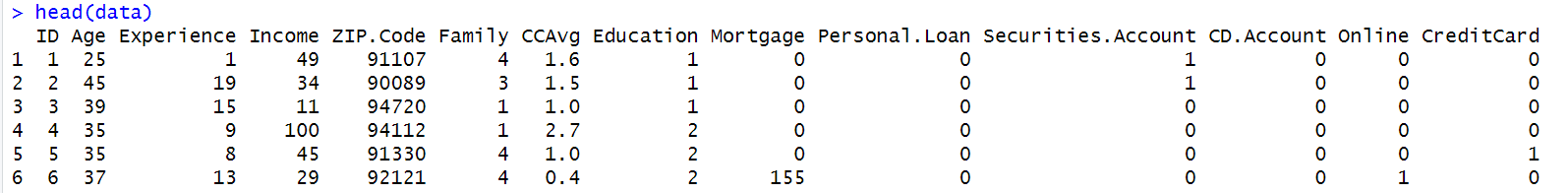
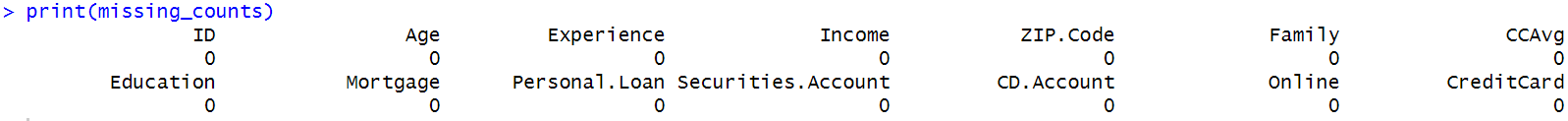
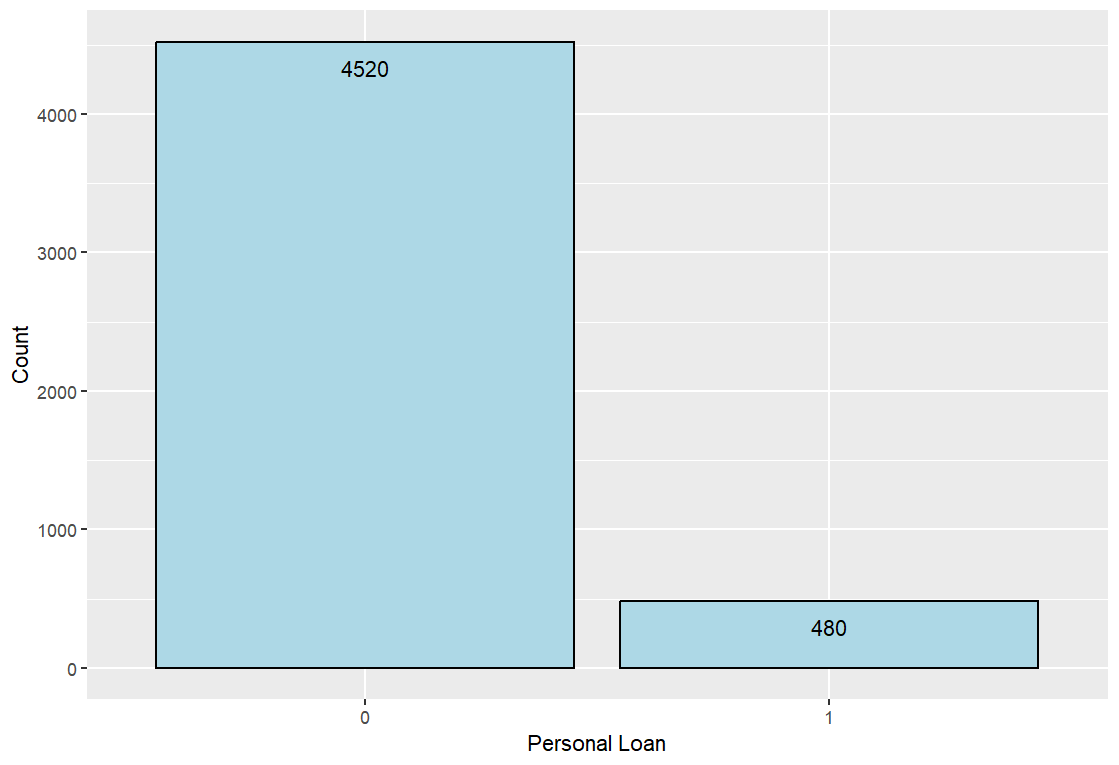
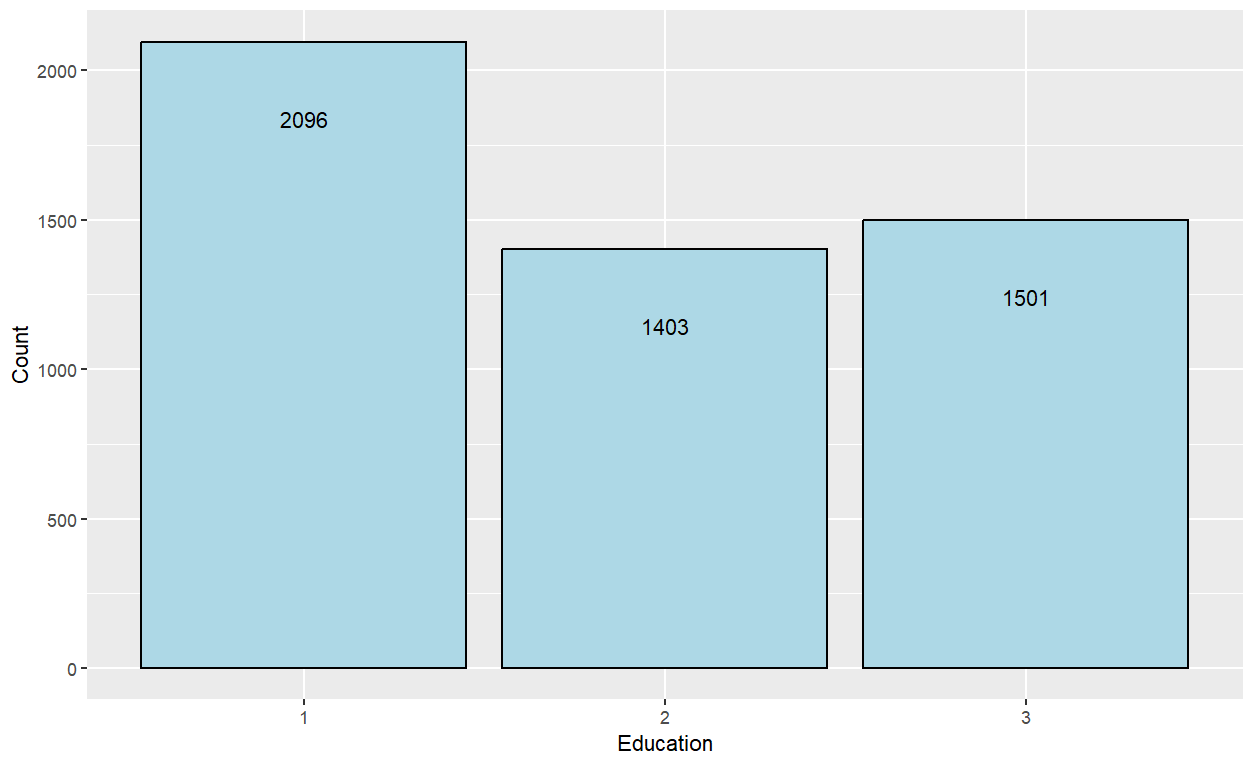
**Logistic Regression**

