Lending Mortage Analysis

Exploratory Data Analysis

Marjorie Blanco, Joe Thompson, Haodi Tu

The data set contains 1989 records. The overall descriptive statistics:

```
##      Married      Meet credit history guidelines
## No     : 678   0  : 171
## Unknown:   3   1  :1816
## Yes    :1308   666:   2
##
##
##
## Other obligations as a percent of total income non-Hispanic Black
## Min.   : 0.00                                    No :1792
## 1st Qu.:28.00                                    Yes: 197
## Median :33.00
## Mean   :32.39
## 3rd Qu.:37.00
## Max.   :95.00
## Hispanic          Male        Mortgage loan approved
## No :1878    No     : 369   No : 244
## Yes: 111    Unknown:  15   Yes:1745
##             Yes    :1605
##
##
##
## Loan amount/purchase price                   Race
## Min.   :  2.105        Hispanic          : 111
## 1st Qu.: 70.000        non-Hispanic Black: 197
## Median : 80.000        non-Hispanic White:1681
## Mean   : 77.064
## 3rd Qu.: 89.894
## Max.   :257.143
```
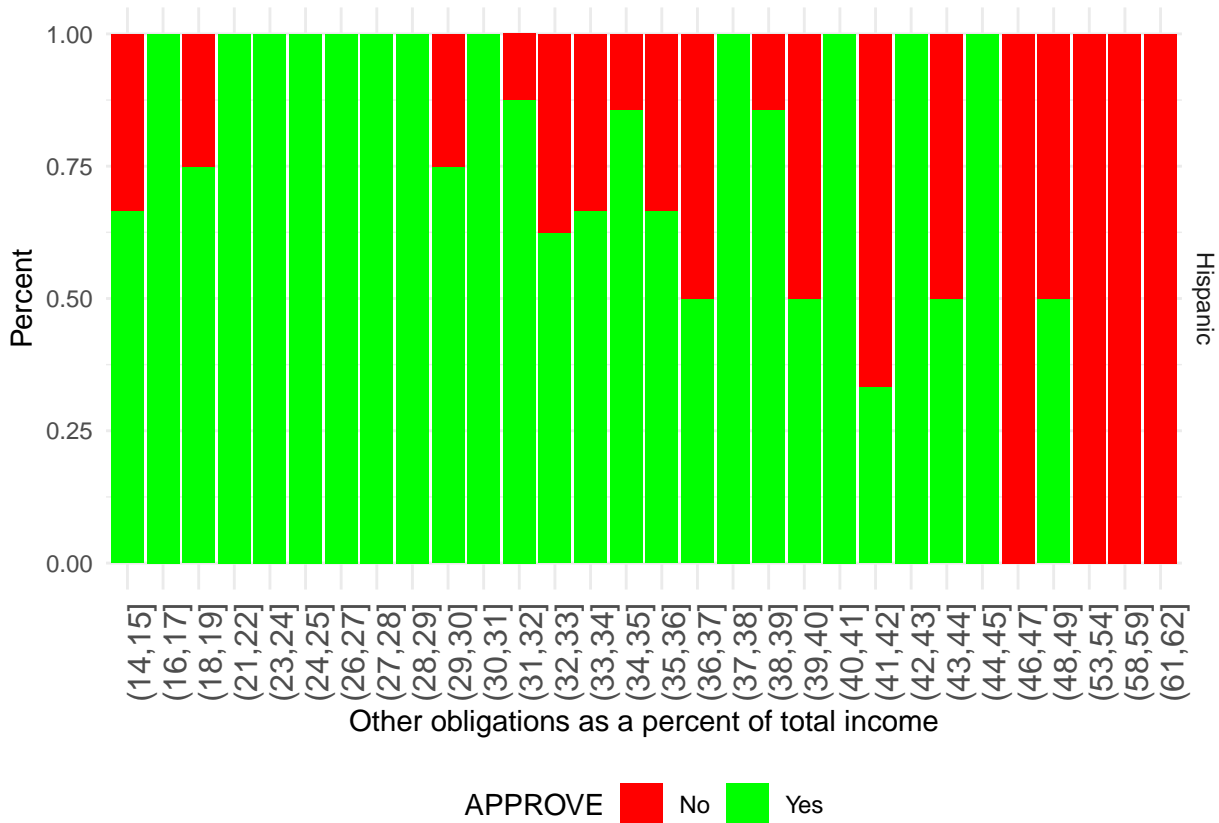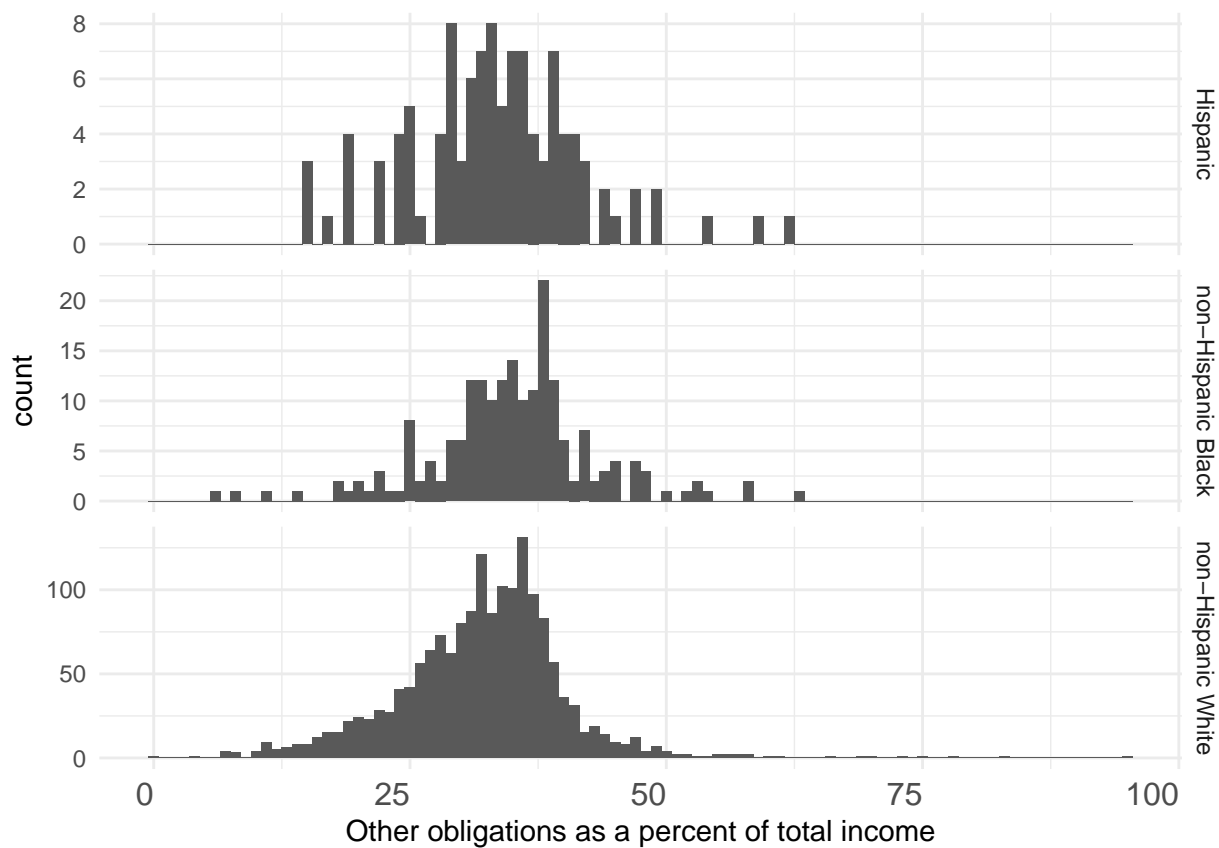
Descriptive statistics by Race:
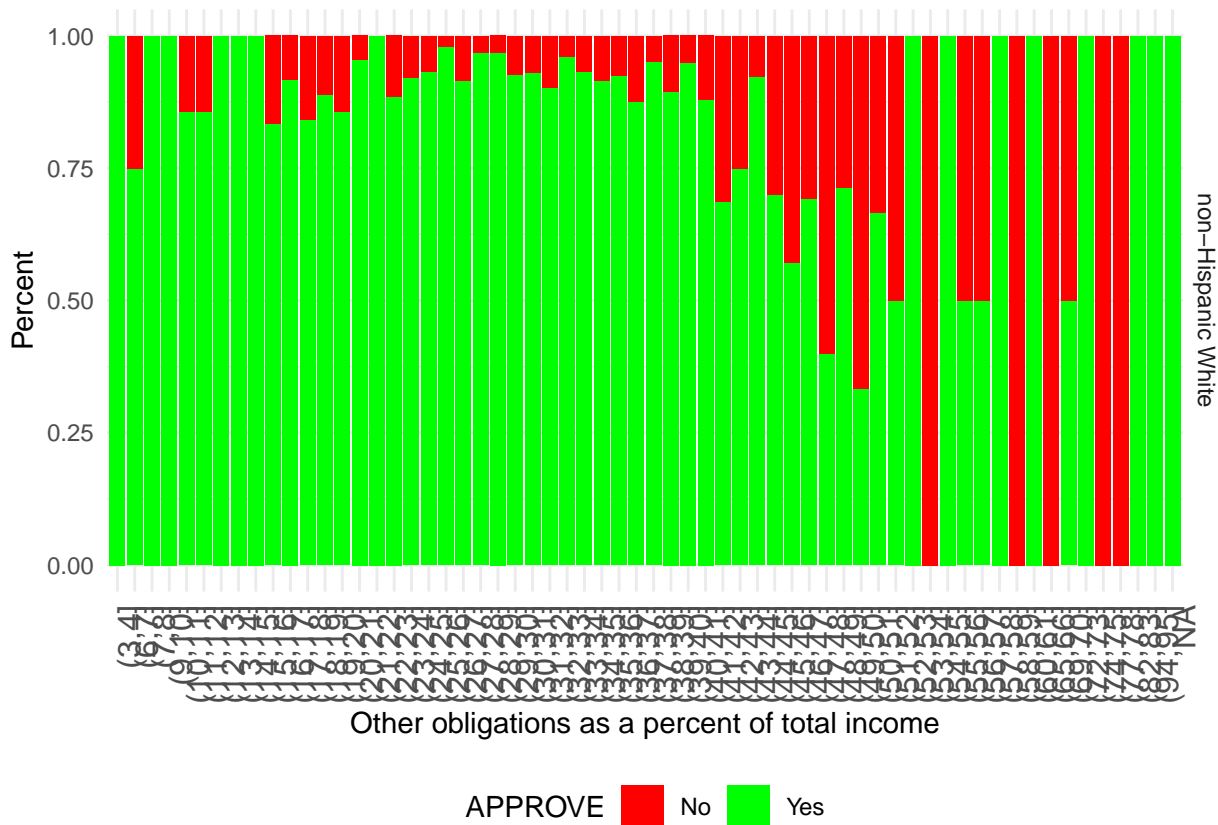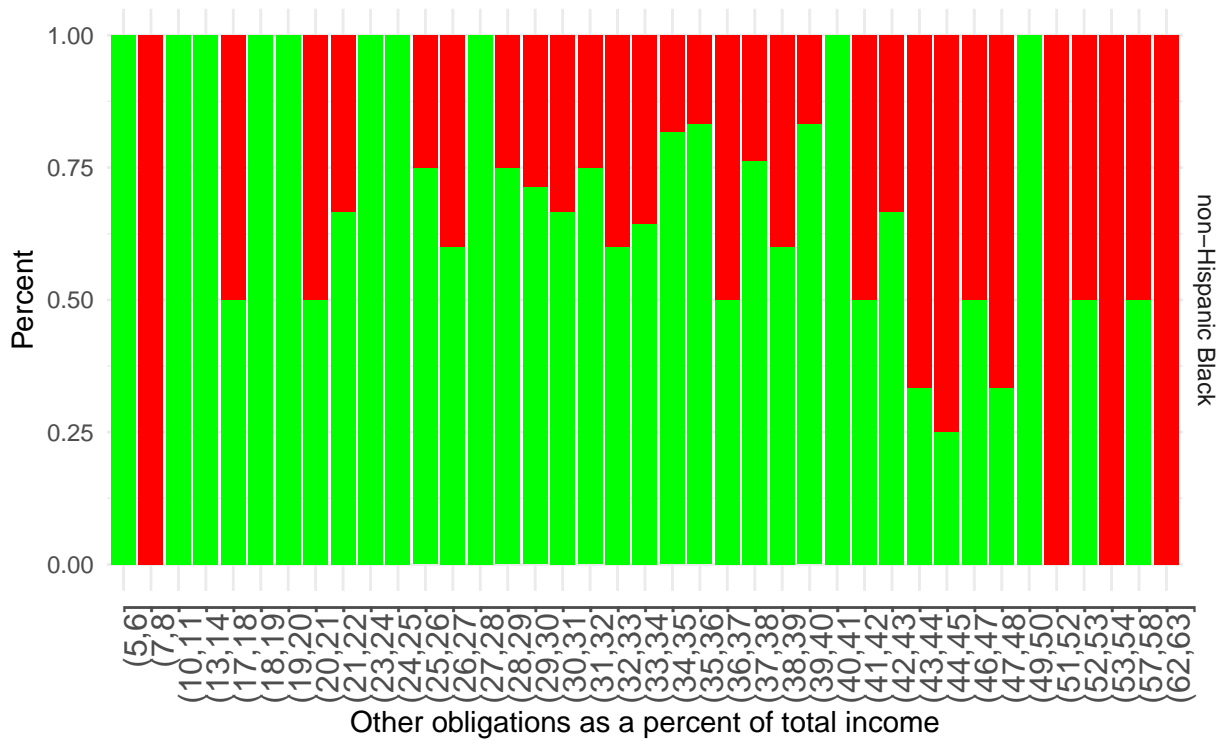
```
## $Hispanic
##     MARRIED    GDLIN        OBRAT           MALE      APPROVE
## No     :31   0  :16   Min.   :14.60   No     :22   No :26
## Unknown: 1   1  :95   1st Qu.:29.00   Unknown: 2   Yes:85
## Yes    :79   666: 0   Median :33.00   Yes    :87
##                       Mean   :33.46
##                       3rd Qu.:38.45
##                       Max.   :62.00
##     LOANPRC                    RACE
## Min.   : 39.39   Hispanic          :111
## 1st Qu.: 80.00   non-Hispanic Black:  0
## Median : 89.39   non-Hispanic White:  0
## Mean   : 85.17
## 3rd Qu.: 90.42
## Max.   :162.63
```
