

# MLD Assignment\_v1

*Maekaila Cena*

*November 12, 2018*

```
rm(list=ls(all=TRUE))
```

## 1. Import Data

```
library(psych)
```

```
## Warning: package 'psych' was built under R version 3.4.4
```

```
MLD <- read.csv("MLD Data File.csv", header=TRUE) # import data
```

```
MLD1 <- read.csv("MLD Data File.csv")
```

## 2. Data Preparation and Cleaning

```
colnames(MLD)
```

```
## [1] "Married"
## [2] "Credit.Guidelines"
## [3] "Other.Obligations..As.Percent..Of..Income"
## [4] "Non.Hispanic.White"
## [5] "Non.Hispanic.Black"
## [6] "Hispanic"
## [7] "Male"
## [8] "Approved"
## [9] "Present.Loan.to.Value"
```

Structure of Data:

```
str(MLD)
```

```
## 'data.frame': 1989 obs. of 9 variables:
## $ Married : int 1 0 1 1 0 0 0 1 1 1 ...
## $ Credit.Guidelines : int 1 1 1 1 1 1 1 1 1 1 ...
## $ Other.Obligations..As.Percent..Of..Income: num 34.5 34.1 26 37 32.1 33 36 37 30.7 49 ...
## $ Non.Hispanic.White : int 1 1 1 1 1 1 1 1 1 1 ...
## $ Non.Hispanic.Black : int 0 0 0 0 0 0 0 0 0 0 ...
## $ Hispanic : int 0 0 0 0 0 0 0 0 0 0 ...
## $ Male : int 1 1 1 1 1 1 1 1 1 1 ...
## $ Approved : int 0 1 1 1 1 1 1 1 1 1 ...
## $ Present.Loan.to.Value : num 0.754 0.8 0.895 0.6 0.896 ...
```

Descriptive statistics:

```
summary(MLD)
```

```
## Married Credit.Guidelines
## Min. :0.0000 Min. :0.0000
## 1st Qu.:0.0000 1st Qu.:1.0000
## Median :1.0000 Median :1.0000
## Mean :0.6592 Mean :0.9132
## 3rd Qu.:1.0000 3rd Qu.:1.0000
```

```
## Max. :1.0000 Max. :1.0000
## NA's :20 NA's :20
## Other.Obligations..As.Percent..Of..Income Non.Hispanic.White
## Min. : 0.00 Min. :0.000
## 1st Qu.:28.00 1st Qu.:1.000
## Median :33.00 Median :1.000
## Mean :32.39 Mean :0.901
## 3rd Qu.:37.00 3rd Qu.:1.000
## Max. :95.00 Max. :1.000
## NA's :20
## Non.Hispanic.Black Hispanic Male Approved
## Min. :0.00000 Min. :0.00000 Min. :0.0000 Min. :0.0000
## 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:1.0000 1st Qu.:1.0000
## Median :0.00000 Median :0.00000 Median :1.0000 Median :1.0000
## Mean :0.09903 Mean :0.05485 Mean :0.8131 Mean :0.8761
## 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:1.0000 3rd Qu.:1.0000
## Max. :1.00000 Max. :1.00000 Max. :1.0000 Max. :1.0000
## NA's :20 NA's :20 NA's :20 NA's :20
## Present.Loan.to.Value
## Min. :0.02105
## 1st Qu.:0.70000
## Median :0.80000
## Mean :0.77064
## 3rd Qu.:0.89894
## Max. :2.57143
##
```

```
MLD <- na.omit(MLD)
summary(MLD)
```

```
## Married Credit.Guidelines
## Min. :0.0000 Min. :0.0000
## 1st Qu.:0.0000 1st Qu.:1.0000
## Median :1.0000 Median :1.0000
## Mean :0.6592 Mean :0.9132
## 3rd Qu.:1.0000 3rd Qu.:1.0000
## Max. :1.0000 Max. :1.0000
## Other.Obligations..As.Percent..Of..Income Non.Hispanic.White
## Min. : 0.00 Min. :0.000
## 1st Qu.:28.00 1st Qu.:1.000
## Median :33.00 Median :1.000
## Mean :32.38 Mean :0.901
## 3rd Qu.:37.00 3rd Qu.:1.000
## Max. :95.00 Max. :1.000
## Non.Hispanic.Black Hispanic Male Approved
## Min. :0.00000 Min. :0.00000 Min. :0.0000 Min. :0.0000
## 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:1.0000 1st Qu.:1.0000
## Median :0.00000 Median :0.00000 Median :1.0000 Median :1.0000
## Mean :0.09903 Mean :0.05485 Mean :0.8131 Mean :0.8761
## 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:1.0000 3rd Qu.:1.0000
## Max. :1.00000 Max. :1.00000 Max. :1.0000 Max. :1.0000
## Present.Loan.to.Value
## Min. :0.02105
## 1st Qu.:0.70000
## Median :0.80000
```

```
## Mean :0.77110
## 3rd Qu.:0.89899
## Max. :2.57143
```

```
names(MLD) <- make.names(names(MLD))
colnames(MLD)
```

```
## [1] "Married"
## [2] "Credit.Guidelines"
## [3] "Other.Obligations..As.Percent..Of..Income"
## [4] "Non.Hispanic.White"
## [5] "Non.Hispanic.Black"
## [6] "Hispanic"
## [7] "Male"
## [8] "Approved"
## [9] "Present.Loan.to.Value"
```

```
SubsettedMLD <- (MLD[,c("Married", "Credit.Guidelines", "Other.Obligations..As.Percent..Of..Income", "Non.Hispanic.White", "Non.Hispanic.Black", "Hispanic", "Male", "Approved", "Present.Loan.to.Value")])
MLDsubsample <- subset(SubsettedMLD, Present.Loan.to.Value >= 1)
MLDsubsample
```