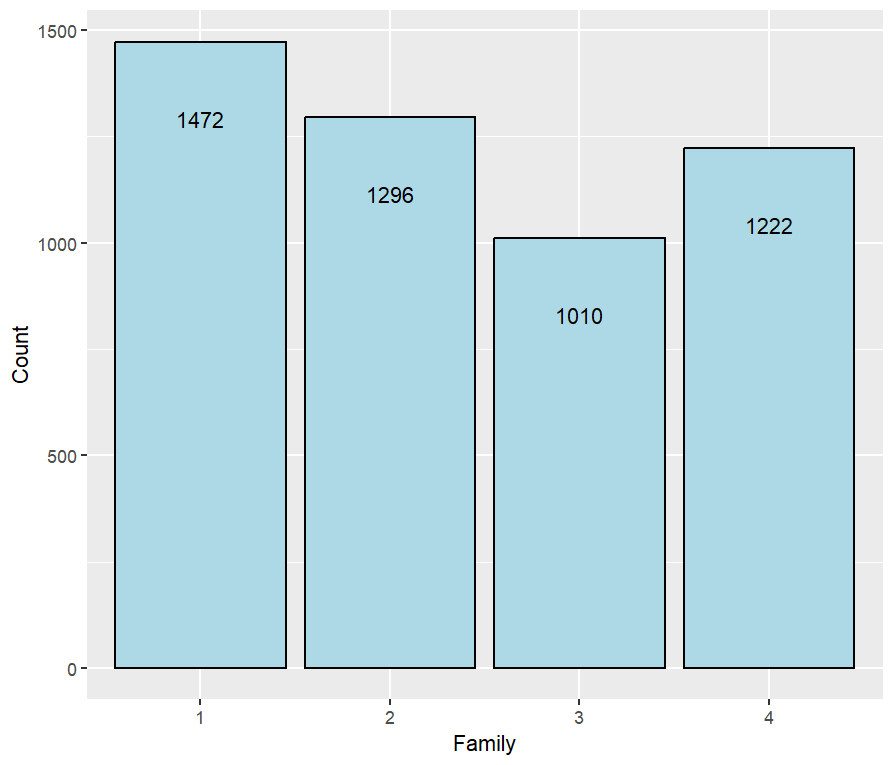
PROJECT.R (80/20)











