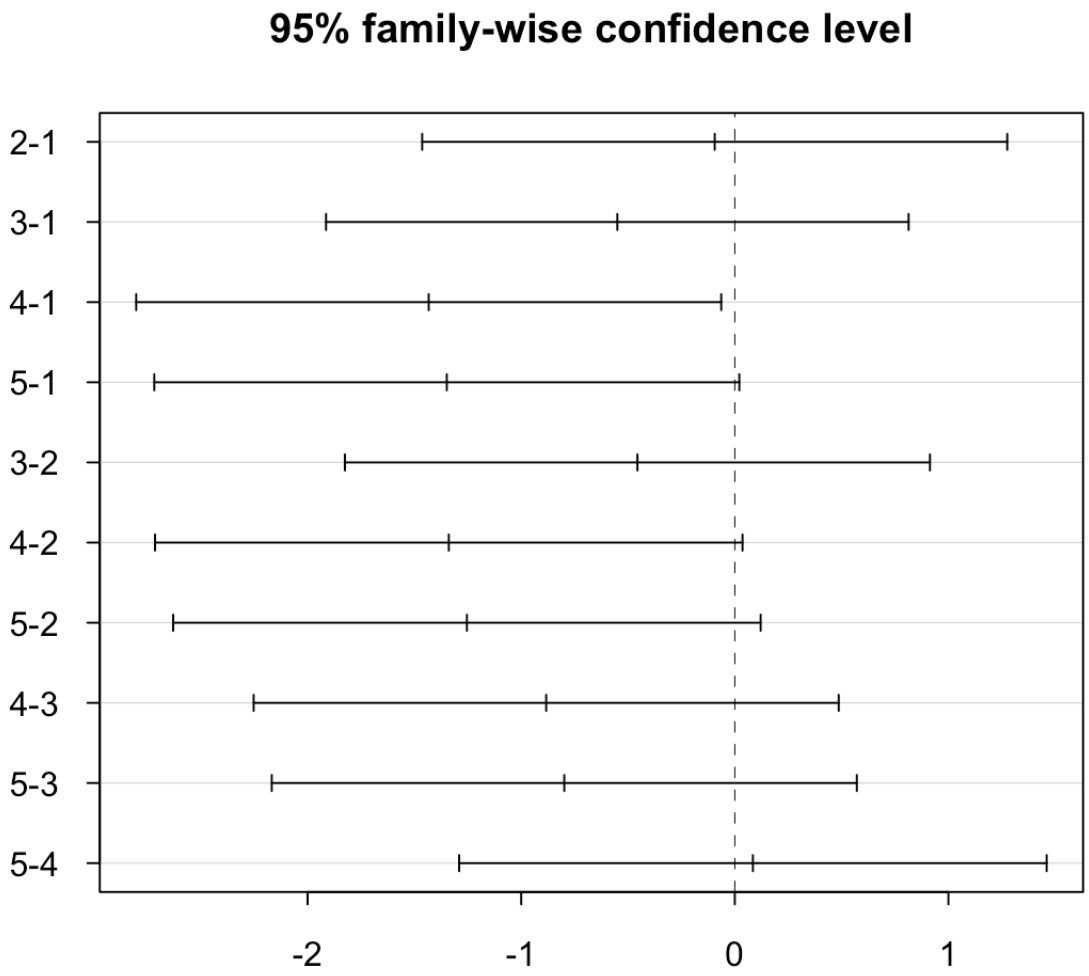


Multiple Comparisons - Age Groups (with Turkey's HSD)

We saw earlier in the reprt the TukeyHSD results give confidence intervals of the sample means. The p-values for the pairwise tests but will be slightly different than pairwise test. The anova analysis below also indicates that pair 4-1 is significant.

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = loyal ~ as.factor(age_bin), data = data_mcd)
##
## $`as.factor(age_bin)`
##             diff       lwr         upr     p adj
## 2-1 -0.09406780 -1.463529  1.27539321 0.9997165
## 3-1 -0.55000000 -1.913695  0.81369483 0.8027616
## 4-1 -1.43305085 -2.802512 -0.06358984 0.0351719
## 5-1 -1.34830508 -2.717766  0.02115592 0.0559850
## 3-2 -0.45593220 -1.825393  0.91352881 0.8914460
## 4-2 -1.33898305 -2.714186  0.03621996 0.0605562
## 5-2 -1.25423729 -2.629440  0.12096572 0.0926325
## 4-3 -0.88305085 -2.252512  0.48641016 0.3931643
## 5-3 -0.79830508 -2.167766  0.57115592 0.4985125
## 5-4  0.08474576 -1.290457  1.45994877 0.9998158
```

TurkeyHSD with adjusted p-value from the previous test for means comparison between subgroups of variable Age(binned) and ‘loyal’ shows that customers between age 43 -51 and 20-31 have significant differences in means.



Business Insights

We concluded that when looking at brand McDonalds age and income are variables that have association between loyalty and are also a significant factor when predicting a customer loyalty for McDonalds brand. We also concluded that within the Age Groups loyalty score is different depending on which age group we target. We could infer that college kids and early adults eat at McDonalds as well well as older customers who have lower income.

McDonalds executive team knows that everyone is familiar with their brand. McDonalds coudl improve Brand Loyalty through more sophisticated target selction. It is not surprising that income was one of the important factors in the customer demographics as McDonalds is known fro their cheap, and unfortunately unheathy meals. For McDonalds, this could be the case that although they have a very high brand equity score they are losing competitivness in the market due the healthy diet being more popular and also healthy food becomes more affordable. McDonalds is also not able to put the premium on their food, due to their initial marketing strategy.

McDonalds Marketing team should in general recognize the challenge it faces in the world that is full of dynamic changes and customer shifts. Peoples' awerness is increasing regarding food and their options.

Discussion Questions:

Ariel created binary variables for familiarity, uniqueness, relevance, loyalty and popularity by splitting responses into “high” and “low.” Why would they choose to do (or not do) this? In other words, what information is gained and what information is lost?

We are able to better group our variable and be able to perform better statistical analysis if our data is binned. How we decided to bin the data often depends on the scope, complexity and contest of our analysis. Good decision was made here to split the data into two groups, only highest scores should count towards brand equity analysis as we want to capture the insight and information the most accurately possible and picking highest values is our one option. We do lose some information when binning the data but as long as we keep the original variables in the data set we can always create multiple models and analysis and see what works best for us.

Do you agree with Ariel’s measure of brand equity?

Ariel focused on the marketing side of brand equity and due to the context it was a good choice of measurement. Like we indicated at the beginning og the report brand equity is hard to measure and is hard to capture. The creative side of brain needs to be working when analyzing brand equity. There could be more variables included that go more into details : Brand Culture(how does brand treat their employees), Brand Reputation (does it matter if the company has good reputation? does it mater if it egages in charity events?), Colors(do certain colors remind you of a certain brand) Smell, Music , Commercials. There is a lot that could be done to measure brand equity but the challenge remins to be able to normalize it at the end and compare to other brands to have benchmart and indicator. It could be challenging if we are peronalizing the surveys too much.

References:

“Brand Equity.” Higgins, Chris A.; Whelan, Jodie. Case No. 9B10E023. Published 2/1/2011, Richard Ivey School of Business, 2010. (3 pages).

“Multiple comparisons with PairViz”" Catherine B. Hurley and R.W. Oldford Published 2018-08-09,