

IBD_Analysis

David Hu

2022-11-24

IBD Analysis

```
#Load in data  
data = read.csv("~/Documents/nsqip_ibd_ic/data/comb5_data.csv", header = TRUE)  
data <- replace(data, data == "", NA)  
#data$ileostomy.formation <- ifelse(test = data$ileostomy.formation  
=="Ileostomy", yes = 1, no = 0)
```

Replaced all missing values with "NA" for imputing values later.

```
#convert all character variables to factor variables  
data[sapply(data, is.character)] <-  
lapply(data[sapply(data, is.character)], as.factor)  
#various methods to check for possible columns to drop  
#sapply(lapply(data, unique), length) #display number of levels  
#sapply(data[, sapply(data, class) == 'factor'], levels) #levels of each  
variables  
#mapply(table, data) #display levels and count for each level
```

Convert the character variables to factor variables to represent categorical data.

```
#colSums(data[, sapply(data, class) == 'factor'] == "")/nrow(data) old method  
#sapply(data[, sapply(data, class) == 'factor'], levels)  
#colSums(is.na(data))/nrow(data)
```

The "sapply(data[, sapply(data, class) == 'factor'], levels)" line is used to display the amount of levels for a categorical variable. We can find which what variables have one level so we can drop them because they will not be useful when building our model.

The "colSums(is.na(data))/nrow(data)" displays the proportion of missing values for each dataset. If a categorical variable has 25% or more of its values missing, then we drop the variable.

```
#from Kevin's email  
data = subset(data, select =  
c(X..of.postop.superficial.incisional.ssi, X..of.postop.deep.incisional.ssi, X..  
.of.postop.organ.space.ssi, X..of.postop.wound.disruption, X..of.postop.pneumon  
ia, X..of.postop.unplanned.intubation, X..of.postop.pulmonary.embolism, X..of.po  
stop.on.ventilator...48.hours, X..of.postop.progressive.renal.insufficiency, X..  
.of.postop.acute.renal.failure, X..of.postop.cva, X..of.postop.uti, X..of.postop  
.cardiac.arrest.requiring.cpr, X..of.postop.myocardial.infarction, X..of.postop  
.transfusion.intraop..postop..72h.of.surgery.start.time., X..of.postop.vein.th
```

```

rombosis.requiring.therapy,X..of.postop.c..diff,X..of.postop.sepsis,X..of.postop.septic.shock,X..of.postop.other.occurrences,surgical.wound.s..closure,hospital.length.of.stay,hospital.discharge.destination,postop.death.w.in.30.days.of.procedure,postoperative.death...30.days.of.procedure,X..of.readmissions.w.in.30.days,X..of.unplanned.readmissions,first.unplanned.return.cpt,first.unplanned.return.to.or,colectomy.postop.anastomotic.leak,colectomy.prolonged.postoperative.npo.or.ngt.use))

```

```

#drop variables that have >=25% missing values and those with only one levels
data = subset(data, select = -
c(malnourishment,colectomy.malignancy..pathologic.m.stage,colectomy.malignancy..pathologic.n.stage,colectomy.malignancy..pathologic.t.stage,
preop.transfusions..rbc.w.in.72.hrs.prior.to.surgery.start.time.,
congestive.heart.failure.w.in.30.days.prior.to.surgery,ascites.w.in.30.days.prior.to.surgery,
colectomy.primary.indication.for.surgery,
colectomy.non.emergent.indication.icd10.code,
currently.requiring.or.on.dialysis,
acute.renal.failure,
case.number
))

```

We've identified the following variables to drop based on whether $\leq 25\%$ of its values were missing and/or had only one level. While case.number didn't fit the criteria, it was dropped because it doesn't provide relevant information since it is the surgery case number.

```

#imputing the missing values
#install.packages("missForest")
library(missForest)
data.imp <- missForest(data)$ximp

```

Impute missing values using random forest nonparametric modeling.

```

#stepwise logistic regression
library(MASS)
model <- glm(ileostomy.formation~.,data=data.imp,family="binomial")
step.model <- stepAIC(model,trace = FALSE,direction = "backward")
summary(model)

##
## Call:
## glm(formula = ileostomy.formation ~ ., family = "binomial", data =
data.imp)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.7703   0.1743   0.3234   0.5175   2.0645
##
## Coefficients: (1 not defined because of singularities)
##

