Data Overview

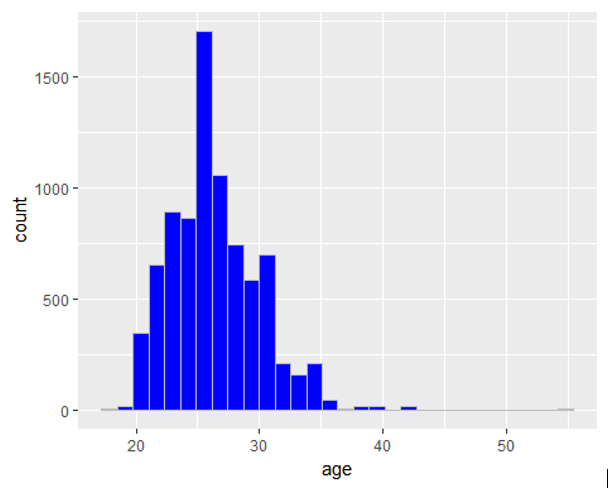
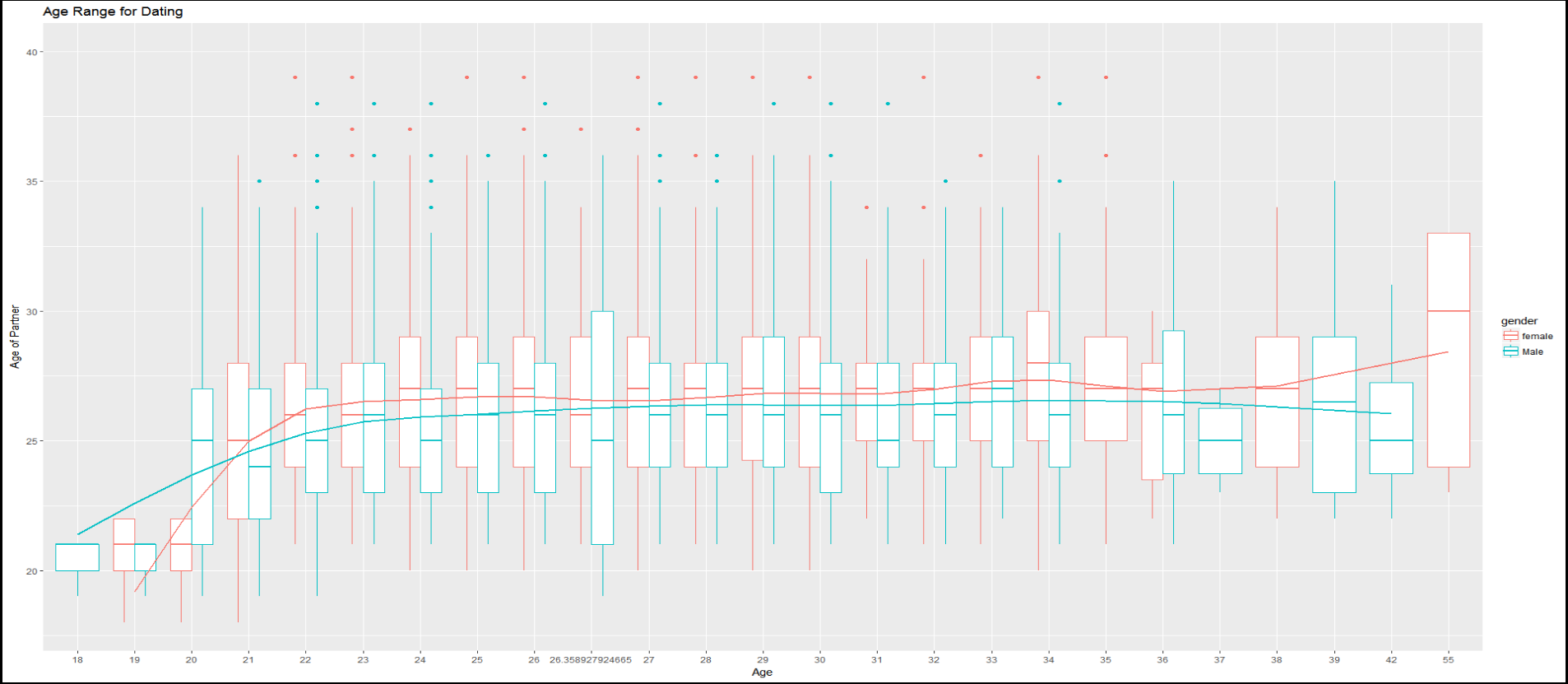