A graph of a graph with blue lines

Description automatically generated A graph showing a graph of experience

Description automatically generated A graph with blue lines and a line

Description automatically generated A graph with a blue line

Description automatically generated A graph with a line graph

Description automatically generated

A screenshot of a computer code

Description automatically generated

> logLik(logit.train)

'log Lik.' -451.3085 (df=15)

> with(logit.train, pchisq(null.deviance - deviance, df.null - df.residual, lower.tail = FALSE))

[1] 0

> with(logit.none, pchisq(null.deviance - deviance, df.null - df.residual, lower.tail = FALSE))

[1] 1

> summary(logit.none)

Call:

glm(formula = Personal.Loan ~ 1, family = "binomial", data = data)

Coefficients:

Estimate Std. Error z value Pr(>|z|)

(Intercept) -2.24248 0.04801 -46.71 <2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 3162 on 4999 degrees of freedom

Residual deviance: 3162 on 4999 degrees of freedom

AIC: 3164

Number of Fisher Scoring iterations: 5

> cftable

predictions.binary 0 1

0 899 33

1 7 61

> accuracy

[1] 0.96

> sensitivity

[1] 0.9922737

> specificity

[1] 0.6489362

> ppv

[1] 0.9645923

> npv

[1] 0.8970588

> cft.train

pred.train.binary 0 1

0 3574 120

1 40 266

> accuracy

[1] 0.96

> sensitivity

[1] 0.9889319

> specificity

[1] 0.6891192

> ppv

[1] 0.9675149

> npv

[1] 0.869281

A graph with numbers and lines

Description automatically generated A graph of a logistic regression

Description automatically generated

A graph showing the number of logistic regression

Description automatically generated A graph of a logistic regression

Description automatically generated

A graph with a blue line

Description automatically generated A graph with a blue line and orange lines

Description automatically generated A bar graph with numbers and text

Description automatically generated

A bar graph with numbers and text

Description automatically generated A graph with numbers and lines

Description automatically generated A graph with numbers and lines

Description automatically generated A graph with a blue line

Description automatically generated A graph of a logistic regression

Description automatically generated A graph with numbers and lines

Description automatically generated A graph of a logistic regression

Description automatically generated A graph showing a logistic regression for classification

Description automatically generated A graph of a logistic regression

Description automatically generated A graph with a blue line

Description automatically generated A graph of a logistic regression

Description automatically generated > roc(train$Personal.Loan, train.prob)

Setting levels: control = 0, case = 1

Setting direction: controls < cases

Call:

roc.default(response = train$Personal.Loan, predictor = train.prob)

Data: train.prob in 3614 controls (train$Personal.Loan 0) < 386 cases (train$Personal.Loan 1).

