ProjectNew

Fatemeh Nosrat

11/8/2018

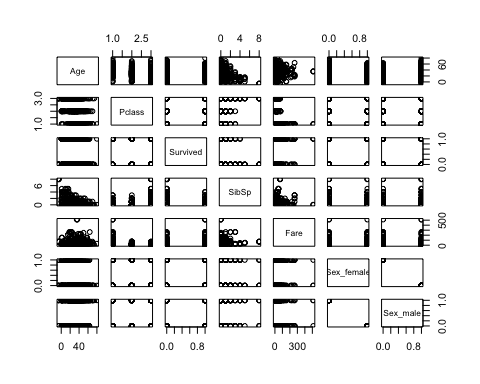
library(dummies)

## dummies-1.5.6 provided by Decision Patterns

Mode <- function(x) {  
 ux <- unique(x)  
 ux[which.max(tabulate(match(x, ux)))]  
}  
  
train <- read.csv(file="/Users/fatemehnosrat/Downloads/train.csv",header = TRUE, na.strings=c("",".","NA"))  
train$Embarked[is.na(train$Embarked)]<-Mode(train$Embarked)  
#train$Sex <- as.integer(train$Sex)  
View(train)  
#train$Embarked <- as.integer(train$Embarked)  
test <- read.csv(file="/Users/fatemehnosrat/Downloads/test.csv",header = TRUE, na.strings=c("",".","NA"))  
#test$Sex <- as.integer(test$Sex)  
#test$Embarked <- as.integer(test$Embarked)  
  
train\_w\_age <- subset(train, !(is.na(train$Age)))  
train\_wo\_age <- subset(train, is.na(train$Age))  
  
#scatterplot3d(train$Pclass,train$Survived,train$Age)  
#abline(fit)  
#aggregate( train$Survived~ train$Pclass, train, sum)  
  
#table(train$Pclass,train$Survived, mean(train$Age))  
fit<-lm(Age~Pclass+Survived+SibSp,data= train)  
summary(fit)

##   
## Call:  
## lm(formula = Age ~ Pclass + Survived + SibSp, data = train)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -37.760 -8.613 -1.066 7.034 44.934   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 51.7407 1.6111 32.114 < 2e-16 \*\*\*  
## Pclass -7.5582 0.5981 -12.636 < 2e-16 \*\*\*  
## Survived -7.0642 1.0180 -6.939 8.9e-12 \*\*\*  
## SibSp -4.4225 0.5032 -8.788 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 12.47 on 710 degrees of freedom  
## (177 observations deleted due to missingness)  
## Multiple R-squared: 0.2667, Adjusted R-squared: 0.2636   
## F-statistic: 86.1 on 3 and 710 DF, p-value: < 2.2e-16

train\_wo\_age$Age <- round(predict(fit,train\_wo\_age))  
updated\_age<-merge(train\_w\_age,train\_wo\_age,all=TRUE)  
View(updated\_age)  
  
drops <- c("Name", "Ticket", "Cabin")  
Cleansed\_data<- updated\_age[,!(names(updated\_age) %in% drops)]  
  
Cleansed<-dummy.data.frame(Cleansed\_data,sep="\_")  
  
View(Cleansed)  
  
pairs(~Age + Pclass + Survived + SibSp + Fare + Sex\_female + Sex\_male, data=Cleansed)