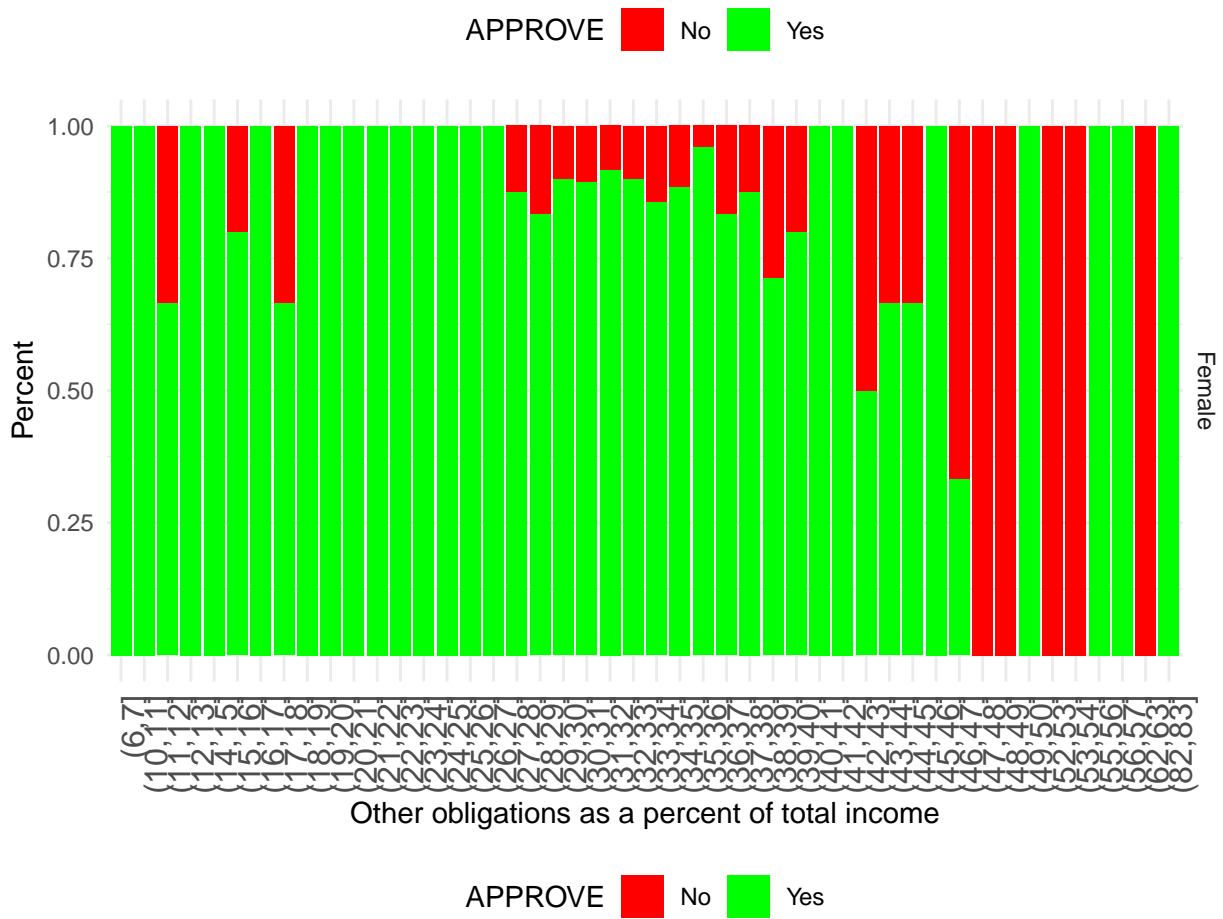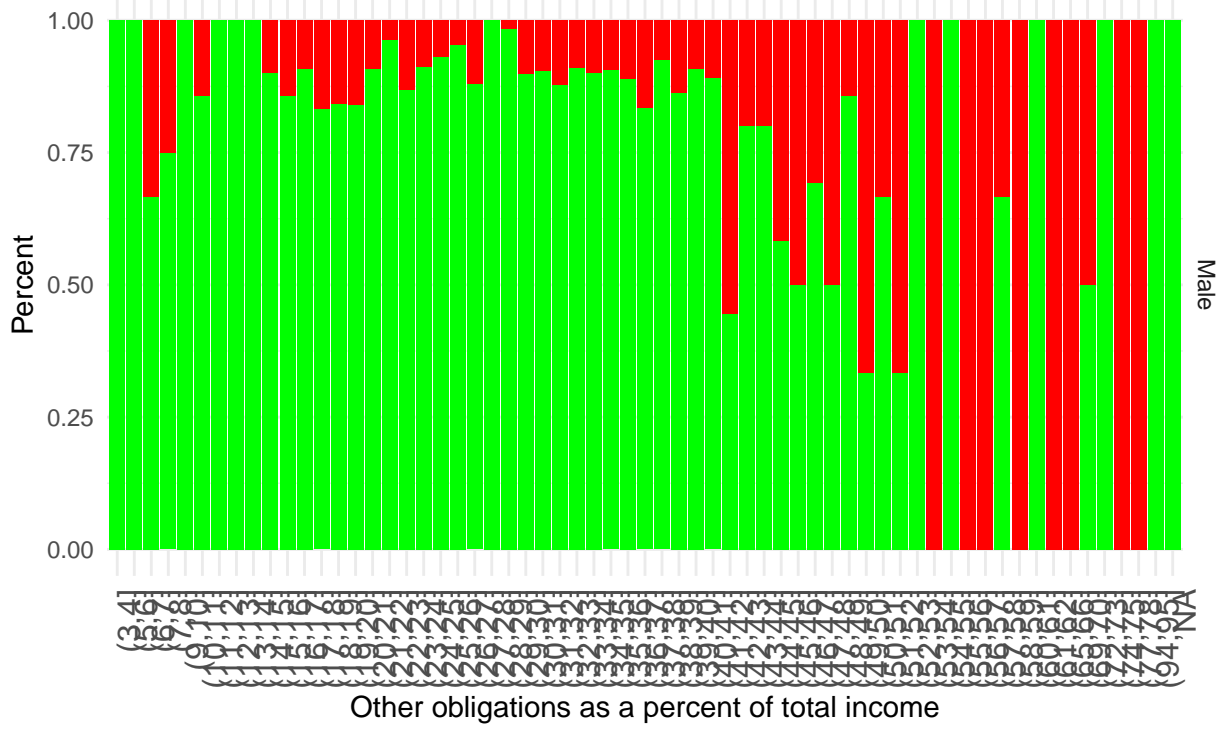
```
## 
## $`non-Hispanic Black`
##     MARRIED      GDLIN          OBRAT            MALE        APPROVE
##  No     : 76   0  : 53   Min.   : 5.60   No     : 51   No : 64
##  Unknown: 0    1  :144   1st Qu.:31.00   Unknown: 2    Yes:133
##  Yes    :121   666: 0    Median :35.00   Yes    :144
##                          Mean   :34.94
##                          3rd Qu.:38.90
##                          Max.   :63.00
##     LOANPRC                    RACE
##  Min.   : 28.99   Hispanic          :  0
##  1st Qu.: 80.00   non-Hispanic Black:197
##  Median : 87.02   non-Hispanic White:  0
##  Mean   : 83.97
##  3rd Qu.: 90.24
##  Max.   :255.52
## 
## $`non-Hispanic White`
##     MARRIED       GDLIN           OBRAT            MALE         APPROVE
##  No     : 571   0  : 102   Min.   : 0.00   No     : 296   No : 154
##  Unknown:   2   1  :1577   1st Qu.:27.60   Unknown:  11   Yes:1527
##  Yes    :1108   666:   2   Median :32.50   Yes    :1374
##                            Mean   :32.02
##                            3rd Qu.:36.50
##                            Max.   :95.00
##     LOANPRC                     RACE
##  Min.   :  2.105   Hispanic          :   0
##  1st Qu.: 68.182   non-Hispanic Black:   0
##  Median : 79.888   non-Hispanic White:1681
##  Mean   : 75.719
##  3rd Qu.: 89.623
##  Max.   :257.143
```
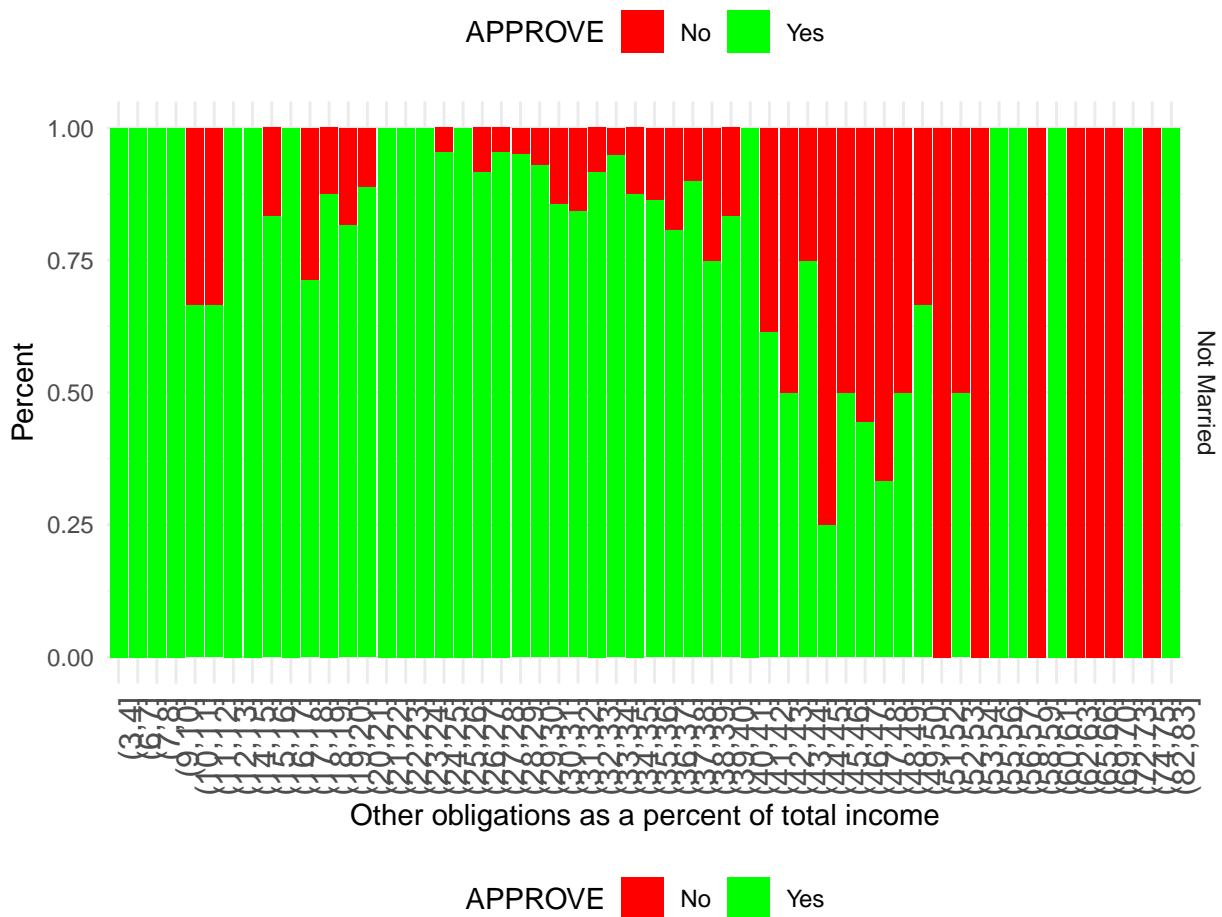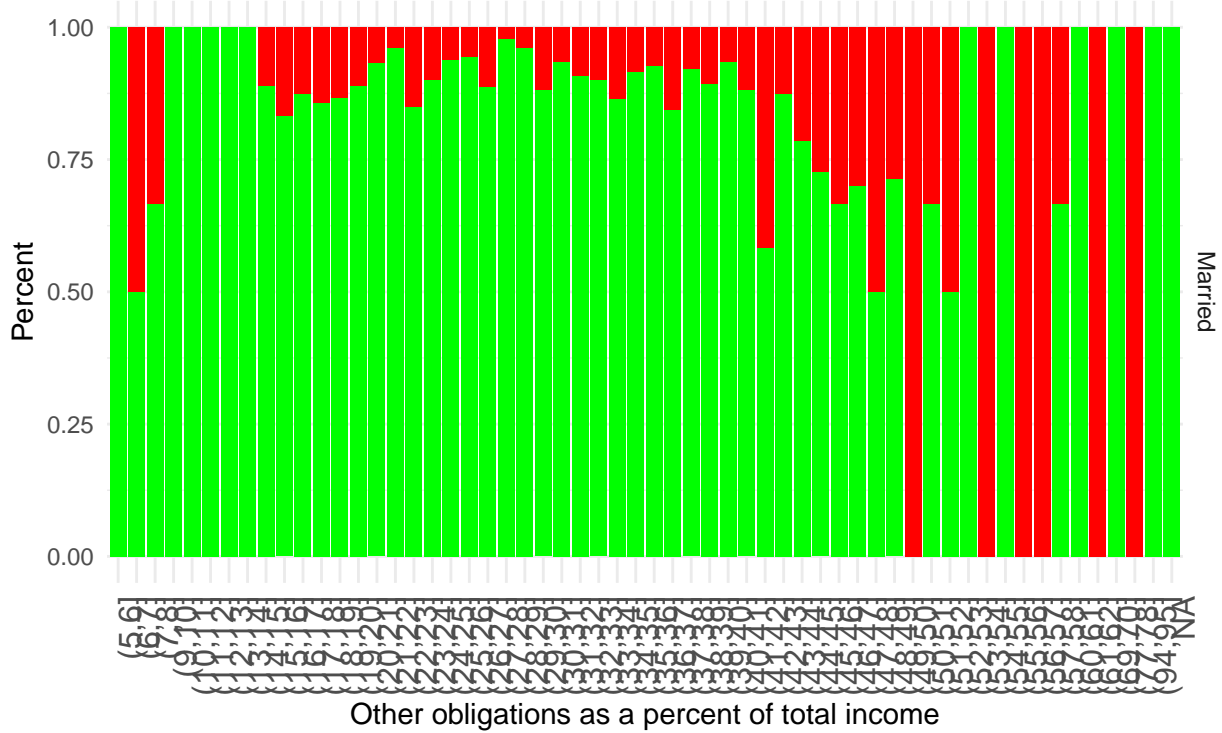
non-Hispanic Black

