simplilearn
your pace, your place

Analyze the Sales Report of a Clothes Manufacturing Outlet

Business Analytic Foundation with R Tools- Solutions

Solutions

Disclaimer: In Business Analytics, there are different ways of solving the same set of problems, we are just presenting one. Feel free to explore other ways of answering these questions.

1. To automate the process of recommendations, the store needs to analyze the given attributes of the product, like style, season, etc., and come up with a model to predict the recommendation of products (in binary output – 0 or 1) accordingly.

   The **Attribute DataSet.csv** file contains all the attributes and the recommendations for each dress in binary. For binary output models, logistic regression can be used to build a suitable model to recommend products (as mentioned earlier, a value of 1 denotes a positive recommendation and 0 denotes a negative recommendation). To perform logistic regression on a given dataset, we need to decide two major attributes of the model – the dependent and independent variables. The required values are:
   Dependent variable: Recommendation
   Independent Variables: All other variables, except the Dress_ID, since it is only an identifier.
   Outlier treatments: None

Code:

```
Mydata  <- read.csv("Attribute DataSet.csv")
Model1 <- glm(Recommendation ~ Style + Price + Rating + Size + Season +
NeckLine +
        SleeveLength + waiseline + Material + FabricType + Decoration +
Pattern.Type,
        data = mydata)
```

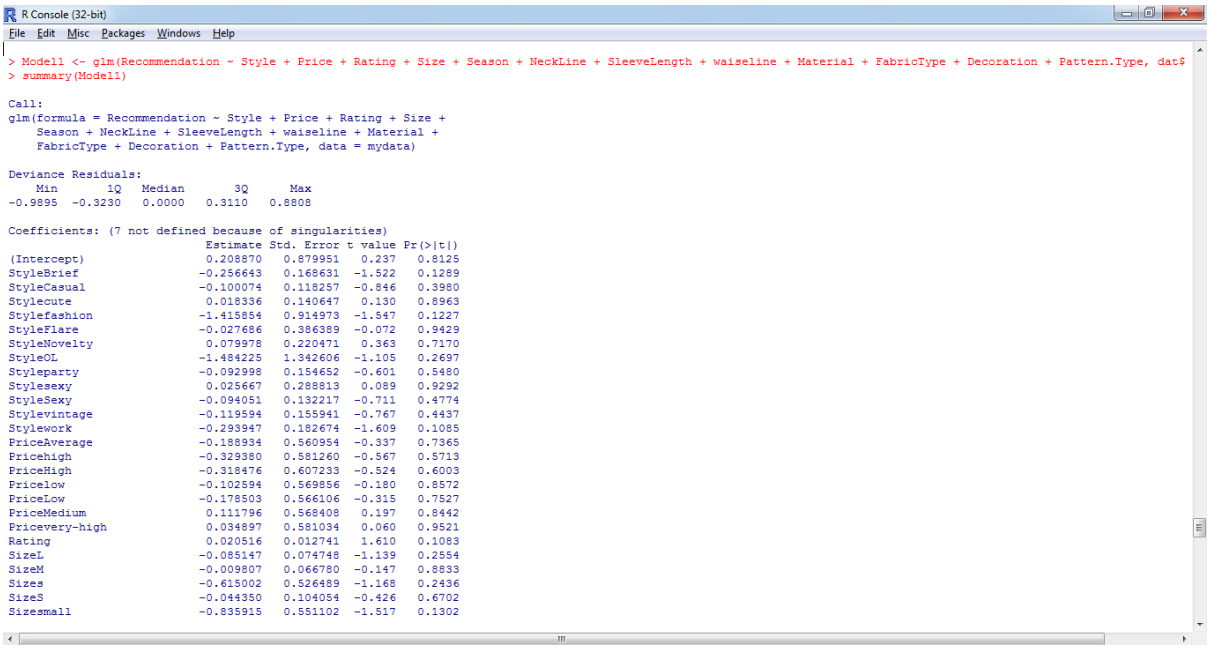
Result:

From the significance codes for each attribute, we can see that Pattern Type and Sleeve Length make an impact on the recommendation, both positively affecting the recommendation. Other than that, we can see that the increased number of factors and comparatively lesser number of entries make the predictions slightly difficult.