	Married	Credit.Guidelines	Other.Obligations..As.Percent..Of..Income
## 9	1	1	30.7
## 23	1	1	36.0
## 33	1	1	73.0
## 37	0	1	35.0
## 53	0	1	38.0
## 85	0	1	75.0
## 126	1	0	58.0
## 132	0	1	27.6
## 160	1	1	43.0
## 162	0	1	25.0
## 170	1	1	37.0
## 233	1	1	18.0
## 258	0	1	45.0
## 262	0	1	56.0
## 276	0	1	45.0
## 294	1	1	37.0
## 300	0	1	47.0
## 324	1	1	16.0
## 341	1	1	30.0
## 448	1	1	26.0
## 467	0	1	40.0
## 472	0	1	26.0
## 481	0	1	26.0
## 490	1	0	25.0
## 517	1	1	15.0
## 544	0	1	52.0
## 560	0	1	33.0
## 565	0	1	10.0
## 567	1	1	35.0
## 589	1	1	38.4
## 655	0	0	40.0
## 694	0	1	32.0
## 725	1	1	34.0

## 734	1	1	30.0
## 750	1	1	35.0
## 793	1	1	37.0
## 798	1	1	36.0
## 812	1	1	32.0
## 821	1	1	27.0
## 841	1	0	41.2
## 904	0	1	24.0
## 914	1	1	40.0
## 940	1	1	38.0
## 947	1	1	18.0
## 950	1	1	20.0
## 957	1	1	35.0
## 963	0	1	41.3
## 1051	1	1	27.6
## 1067	0	1	34.0
## 1112	1	1	35.5
## 1158	1	1	14.0
## 1178	1	1	29.0
## 1193	1	1	12.0
## 1206	1	1	41.0
## 1226	1	1	14.0
## 1229	1	1	26.0
## 1236	0	1	30.0
## 1279	0	1	30.0
## 1322	1	1	32.0
## 1340	1	1	49.0
## 1364	1	1	25.0
## 1378	1	1	18.0
## 1385	1	1	36.5
## 1388	1	1	33.0
## 1392	1	1	13.0
## 1394	0	1	37.0
## 1413	1	1	21.0
## 1430	1	0	35.3
## 1441	1	1	33.5
## 1457	1	1	19.0
## 1499	1	1	31.0
## 1519	0	1	27.6
## 1545	1	1	8.0
## 1623	1	1	29.0
## 1628	1	1	31.0
## 1632	0	1	36.8
## 1667	1	1	43.0
## 1683	1	1	25.0
## 1699	1	1	30.1
## 1755	0	1	36.0
## 1759	1	1	46.0
## 1763	0	1	35.0
## 1923	1	1	27.0
## 1966	1	1	25.0