Area under the curve: 0.9636

> test.roc = roc(test$Personal.Loan, test.prob)

Setting levels: control = 0, case = 1

Setting direction: controls < cases

A graph with a line

Description automatically generated

> test.roc

Call:

roc.default(response = test$Personal.Loan, predictor = test.prob)

Data: test.prob in 906 controls (test$Personal.Loan 0) < 94 cases (test$Personal.Loan 1).

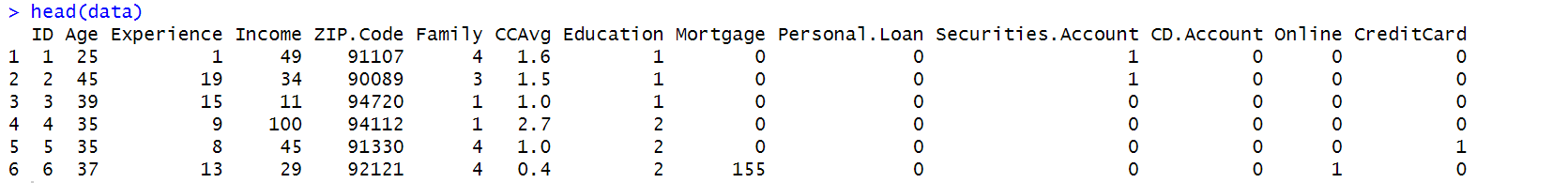
Area under the curve: 0.9631

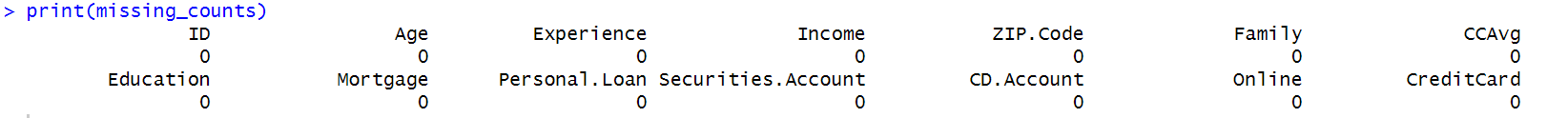
A graph with a line

Description automatically generated A graph with a line

Description automatically generated

Fifty\_fifty.R (50/50)





> summary(logit.train)

Call:

glm(formula = Personal.Loan ~ Age + Experience + Income + Family +

CCAvg + Education + Mortgage + Securities.Account + CD.Account +

Online + CreditCard, family = "binomial", data = train)

Coefficients:

Estimate Std. Error z value Pr(>|z|)

(Intercept) -1.338e+01 2.793e+00 -4.791 1.66e-06 \*\*\*

Age -2.467e-02 1.034e-01 -0.238 0.81152

Experience 3.384e-02 1.030e-01 0.328 0.74255

Income 6.856e-02 5.010e-03 13.683 < 2e-16 \*\*\*

Family2 -1.311e-01 3.511e-01 -0.373 0.70890

Family3 2.058e+00 3.762e-01 5.470 4.50e-08 \*\*\*

Family4 1.511e+00 3.535e-01 4.276 1.90e-05 \*\*\*

CCAvg 1.719e-01 7.144e-02 2.406 0.01614 \*

Education2 4.556e+00 4.439e-01 10.264 < 2e-16 \*\*\*

Education3 4.563e+00 4.378e-01 10.422 < 2e-16 \*\*\*

Mortgage 6.474e-04 8.700e-04 0.744 0.45679

Securities.Account -9.145e-01 4.948e-01 -1.848 0.06458 .