APPROVE No Yes

non-Hispanic White

APPROVE No Yes

APPROVE No Yes

Male

Other obligations as a percent of total income

APPROVE No Yes

Female

Other obligations as a percent of total income

Married

Other obligations as a percent of total income

APPROVE ■ No ■ Yes

Not Married

Other obligations as a percent of total income

APPROVE ■ No ■ Yes

6

Percent

Hispanic

Loan amount/purchase price %

APPROVE   No   Yes

Percent

non-Hispanic Black

Loan amount/purchase price %

APPROVE   No   Yes

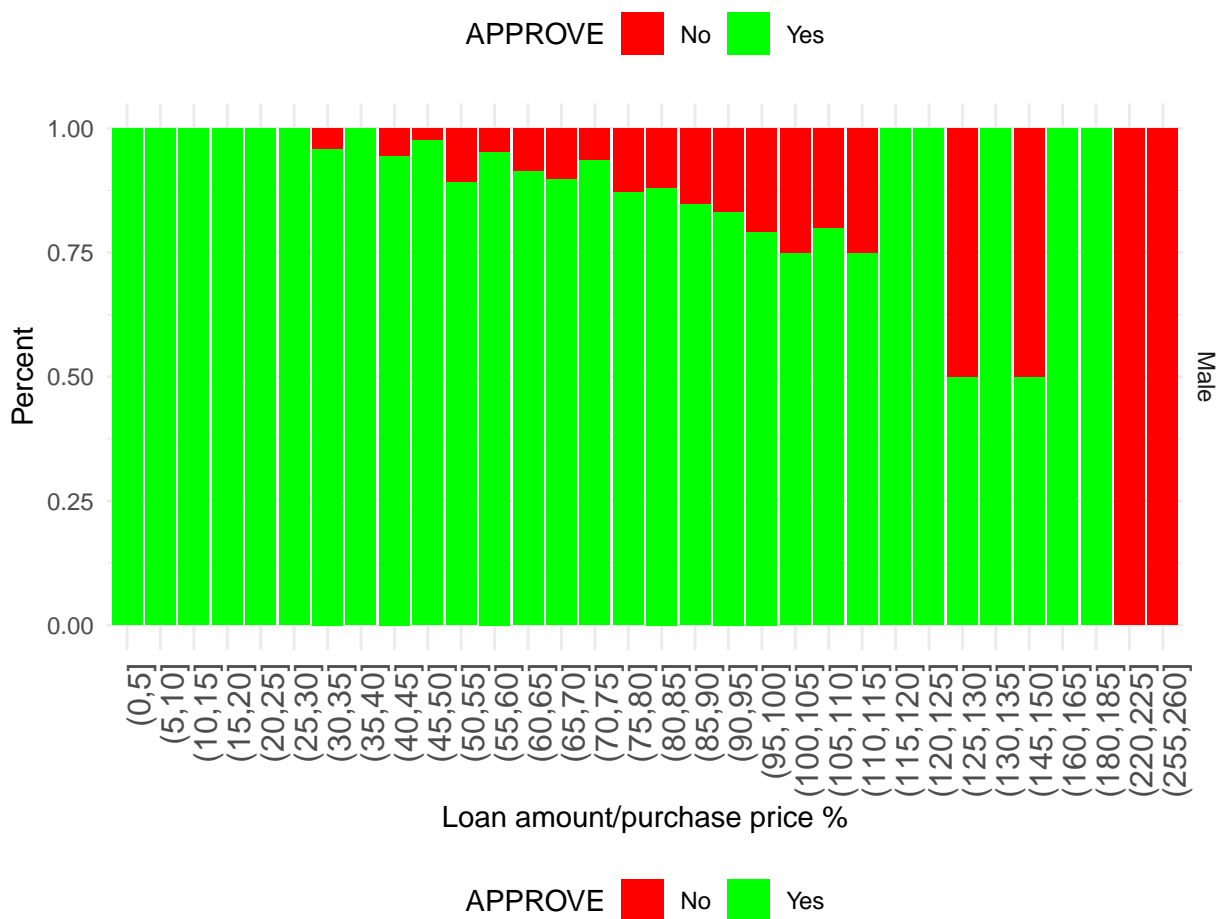Loan amount/purchase price %

APPROVE ■ No ■ Yes

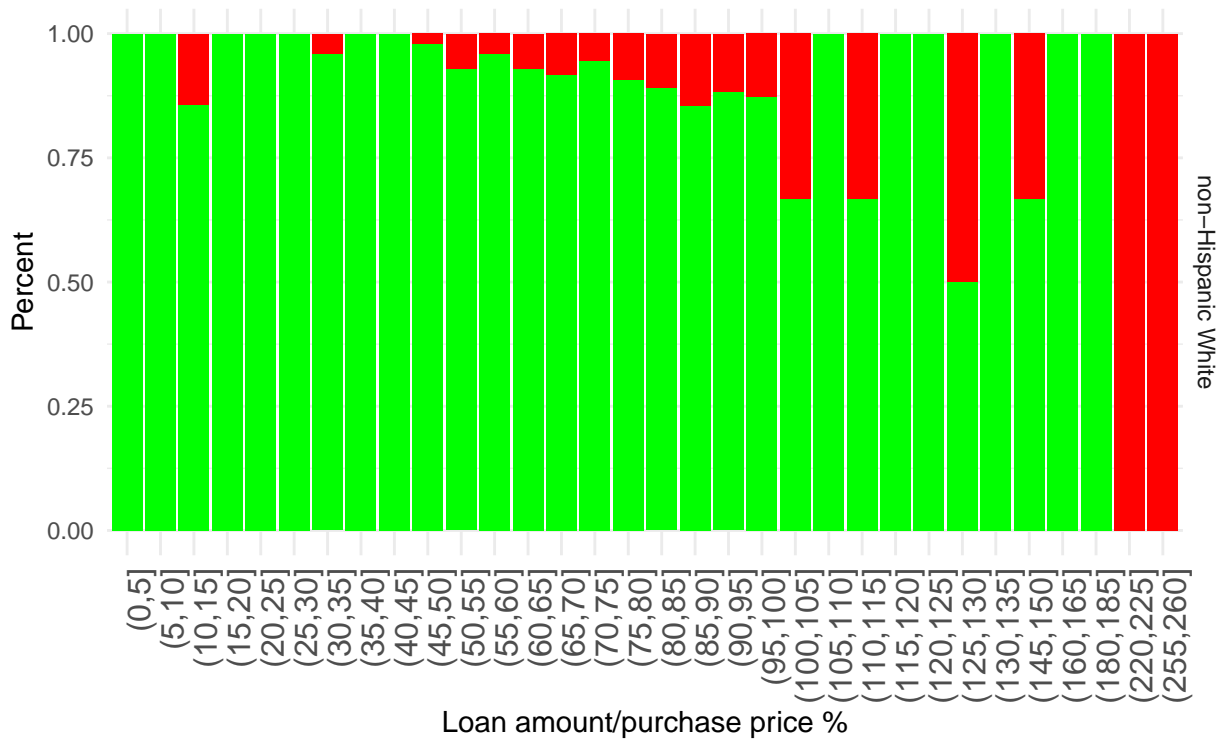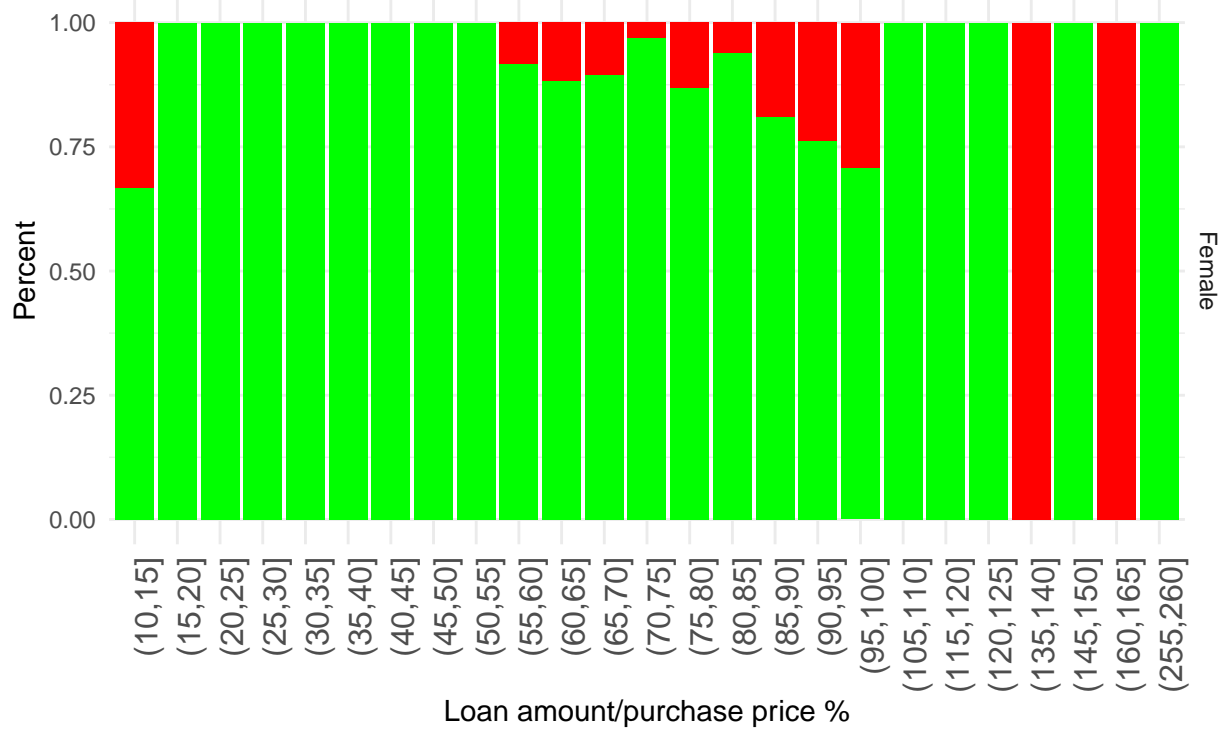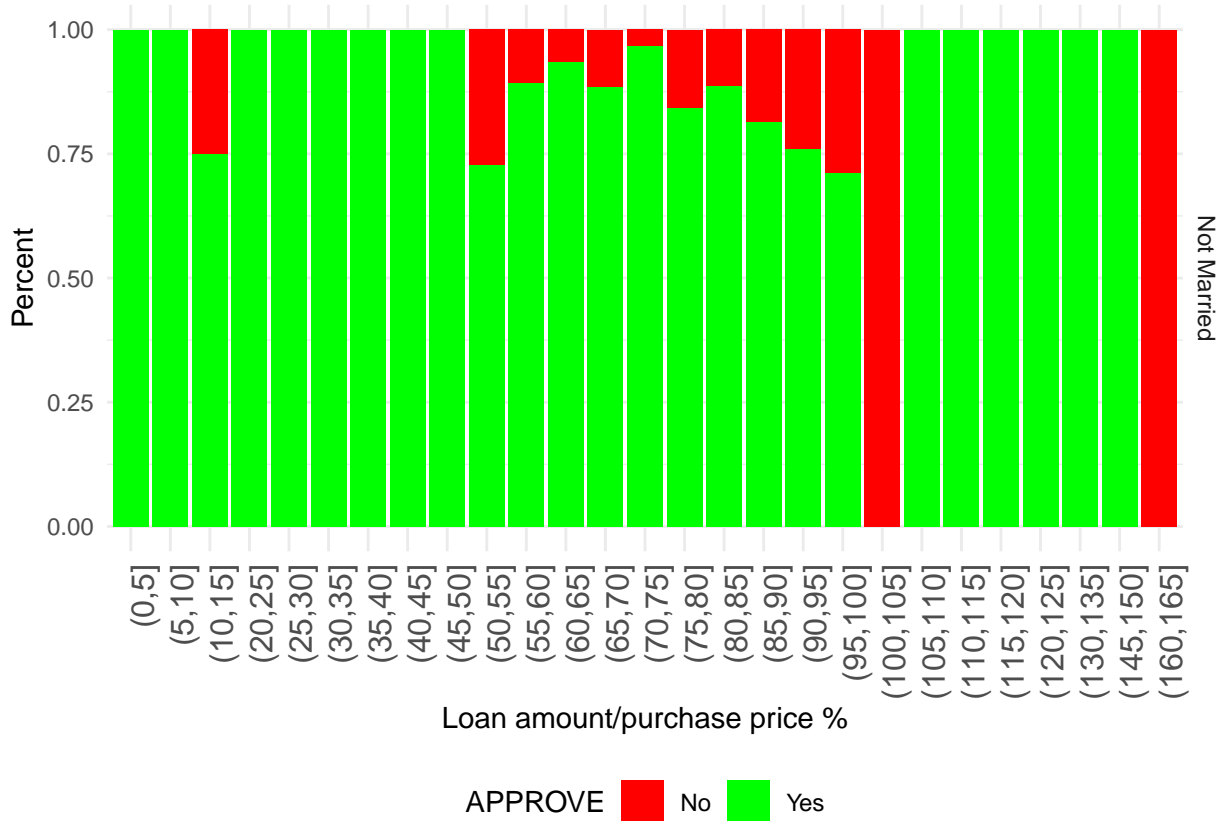Loan amount/purchase price %