```

Estimate

## (Intercept)	9.410e+01
## age.at.time.of.surgery	2.654e-02
## genderMale	-6.233e-02
## raceBlack or African American	-1.459e+01
## raceNative Hawaiian or Other Pacific Islander	3.289e-01
## raceSome Other Race	-1.733e+01
## raceUnknown/Not Reported	-1.369e+01
## raceWhite	-1.511e+01
## hispanic.ethnicityYes	-5.710e-01
## cpt.code	-1.525e-03
## in.out.patient.statusOutpatient	1.392e+01
## elective.surgeryYes	-2.568e-01
## origin.statusFrom acute care hospital inpatient	5.400e-01
## origin.statusHome/Permanent residence	-3.460e-01
## origin.statusNot transferred (admitted from home)	2.456e-01
## origin.statusNursing home - Chronic care - Intermediate care	-1.644e+01
## origin.statusOther facility	1.519e+01
## origin.statusOutside emergency department	-6.563e-01
## origin.statusTransfer from other	-1.782e+00
## height	2.464e-02
## weight	-9.990e-03
## bmi	6.177e-02
## diabetes.mellitusNo	-3.037e+00
## diabetes.mellitusNon-Insulin	-2.747e+00
## current.smoker.within.1.yearYes	-2.304e-01
## dyspneaNo	1.600e-01
## functional.heath.statusPartially Dependent	-1.614e+00
## functional.heath.statusTotally Dependent	-1.711e+01
## ventilator.dependentYes	NA
## history.of.severe.copdYes	4.595e-01
## hypertension.requiring.medicationYes	-1.000e-02
## disseminated.cancerYes	5.046e-02
## open.wound.with.or.without.infectionYes	6.059e-02
## steroid.immunosuppressant.use.for.chronic.conditionYes	-4.254e-01
## bleeding.disorderYes	-8.470e-01
## sepsis..sirs.sepsis.septic.shock...48h.Sepsis	3.838e-01
## sepsis..sirs.sepsis.septic.shock...48h.SIRS	-5.228e-01
## serum.sodium	1.972e-02
## bun	-4.032e-03
## serum.creatinine	-3.923e-01
## albumin	7.792e-02
## wbc	-3.429e-02
## hematocrit	2.146e-02
## platelet.count	-1.466e-03
## emergency.caseYes	-1.114e+00
## wound.classificationClean/Contaminated	1.659e+00
## wound.classificationContaminated	7.998e-01
## wound.classificationDirty/Infected	1.344e+00
## asa.classificationASA 2 - Mild Disturb	5.113e-01
## asa.classificationASA 3 - Severe Disturb	9.828e-03

## asa.classificationASA 4 - Life Threat	-6.888e-01
## X..of.other.procedures	-4.696e-01
## X..of.concurrent.procedures	2.808e-01
## duration.of.surgical.procedure..in.minutes.	-5.028e-03
## colectomy.steroid.immunosuppressant.for.inflammatory.bowelYes	3.406e-01
## colectomy.mechanical.bowel.prepYes	6.744e-01
## colectomy.oral.antibiotic.prepYes	2.758e-01
## colectomy.chemotherapy.w.in.90.daysYes	1.368e+00
## colectomy.operative.approachMIS, unplanned conversion to open	-1.481e-01
## colectomy.operative.approachOpen	-7.482e-01
## colectomy.number.of.nodes.evaluated	-1.601e-03
## biologic.agent.60.days.prior.to.surgeryNo	7.151e-01
## biologic.agent.60.days.prior.to.surgeryYes	6.867e-01
## immune.modulation.60.days.prior.to.surgeryNo	-1.518e+01
## immune.modulation.60.days.prior.to.surgeryYes	-1.507e+01
## colonic.dysplasiaHigh Grade Dysplasia	1.048e+01
## colonic.dysplasiaLow grade Dysplasia	-2.727e+00
## colonic.dysplasiaNo	1.339e+01
## colonic.dysplasiaNo Dysplasia	-1.638e+00
## colonic.dysplasiaNot applicable	-1.299e+00
## colonic.dysplasiaUnable to determine	-1.346e+00
##	Std. Error
## (Intercept)	1.351e+03
## age.at.time.of.surgery	7.364e-03
## genderMale	2.507e-01
## raceBlack or African American	5.436e+02
## raceNative Hawaiian or Other Pacific Islander	2.460e+03
## raceSome Other Race	5.436e+02
## raceUnknown/Not Reported	5.436e+02
## raceWhite	5.436e+02
## hispanic.ethnicityYes	4.710e-01
## cpt.code	8.802e-03
## in.out.patient.statusOutpatient	1.025e+03
## elective.surgeryYes	2.544e-01
## origin.statusFrom acute care hospital inpatient	7.857e-01
## origin.statusHome/Permanent residence	6.981e-01
## origin.statusNot transferred (admitted from home)	6.890e-01
## origin.statusNursing home - Chronic care - Intermediate care	2.400e+03
## origin.statusOther facility	1.661e+03
## origin.statusOutside emergency department	8.538e-01
## origin.statusTransfer from other	1.786e+00
## height	1.158e-02
## weight	2.152e-03
## bmi	1.898e-02
## diabetes.mellitusNo	2.046e+00
## diabetes.mellitusNon-Insulin	2.229e+00
## current.smoker.within.1.yearYes	2.189e-01
## dyspneaNo	6.055e-01
## functional.heath.statusPartially Dependent	1.448e+00
## functional.heath.statusTotally Dependent	2.400e+03