CD.Account 3.362e+00 5.566e-01 6.040 1.54e-09 \*\*\*

Online -5.547e-01 2.500e-01 -2.219 0.02652 \*

CreditCard -9.786e-01 3.222e-01 -3.037 0.00239 \*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 1526.53 on 2499 degrees of freedom

Residual deviance: 507.82 on 2485 degrees of freedom

AIC: 537.82

Number of Fisher Scoring iterations: 8

> logLik(logit.train)

'log Lik.' -253.9103 (df=15)

> with(logit.train, pchisq(null.deviance - deviance, df.null - df.residual, lower.tail = FALSE))

[1] 1.515657e-208

> with(logit.none, pchisq(null.deviance - deviance, df.null - df.residual, lower.tail = FALSE))

[1] 1

> summary(logit.none)

Call:

glm(formula = Personal.Loan ~ 1, family = "binomial", data = data)

Coefficients:

Estimate Std. Error z value Pr(>|z|)

(Intercept) -2.24248 0.04801 -46.71 <2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 3162 on 4999 degrees of freedom

Residual deviance: 3162 on 4999 degrees of freedom

AIC: 3164

Number of Fisher Scoring iterations: 5

> cftable

predictions.binary 0 1

0 2224 82

1 24 170

>accuracy

[1] 0.9576

> sensitivity

[1] 0.9893238

> specificity

[1] 0.6746032

> ppv

[1] 0.9644406

> npv

[1] 0.8762887

> cft.train

pred.train.binary 0 1

0 2250 66

1 22 162

> accuracy

[1] 0.9648

> sensitivity

[1] 0.9903169

> specificity

[1] 0.7105263

> ppv

[1] 0.9715026

> npv

[1] 0.8804348

A graph of a logistic regression

Description automatically generated A graph of a logistic regression

Description automatically generated A graph of a logistic regression

Description automatically generated A graph of a logistic regression

Description automatically generated A graph of a logistic regression

Description automatically generated A graph of a logistic regression

Description automatically generated A bar graph with numbers and a bar code

Description automatically generated A graph with numbers and a bar code

Description automatically generated A graph with numbers and lines

Description automatically generated A graph of logistic regression

Description automatically generated A graph of a logistic regression

Description automatically generated A graph of a logistic regression

Description automatically generated A graph with numbers and symbols

Description automatically generated A graph with numbers and symbols

Description automatically generated A graph with numbers and symbols

Description automatically generated A graph with numbers and symbols

Description automatically generated A graph with blue line

Description automatically generated A graph of a logistic regression

Description automatically generated

> roc(train$Personal.Loan, train.prob)

Setting levels: control = 0, case = 1

Setting direction: controls < cases

Call:

roc.default(response = train$Personal.Loan, predictor = train.prob)

Data: train.prob in 2272 controls (train$Personal.Loan 0) < 228 cases (train$Personal.Loan 1).

Area under the curve: 0.9676

A graph with a line

Description automatically generated with medium confidence > test.roc

Call:

roc.default(response = test$Personal.Loan, predictor = test.prob)

Data: test.prob in 2248 controls (test$Personal.Loan 0) < 252 cases (test$Personal.Loan 1).

Area under the curve: 0.956

A graph of a curve

Description automatically generated A graph of a normalized curve

Description automatically generated with medium confidence

**Decision Tree**

50-50.R

A table with numbers and numbers

Description automatically generated

A screenshot of a computer screen

Description automatically generated

A diagram of a family tree

Description automatically generated

A black and white text with numbers

Description automatically generated with medium confidence

A number of numbers on a white background

Description automatically generated

A graph with a line and a number of points

Description automatically generated with medium confidence

A graph of a tree

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A black and blue numbers

Description automatically generated

A computer code with blue text

Description automatically generated

80-20.R

A screenshot of a computer

Description automatically generated

A screenshot of a computer screen

Description automatically generated

A diagram of a family tree

Description automatically generated

A number of numbers on a white background

Description automatically generated

A number of numbers on a white background

Description automatically generated

A graph of a number of lines

Description automatically generated with medium confidence

A graph with a line and numbers

Description automatically generated

A close-up of numbers

Description automatically generated

A screenshot of a computer

Description automatically generated

A close-up of a number

Description automatically generated

A computer code with blue text

Description automatically generated