APPROVE ■ No ■ Yes

Descriptive statistics by Marital Status:

```
## $No
##     MARRIED      GDLIN         OBRAT          MALE         APPROVE
##   No    :678   0  : 64   Min.   : 4.00   No     :252   No :102
##   Unknown:  0   1  :614   1st Qu.:28.00   Unknown:  7   Yes:576
##   Yes   :  0   666:  0   Median :33.00   Yes    :419
##                          Mean   :32.74
##                          3rd Qu.:37.00
##                          Max.   :83.00
##     LOANPRC                      RACE
##   Min.   :  2.105   Hispanic          : 31
##   1st Qu.: 72.426   non-Hispanic Black: 76
##   Median : 80.000   non-Hispanic White:571
##   Mean   : 77.967
##   3rd Qu.: 89.978
##   Max.   :162.626
##
## $Unknown
##     MARRIED   GDLIN        OBRAT          MALE      APPROVE    LOANPRC
##   No    :0   0  :0   Min.   :13.0   No     :1   No :0   Min.   : 86.96
##   Unknown:3   1  :3   1st Qu.:23.3   Unknown:0   Yes:3   1st Qu.: 88.62
##   Yes   :0   666:0   Median :33.6   Yes    :2           Median : 90.29
##                      Mean   :27.2                       Mean   : 98.16
##                      3rd Qu.:34.3                       3rd Qu.:103.76
##                      Max.   :35.0                       Max.   :117.24
##                  RACE
##   Hispanic          :1
##   non-Hispanic Black:0
```

10

```
##  non-Hispanic White:2
##
##
##
##
## $Yes
##    MARRIED     GDLIN        OBRAT          MALE      APPROVE
## No    :   0   0 : 107   Min.   : 0.00   No    : 116   No : 142
## Unknown:   0   1 :1199   1st Qu.:28.00   Unknown:  8   Yes:1166
## Yes    :1308   666:  2   Median :33.00   Yes    :1184
##                         Mean   :32.22
##                         3rd Qu.:37.00
##                         Max.   :95.00
##     LOANPRC                    RACE
## Min.   :  8.772   Hispanic         :  79
## 1st Qu.: 68.857   non-Hispanic Black: 121
## Median : 80.000   non-Hispanic White:1108
## Mean   : 76.547
## 3rd Qu.: 89.866
## Max.   :257.143
```

Descriptive statistics by Gender:

```
## $No
##    MARRIED     GDLIN        OBRAT          MALE      APPROVE
## No    :252   0 : 31   Min.   : 6.99   No    :369   No : 50
## Unknown: 1   1 :338   1st Qu.:28.00   Unknown:  0   Yes:319
## Yes    :116   666: 0   Median :33.00   Yes    :  0
##                       Mean   :32.64
##                       3rd Qu.:37.00
##                       Max.   :83.00
##     LOANPRC                    RACE
## Min.   : 11.01   Hispanic         : 22
## 1st Qu.: 70.83   non-Hispanic Black: 51
## Median : 80.00   non-Hispanic White:296
## Mean   : 77.66
## 3rd Qu.: 90.00
## Max.   :255.52
##
## $Unknown
##    MARRIED   GDLIN       OBRAT          MALE     APPROVE
## No    :7   0 : 0   Min.   :24.00   No    : 0   No : 0
## Unknown:0   1 :15   1st Qu.:29.95   Unknown:15   Yes:15
## Yes    :8   666: 0   Median :34.50   Yes    : 0
##                     Mean   :33.33
##                     3rd Qu.:37.65
##                     Max.   :40.30
##     LOANPRC                  RACE
## Min.   :39.39   Hispanic         : 2
## 1st Qu.:74.93   non-Hispanic Black: 2
## Median :75.42   non-Hispanic White:11
## Mean   :75.59
## 3rd Qu.:80.43
## Max.   :92.90
##
## $Yes
```

```
##     MARRIED      GDLIN         OBRAT          MALE       APPROVE
## No     : 419   0 : 140   Min.   : 0.00   No     :   0   No : 194
## Unknown:   2   1 :1463   1st Qu.:28.00   Unknown:   0   Yes:1411
## Yes    :1184   666:   2   Median :33.00   Yes    :1605
##                          Mean   :32.32
##                          3rd Qu.:37.00
##                          Max.   :95.00
##     LOANPRC                      RACE
## Min.   :  2.105   Hispanic          :  87
## 1st Qu.: 69.655   non-Hispanic Black: 144
## Median : 80.000   non-Hispanic White:1374
## Mean   : 76.942
## 3rd Qu.: 89.881
## Max.   :257.143
```

There are 3 records are missing married (MARRIED) field.

ID	MARRIED	GDLIN	OBRAT	BLACK	HISPAN	MALE	APPROVE	LOANPRC	RACE
356	Unknown	1	35.0	No	Yes	Yes	Yes	86.95652	Hispanic
759	Unknown	1	33.6	No	No	Yes	Yes	90.28571	non-Hispanic White
1392	Unknown	1	13.0	No	No	No	Yes	117.24140	non-Hispanic White

There are 3 records are missing married (GDLIN) field.

ID	MARRIED	GDLIN	OBRAT	BLACK	HISPAN	MALE	APPROVE	LOANPRC	RACE
881	Yes	666	35	No	No	Yes	Yes	75.82939	non-Hispanic White
1229	Yes	666	26	No	No	Yes	Yes	100.00000	non-Hispanic White

There are 15 records are missing gender (MALE) field.

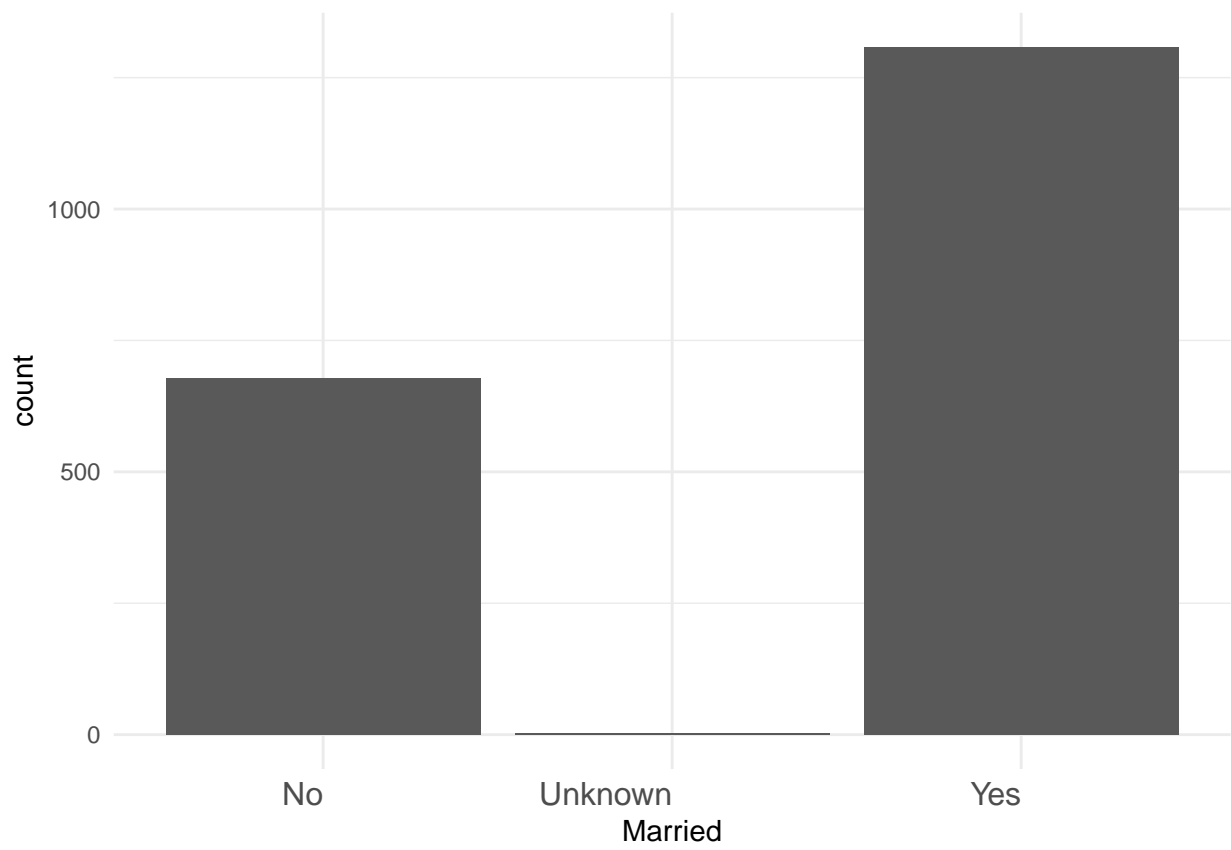| ID	MARRIED	GDLIN	OBRAT	BLACK	HISPAN	MALE	APPROVE	LOANPRC	RACE
1	No	1	34.5	No	No	Unknown	Yes	75.42373	non-Hispanic White
127	No	1	31.6	No	No	Unknown	Yes	80.80000	non-Hispanic White
286	Yes	1	37.3	No	No	Unknown	Yes	80.05337	non-Hispanic White
452	Yes	1	40.1	Yes	No	Unknown	Yes	75.00000	non-Hispanic Black
618	Yes	1	38.5	No	No	Unknown	Yes	92.90323	non-Hispanic White
695	Yes	1	25.0	No	No	Unknown	Yes	64.48276	non-Hispanic White
762	Yes	1	27.6	No	No	Unknown	Yes	75.55556	non-Hispanic White
768	No	1	35.6	No	No	Unknown	Yes	64.74397	non-Hispanic White
833	Yes	1	24.0	No	Yes	Unknown	Yes	79.80769	Hispanic
979	No	1	31.7	No	No	Unknown	Yes	74.86033	non-Hispanic White
1040	No	1	38.0	Yes	No	Unknown	Yes	75.38462	non-Hispanic Black
1070	Yes	1	40.3	No	Yes	Unknown	Yes	39.39394	Hispanic
1092	Yes	1	29.7	No	No	Unknown	Yes	90.10239	non-Hispanic White
1613	No	1	30.2	No	No	Unknown	Yes	90.00000	non-Hispanic White
1924	No	1	35.8	No	No	Unknown	Yes	75.32051	non-Hispanic White

```
##                        APPROVE
## RACE                  No   Yes
##   Hispanic            26    85
##   non-Hispanic Black  64   133
##   non-Hispanic White 154  1527
```
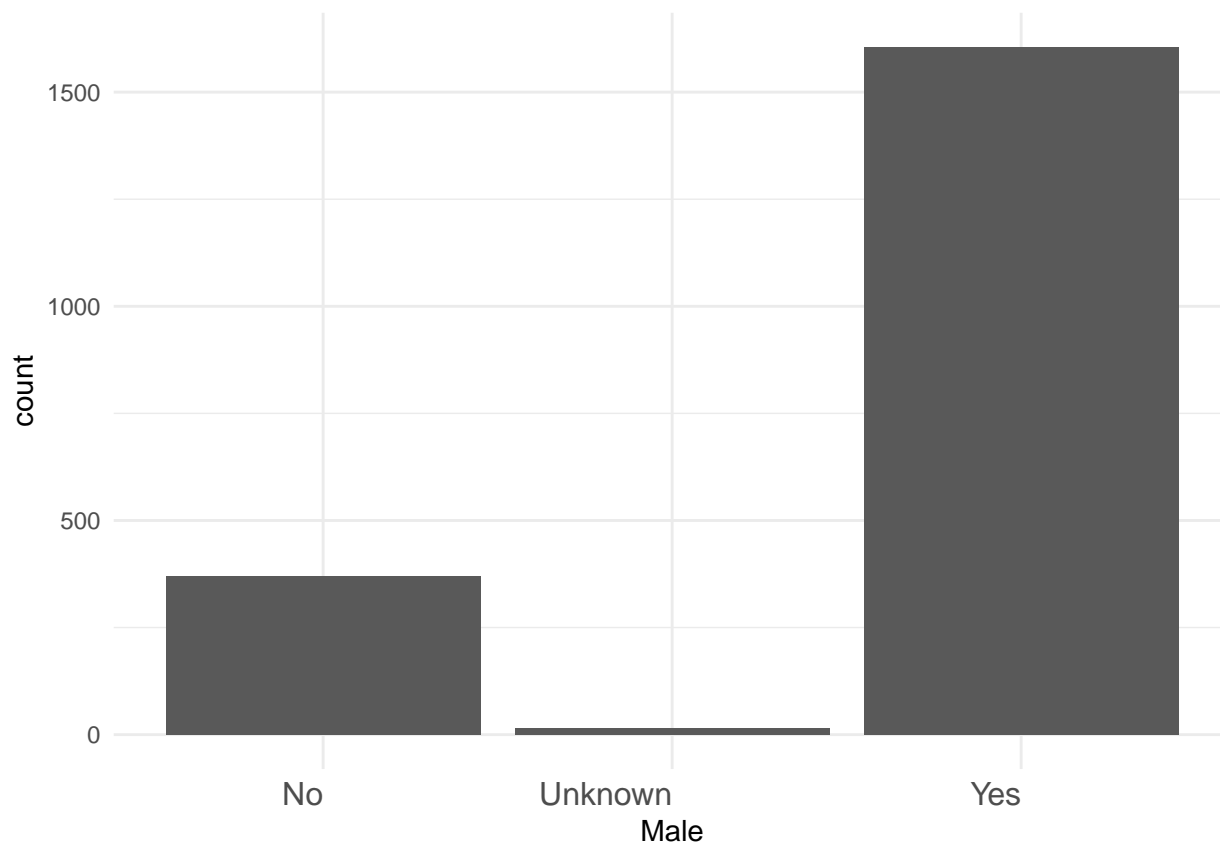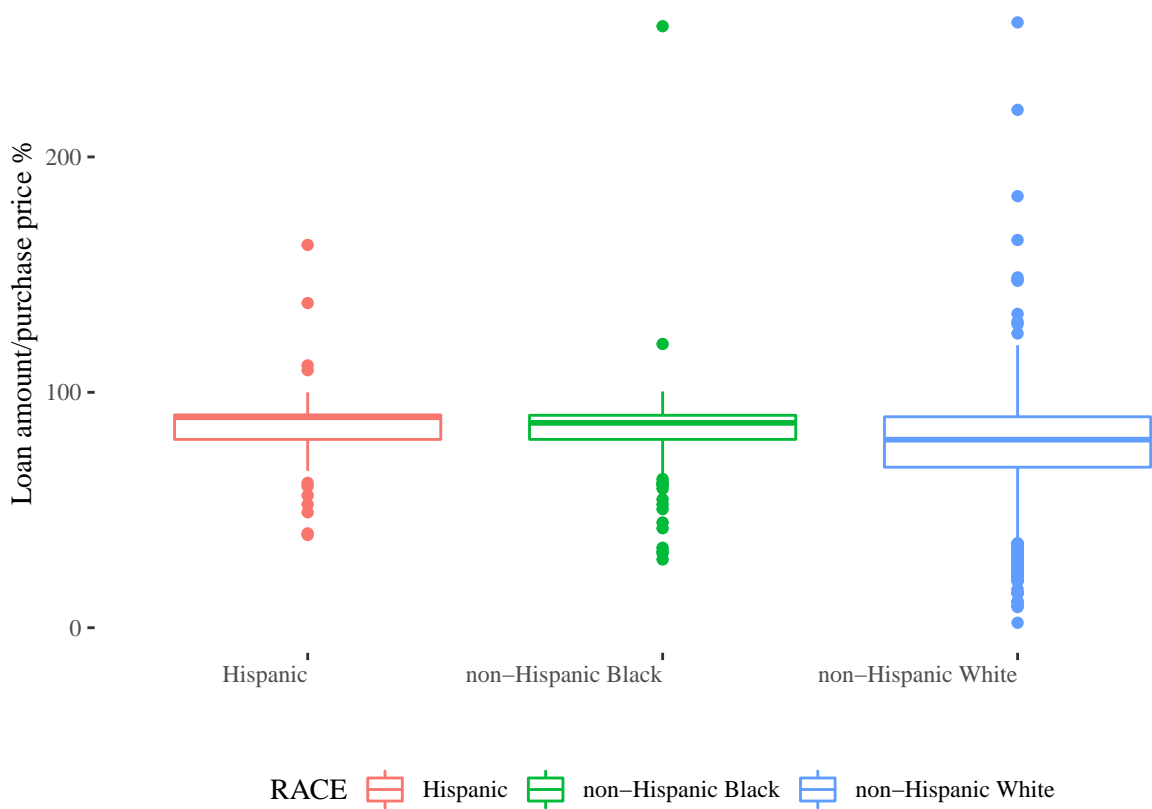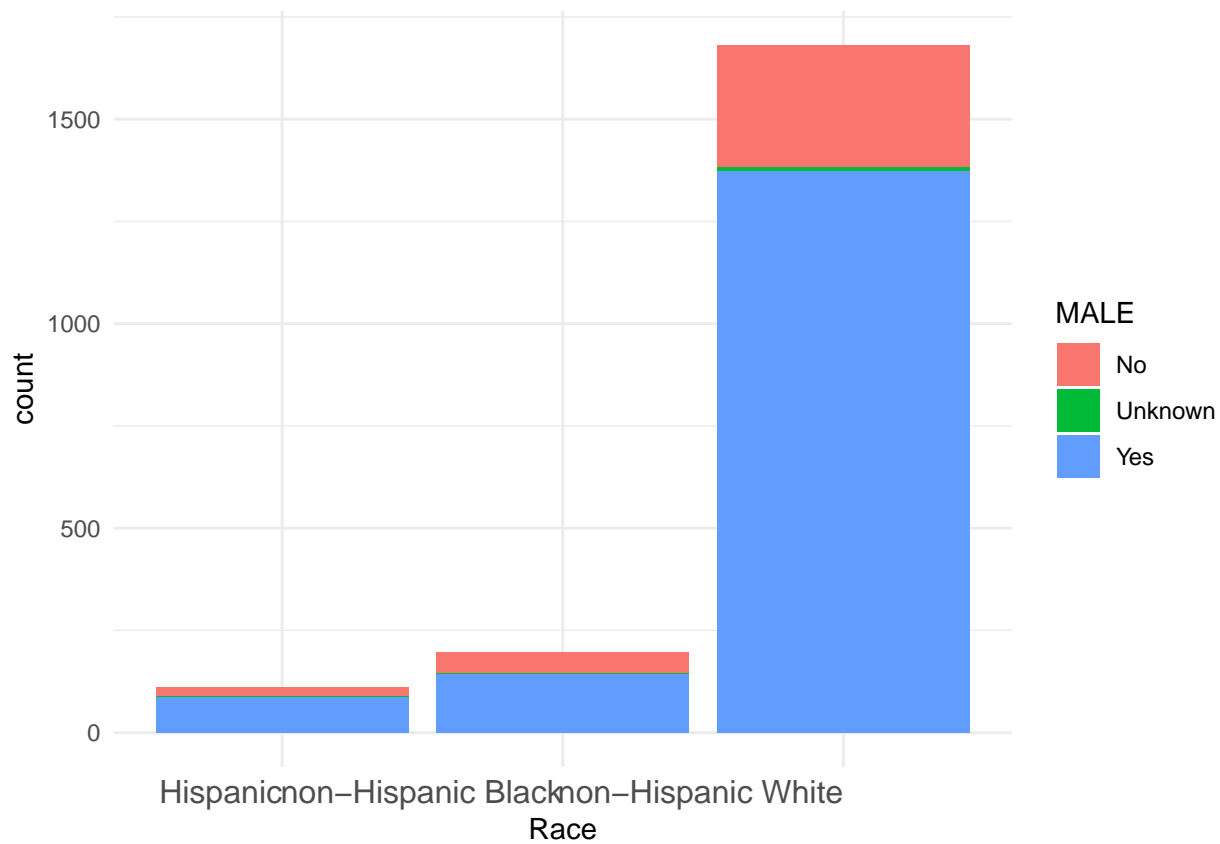
```
##           APPROVE
## MARRIED     No  Yes
##    No      102  576
##    Unknown   0    3
##    Yes     142 1166
```

```
##           APPROVE
## MALE       No  Yes
##   No       50  319
##   Unknown   0   15
##   Yes     194 1411
```

```
## , , RACE = Hispanic
##
##          APPROVE
## MARRIED   No  Yes
##   No        7   24
##   Unknown   0    1
##   Yes      19   60
##
## , , RACE = non-Hispanic Black
##
##          APPROVE
## MARRIED   No  Yes
##   No       27   49
##   Unknown   0    0
##   Yes      37   84
##
## , , RACE = non-Hispanic White
##
##          APPROVE
## MARRIED   No  Yes
##   No       68  503
##   Unknown   0    2
##   Yes      86 1022
```