FitModel1 <- lm(Survived ~ Pclass + Age + SibSp + Fare + Sex\_female + Sex\_male + Pclass \* Age + Pclass \* SibSp + Pclass \* Fare + Pclass \* Sex\_female +Pclass \* Sex\_male + Age \* SibSp + Age \* Sex\_female + Age \* Fare + Age \* Sex\_male + SibSp \* Fare + SibSp \* Sex\_female +SibSp\* Sex\_male + Fare \* Sex\_male + Fare \* Sex\_female + Sex\_male \* Sex\_female+ Pclass \* Age \* SibSp + Pclass \* Age \* Sex\_female+ Pclass \* Age \* Sex\_male+ Pclass \* Age \* Fare + Pclass \* SibSp \* Fare+ Pclass \* SibSp \* Sex\_female +Pclass \* SibSp \* Sex\_male+ Pclass \* Fare \* Sex\_female+Pclass \* Fare \* Sex\_male+ Pclass \* Sex\_male \* Sex\_female+ Age \* SibSp \* Sex\_female  
+Age \* SibSp \* Sex\_male+Age\*Sex\_male\*Sex\_male+ Age \* Fare\* Sex\_female+Age\* Fare\* Sex\_male  
 +Age \* SibSp \* Fare+SibSp\*Fare\*Sex\_male +SibSp\*Fare\*Sex\_female+SibSp\*Sex\_male\*Sex\_female +Fare\*Sex\_male\*Sex\_female+ Pclass \* Age \* SibSp \* Sex\_female+Pclass \* Age \* SibSp \* Fare+ Pclass \* Age \* SibSp \* Sex\_male + Pclass \* Age \* Fare \* Sex\_female+ Pclass \* Age \* Fare \* Sex\_male+Pclass\*Fare\*SibSp\*Sex\_male+Pclass\*Fare\*SibSp\*Sex\_female+Pclass\*SibSp\*Sex\_female\*Sex\_male+Pclass\*Fare\*Sex\_male\*Sex\_female+Pclass\*Age\*Sex\_female\*Sex\_male + Age \* SibSp \* Fare \* Sex\_female + Age \* SibSp \* Fare \* Sex\_male+ Age \* SibSp \* Sex\_male \* Sex\_female+ Age \* Sex\_male \* Fare \* Sex\_female + Fare \* Sex\_male\* Sex\_female\* SibSp+Fare \* Sex\_male\* Sex\_female\* Pclass \*Age+Fare \* Sex\_male\* Sex\_female\* SibSp \*Pclass+Fare \* Sex\_male\* Sex\_female\* SibSp \*Age+Fare \* Sex\_male\* Pclass\* SibSp \*Age+Fare \* Pclass\* Sex\_female\* SibSp \*Age+Pclass\* Sex\_female\* SibSp \*Age \*Sex\_male\*Fare, data = Cleansed)  
summary(FitModel1)