## ventilator.dependentYes	NA
## history.of.severe.copdYes	9.691e-01
## hypertension.requiring.medicationYes	2.889e-01
## disseminated.cancerYes	1.424e+00
## open.wound.with.or.without.infectionYes	4.698e-01
## steroid.immunosuppressant.use.for.chronic.conditionYes	5.915e-01
## bleeding.disorderYes	6.134e-01
## sepsis..sirs.sepsis.septic.shock...48h.Sepsis	4.761e-01
## sepsis..sirs.sepsis.septic.shock...48h.SIRS	4.577e-01
## serum.sodium	3.775e-02
## bun	1.704e-02
## serum.creatinine	4.385e-01
## albumin	1.977e-01
## wbc	2.983e-02
## hematocrit	2.089e-02
## platelet.count	9.043e-04
## emergency.caseYes	4.786e-01
## wound.classificationClean/Contaminated	1.384e+00
## wound.classificationContaminated	1.387e+00
## wound.classificationDirty/Infected	1.391e+00
## asa.classificationASA 2 - Mild Disturb	8.619e-01
## asa.classificationASA 3 - Severe Disturb	8.620e-01
## asa.classificationASA 4 - Life Threat	1.078e+00
## X..of.other.procedures	7.229e-02
## X..of.concurrent.procedures	1.220e-01
## duration.of.surgical.procedure..in.minutes.	1.715e-03
## colectomy.steroid.immunosuppressant.for.inflammatory.bowelYes	6.121e-01
## colectomy.mechanical.bowel.prepYes	2.148e-01
## colectomy.oral.antibiotic.prepYes	2.400e-01
## colectomy.chemotherapy.w.in.90.daysYes	1.761e+00
## colectomy.operative.approachMIS, unplanned conversion to open	4.253e-01
## colectomy.operative.approachOpen	4.176e-01
## colectomy.number.of.nodes.evaluated	2.745e-02
## biologic.agent.60.days.prior.to.surgeryNo	6.477e-01
## biologic.agent.60.days.prior.to.surgeryYes	6.361e-01
## immune.modulation.60.days.prior.to.surgeryNo	1.174e+03
## immune.modulation.60.days.prior.to.surgeryYes	1.174e+03
## colonic.dysplasiaHigh Grade Dysplasia	9.065e+02
## colonic.dysplasiaLow grade Dysplasia	1.866e+00
## colonic.dysplasiaNo	2.400e+03
## colonic.dysplasiaNo Dysplasia	1.612e+00
## colonic.dysplasiaNot applicable	1.816e+00
## colonic.dysplasiaUnable to determine	1.673e+00
##	z value
Pr(> z)	
## (Intercept)	0.070
0.944488	
## age.at.time.of.surgery	3.603
0.000314	
## genderMale	-0.249

0.803671	
## raceBlack or African American	-0.027
0.978594	
## raceNative Hawaiian or Other Pacific Islander	0.000
0.999893	
## raceSome Other Race	-0.032
0.974563	
## raceUnknown/Not Reported	-0.025
0.979911	
## raceWhite	-0.028
0.977823	
## hispanic.ethnicityYes	-1.212
0.225409	
## cpt.code	-0.173
0.862492	
## in.out.patient.statusOutpatient	0.014
0.989167	
## elective.surgeryYes	-1.009
0.312847	
## origin.statusFrom acute care hospital inpatient	0.687
0.491857	
## origin.statusHome/Permanent residence	-0.496
0.620110	
## origin.statusNot transferred (admitted from home)	0.356
0.721476	
## origin.statusNursing home - Chronic care - Intermediate care	-0.007
0.994533	
## origin.statusOther facility	0.009
0.992703	
## origin.statusOutside emergency department	-0.769
0.442094	
## origin.statusTransfer from other	-0.998
0.318342	
## height	2.128
0.033302	
## weight	-4.642
3.45e-06	
## bmi	3.254
0.001137	
## diabetes.mellitusNo	-1.484
0.137732	
## diabetes.mellitusNon-Insulin	-1.232
0.217874	
## current.smoker.within.1.yearYes	-1.053
0.292425	
## dyspneaNo	0.264
0.791545	
## functional.heath.statusPartially Dependent	-1.115
0.264871	
## functional.heath.statusTotally Dependent	-0.007