Logistic Regression Full Model

$log(p/1-p) = b_0 + b_1 * GDLIN + b_2 * OBRAT + b_3 * BLACK + b_4 * HISPAN + b_5 * MALE + b_6 * LOANPRC + b_7 * MARRIED$

```
##
## Call:  glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MALE +
##     LOANPRC + MARRIED, family = "binomial", data = data)
##
## Coefficients:
## (Intercept)       GDLIN1        OBRAT       BLACK1      HISPAN1
##    13.85208      3.71927     -0.03407     -0.81569     -0.90001
##       MALE1      LOANPRC      MARRIED0     MARRIED1
##    -0.05395     -0.01681    -12.47055    -11.99481
##
## Degrees of Freedom: 1971 Total (i.e. Null);  1963 Residual
## Null Deviance:       1476
## Residual Deviance: 959.4     AIC: 977.4
```

For every one unit change in `OBRAT`, the log odds of loan approval (versus non loan approval) increases by 3.719269.

For every one unit change in `LOANPRC`, the log odds of loan approval (versus non loan approval) decreases by 3.719269.

The log odds of loan approval for applicants that meet credit guidelines is 3.719269.

The log odds of loan approval for married applicants is NA.

The log odds of loan approval for Black applicants is -0.8156932.

The log odds of loan approval for Hispanic applicants is -0.9000102.

Statistical Tests for Individual Predictors

```
##            Overall
## GDLIN1   17.1261604
## OBRAT     3.3050712
## BLACK1    3.3962204
## HISPAN1   2.8977950
## MALE1     0.2299810
## LOANPRC   3.3133889
## MARRIED0  0.0249966
## MARRIED1  0.0240430
```

```
## Wald test for GDLIN
##  in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MALE +
##     LOANPRC + MARRIED, family = "binomial", data = data)
## F =  293.3054  on  1  and  1963  df: p= < 0.000000000000000222
```

```
## Wald test for OBRAT
##  in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MALE +
##     LOANPRC + MARRIED, family = "binomial", data = data)
## F =  10.9235  on  1  and  1963  df: p= 0.00096662
```

```
## Wald test for BLACK
##  in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MALE +
##     LOANPRC + MARRIED, family = "binomial", data = data)
## F =  11.53431  on  1  and  1963  df: p= 0.00069685
```

```
## Wald test for HISPAN
##  in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MALE +
##     LOANPRC + MARRIED, family = "binomial", data = data)
## F =  8.397216  on  1  and  1963  df: p= 0.0037996
```

```
## Wald test for MALE
##  in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MALE +
##     LOANPRC + MARRIED, family = "binomial", data = data)
## F =  0.05289126  on  1  and  1963  df: p= 0.81813
```

```
## Wald test for LOANPRC
##  in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MALE +
##     LOANPRC + MARRIED, family = "binomial", data = data)
## F =  10.97855  on  1  and  1963  df: p= 0.00093848
```

```
## Wald test for MARRIED
##  in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MALE +
##     LOANPRC + MARRIED, family = "binomial", data = data)
## F =  3.069927  on  2  and  1963  df: p= 0.046647
```

```
##          llh       llhNull          G2     McFadden         r2ML
## -479.7016525 -738.1114519  516.8195988    0.3500959    0.2305497
##         r2CU
##    0.4375025
```

Logistic Regression Reduced Model 1

```
##
## Call:  glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MARRIED +
##     LOANPRC, family = "binomial", data = data)
##
## Coefficients:
## (Intercept)        GDLIN1         OBRAT        BLACK1       HISPAN1
##    13.80364       3.72139      -0.03410      -0.81143      -0.89733
##     MARRIED0      MARRIED1       LOANPRC
##   -12.46126     -12.00033      -0.01677
##
## Degrees of Freedom: 1971 Total (i.e. Null);  1964 Residual
## Null Deviance:       1476
## Residual Deviance: 959.5     AIC: 975.5


## Wald test for GDLIN
##  in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MARRIED +
##     LOANPRC, family = "binomial", data = data)
## F =  294.1283  on  1  and  1964  df: p= < 0.000000000000000222


## Wald test for OBRAT
##  in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MARRIED +
##     LOANPRC, family = "binomial", data = data)
## F =  10.93867  on  1  and  1964  df: p= 0.00095877


## Wald test for BLACK
##  in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MARRIED +
##     LOANPRC, family = "binomial", data = data)
## F =  11.47424  on  1  and  1964  df: p= 0.0007196


## Wald test for HISPAN
##  in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MARRIED +
##     LOANPRC, family = "binomial", data = data)
## F =  8.358632  on  1  and  1964  df: p= 0.0038808


## Wald test for MARRIED
##  in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MARRIED +
##     LOANPRC, family = "binomial", data = data)
## F =  3.24141  on  2  and  1964  df: p= 0.039318


## Wald test for LOANPRC
##  in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MARRIED +
##     LOANPRC, family = "binomial", data = data)
## F =  10.92337  on  1  and  1964  df: p= 0.00096667
```

For every one unit change in `OBRAT`, the log odds of loan approval (versus non loan approval) increases by 3.721387.

For every one unit change in `LOANPRC`, the log odds of loan approval (versus non loan approval) decreases by 3.721387.

The log odds of loan approval for applicants that meet credit guidelines is 3.721387.

The log odds of loan approval for Black applicants is -0.8114263.

The log odds of loan approval for Hispanic applicants is -0.8973309.