##   
## Call:  
## lm(formula = Survived ~ Pclass + Age + SibSp + Fare + Sex\_female +   
## Sex\_male + Pclass \* Age + Pclass \* SibSp + Pclass \* Fare +   
## Pclass \* Sex\_female + Pclass \* Sex\_male + Age \* SibSp + Age \*   
## Sex\_female + Age \* Fare + Age \* Sex\_male + SibSp \* Fare +   
## SibSp \* Sex\_female + SibSp \* Sex\_male + Fare \* Sex\_male +   
## Fare \* Sex\_female + Sex\_male \* Sex\_female + Pclass \* Age \*   
## SibSp + Pclass \* Age \* Sex\_female + Pclass \* Age \* Sex\_male +   
## Pclass \* Age \* Fare + Pclass \* SibSp \* Fare + Pclass \* SibSp \*   
## Sex\_female + Pclass \* SibSp \* Sex\_male + Pclass \* Fare \*   
## Sex\_female + Pclass \* Fare \* Sex\_male + Pclass \* Sex\_male \*   
## Sex\_female + Age \* SibSp \* Sex\_female + Age \* SibSp \* Sex\_male +   
## Age \* Sex\_male \* Sex\_male + Age \* Fare \* Sex\_female + Age \*   
## Fare \* Sex\_male + Age \* SibSp \* Fare + SibSp \* Fare \* Sex\_male +   
## SibSp \* Fare \* Sex\_female + SibSp \* Sex\_male \* Sex\_female +   
## Fare \* Sex\_male \* Sex\_female + Pclass \* Age \* SibSp \* Sex\_female +   
## Pclass \* Age \* SibSp \* Fare + Pclass \* Age \* SibSp \* Sex\_male +   
## Pclass \* Age \* Fare \* Sex\_female + Pclass \* Age \* Fare \*   
## Sex\_male + Pclass \* Fare \* SibSp \* Sex\_male + Pclass \* Fare \*   
## SibSp \* Sex\_female + Pclass \* SibSp \* Sex\_female \* Sex\_male +   
## Pclass \* Fare \* Sex\_male \* Sex\_female + Pclass \* Age \* Sex\_female \*   
## Sex\_male + Age \* SibSp \* Fare \* Sex\_female + Age \* SibSp \*   
## Fare \* Sex\_male + Age \* SibSp \* Sex\_male \* Sex\_female + Age \*   
## Sex\_male \* Fare \* Sex\_female + Fare \* Sex\_male \* Sex\_female \*   
## SibSp + Fare \* Sex\_male \* Sex\_female \* Pclass \* Age + Fare \*   
## Sex\_male \* Sex\_female \* SibSp \* Pclass + Fare \* Sex\_male \*   
## Sex\_female \* SibSp \* Age + Fare \* Sex\_male \* Pclass \* SibSp \*   
## Age + Fare \* Pclass \* Sex\_female \* SibSp \* Age + Pclass \*   
## Sex\_female \* SibSp \* Age \* Sex\_male \* Fare, data = Cleansed)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.00104 -0.18816 -0.09629 0.18615 1.00129   
##   
## Coefficients: (32 not defined because of singularities)  
## Estimate Std. Error t value  
## (Intercept) 4.766e-01 2.520e-01 1.892  
## Pclass -1.130e-01 9.752e-02 -1.159  
## Age -1.718e-03 6.285e-03 -0.273  
## SibSp 5.223e-01 3.462e-01 1.508  
## Fare -9.336e-04 6.014e-03 -0.155  
## Sex\_female 9.800e-01 4.134e-01 2.371  
## Sex\_male NA NA NA  
## Pclass:Age -2.950e-04 2.690e-03 -0.110  
## Pclass:SibSp -1.664e-01 1.203e-01 -1.383  
## Pclass:Fare 8.682e-03 4.243e-03 2.046  
## Pclass:Sex\_female -2.329e-02 1.662e-01 -0.140  
## Pclass:Sex\_male NA NA NA  
## Age:SibSp -8.998e-03 9.935e-03 -0.906  
## Age:Sex\_female 5.268e-04 1.157e-02 0.046  
## Age:Fare 1.192e-04 1.732e-04 0.688  
## Age:Sex\_male NA NA NA  
## SibSp:Fare -2.066e-04 3.701e-03 -0.056  
## SibSp:Sex\_female -3.669e-01 5.339e-01 -0.687  
## SibSp:Sex\_male NA NA NA  
## Fare:Sex\_male -3.628e-03 7.993e-03 -0.454  
## Fare:Sex\_female NA NA NA  
## Sex\_female:Sex\_male NA NA NA  
## Pclass:Age:SibSp 3.119e-03 3.705e-03 0.842  
## Pclass:Age:Sex\_female -5.159e-03 5.185e-03 -0.995  
## Pclass:Age:Sex\_male NA NA NA  
## Pclass:Age:Fare -2.189e-04 1.502e-04 -1.457  
## Pclass:SibSp:Fare 4.973e-04 1.505e-03 0.330  
## Pclass:SibSp:Sex\_female 5.636e-02 1.857e-01 0.303  
## Pclass:SibSp:Sex\_male NA NA NA  
## Pclass:Fare:Sex\_female -1.136e-02 6.932e-03 -1.639  
## Pclass:Fare:Sex\_male NA NA NA  
## Pclass:Sex\_female:Sex\_male NA NA NA  
## Age:SibSp:Sex\_female 5.440e-03 1.637e-02 0.332  
## Age:SibSp:Sex\_male NA NA NA  
## Age:Fare:Sex\_female -1.479e-04 2.533e-04 -0.584  