Figure 1 (Age Histogram)

Figure 2 (Boxplots of Age vs. gender)

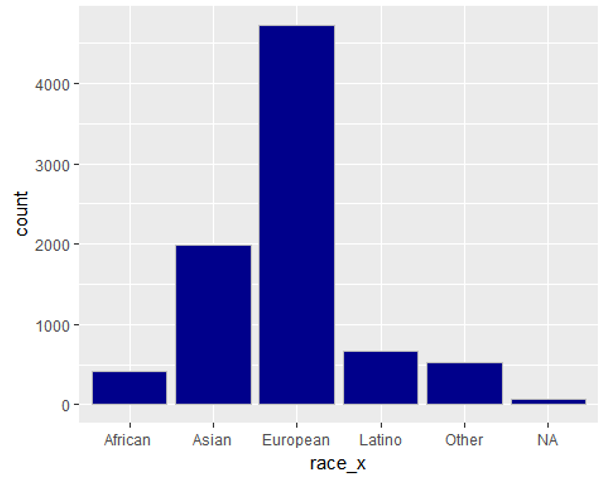


Figure 3 (Race Histogram)

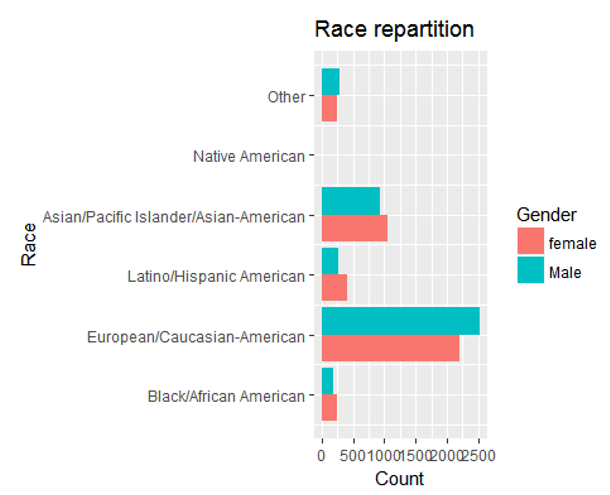


Figure 4 (Race Histogram by Gender)

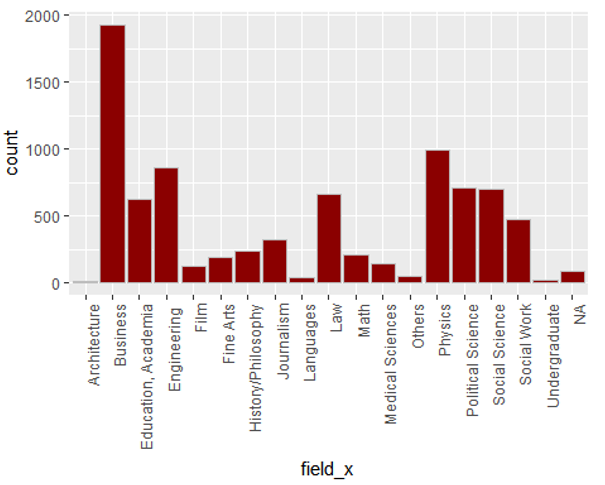


Figure 5 (Field of Study Histogram)

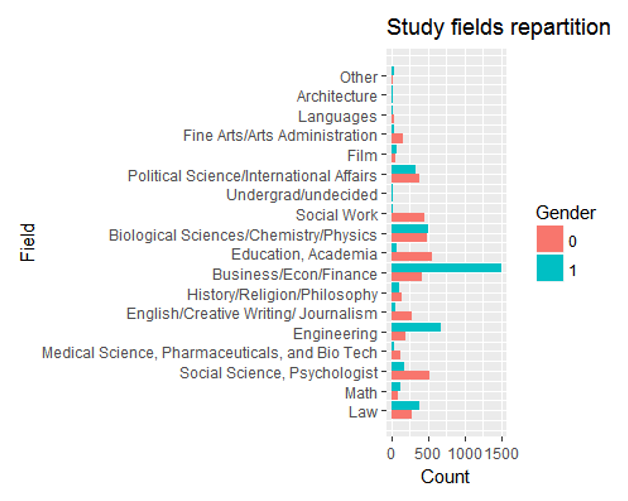


Figure 6 (Field of Study by Gender)