However, the residual deviance is lower than the null deviance, which implies that using the independent variables makes it closer to predicting the actual values of recommendation.

With the given model, the new data or attributes can be fed into the model to get recommendations.

Output:



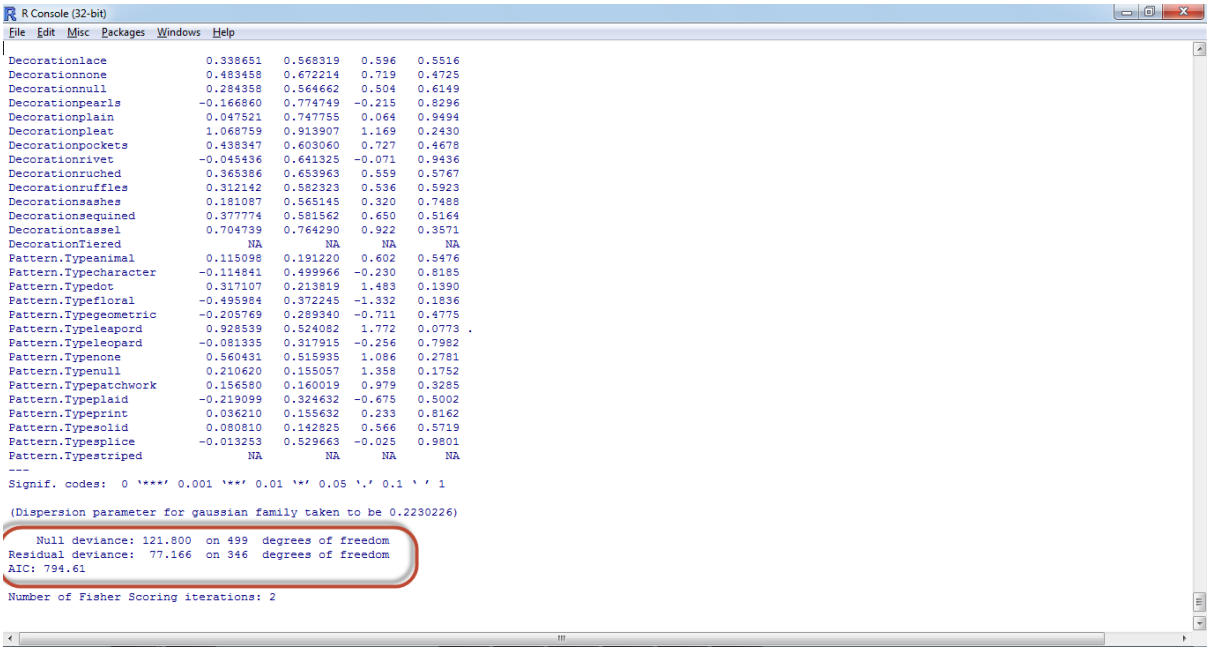
```
R Console (32-bit)
File  Edit  Misc  Packages  Windows  Help

> Model1 <- glm(Recommendation ~ Style + Price + Rating + Size + Season + NeckLine + SleeveLength + waiseline + Material + FabricType + Decoration + Pattern.Type, dat$
> summary(Model1)

Call:
glm(formula = Recommendation ~ Style + Price + Rating + Size +
    Season + NeckLine + SleeveLength + waiseline + Material +
    FabricType + Decoration + Pattern.Type, data = mydata)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.9895  -0.3230   0.0000   0.3110   0.8808

Coefficients: (7 not defined because of singularities)
                        Estimate Std. Error t value Pr(>|t|)
(Intercept)             0.208870   0.879951   0.237   0.8125
StyleBrief             -0.256643   0.168631  -1.522   0.1289
StyleCasual            -0.100074   0.118257  -0.846   0.3980
Stylecute               0.018336   0.140647   0.130   0.8963
Stylefashion           -1.415854   0.914973  -1.547   0.1227
StyleFlare             -0.027686   0.386389  -0.072   0.9429
StyleNovelty            0.079978   0.220471   0.363   0.7170
StyleOL                -1.484225   1.342606  -1.105   0.2697
Styleparty             -0.092998   0.154652  -0.601   0.5480
Stylesexy               0.025667   0.288813   0.089   0.9292
StyleSexy              -0.094051   0.132217  -0.711   0.4774
Stylevintage           -0.119594   0.155941  -0.767   0.4437
Stylework              -0.293947   0.182674  -1.609   0.1085
PriceAverage           -0.188934   0.560954  -0.337   0.7365
Pricehigh              -0.329380   0.581260  -0.567   0.5713
PriceHigh              -0.318476   0.607233  -0.524   0.6003
Pricelow               -0.102594   0.569856  -0.180   0.8572
PriceLow               -0.178503   0.566106  -0.315   0.7527
PriceMedium             0.111796   0.568408   0.197   0.8442
Pricevery-high          0.034897   0.581034   0.060   0.9521
Rating                  0.020516   0.012741   1.610   0.1083
SizeL                  -0.085147   0.074748  -1.139   0.2554