0.994309	
## ventilator.dependentYes	NA
NA	
## history.of.severe.copdYes	0.474
0.635385	
## hypertension.requiring.medicationYes	-0.035
0.972386	
## disseminated.cancerYes	0.035
0.971738	
## open.wound.with.or.without.infectionYes	0.129
0.897386	
## steroid.immunosuppressant.use.for.chronic.conditionYes	-0.719
0.472026	
## bleeding.disorderYes	-1.381
0.167315	
## sepsis..sirs.sepsis.septic.shock...48h.Sepsis	0.806
0.420190	
## sepsis..sirs.sepsis.septic.shock...48h.SIRS	-1.142
0.253278	
## serum.sodium	0.522
0.601440	
## bun	-0.237
0.812933	
## serum.creatinine	-0.895
0.370907	
## albumin	0.394
0.693480	
## wbc	-1.149
0.250428	
## hematocrit	1.027
0.304413	
## platelet.count	-1.621
0.104969	
## emergency.caseYes	-2.327
0.019957	
## wound.classificationClean/Contaminated	1.199
0.230635	
## wound.classificationContaminated	0.577
0.564091	
## wound.classificationDirty/Infected	0.966
0.334023	
## asa.classificationASA 2 - Mild Disturb	0.593
0.552976	
## asa.classificationASA 3 - Severe Disturb	0.011
0.990903	
## asa.classificationASA 4 - Life Threat	-0.639
0.522947	
## X..of.other.procedures	-6.497
8.22e-11	
## X..of.concurrent.procedures	2.301

0.021384	
## duration.of.surgical.procedure..in.minutes.	-2.932
0.003373	
## colectomy.steroid.immunosuppressant.for.inflammatory.bowelYes	0.556
0.577920	
## colectomy.mechanical.bowel.prepYes	3.139
0.001695	
## colectomy.oral.antibiotic.prepYes	1.149
0.250662	
## colectomy.chemotherapy.w.in.90.daysYes	0.777
0.437041	
## colectomy.operative.approachMIS, unplanned conversion to open	-0.348
0.727715	
## colectomy.operative.approachOpen	-1.792
0.073183	
## colectomy.number.of.nodes.evaluated	-0.058
0.953493	
## biologic.agent.60.days.prior.to.surgeryNo	1.104
0.269556	
## biologic.agent.60.days.prior.to.surgeryYes	1.079
0.280376	
## immune.modulation.60.days.prior.to.surgeryNo	-0.013
0.989690	
## immune.modulation.60.days.prior.to.surgeryYes	-0.013
0.989761	
## colonic.dysplasiaHigh Grade Dysplasia	0.012
0.990772	
## colonic.dysplasiaLow grade Dysplasia	-1.461
0.143929	
## colonic.dysplasiaNo	0.006
0.995548	
## colonic.dysplasiaNo Dysplasia	-1.016
0.309662	
## colonic.dysplasiaNot applicable	-0.716
0.474194	
## colonic.dysplasiaUnable to determine	-0.805
0.421074	
##	
## (Intercept)	
## age.at.time.of.surgery	***
## genderMale	
## raceBlack or African American	
## raceNative Hawaiian or Other Pacific Islander	
## raceSome Other Race	
## raceUnknown/Not Reported	
## raceWhite	
## hispanic.ethnicityYes	
## cpt.code	
## in.out.patient.statusOutpatient	
## elective.surgeryYes	


```

## origin.statusFrom acute care hospital inpatient
## origin.statusHome/Permanent residence
## origin.statusNot transferred (admitted from home)
## origin.statusNursing home - Chronic care - Intermediate care
## origin.statusOther facility
## origin.statusOutside emergency department
## origin.statusTransfer from other
## height *
## weight ***
## bmi **
## diabetes.mellitusNo
## diabetes.mellitusNon-Insulin
## current.smoker.within.1.yearYes
## dyspneaNo
## functional.heath.statusPartially Dependent
## functional.heath.statusTotally Dependent
## ventilator.dependentYes
## history.of.severe.copdYes
## hypertension.requiring.medicationYes
## disseminated.cancerYes
## open.wound.with.or.without.infectionYes
## steroid.immunosuppressant.use.for.chronic.conditionYes
## bleeding.disorderYes
## sepsis..sirs.sepsis.septic.shock...48h.Sepsis
## sepsis..sirs.sepsis.septic.shock...48h.SIRS
## serum.sodium
## bun
## serum.creatinine
## albumin
## wbc
## hematocrit
## platelet.count
## emergency.caseYes *
## wound.classificationClean/Contaminated
## wound.classificationContaminated
## wound.classificationDirty/Infected
## asa.classificationASA 2 - Mild Disturb
## asa.classificationASA 3 - Severe Disturb
## asa.classificationASA 4 - Life Threat
## X..of.other.procedures ***
## X..of.concurrent.procedures *
## duration.of.surgical.procedure..in.minutes. **
## colectomy.steroid.immunosuppressant.for.inflammatory.bowelYes
## colectomy.mechanical.bowel.prepYes **
## colectomy.oral.antibiotic.prepYes
## colectomy.chemotherapy.w.in.90.daysYes
## colectomy.operative.approachMIS, unplanned conversion to open
## colectomy.operative.approachOpen .
## colectomy.number.of.nodes.evaluated
## biologic.agent.60.days.prior.to.surgeryNo