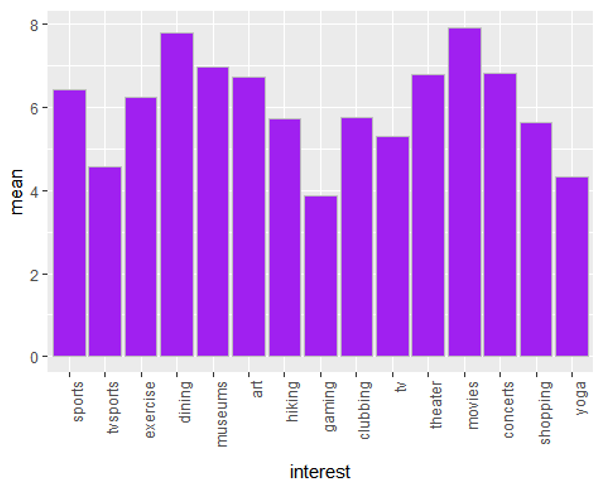


Figure 7 (Interest Histogram)

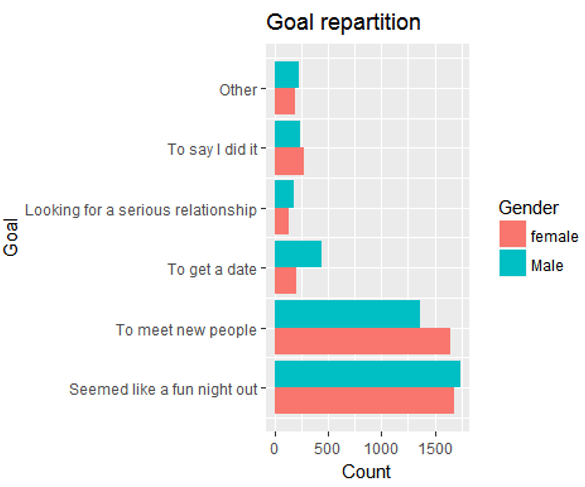


Figure 8 (Goal by Gender)

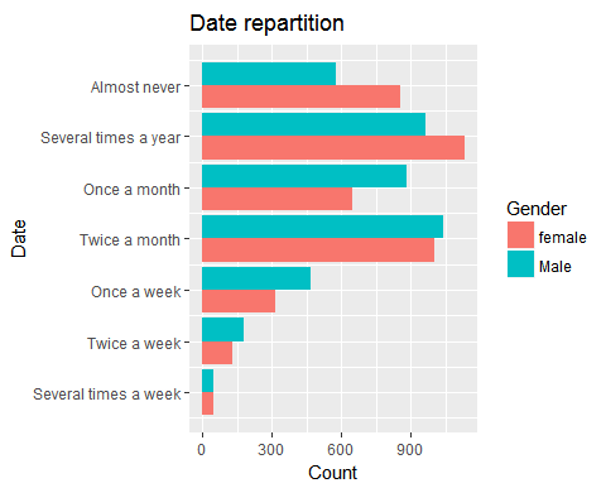
****

Figure 9 (Date Frequency by Gender)