## Age:Fare:Sex\_male NA NA NA  
## Age:SibSp:Fare 1.928e-04 1.041e-04 1.852  
## SibSp:Fare:Sex\_male -3.462e-03 4.675e-03 -0.740  
## SibSp:Fare:Sex\_female NA NA NA  
## SibSp:Sex\_female:Sex\_male NA NA NA  
## Fare:Sex\_female:Sex\_male NA NA NA  
## Age:Sex\_female:Sex\_male NA NA NA  
## Pclass:Age:SibSp:Sex\_female -1.678e-05 6.118e-03 -0.003  
## Pclass:Age:SibSp:Fare -7.702e-05 7.426e-05 -1.037  
## Pclass:Age:SibSp:Sex\_male NA NA NA  
## Pclass:Age:Fare:Sex\_female 3.289e-04 2.265e-04 1.452  
## Pclass:Age:Fare:Sex\_male NA NA NA  
## Pclass:SibSp:Fare:Sex\_male -6.948e-04 1.897e-03 -0.366  
## Pclass:SibSp:Fare:Sex\_female NA NA NA  
## Pclass:SibSp:Sex\_female:Sex\_male NA NA NA  
## Pclass:Fare:Sex\_female:Sex\_male NA NA NA  
## Pclass:Age:Sex\_female:Sex\_male NA NA NA  
## Age:SibSp:Fare:Sex\_female -1.222e-04 1.837e-04 -0.665  
## Age:SibSp:Fare:Sex\_male NA NA NA  
## Age:SibSp:Sex\_female:Sex\_male NA NA NA  
## Age:Fare:Sex\_female:Sex\_male NA NA NA  
## SibSp:Fare:Sex\_female:Sex\_male NA NA NA  
## Pclass:Age:Fare:Sex\_female:Sex\_male NA NA NA  
## Pclass:SibSp:Fare:Sex\_female:Sex\_male NA NA NA  
## Age:SibSp:Fare:Sex\_female:Sex\_male NA NA NA  
## Pclass:Age:SibSp:Fare:Sex\_male -1.754e-05 8.884e-05 -0.197  
## Pclass:Age:SibSp:Fare:Sex\_female NA NA NA  
## Pclass:Age:SibSp:Sex\_female:Sex\_male NA NA NA  
## Pclass:Age:SibSp:Fare:Sex\_female:Sex\_male NA NA NA  
## Pr(>|t|)   
## (Intercept) 0.0589 .  
## Pclass 0.2469   
## Age 0.7846   
## SibSp 0.1318   
## Fare 0.8767   
## Sex\_female 0.0180 \*  
## Sex\_male NA   
## Pclass:Age 0.9127   
## Pclass:SibSp 0.1670   
## Pclass:Fare 0.0410 \*  
## Pclass:Sex\_female 0.8886   
## Pclass:Sex\_male NA   
## Age:SibSp 0.3654   
## Age:Sex\_female 0.9637   
## Age:Fare 0.4914   
## Age:Sex\_male NA   
## SibSp:Fare 0.9555   
## SibSp:Sex\_female 0.4921   
## SibSp:Sex\_male NA   
## Fare:Sex\_male 0.6501   
## Fare:Sex\_female NA   
## Sex\_female:Sex\_male NA   
## Pclass:Age:SibSp 0.4002   
## Pclass:Age:Sex\_female 0.3201   
## Pclass:Age:Sex\_male NA   
## Pclass:Age:Fare 0.1454   
## Pclass:SibSp:Fare 0.7411   
## Pclass:SibSp:Sex\_female 0.7616   
## Pclass:SibSp:Sex\_male NA   
## Pclass:Fare:Sex\_female 0.1016   
## Pclass:Fare:Sex\_male NA   
## Pclass:Sex\_female:Sex\_male NA   
## Age:SibSp:Sex\_female 0.7398   
## Age:SibSp:Sex\_male NA   
## Age:Fare:Sex\_female 0.5594   
## Age:Fare:Sex\_male NA   
## Age:SibSp:Fare 0.0644 .  
## SibSp:Fare:Sex\_male 0.4592   
## SibSp:Fare:Sex\_female NA   
## SibSp:Sex\_female:Sex\_male NA   
## Fare:Sex\_female:Sex\_male NA   
## Age:Sex\_female:Sex\_male NA   
## Pclass:Age:SibSp:Sex\_female 0.9978   
## Pclass:Age:SibSp:Fare 0.3000   
## Pclass:Age:SibSp:Sex\_male NA   
## Pclass:Age:Fare:Sex\_female 0.1469   
## Pclass:Age:Fare:Sex\_male NA   
## Pclass:SibSp:Fare:Sex\_male 0.7143   
## Pclass:SibSp:Fare:Sex\_female NA   
## Pclass:SibSp:Sex\_female:Sex\_male NA   
## Pclass:Fare:Sex\_female:Sex\_male NA   
## Pclass:Age:Sex\_female:Sex\_male NA   
## Age:SibSp:Fare:Sex\_female 0.5062   
## Age:SibSp:Fare:Sex\_male NA   
## Age:SibSp:Sex\_female:Sex\_male NA   
## Age:Fare:Sex\_female:Sex\_male NA   
## SibSp:Fare:Sex\_female:Sex\_male NA   
## Pclass:Age:Fare:Sex\_female:Sex\_male NA   
## Pclass:SibSp:Fare:Sex\_female:Sex\_male NA   
## Age:SibSp:Fare:Sex\_female:Sex\_male NA   
## Pclass:Age:SibSp:Fare:Sex\_male 0.8435   
## Pclass:Age:SibSp:Fare:Sex\_female NA   
## Pclass:Age:SibSp:Sex\_female:Sex\_male NA   
## Pclass:Age:SibSp:Fare:Sex\_female:Sex\_male NA   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.3653 on 859 degrees of freedom  
## Multiple R-squared: 0.456, Adjusted R-squared: 0.4364   
## F-statistic: 23.23 on 31 and 859 DF, p-value: < 2.2e-16