CIs using profiled log-likelihood

```
##                   2.5 %       97.5 %
## (Intercept) -50.94948215          NA
## GDLIN1         3.30547330  4.15768272
## OBRAT         -0.05017744 -0.01392120
## BLACK1        -1.27292236 -0.33254331
## HISPAN1       -1.48462743 -0.26459344
## MARRIED0               NA 52.66251434
## MARRIED1               NA 53.13273781
## LOANPRC       -0.02697832 -0.00693732
```

CIs using standard errors

```
##                    2.5 %        97.5 %
## (Intercept) -965.11669932 992.723986084
## GDLIN1          3.29609767   4.146676366
## OBRAT          -0.05430447  -0.013891282
## BLACK1         -1.28092588  -0.341926689
## HISPAN1        -1.50565233  -0.289009459
## MARRIED0     -991.38099858 966.458485471
## MARRIED1     -990.92005217 966.919399273
## LOANPRC        -0.02672039  -0.006826432
```

Odds ratios only

```
##         (Intercept)                 GDLIN1                  OBRAT
## 988202.962586327689      41.321668209943        0.966476904510
##              BLACK1                HISPAN1               MARRIED0
##     0.444224024264       0.407656286038        0.000003873870
##            MARRIED1                LOANPRC
##     0.000006142207       0.983366480966
```

Odds ratios and 95% CI

```
##                            OR                                     2.5 %
## (Intercept) 988202.962586327689   0.00000000000000000000007463131
## GDLIN1           41.321668209943 27.261441344313702472845761 80391
## OBRAT             0.966476904510  0.951060656186796471445177 30339
## BLACK1            0.444224024264  0.280012129299445189367645 52564
## HISPAN1           0.407656286038  0.226586743874862023284677 54796
## MARRIED0          0.000003873870                                NA
## MARRIED1          0.000006142207                                NA
## LOANPRC           0.983366480966  0.973382342982328729874552 66411
##                                              97.5 %
## (Intercept)                                      NA
## GDLIN1                                    63.9232226
## OBRAT                                      0.9861753
## BLACK1                                     0.7170976
## HISPAN1                                    0.7675179
## MARRIED0      7430865168401802933 0432.0000000
## MARRIED1     11891998342204516361 8304.0000000
## LOANPRC                                    0.9930867
```

Model comparison

```
## Analysis of Deviance Table
##
## Model 1: APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MALE + LOANPRC + MARRIED
## Model 2: APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MARRIED + LOANPRC
##   Resid. Df Resid. Dev Df  Deviance Pr(>Chi)
## 1      1963     959.40
## 2      1964     959.46 -1 -0.053132   0.8177
```

```
## Likelihood ratio test
##
## Model 1: APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MALE + LOANPRC + MARRIED
## Model 2: APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MARRIED + LOANPRC
##   #Df  LogLik Df  Chisq Pr(>Chisq)
## 1   9 -479.70
## 2   8 -479.73 -1 0.0531     0.8177
```

Logistic Regression Reduced Model 2

```
##
## Call:  glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC,
##     family = "binomial", data = data)
##
## Coefficients:
## (Intercept)        GDLIN1         OBRAT       BLACK1       HISPAN1
##     1.66644       3.69659      -0.03524      -0.81891      -0.84836
##     LOANPRC
##    -0.01652
##
## Degrees of Freedom: 1971 Total (i.e. Null);  1966 Residual
## Null Deviance:        1476
## Residual Deviance: 966.4     AIC: 978.4
```

```
## Wald test for GDLIN
##  in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MARRIED +
##     LOANPRC, family = "binomial", data = data)
## F =  294.1283  on  1  and  1964  df: p= < 0.000000000000000222
```

```
## Wald test for OBRAT
##  in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MARRIED +
##     LOANPRC, family = "binomial", data = data)
## F =  10.93867  on  1  and  1964  df: p= 0.00095877
```

```
## Wald test for BLACK
##  in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MARRIED +
##     LOANPRC, family = "binomial", data = data)
## F =  11.47424  on  1  and  1964  df: p= 0.0007196
```

```
## Wald test for HISPAN
##  in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MARRIED +
##     LOANPRC, family = "binomial", data = data)
## F =  8.358632  on  1  and  1964  df: p= 0.0038808
```

```
## Wald test for LOANPRC
##  in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MARRIED +
##     LOANPRC, family = "binomial", data = data)
## F =  10.92337  on  1  and  1964  df: p= 0.00096667
```

For every one unit change in `OBRAT`, the log odds of loan approval (versus non loan approval) increases by 3.6965921.

For every one unit change in `LOANPRC`, the log odds of loan approval (versus non loan approval) decreases by 3.6965921.

The log odds of loan approval for applicants that meet credit guidelines is 3.6965921.

The log odds of loan approval for Black applicants is -0.8189078.

The log odds of loan approval for Hispanic applicants is -0.8483646.

CIs using profiled log-likelihood

```
##                     2.5 %        97.5 %
## (Intercept)  0.59172618   2.763074205
## GDLIN1       3.28432234   4.128687936
## OBRAT       -0.05548902  -0.015026037
## BLACK1      -1.27982769  -0.340579521
## HISPAN1     -1.43239130  -0.218385360
## LOANPRC     -0.02669073  -0.006764009
```

CIs using standard errors

```
##                     2.5 %        97.5 %
## (Intercept)  0.58428598   2.74859841
## GDLIN1       3.27522387   4.11796030
## OBRAT       -0.05547238  -0.01500817
## BLACK1      -1.28785992  -0.34995576
## HISPAN1     -1.45374181  -0.24298746
## LOANPRC     -0.02641452  -0.00662186
```

Odds ratios only

```
## (Intercept)      GDLIN1       OBRAT       BLACK1      HISPAN1      LOANPRC
##   5.2933017   40.3096979   0.9653734    0.4409129    0.4281145    0.9836175
```

Odds ratios and 95% CI

```
##                   OR        2.5 %        97.5 %
## (Intercept)  5.2933017   1.8071051   15.8484896
## GDLIN1      40.3096979  26.6908908   62.0963950
## OBRAT        0.9653734   0.9460224    0.9850863
## BLACK1       0.4409129   0.2780852    0.7113580
## HISPAN1      0.4281145   0.2387373    0.8038156
## LOANPRC      0.9836175   0.9736623    0.9932588
```

Model comparison

```
## Analysis of Deviance Table
##
## Model 1: APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MALE + LOANPRC + MARRIED
## Model 2: APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1      1963       959.4
## 2      1966       966.4 -3  -6.9969    0.072 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


## Likelihood ratio test
##
## Model 1: APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MALE + LOANPRC + MARRIED
## Model 2: APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1   9 -479.7
## 2   6 -483.2 -3 6.9969      0.072 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Probit Regression Full Model

```
##
## Call:  glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MALE +
##     LOANPRC + MARRIED, family = binomial(link = "probit"), data = data)
##
## Coefficients:
## (Intercept)       GDLIN1        OBRAT       BLACK1      HISPAN1
##    4.258487     2.142460    -0.016400    -0.425865    -0.463476
##       MALE1      LOANPRC      MARRIED0      MARRIED1
##   -0.033268    -0.008409    -3.692276    -3.454382
##
## Degrees of Freedom: 1971 Total (i.e. Null);  1963 Residual
## Null Deviance:        1476
## Residual Deviance: 958.8     AIC: 976.8
```

Statistical Tests for Individual Predictors

```
## Wald test for GDLIN
##  in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MALE +
##     LOANPRC + MARRIED, family = binomial(link = "probit"), data = data)
## F =  312.1497  on  1  and  1963  df: p= < 0.000000000000000222


## Wald test for OBRAT
##  in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MALE +
##     LOANPRC + MARRIED, family = binomial(link = "probit"), data = data)
## F =  9.418963  on  1  and  1963  df: p= 0.0021769


## Wald test for BLACK
##  in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MALE +
##     LOANPRC + MARRIED, family = binomial(link = "probit"), data = data)
## F =  11.25189  on  1  and  1963  df: p= 0.00081055
```

```
## Wald test for HISPAN
##  in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MALE +
##      LOANPRC + MARRIED, family = binomial(link = "probit"), data = data)
## F =  8.034942  on  1  and  1963  df: p= 0.0046353


## Wald test for MALE
##  in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MALE +
##      LOANPRC + MARRIED, family = binomial(link = "probit"), data = data)
## F =  0.0795602  on  1  and  1963  df: p= 0.77792


## Wald test for LOANPRC
##  in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MALE +
##      LOANPRC + MARRIED, family = binomial(link = "probit"), data = data)
## F =  10.54323  on  1  and  1963  df: p= 0.0011858


## Wald test for MARRIED
##  in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MALE +
##      LOANPRC + MARRIED, family = binomial(link = "probit"), data = data)
## F =  3.062777  on  2  and  1963  df: p= 0.046981
```

Probit Regression Reduced Model 1

```
##
## Call:  glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MARRIED +
##      LOANPRC, family = binomial(link = "probit"), data = data)
##
## Coefficients:
## (Intercept)        GDLIN1         OBRAT        BLACK1       HISPAN1
##    4.223559      2.143873     -0.016411     -0.422677     -0.461723
##     MARRIED0       MARRIED1       LOANPRC
##   -3.681872      -3.452946     -0.008386
##
## Degrees of Freedom: 1971 Total (i.e. Null);   1964 Residual
## Null Deviance:        1476
## Residual Deviance: 958.9      AIC: 974.9


##             Overall
## GDLIN1    17.69103973
## OBRAT      3.07063777
## BLACK1     3.33759452
## HISPAN1    2.82531621
## MARRIED0   0.04454105
## MARRIED1   0.04177164
## LOANPRC    3.23991979


##          llh       llhNull            G2      McFadden          r2ML
## -479.4603593 -738.1114519  517.3021851     0.3504228     0.2307380
##         r2CU
##     0.4378597


## Analysis of Deviance Table
##
## Model 1: APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MALE + LOANPRC + MARRIED
## Model 2: APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MARRIED + LOANPRC
```

```
##    Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1      1963     958.84
## 2      1964     958.92 -1 -0.07991   0.7774


## Likelihood ratio test
##
## Model 1: APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MALE + LOANPRC + MARRIED
## Model 2: APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MARRIED + LOANPRC
##   #Df  LogLik Df  Chisq Pr(>Chisq)
## 1   9 -479.42
## 2   8 -479.46 -1 0.0799     0.7774
```

Probit Regression Reduced Model 2

```
##
## Call:  glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC,
##     family = binomial(link = "probit"), data = data)
##
## Coefficients:
## (Intercept)        GDLIN1         OBRAT        BLACK1       HISPAN1
##    0.706553      2.136046     -0.017060     -0.427179     -0.432845
##     LOANPRC
##   -0.008317
##
## Degrees of Freedom: 1971 Total (i.e. Null);  1966 Residual
## Null Deviance:       1476
## Residual Deviance: 965.8     AIC: 977.8


##          Overall
## GDLIN1  17.702452
## OBRAT    3.197948
## BLACK1   3.383556
## HISPAN1  2.662084
## LOANPRC  3.236990


##          llh       llhNull          G2      McFadden          r2ML
## -482.9023693 -738.1114519  510.4181652     0.3457595     0.2280479
##         r2CU
##    0.4327549


## Analysis of Deviance Table
##
## Model 1: APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MALE + LOANPRC + MARRIED
## Model 2: APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1      1963     958.84
## 2      1966     965.80 -3  -6.9639  0.07306 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