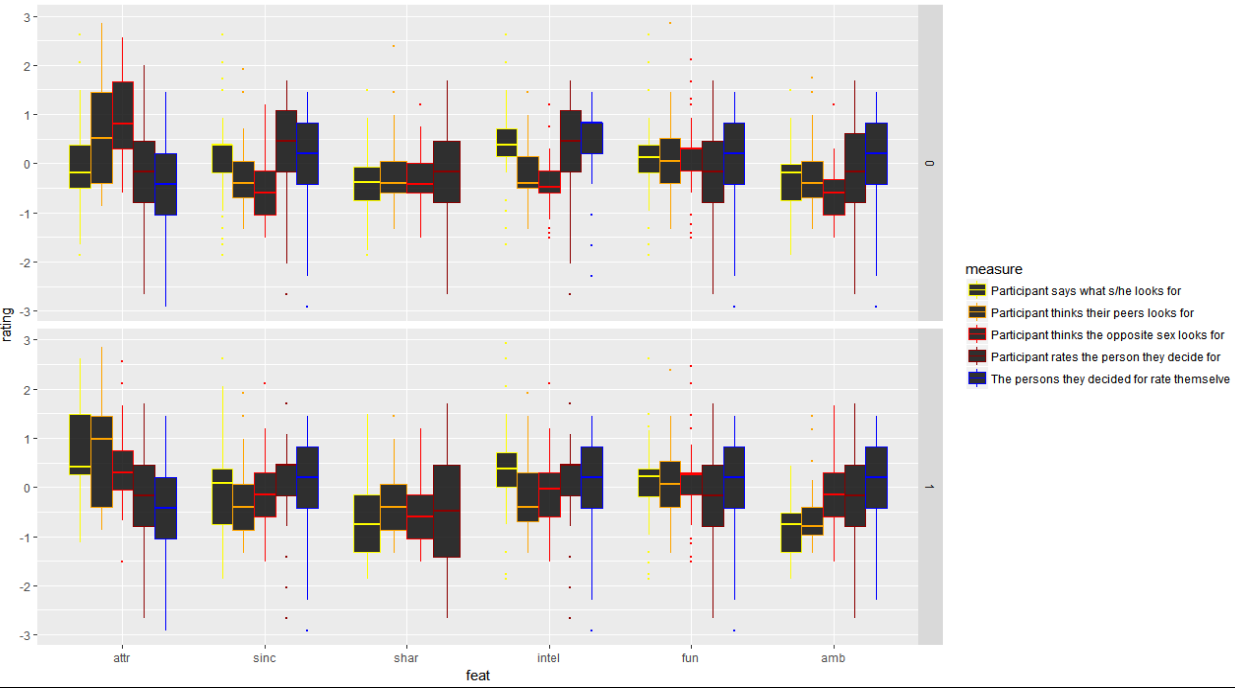


Figure 10 (Six Attributes by Different Questions & Gender)

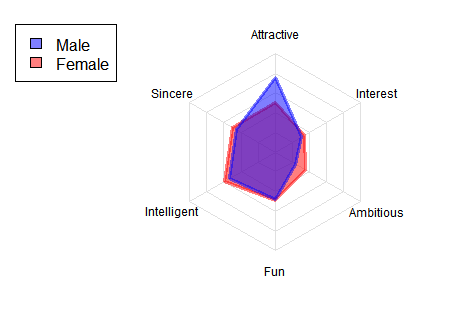
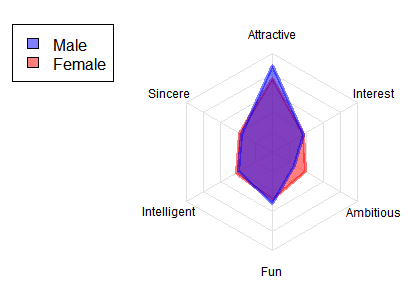
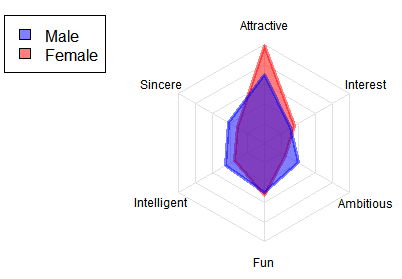
  

Figure 11Figure 12Figure 13

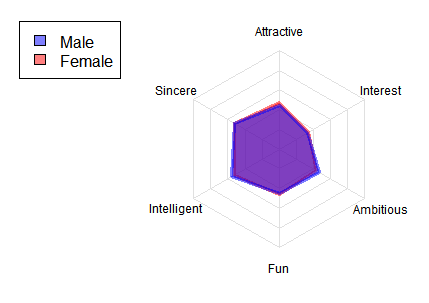
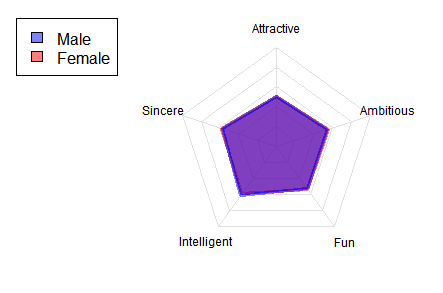
**** ****

Figure 14 Figure 15

\*Figure 11-15 are same six attributes by gender and five different questions same as Figure 10

Appendix B

Some significant R output

B1. Model 1 (General Model)