```

```
## biologic.agent.60.days.prior.to.surgeryYes
## immune.modulation.60.days.prior.to.surgeryNo
## immune.modulation.60.days.prior.to.surgeryYes
## colonic.dysplasiaHigh Grade Dysplasia
## colonic.dysplasiaLow grade Dysplasia
## colonic.dysplasiaNo
## colonic.dysplasiaNo Dysplasia
## colonic.dysplasiaNot applicable
## colonic.dysplasiaUnable to determine
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 1385.6  on 1557  degrees of freedom
## Residual deviance: 1015.4  on 1488  degrees of freedom
## AIC: 1155.4
##
## Number of Fisher Scoring iterations: 15
```

Use step wise AIC logistic regression with backward selection to select the variables that are the most useful to help the explain the probability of an ileostomy after logistic regression (will be referred to as original logistic model) on all the variables throws a warning.

```
library(car)

## Loading required package: carData

#vif(model)
```

Unable to check multicollinearity using vif (Variance Inflation Factor) on the original logistic model. Throws warning "Error in vif.default(model) : there are aliased coefficients in the model".

```
probabilities <- predict(model, data.imp, type = "response")

## Warning in predict.lm(object, newdata, se.fit, scale = 1, type = if (type
== :
## prediction from a rank-deficient fit may be misleading

predicted.classes <- ifelse(probabilities > 0.50, "None", "Ileostomy")
# Prediction accuracy
observed.classes <- data.imp$ileostomy.formation
mean(predicted.classes == observed.classes)

## [1] 0.8677792
```

Although original logistic model throws a warning, the model does predict the correct outcome for 87% of the observations.

```
summary(step.model)
```

```
##
## Call:
## glm(formula = ileostomy.formation ~ age.at.time.of.surgery +
##      race + origin.status + height + weight + bmi + platelet.count +
##      emergency.case + wound.classification + asa.classification +
##      X..of.other.procedures + X..of.concurrent.procedures +
duration.of.surgical.procedure..in.minutes. +
##      colectomy.mechanical.bowel.prep + colectomy.operative.approach,
##      family = "binomial", data = data.imp)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q        Max
## -2.7718   0.1950   0.3335   0.5257   2.0192
##
## Coefficients:
##                                     Estimate
## (Intercept)                        1.155e+01
## age.at.time.of.surgery              2.435e-02
## raceBlack or African American      -1.450e+01
## raceNative Hawaiian or Other Pacific Islander  1.080e+00
## raceSome Other Race                -1.698e+01
## raceUnknown/Not Reported          -1.378e+01
## raceWhite                         -1.496e+01
## origin.statusFrom acute care hospital inpatient  4.065e-01
## origin.statusHome/Permanent residence          -1.717e-01
## origin.statusNot transferred (admitted from home)  3.897e-01
## origin.statusNursing home - Chronic care - Intermediate care -1.839e+01
## origin.statusOther facility              1.482e+01
## origin.statusOutside emergency department        -5.865e-01
## origin.statusTransfer from other          -1.866e+00
## height                               1.999e-02
## weight                              -9.527e-03
## bmi                                 5.659e-02
## platelet.count                     -2.073e-03
## emergency.caseYes                  -9.576e-01
## wound.classificationClean/Contaminated          1.877e+00
## wound.classificationContaminated              1.003e+00
## wound.classificationDirty/Infected             1.599e+00
## asa.classificationASA 2 - Mild Disturb          3.556e-01
## asa.classificationASA 3 - Severe Disturb        -1.536e-01
## asa.classificationASA 4 - Life Threat           -1.176e+00
## X..of.other.procedures                 -4.764e-01
## X..of.concurrent.procedures              2.804e-01
## duration.of.surgical.procedure..in.minutes.    -5.048e-03
## colectomy.mechanical.bowel.prepYes             7.869e-01
## colectomy.operative.approachMIS, unplanned conversion to open -8.733e-02
## colectomy.operative.approachOpen             -7.275e-01
##                                     Std. Error
## (Intercept)                        5.495e+02
## age.at.time.of.surgery              6.273e-03
```