SizeM                  -0.009807   0.066780  -0.147   0.8833
Sizes                  -0.615002   0.526489  -1.168   0.2436
SizeS                  -0.044350   0.104054  -0.426   0.6702
Sizesmall              -0.835915   0.551102  -1.517   0.1302
```



```
R Console (32-bit)
File  Edit  Misc  Packages  Windows  Help

Decorationlace          0.338651   0.568319   0.596   0.5516
Decorationnone          0.483458   0.672214   0.719   0.4725
Decorationnull          0.284358   0.564662   0.504   0.6149
Decorationpearls       -0.166860   0.774749  -0.215   0.8296
Decorationplain         0.047521   0.747755   0.064   0.9494
Decorationpleat         1.068759   0.913907   1.169   0.2430
Decorationpockets       0.438347   0.603060   0.727   0.4678
Decorationrivet        -0.045436   0.641325  -0.071   0.9436
Decorationruched        0.365386   0.653963   0.559   0.5767
Decorationruffles       0.312142   0.582323   0.536   0.5923
Decorationsashes        0.181087   0.565145   0.320   0.7488
Decorationsequined      0.377774   0.581562   0.650   0.5164
Decorationtassel        0.704739   0.764290   0.922   0.3571
DecorationTiered              NA         NA      NA       NA
Pattern.Typeanimal      0.115098   0.191220   0.602   0.5476
Pattern.Typecharacter  -0.114841   0.499966  -0.230   0.8185
Pattern.Typedot         0.317107   0.213819   1.483   0.1390
Pattern.Typefloral     -0.495984   0.372245  -1.332   0.1836
Pattern.Typegeometric  -0.205769   0.289340  -0.711   0.4775
Pattern.Typeleapord     0.928539   0.524082   1.772   0.0773 .
Pattern.Typeleopard    -0.081335   0.317915  -0.256   0.7982
Pattern.Typenone        0.560431   0.515935   1.086   0.2781
Pattern.Typenull        0.210620   0.155057   1.358   0.1752
Pattern.Typepatchwork   0.156580   0.160019   0.979   0.3285
Pattern.Typeplaid      -0.219099   0.324632  -0.675   0.5002
Pattern.Typeprint       0.036210   0.155632   0.233   0.8162
Pattern.Typesolid       0.080810   0.142825   0.566   0.5719
Pattern.Typesplice     -0.013253   0.529663  -0.025   0.9801
Pattern.Typestriped           NA         NA      NA       NA
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for gaussian family taken to be 0.2230226)

    Null deviance: 121.800  on 499  degrees of freedom
Residual deviance:  77.166  on 346  degrees of freedom
AIC: 794.61

Number of Fisher Scoring iterations: 2
```