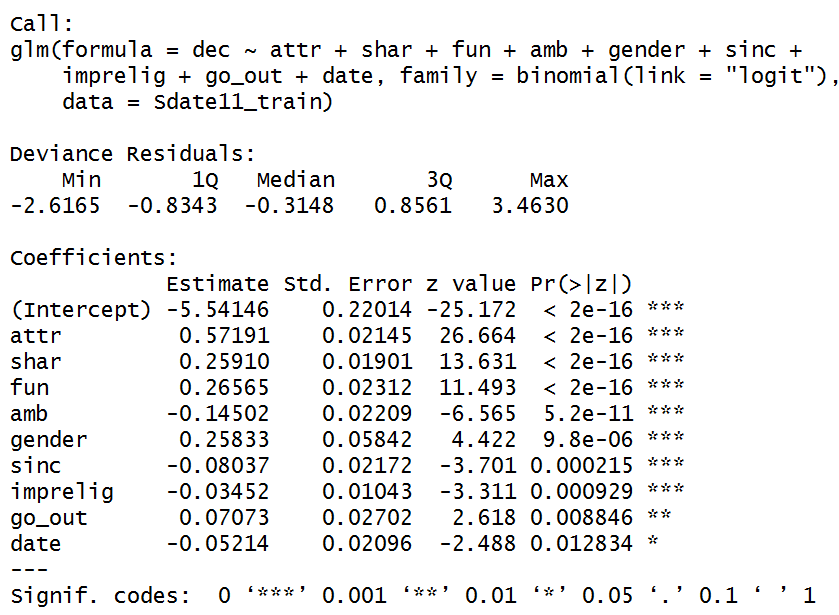


Figure 16 (Model 1 Output)

|  |  |  |
| --- | --- | --- |
| Desicion | Predicted Values | |
| Actual Values | No | Yes |
| No | 395 | 141 |
| Yes | 101 | 201 |

Table 1 (Confusion Matrix for Model 1)

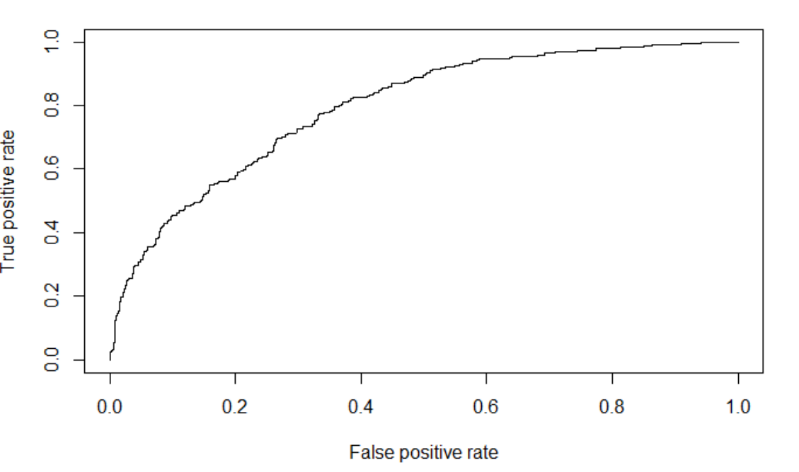


Figure 17 (ROC Plot for Model 1)

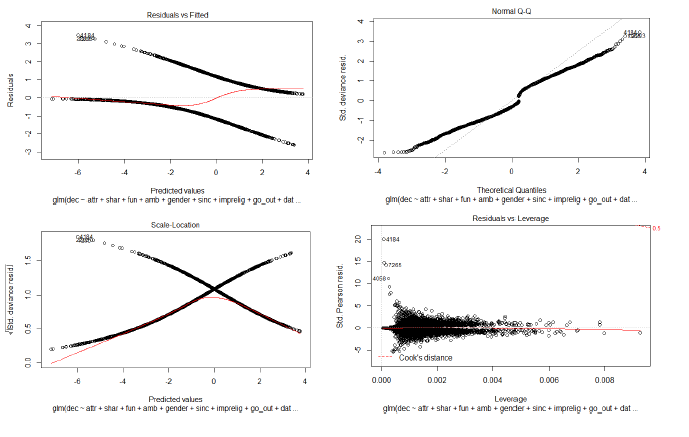


Figure 18 (Residual Plot for Model 1)