## raceBlack or African American	5.495e+02
## raceNative Hawaiian or Other Pacific Islander	2.462e+03
## raceSome Other Race	5.495e+02
## raceUnknown/Not Reported	5.495e+02
## raceWhite	5.495e+02
## origin.statusFrom acute care hospital inpatient	7.484e-01
## origin.statusHome/Permanent residence	6.716e-01
## origin.statusNot transferred (admitted from home)	6.618e-01
## origin.statusNursing home - Chronic care - Intermediate care	2.400e+03
## origin.statusOther facility	1.681e+03
## origin.statusOutside emergency department	8.092e-01
## origin.statusTransfer from other	1.491e+00
## height	8.680e-03
## weight	1.988e-03
## bmi	1.763e-02
## platelet.count	7.750e-04
## emergency.caseYes	4.384e-01
## wound.classificationClean/Contaminated	1.409e+00
## wound.classificationContaminated	1.411e+00
## wound.classificationDirty/Infected	1.415e+00
## asa.classificationASA 2 - Mild Disturb	8.697e-01
## asa.classificationASA 3 - Severe Disturb	8.693e-01
## asa.classificationASA 4 - Life Threat	1.025e+00
## X..of.other.procedures	6.967e-02
## X..of.concurrent.procedures	1.136e-01
## duration.of.surgical.procedure..in.minutes.	1.638e-03
## colectomy.mechanical.bowel.prepYes	1.722e-01
## colectomy.operative.approachMIS, unplanned conversion to open	2.892e-01
## colectomy.operative.approachOpen	1.939e-01
##	z value
Pr(> z)	
## (Intercept)	0.021
0.983228	
## age.at.time.of.surgery	3.881
0.000104	
## raceBlack or African American	-0.026
0.978952	
## raceNative Hawaiian or Other Pacific Islander	0.000
0.999650	
## raceSome Other Race	-0.031
0.975345	
## raceUnknown/Not Reported	-0.025
0.980000	
## raceWhite	-0.027
0.978285	
## origin.statusFrom acute care hospital inpatient	0.543
0.586997	
## origin.statusHome/Permanent residence	-0.256
0.798250	
## origin.statusNot transferred (admitted from home)	0.589

0.555989	
## origin.statusNursing home - Chronic care - Intermediate care	-0.008
0.993886	
## origin.statusOther facility	0.009
0.992965	
## origin.statusOutside emergency department	-0.725
0.468538	
## origin.statusTransfer from other	-1.252
0.210658	
## height	2.303
0.021261	
## weight	-4.792
1.65e-06	
## bmi	3.210
0.001327	
## platelet.count	-2.675
0.007468	
## emergency.caseYes	-2.185
0.028921	
## wound.classificationClean/Contaminated	1.332
0.182838	
## wound.classificationContaminated	0.711
0.477211	
## wound.classificationDirty/Infected	1.130
0.258330	
## asa.classificationASA 2 - Mild Disturb	0.409
0.682644	
## asa.classificationASA 3 - Severe Disturb	-0.177
0.859742	
## asa.classificationASA 4 - Life Threat	-1.147
0.251494	
## X..of.other.procedures	-6.838
8.05e-12	
## X..of.concurrent.procedures	2.468
0.013573	
## duration.of.surgical.procedure..in.minutes.	-3.082
0.002054	
## colectomy.mechanical.bowel.prepYes	4.571
4.85e-06	
## colectomy.operative.approachMIS, unplanned conversion to open	-0.302
0.762657	
## colectomy.operative.approachOpen	-3.751
0.000176	
##	
## (Intercept)	
## age.at.time.of.surgery	***
## raceBlack or African American	
## raceNative Hawaiian or Other Pacific Islander	
## raceSome Other Race	
## raceUnknown/Not Reported	

```

## raceWhite
## origin.statusFrom acute care hospital inpatient
## origin.statusHome/Permanent residence
## origin.statusNot transferred (admitted from home)
## origin.statusNursing home - Chronic care - Intermediate care
## origin.statusOther facility
## origin.statusOutside emergency department
## origin.statusTransfer from other
## height *
## weight ***
## bmi **
## platelet.count **
## emergency.caseYes *
## wound.classificationClean/Contaminated
## wound.classificationContaminated
## wound.classificationDirty/Infected
## asa.classificationASA 2 - Mild Disturb
## asa.classificationASA 3 - Severe Disturb
## asa.classificationASA 4 - Life Threat
## X..of.other.procedures ***
## X..of.concurrent.procedures *
## duration.of.surgical.procedure..in.minutes. **
## colectomy.mechanical.bowel.prepYes ***
## colectomy.operative.approachMIS, unplanned conversion to open
## colectomy.operative.approachOpen ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 1385.6 on 1557 degrees of freedom
## Residual deviance: 1041.7 on 1527 degrees of freedom
## AIC: 1103.7
##
## Number of Fisher Scoring iterations: 15