## Likelihood ratio test
##
## Model 1: APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + MALE + LOANPRC + MARRIED
## Model 2: APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC
```
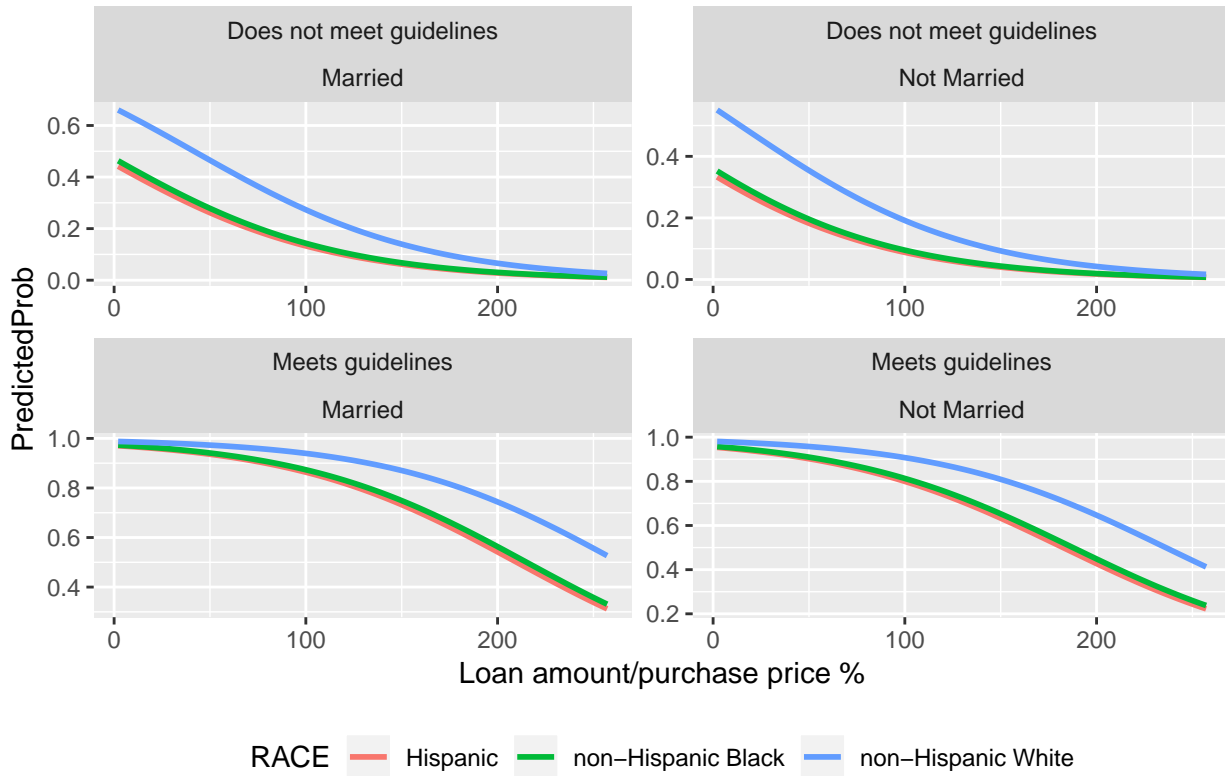
```
##    #Df  LogLik Df  Chisq Pr(>Chisq)
## 1   9 -479.42
## 2   6 -482.90 -3 6.9639    0.07306 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```
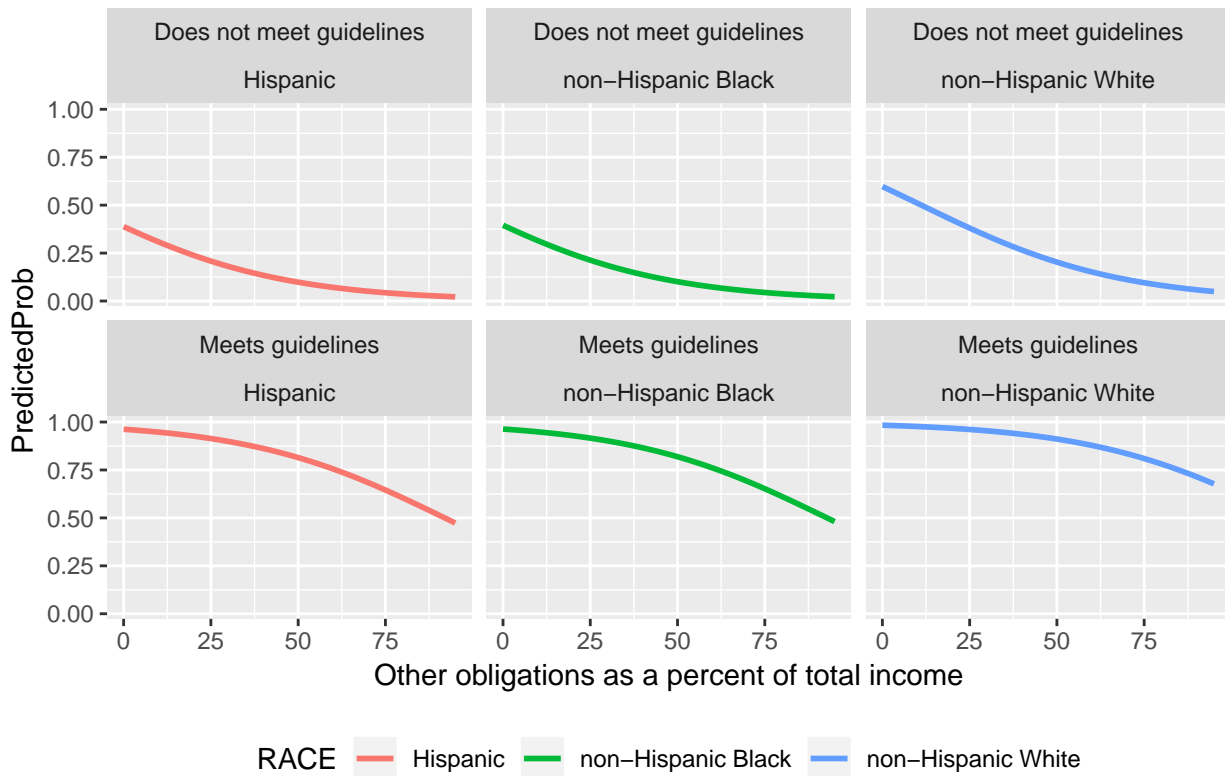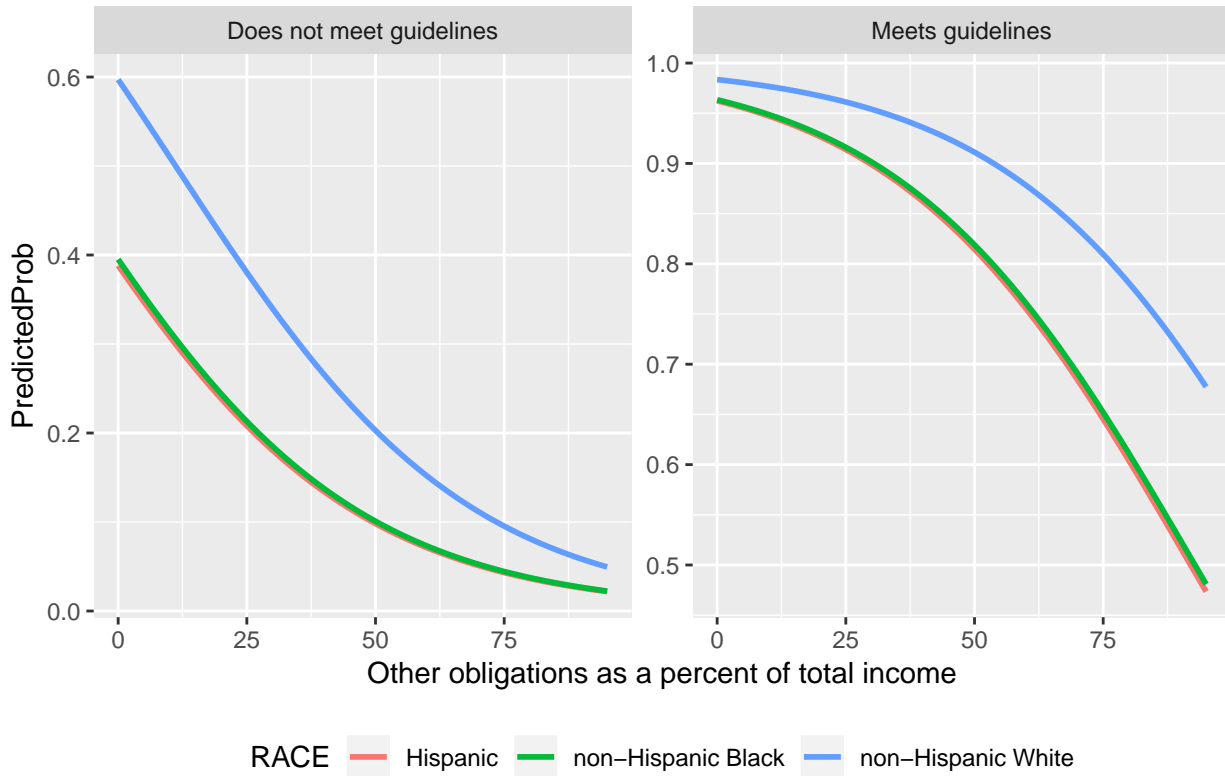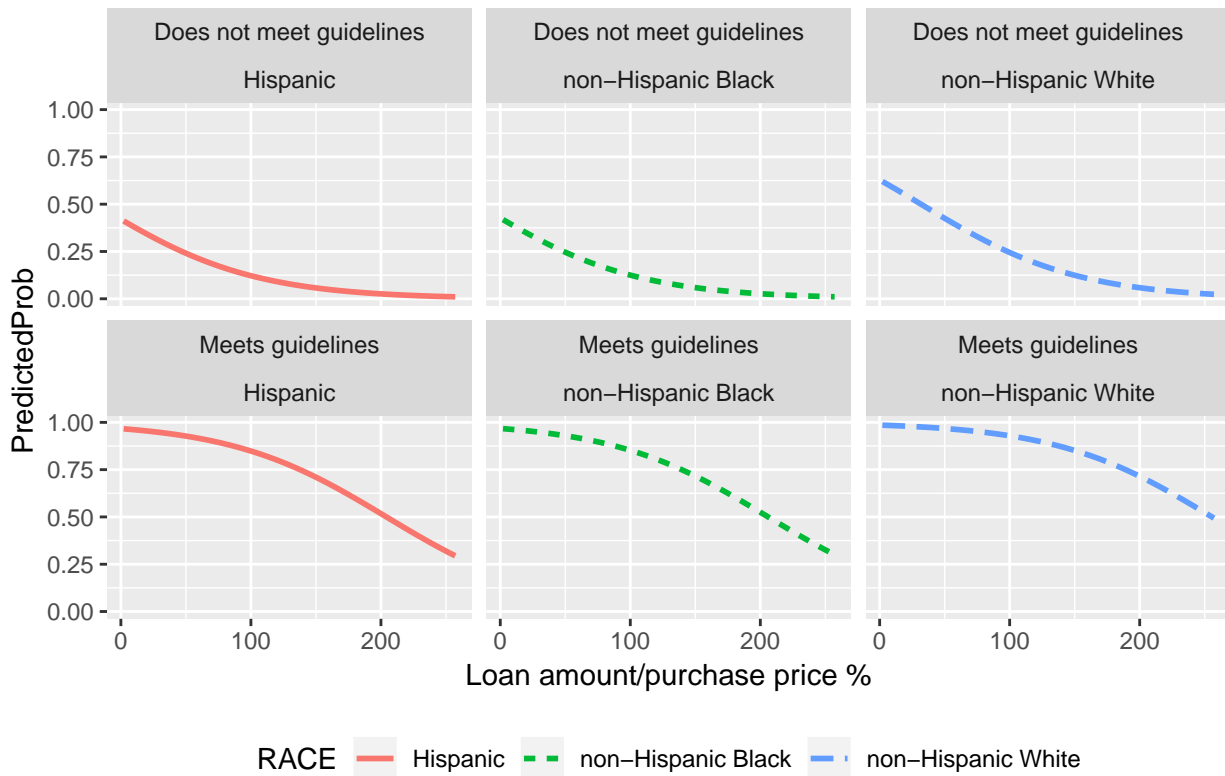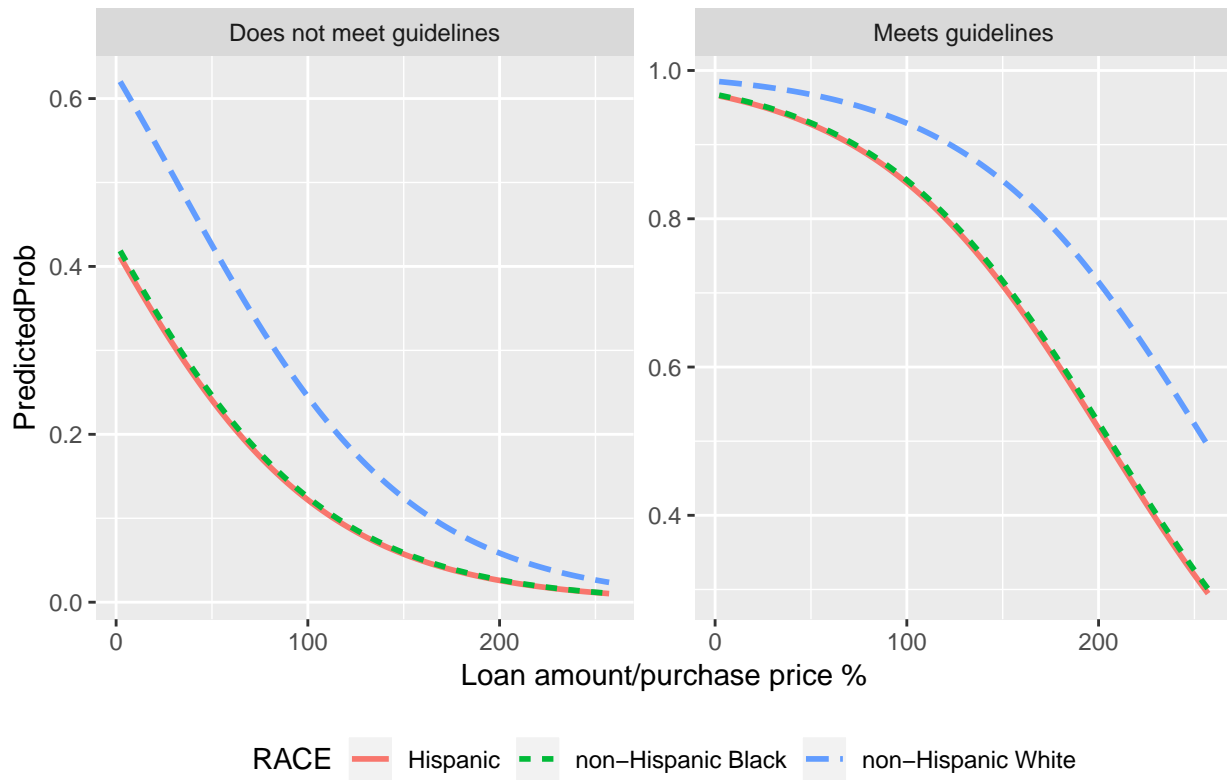
TBD