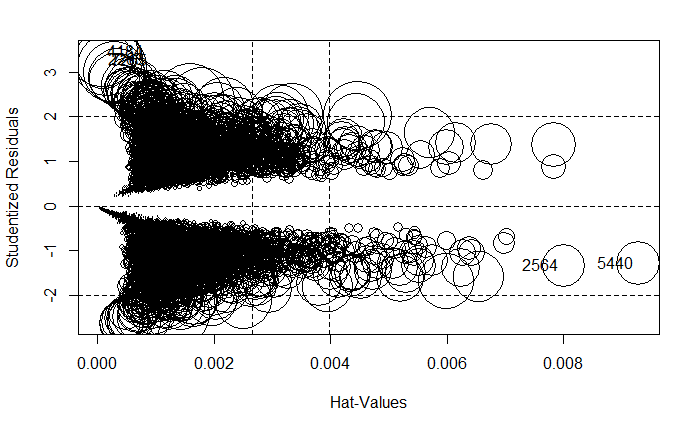


Figure 19 (Bubble-Plot for Model 1)

B2. Model 3 (Female Model)

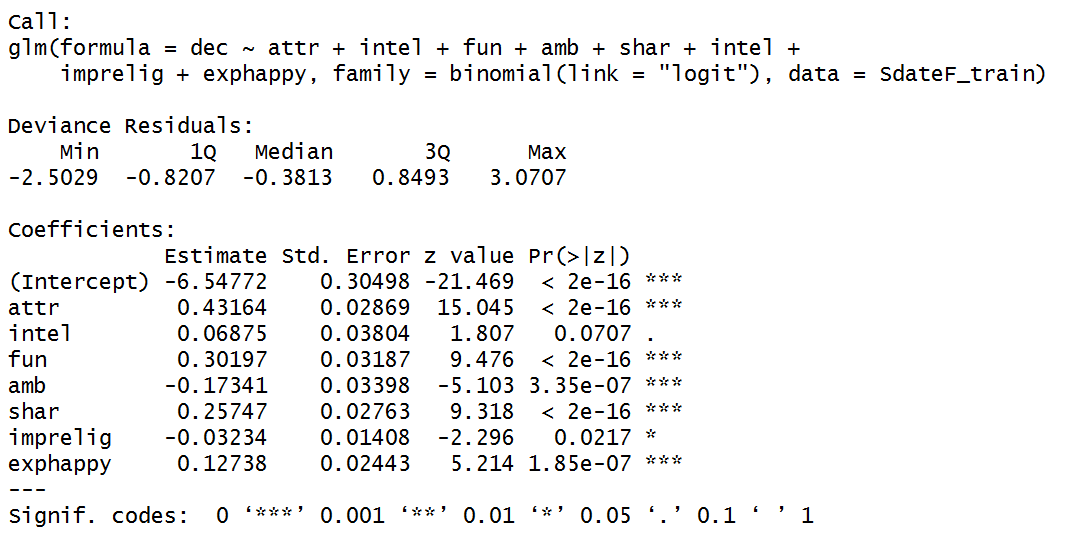


Figure 20 (Model 3 Output)

|  |  |  |
| --- | --- | --- |
| Desicion | Predicted Values | |
| Actual Values | No | Yes |
| No | 199 | 62 |
| Yes | 55 | 103 |

Table 2 (Confusion Matrix for Model 3)

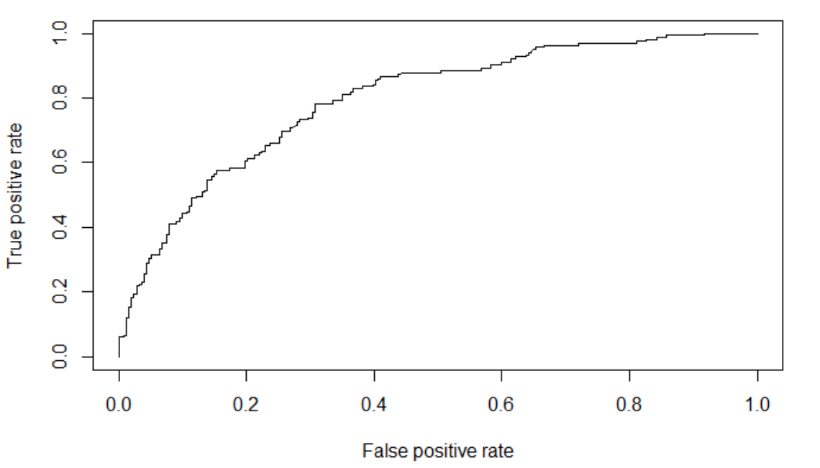


Figure 21 (ROC Plot for Model 3)