##	Non.Hispanic.White	Non.Hispanic.Black	Hispanic	Approved
## 9	1	0	0	1
## 23	1	0	0	1

## 33	1	0	0	1
## 37	1	0	0	1
## 53	1	0	1	1
## 85	0	1	0	1
## 126	1	0	0	1
## 132	1	0	0	1
## 160	1	0	0	1
## 162	1	0	0	1
## 170	0	1	0	1
## 233	1	0	0	1
## 258	1	0	0	1
## 262	1	0	0	1
## 276	1	0	1	1
## 294	1	0	0	1
## 300	1	0	0	1
## 324	0	1	0	1
## 341	1	0	0	1
## 448	1	0	0	1
## 467	1	0	0	1
## 472	1	0	0	1
## 481	1	0	1	1
## 490	1	0	0	1
## 517	1	0	0	1
## 544	1	0	0	1
## 560	1	0	0	1
## 565	1	0	0	1
## 567	1	0	0	1
## 589	0	1	0	1
## 655	1	0	0	1
## 694	1	0	0	0
## 725	1	0	0	1
## 734	1	0	0	1
## 750	1	0	0	1
## 793	1	0	0	1
## 798	0	1	0	1
## 812	1	0	0	1
## 821	1	0	0	1
## 841	0	1	0	0
## 904	1	0	1	1
## 914	1	0	0	1
## 940	1	0	0	1
## 947	1	0	0	1
## 950	0	1	0	0
## 957	1	0	0	1
## 963	1	0	0	1
## 1051	1	0	0	1
## 1067	0	1	0	1
## 1112	1	0	0	1
## 1158	1	0	0	1
## 1178	1	0	1	1
## 1193	1	0	0	1
## 1206	1	0	0	1
## 1226	1	0	0	1
## 1229	1	0	0	1

## 1236	0	1	0	1
## 1279	1	0	0	1
## 1322	1	0	0	1
## 1340	1	0	0	1
## 1364	1	0	0	1
## 1378	1	0	0	1
## 1385	1	0	0	1
## 1388	1	0	0	1
## 1392	1	0	0	1
## 1394	1	0	0	1
## 1413	1	0	0	1
## 1430	1	0	0	1
## 1441	1	0	0	1
## 1457	1	0	1	1
## 1499	1	0	1	1
## 1519	1	0	0	1
## 1545	1	0	0	1
## 1623	1	0	0	1
## 1628	1	0	0	0
## 1632	0	1	0	0
## 1667	1	0	0	1
## 1683	1	0	0	1
## 1699	1	0	0	1
## 1755	1	0	0	1
## 1759	1	0	1	1
## 1763	1	0	0	1
## 1923	0	1	0	1
## 1966	1	0	0	1
##	Present.Loan.to.Value			
## 9	1.003009			
## 23	1.000000			
## 33	1.000000			
## 37	1.127119			
## 53	1.085714			
## 85	1.000000			
## 126	1.000000			
## 132	1.011905			
## 160	1.000000			
## 162	1.647059			
## 170	1.000000			
## 233	1.055556			
## 258	2.200000			
## 262	2.571429			
## 276	1.000000			
## 294	1.488095			
## 300	1.000000			
## 324	1.000000			
## 341	1.000000			
## 448	1.000000			
## 467	1.000000			
## 472	1.000000			
## 481	1.000000			
## 490	1.205480			
## 517	1.000000			