FitModel8 <- lm(Survived ~ Pclass + Age + SibSp + Fare + Sex\_female + Sex\_male + Pclass \* Age + Pclass \* SibSp + Pclass \* Fare + Pclass \* Sex\_female +Pclass \* Sex\_male + Age \* SibSp , data = Cleansed)  
summary(FitModel8)

##   
## Call:  
## lm(formula = Survived ~ Pclass + Age + SibSp + Fare + Sex\_female +   
## Sex\_male + Pclass \* Age + Pclass \* SibSp + Pclass \* Fare +   
## Pclass \* Sex\_female + Pclass \* Sex\_male + Age \* SibSp, data = Cleansed)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.1774 -0.2311 -0.1006 0.2021 1.0451   
##   
## Coefficients: (2 not defined because of singularities)  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.5917336 0.1246097 4.749 2.39e-06 \*\*\*  
## Pclass -0.0590278 0.0494821 -1.193 0.2332   
## Age -0.0020838 0.0027906 -0.747 0.4554   
## SibSp 0.0750828 0.0785259 0.956 0.3393   
## Fare -0.0020152 0.0010591 -1.903 0.0574 .   
## Sex\_female 0.7902613 0.0780503 10.125 < 2e-16 \*\*\*  
## Sex\_male NA NA NA NA   
## Pclass:Age -0.0029342 0.0012440 -2.359 0.0186 \*   
## Pclass:SibSp -0.0626710 0.0266332 -2.353 0.0188 \*   
## Pclass:Fare 0.0019179 0.0009780 1.961 0.0502 .   
## Pclass:Sex\_female -0.1342525 0.0323694 -4.148 3.69e-05 \*\*\*  
## Pclass:Sex\_male NA NA NA NA   
## Age:SibSp 0.0005231 0.0009681 0.540 0.5891   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.37 on 880 degrees of freedom  
## Multiple R-squared: 0.4282, Adjusted R-squared: 0.4217   
## F-statistic: 65.89 on 10 and 880 DF, p-value: < 2.2e-16

## P-value for the Age\*SibSp is not significant, so the NH hypothesis is not rejected.  
FitModel9 <- lm(Survived ~ Pclass + Age + SibSp + Fare + Sex\_female + Sex\_male + Pclass \* Age + Pclass \* SibSp + Pclass \* Fare + Pclass \* Sex\_female +Pclass \* Sex\_male , data = Cleansed)  
summary(FitModel9)

##   
## Call:  
## lm(formula = Survived ~ Pclass + Age + SibSp + Fare + Sex\_female +   
## Sex\_male + Pclass \* Age + Pclass \* SibSp + Pclass \* Fare +   
## Pclass \* Sex\_female + Pclass \* Sex\_male, data = Cleansed)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.18740 -0.23289 -0.09881 0.20039 1.03941   
##   
## Coefficients: (2 not defined because of singularities)  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.5818100 0.1231995 4.723 2.71e-06 \*\*\*  
## Pclass -0.0569090 0.0493067 -1.154 0.24874   
## Age -0.0018757 0.0027628 -0.679 0.49737   
## SibSp 0.1024908 0.0599248 1.710 0.08756 .   
## Fare -0.0019946 0.0010579 -1.885 0.05972 .   
## Sex\_female 0.7895768 0.0780086 10.122 < 2e-16 \*\*\*  
## Sex\_male NA NA NA NA   
## Pclass:Age -0.0029306 0.0012435 -2.357 0.01865 \*   
## Pclass:SibSp -0.0708036 0.0219647 -3.224 0.00131 \*\*   
## Pclass:Fare 0.0018884 0.0009761 1.935 0.05336 .   
## Pclass:Sex\_female -0.1330697 0.0322823 -4.122 4.11e-05 \*\*\*  
## Pclass:Sex\_male NA NA NA NA   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.3699 on 881 degrees of freedom  
## Multiple R-squared: 0.428, Adjusted R-squared: 0.4221   
## F-statistic: 73.24 on 9 and 881 DF, p-value: < 2.2e-16