Predicted probabilities
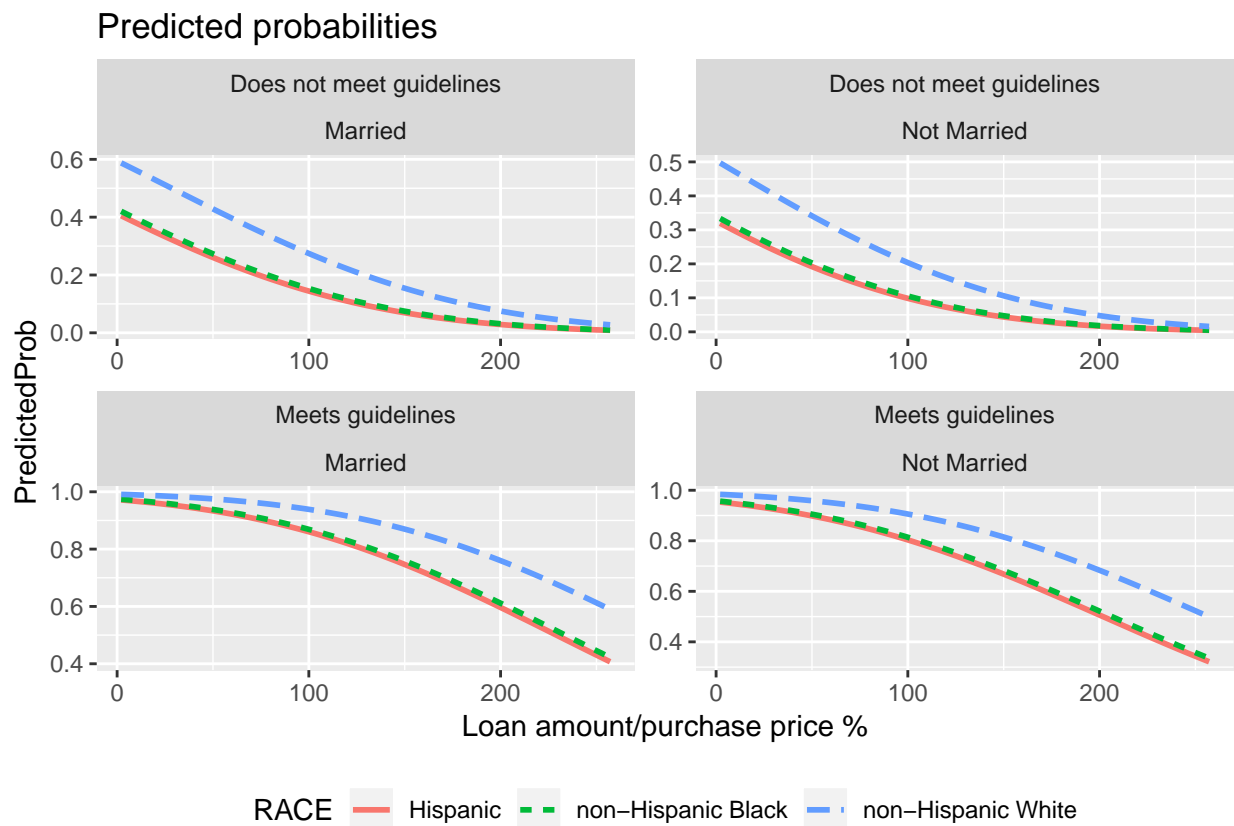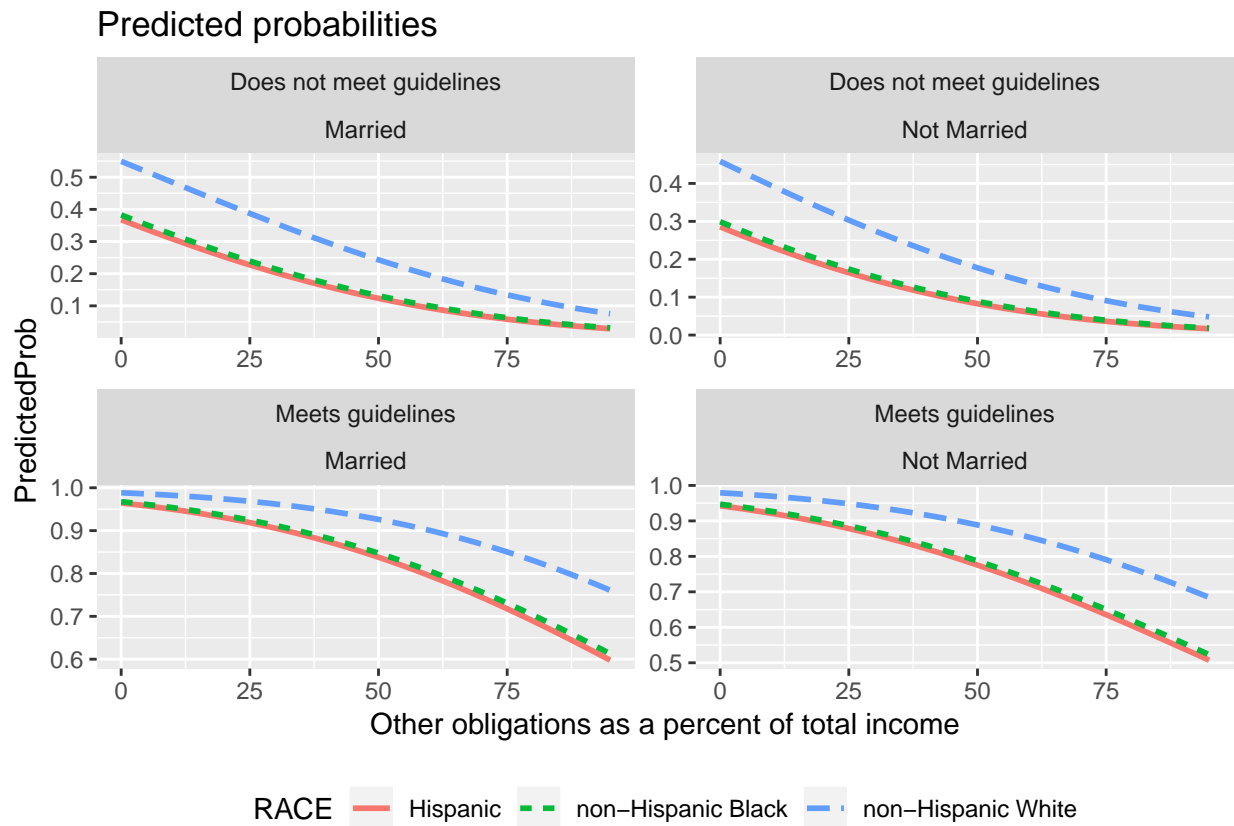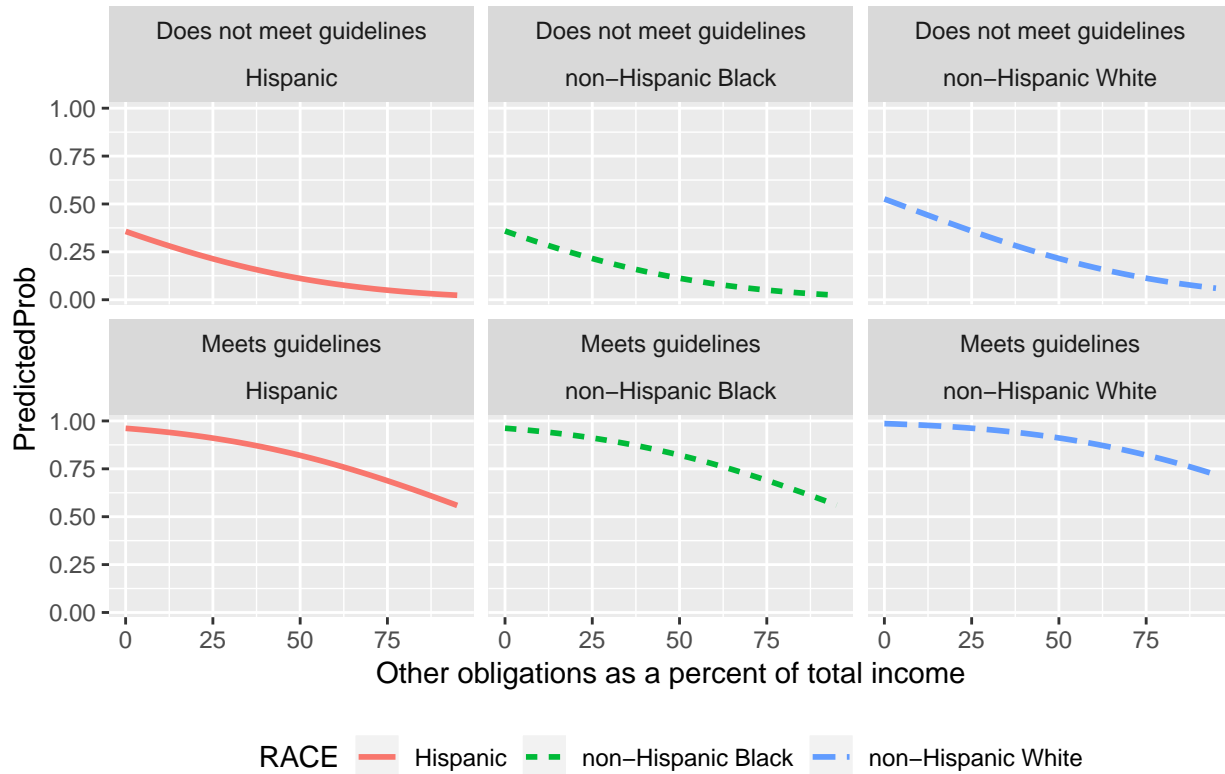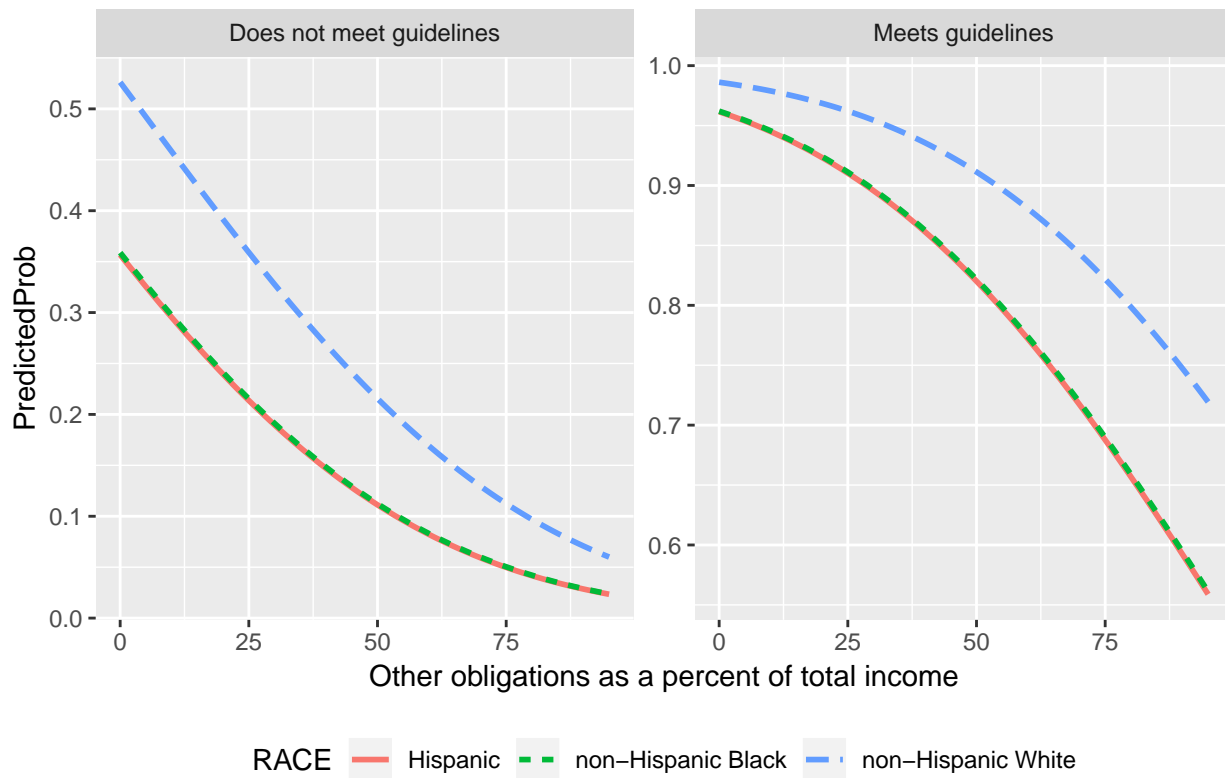
Predicted probabilities

Predicted probabilities

Predicted probabilities

Does not meet guidelines | Meets guidelines

Other obligations as a percent of total income

RACE — Hispanic — non–Hispanic Black — non–Hispanic White

Predicted probabilities

Loan amount/purchase price %

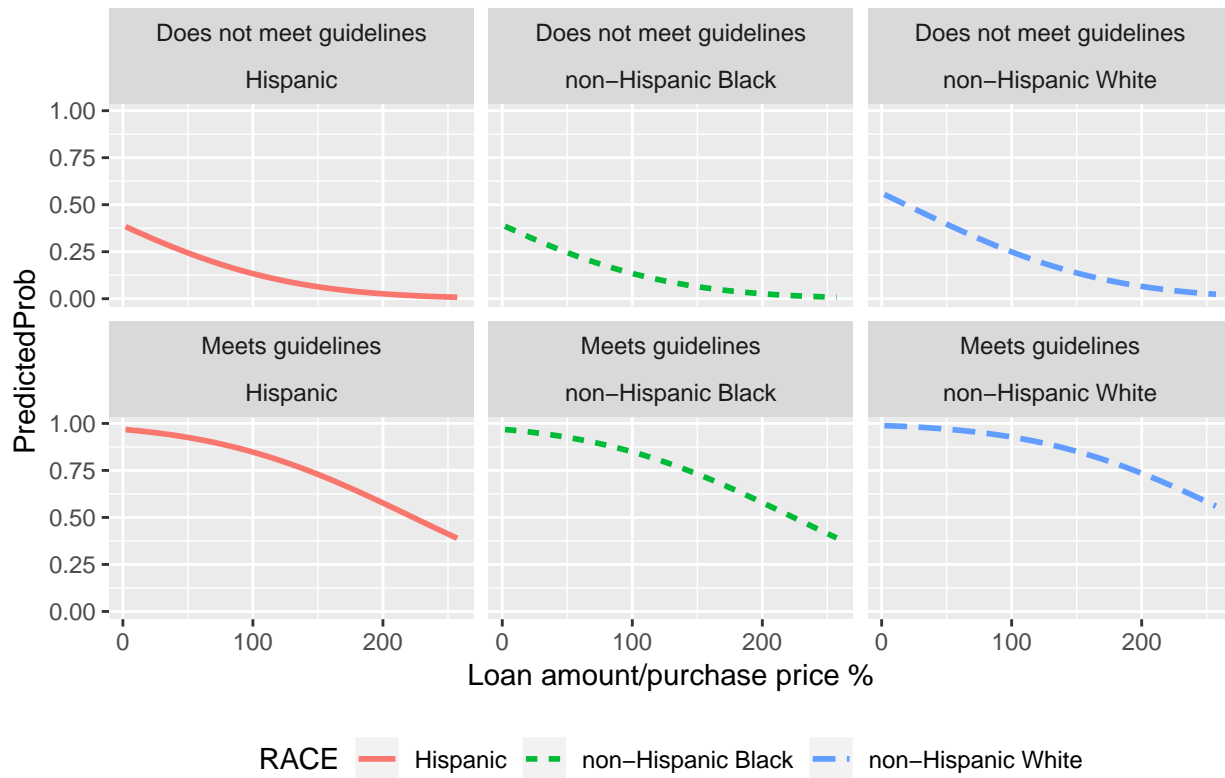RACE — Hispanic — — non–Hispanic Black — — non–Hispanic White

Predicted probabilities

Predicted probabilities

Predicted probabilities

Predicted probabilities

Predicted probabilities

Predicted probabilities

Predicted probabilities