## 544	1.176471
## 560	1.379310
## 565	1.333333
## 567	1.000000
## 589	1.000000
## 655	1.000000
## 694	1.000000
## 725	1.000000
## 734	1.000000
## 750	1.478261
## 793	1.300000
## 798	1.473684
## 812	1.000000
## 821	1.000000
## 841	1.000676
## 904	1.833333
## 914	1.114286
## 940	1.289474
## 947	1.003521
## 950	1.000000
## 957	1.000000
## 963	1.094118
## 1051	1.000000
## 1067	1.000000
## 1112	1.626263
## 1158	1.000000
## 1178	1.000000
## 1193	1.000000
## 1206	1.250000
## 1226	1.000000
## 1229	1.000000
## 1236	1.000000
## 1279	1.000000
## 1322	1.105263
## 1340	1.000000
## 1364	1.000000
## 1378	1.000000
## 1385	1.000000
## 1388	1.000000
## 1392	1.172414
## 1394	1.140845
## 1413	1.000000
## 1430	1.000000
## 1441	1.000000
## 1457	1.080000
## 1499	1.000000
## 1519	1.000000
## 1545	1.000000
## 1623	1.090909
## 1628	1.000000
## 1632	2.555248
## 1667	1.200000
## 1683	1.187500
## 1699	1.000000

```
## 1755          1.000000
## 1759          1.072000
## 1763          1.000000
## 1923          1.000000
## 1966          1.000000
```

## Logit Models

### *#Estimate Logit Model*

```
BlackLogit = glm(Approved ~ Non.Hispanic.Black + Married + Credit.Guidelines + Other.Obligations..As.Percent..Of..Income + Present.Loan.to.Value,
summary(BlackLogit)
```

```
##
## Call:
## glm(formula = Approved ~ Non.Hispanic.Black + Married + Credit.Guidelines +
##      Other.Obligations..As.Percent..Of..Income + Present.Loan.to.Value,
##      family = "binomial", data = SubsettedMLD)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.5448  0.3047  0.3213  0.3874  2.1721
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -0.888561   0.507612  -1.750    0.080037 .
## Non.Hispanic.Black    -0.859661   0.236793  -3.630    0.000283 ***
## Married           0.451772   0.178522   2.531    0.011386 *
## Credit.Guidelines    3.869217   0.213037  18.162    < 2e-16 ***
## Other.Obligations..As.Percent..Of..Income -0.001007   0.010935  -0.092    0.926660
## Present.Loan.to.Value -0.523259   0.444702  -1.177    0.239335
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 1475.43  on 1968  degrees of freedom
## Residual deviance:  993.58  on 1963  degrees of freedom
## AIC: 1005.6
##
## Number of Fisher Scoring iterations: 5
```

### *#Estimate Logit Model*

```
HispanicLogit = glm(Approved ~ Hispanic + Married + Credit.Guidelines + Other.Obligations..As.Percent..Of..Income + Present.Loan.to.Value,
summary(HispanicLogit)
```

```
##
## Call:
## glm(formula = Approved ~ Hispanic + Married + Credit.Guidelines +
```



```
##      Other.Obligations..As.Percent..Of..Income + Present.Loan.to.Value,
##      family = "binomial", data = SubsettedMLD)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q        Max
## -2.5486   0.3076   0.3231   0.3943   2.2567
##
## Coefficients:
##                                Estimate Std. Error z value
## (Intercept)                 -1.0806771   0.5101955  -2.118
## Hispanic                   -0.8894779   0.3056345  -2.910
## Married                     0.4867328   0.1781279   2.732
## Credit.Guidelines           3.9951795   0.2118168  18.861
## Other.Obligations..As.Percent..Of..Income -0.0006243   0.0109501  -0.057
## Present.Loan.to.Value      -0.5209999   0.4501911  -1.157
##                                Pr(>|z|)
## (Intercept)                  0.03416 *
## Hispanic                     0.00361 **
## Married                      0.00629 **
## Credit.Guidelines            < 2e-16 ***
## Other.Obligations..As.Percent..Of..Income 0.95453
## Present.Loan.to.Value        0.24716
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 1475.43  on 1968  degrees of freedom
## Residual deviance:  998.23  on 1963  degrees of freedom
## AIC: 1010.2
##
## Number of Fisher Scoring iterations: 5
```