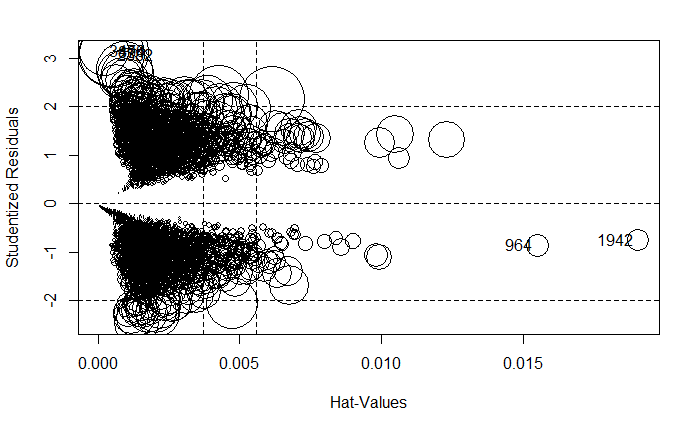


Figure 22 (Bubble-Plot for Model 3)

B3. Model 4 (Male Model)

![C:\Users\Flora\AppData\Roaming\Tencent\Users\925610963\QQ\WinTemp\RichOle\GVXHT]((JP66MM){EG(KZ6P.png](data:image/png;base64,)

Figure 23 (Model 4 Output)

|  |  |  |
| --- | --- | --- |
| Desicion | Predicted Values | |
| Actual Values | No | Yes |
| No | 164 | 54 |
| Yes | 57 | 145 |

Table 3 (Confusion Matrix for Model 4)

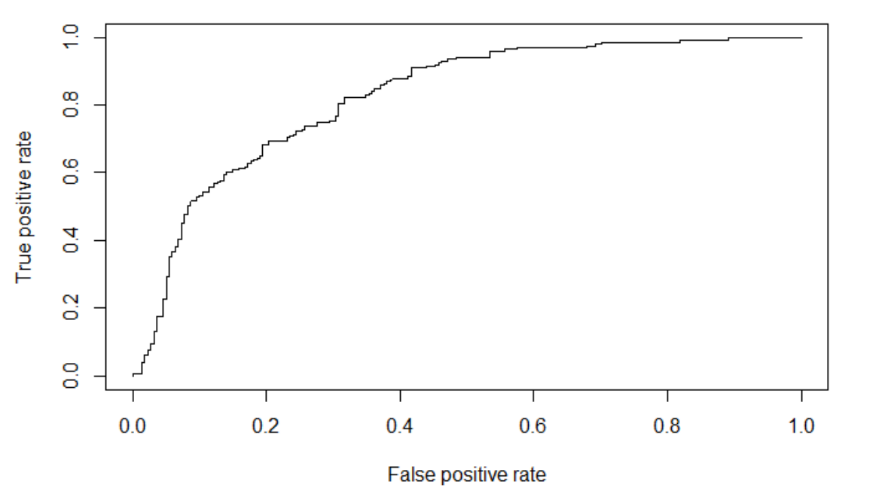


Figure 24 (ROC Plot for Model 4)

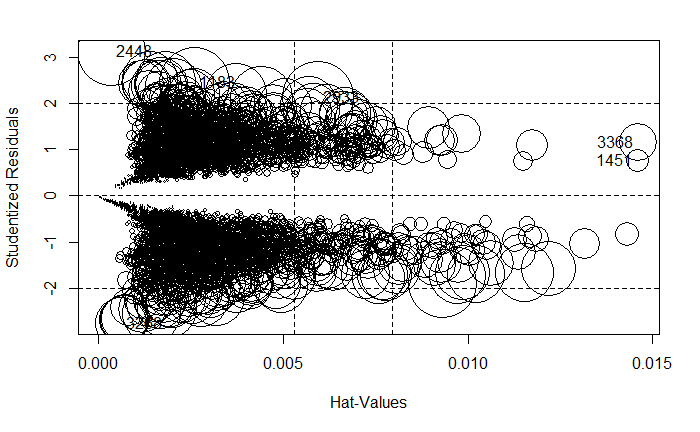


Figure 25 (Bubble-Plot for Model 4)