```

Coefficients for the stepwise AIC logistic model.

```

vif(step.model)

```

	GVIF	Df	GVIF ^{1/(2*Df)}
## age.at.time.of.surgery	1.438144	1	1.199226
## race	1.216491	5	1.019790
## origin.status	1.358051	7	1.022101
## height	1.217736	1	1.103511
## weight	2.139169	1	1.462590
## bmi	1.780436	1	1.334330
## platelet.count	1.197883	1	1.094478
## emergency.case	1.104460	1	1.050933
## wound.classification	1.599576	3	1.081436

## asa.classification	1.299578	3	1.044641
## X..of.other.procedures	1.587241	1	1.259858
## X..of.concurrent.procedures	1.236299	1	1.111890
## duration.of.surgical.procedure..in.minutes.	1.481382	1	1.217121
## colectomy.mechanical.bowel.prep	1.107559	1	1.052406
## colectomy.operative.approach	1.372417	2	1.082360

Using VIF, no multicollinearity detected (VIF>=5)

```

probabilities <- predict(step.model, data.imp, type = "response")
predicted.classes <- ifelse(probabilities > 0.5, "None", "Ileostomy")
# Prediction accuracy
observed.classes <- data.imp$ileostomy.formation
mean(predicted.classes == observed.classes)

## [1] 0.8581515

```

The accuracy does go down by about one percent using the step wise AIC logistic model.

```

#the coefficients for each model ran in here
as.data.frame(coef(model))

##
coef(model)
## (Intercept)
94.096941295
## age.at.time.of.surgery
0.026536358
## genderMale
0.062325076
## raceBlack or African American
14.585781139
## raceNative Hawaiian or Other Pacific Islander
0.328868778
## raceSome Other Race
17.333141838
## raceUnknown/Not Reported
13.688590060
## raceWhite
15.111659006
## hispanic.ethnicityYes
0.571006695
## cpt.code
0.001524562
## in.out.patient.statusOutpatient
13.921559263
## elective.surgeryYes
0.256774211
## origin.statusFrom acute care hospital inpatient
0.540028912
## origin.statusHome/Permanent residence

```

0.346040184	
## origin.statusNot transferred (admitted from home)	
0.245617695	
## origin.statusNursing home - Chronic care - Intermediate care	-
16.442702885	
## origin.statusOther facility	
15.191925356	
## origin.statusOutside emergency department	-
0.656262189	
## origin.statusTransfer from other	-
1.782037433	
## height	
0.024639167	
## weight	-
0.009990095	
## bmi	
0.061767208	
## diabetes.mellitusNo	-
3.036859953	
## diabetes.mellitusNon-Insulin	-
2.746595725	
## current.smoker.within.1.yearYes	-
0.230443485	
## dyspneaNo	
0.160044062	
## functional.heath.statusPartially Dependent	-
1.613991833	
## functional.heath.statusTotally Dependent	-
17.114021596	
## ventilator.dependentYes	
NA	
## history.of.severe.copdYes	
0.459494983	
## hypertension.requiring.medicationYes	-
0.010000818	
## disseminated.cancerYes	
0.050462259	
## open.wound.with.or.without.infectionYes	
0.060585738	
## steroid.immunosuppressant.use.for.chronic.conditionYes	-
0.425375478	
## bleeding.disorderYes	-
0.846991606	
## sepsis..sirs.sepsis.septic.shock...48h.Sepsis	
0.383760577	
## sepsis..sirs.sepsis.septic.shock...48h.SIRS	-
0.522839633	
## serum.sodium	
0.019719307	
## bun	-