##The p-value for the Pclass \* Fare is not significant, so the NH hypothesis is not rejected.  
FitModel11 <- lm(Survived ~ Pclass + Age + SibSp + Fare + Sex\_female + Sex\_male + Pclass \* Age + Pclass \* SibSp + Pclass \* Sex\_female , data = Cleansed)  
summary(FitModel11)

##   
## Call:  
## lm(formula = Survived ~ Pclass + Age + SibSp + Fare + Sex\_female +   
## Sex\_male + Pclass \* Age + Pclass \* SibSp + Pclass \* Sex\_female,   
## data = Cleansed)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.2046 -0.2300 -0.1032 0.2049 1.0295   
##   
## Coefficients: (1 not defined because of singularities)  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 5.833e-01 1.234e-01 4.727 2.65e-06 \*\*\*  
## Pclass -4.691e-02 4.911e-02 -0.955 0.3398   
## Age -2.366e-03 2.755e-03 -0.859 0.3908   
## SibSp 9.179e-02 5.976e-02 1.536 0.1249   
## Fare -4.376e-05 3.206e-04 -0.137 0.8915   
## Sex\_female 7.803e-01 7.798e-02 10.006 < 2e-16 \*\*\*  
## Sex\_male NA NA NA NA   
## Pclass:Age -2.706e-03 1.240e-03 -2.182 0.0294 \*   
## Pclass:SibSp -5.890e-02 2.112e-02 -2.789 0.0054 \*\*   
## Pclass:Sex\_female -1.282e-01 3.223e-02 -3.976 7.58e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.3705 on 882 degrees of freedom  
## Multiple R-squared: 0.4255, Adjusted R-squared: 0.4203   
## F-statistic: 81.67 on 8 and 882 DF, p-value: < 2.2e-16

##The p-value for the Pclass \* Age is not significant, so the NH hypothesis is not rejected.  
FitModel12 <- lm(Survived ~ Pclass + Age + SibSp + Fare + Sex\_female + Sex\_male + Pclass \* SibSp+Pclass \* Sex\_female , data = Cleansed)  
summary(FitModel12)

##   
## Call:  
## lm(formula = Survived ~ Pclass + Age + SibSp + Fare + Sex\_female +   
## Sex\_male + Pclass \* SibSp + Pclass \* Sex\_female, data = Cleansed)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.2875 -0.2157 -0.1132 0.1978 0.9894   
##   
## Coefficients: (1 not defined because of singularities)  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 7.917e-01 7.828e-02 10.114 < 2e-16 \*\*\*  
## Pclass -1.418e-01 2.283e-02 -6.213 8.02e-10 \*\*\*  
## Age -7.890e-03 1.090e-03 -7.235 1.01e-12 \*\*\*  
## SibSp 7.575e-02 5.943e-02 1.275 0.202804   
## Fare -6.412e-05 3.212e-04 -0.200 0.841804   
## Sex\_female 7.538e-01 7.719e-02 9.765 < 2e-16 \*\*\*  
## Sex\_male NA NA NA NA   
## Pclass:SibSp -4.969e-02 2.074e-02 -2.396 0.016775 \*   
## Pclass:Sex\_female -1.167e-01 3.187e-02 -3.662 0.000265 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.3713 on 883 degrees of freedom  
## Multiple R-squared: 0.4224, Adjusted R-squared: 0.4179   
## F-statistic: 92.26 on 7 and 883 DF, p-value: < 2.2e-16