## Logit Odd Ratios

```
#BLACK
#Generate Odds Ratios
exp(coef(BlackLogit))
```

```
##                                (Intercept)
##                                0.4112471
##                                Non.Hispanic.Black
##                                0.4233055
##                                Married
##                                1.5710930
##                                Credit.Guidelines
##                                47.9048722
## Other.Obligations..As.Percent..Of..Income
##                                0.9989939
##                                Present.Loan.to.Value
##                                0.5925861
```

```
#HISPANIC
exp(coef(HispanicLogit))
```

```
## (Intercept)
## 0.3393657
## Hispanic
## 0.4108702
## Married
## 1.6269919
## Credit.Guidelines
## 54.3355931
## Other.Obligations..As.Percent..Of..Income
## 0.9993759
## Present.Loan.to.Value
## 0.5939264
```

## Probit Models

```
#Estimate Probit Model
```

```
BlackProbit = glm(Approved ~ Non.Hispanic.Black + Married + Credit.Guidelines + Other.Obligations..As.P
family = "binomial" (link = "probit"))
summary(BlackProbit)
```

```
##
## Call:
## glm(formula = Approved ~ Non.Hispanic.Black + Married + Credit.Guidelines +
## Other.Obligations..As.Percent..Of..Income + Present.Loan.to.Value,
## family = binomial(link = "probit"), data = SubsettedMLD)
##
## Deviance Residuals:
## Min 1Q Median 3Q Max
## -2.5579 0.3019 0.3200 0.3888 2.1374
##
## Coefficients:
## Estimate Std. Error z value
## (Intercept) -0.5575274 0.2609781 -2.136
## Non.Hispanic.Black -0.4419577 0.1246676 -3.545
## Married 0.2243349 0.0886266 2.531
## Credit.Guidelines 2.2384485 0.1192106 18.777
## Other.Obligations..As.Percent..Of..Income -0.0008909 0.0053865 -0.165
## Present.Loan.to.Value -0.2622344 0.2272630 -1.154
## Pr(>|z|)
## (Intercept) 0.032655 *
## Non.Hispanic.Black 0.000392 ***
## Married 0.011366 *
## Credit.Guidelines < 2e-16 ***
## Other.Obligations..As.Percent..Of..Income 0.868638
## Present.Loan.to.Value 0.248549
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 1475.4 on 1968 degrees of freedom
## Residual deviance: 993.7 on 1963 degrees of freedom
## AIC: 1005.7
##
```

```
## Number of Fisher Scoring iterations: 5
#Estimate Probit Model
HispanicProbit = glm(Approved ~ Hispanic + Married + Credit.Guidelines + Other.Obligations..As.Percent.
                      family = "binomial" (link = "probit"))
summary(HispanicProbit)

##
## Call:
## glm(formula = Approved ~ Hispanic + Married + Credit.Guidelines +
##      Other.Obligations..As.Percent..Of..Income + Present.Loan.to.Value,
##      family = binomial(link = "probit"), data = SubsettedMLD)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.5675   0.3039   0.3211   0.3968   2.2261
##
## Coefficients:
##                  Estimate Std. Error z value
## (Intercept)      -0.6662901  0.2600685  -2.562
## Hispanic         -0.4598912  0.1605739  -2.864
## Married           0.2455749  0.0884654   2.776
## Credit.Guidelines 2.3053550  0.1174757  19.624
## Other.Obligations..As.Percent..Of..Income -0.0004224  0.0053899  -0.078
## Present.Loan.to.Value -0.2620026  0.2285612  -1.146
##
## Pr(>|z|)
## (Intercept)      0.01041 *
## Hispanic         0.00418 **
## Married          0.00550 **
## Credit.Guidelines < 2e-16 ***
## Other.Obligations..As.Percent..Of..Income 0.93754
## Present.Loan.to.Value 0.25167
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 1475.43  on 1968  degrees of freedom
## Residual deviance:  997.63  on 1963  degrees of freedom
## AIC: 1009.6
##
## Number of Fisher Scoring iterations: 5
```