0.004032410	
## serum.creatinine	-
0.392346376	
## albumin	
0.077922982	
## wbc	-
0.034288873	
## hematocrit	
0.021459253	
## platelet.count	-
0.001466118	
## emergency.caseYes	-
1.113752170	
## wound.classificationClean/Contaminated	
1.658943422	
## wound.classificationContaminated	
0.799762325	
## wound.classificationDirty/Infected	
1.344093350	
## asa.classificationASA 2 - Mild Disturb	
0.511342930	
## asa.classificationASA 3 - Severe Disturb	
0.009828161	
## asa.classificationASA 4 - Life Threat	-
0.688794665	
## X..of.other.procedures	-
0.469605271	
## X..of.concurrent.procedures	
0.280803022	
## duration.of.surgical.procedure..in.minutes.	-
0.005028507	
## colectomy.steroid.immunosuppressant.for.inflammatory.bowelYes	
0.340592230	
## colectomy.mechanical.bowel.prepYes	
0.674355240	
## colectomy.oral.antibiotic.prepYes	
0.275750477	
## colectomy.chemotherapy.w.in.90.daysYes	
1.368493788	
## colectomy.operative.approachMIS, unplanned conversion to open	-
0.148065845	
## colectomy.operative.approachOpen	-
0.748188174	
## colectomy.number.of.nodes.evaluated	-
0.001601058	
## biologic.agent.60.days.prior.to.surgeryNo	
0.715119182	
## biologic.agent.60.days.prior.to.surgeryYes	
0.686663087	
## immune.modulation.60.days.prior.to.surgeryNo	-

```

15.175531929
## immune.modulation.60.days.prior.to.surgeryYes -
15.071534631
## colonic.dysplasiaHigh Grade Dysplasia
10.484949267
## colonic.dysplasiaLow grade Dysplasia -
2.726717267
## colonic.dysplasiaNo
13.388638827
## colonic.dysplasiaNo Dysplasia -
1.637504499
## colonic.dysplasiaNot applicable -
1.299328859
## colonic.dysplasiaUnable to determine -
1.345907823

as.data.frame(coef(step.model))

##
coef(step.model)
## (Intercept)
11.552326930
## age.at.time.of.surgery
0.024346884
## raceBlack or African American -
14.498146041
## raceNative Hawaiian or Other Pacific Islander
1.080258065
## raceSome Other Race -
16.983440714
## raceUnknown/Not Reported -
13.776062927
## raceWhite -
14.957566520
## origin.statusFrom acute care hospital inpatient
0.406508817
## origin.statusHome/Permanent residence -
0.171662592
## origin.statusNot transferred (admitted from home)
0.389695180
## origin.statusNursing home - Chronic care - Intermediate care -
18.386990104
## origin.statusOther facility
14.820835511
## origin.statusOutside emergency department -
0.586535624
## origin.statusTransfer from other -
1.866010048
## height
0.019992007

```

```

## weight -
0.009526571
## bmi
0.056592578
## platelet.count -
0.002073414
## emergency.caseYes -
0.957604475
## wound.classificationClean/Contaminated
1.877248958
## wound.classificationContaminated
1.003064675
## wound.classificationDirty/Infected
1.599273795
## asa.classificationASA 2 - Mild Disturb
0.355566323
## asa.classificationASA 3 - Severe Disturb -
0.153599277
## asa.classificationASA 4 - Life Threat -
1.175802418
## X..of.other.procedures -
0.476360409
## X..of.concurrent.procedures
0.280394423
## duration.of.surgical.procedure..in.minutes. -
0.005047635
## colectomy.mechanical.bowel.prepYes
0.786944090
## colectomy.operative.approachMIS, unplanned conversion to open -
0.087325122
## colectomy.operative.approachOpen -
0.727497648

#other stuff to experiment with it
#Levels(data.imp$race) <- c("Other", "Other", "Black or African
American", "Other", "Other", "Unknown/Not Reported", "Other")
#Levels(data.imp$origin.status) <- c("Other", "From acute care hospital
inpatient", "Other", "Not transferred (admitted from
home)", "Other", "Other", "Other", "Other")
#Levels(data.imp$wound.classification) <-
c("Other", "Other", "Other", "Contaminated", "Other")

```

Bottom was my attempt at building a model if we wanted to some sort of modeling for future predictions but not needed for this context. Tried to combine levels to make all the variables in model significant.