## The p-value for the Pclass \* SibSp is not signifacant, so the NH hypothesis is not rejected.  
FitModel13 <- lm(Survived ~ Pclass + Age + SibSp + Fare + Sex\_female + Sex\_male +Pclass \* Sex\_female , data = Cleansed)  
summary(FitModel13)

##   
## Call:  
## lm(formula = Survived ~ Pclass + Age + SibSp + Fare + Sex\_female +   
## Sex\_male + Pclass \* Sex\_female, data = Cleansed)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.2497 -0.2099 -0.1145 0.1992 0.9918   
##   
## Coefficients: (1 not defined because of singularities)  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 8.204e-01 7.756e-02 10.577 < 2e-16 \*\*\*  
## Pclass -1.545e-01 2.227e-02 -6.939 7.64e-12 \*\*\*  
## Age -7.757e-03 1.092e-03 -7.104 2.50e-12 \*\*\*  
## SibSp -6.344e-02 1.259e-02 -5.038 5.71e-07 \*\*\*  
## Fare 5.562e-05 3.181e-04 0.175 0.861   
## Sex\_female 7.815e-01 7.652e-02 10.212 < 2e-16 \*\*\*  
## Sex\_male NA NA NA NA   
## Pclass:Sex\_female -1.271e-01 3.166e-02 -4.016 6.43e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.3723 on 884 degrees of freedom  
## Multiple R-squared: 0.4187, Adjusted R-squared: 0.4147   
## F-statistic: 106.1 on 6 and 884 DF, p-value: < 2.2e-16

## The P-value for the Pclass \* SibSp is not significant, so the NH hypothesis is not rejected.  
FitModel14 <- lm(Survived ~ Pclass + Age + SibSp + Fare + Sex\_female +Pclass \* Sex\_female , data = Cleansed)  
summary(FitModel14)

##   
## Call:  
## lm(formula = Survived ~ Pclass + Age + SibSp + Fare + Sex\_female +   
## Pclass \* Sex\_female, data = Cleansed)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.2497 -0.2099 -0.1145 0.1992 0.9918   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 8.204e-01 7.756e-02 10.577 < 2e-16 \*\*\*  
## Pclass -1.545e-01 2.227e-02 -6.939 7.64e-12 \*\*\*  
## Age -7.757e-03 1.092e-03 -7.104 2.50e-12 \*\*\*  
## SibSp -6.344e-02 1.259e-02 -5.038 5.71e-07 \*\*\*  
## Fare 5.562e-05 3.181e-04 0.175 0.861   
## Sex\_female 7.815e-01 7.652e-02 10.212 < 2e-16 \*\*\*  
## Pclass:Sex\_female -1.271e-01 3.166e-02 -4.016 6.43e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.3723 on 884 degrees of freedom  
## Multiple R-squared: 0.4187, Adjusted R-squared: 0.4147   
## F-statistic: 106.1 on 6 and 884 DF, p-value: < 2.2e-16

##Sex\_male is NA.  
  
##The p-value for Fare is not significant, so the NH hypothesis is not rejected.  
FitModel14 <- lm(Survived ~ Pclass + Age + SibSp + Sex\_female +Pclass \* Sex\_female , data = Cleansed)  
summary(FitModel14)

##   
## Call:  
## lm(formula = Survived ~ Pclass + Age + SibSp + Sex\_female + Pclass \*   
## Sex\_female, data = Cleansed)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.2471 -0.2099 -0.1095 0.1980 0.9921   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.825641 0.071502 11.547 < 2e-16 \*\*\*  
## Pclass -0.156051 0.020478 -7.621 6.50e-14 \*\*\*  
## Age -0.007769 0.001089 -7.133 2.04e-12 \*\*\*  
## SibSp -0.062993 0.012319 -5.113 3.88e-07 \*\*\*  
## Sex\_female 0.784236 0.074840 10.479 < 2e-16 \*\*\*  
## Pclass:Sex\_female -0.128154 0.031091 -4.122 4.11e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.3721 on 885 degrees of freedom  
## Multiple R-squared: 0.4187, Adjusted R-squared: 0.4154   
## F-statistic: 127.5 on 5 and 885 DF, p-value: < 2.2e-16

##Fitmodel15<-lm (Survived ~ Pclass + Age + SibSp + Sex\_female+ Sex\_male, data=Cleansed)  
##summary(Fitmodel15)