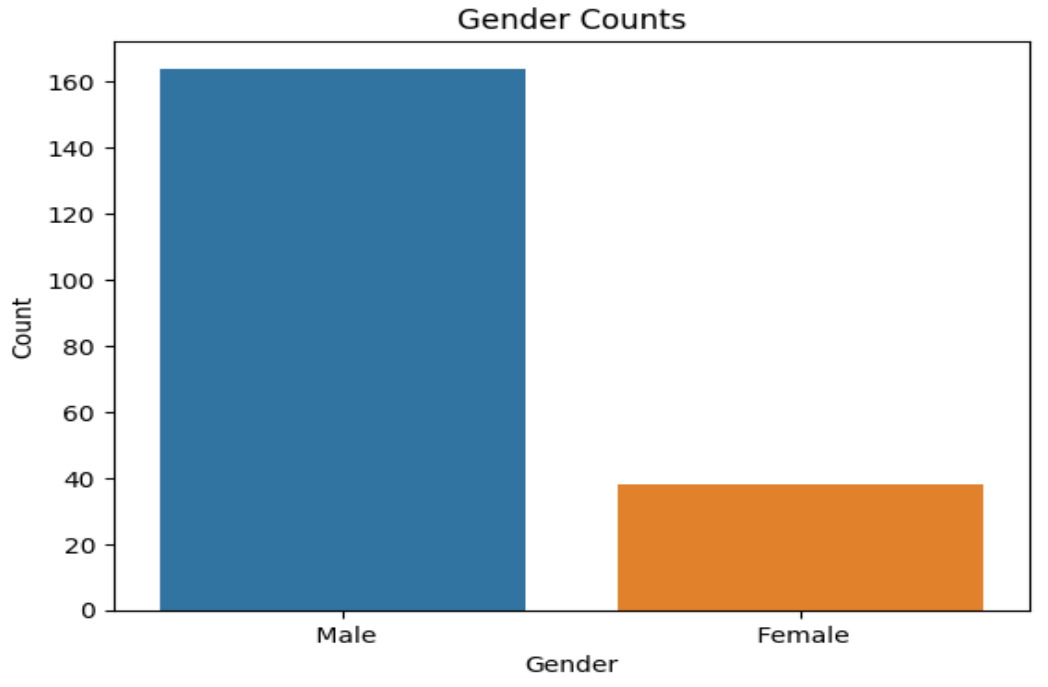
Appendix A1 - Final Dataset including all variables (202 rows x 36 columns)

```
> str(data)
'data.frame':   202 obs. of  36 variables:
 $ X                                : int  0 1 2 3 4 5 6 7 8 9 ...
 $ Doctor.Name                      : chr  "Dr. ABDULMASSIH Abdulmassih" "Dr. Ahmad A. Shaher" "Dr. Alan L. Sisson
i" ...
 $ State                            : chr  "Illinois" "Illinois" "Illinois" "California" ...
 $ Text_of_sanction                 : chr  "physician and surgeon license reprimanded and fined $5,000 and must ta
five topic areas of Ethic"| __truncated__ "physician and surgeon license placed on indefinite probation for a min
rs," "physician and surgeon license and controlled substance license placed on permanent relinquished status afte
ted__ "Continuing Education" ...
 $ Gender                           : num  0 0 0 0 0 0 1 0 0 0 ...
 $ Overall_Rating                   : num  4.58 4.75 1 1 3.83 1.84 5 3.75 3.65 3.63 ...
 $ Negative_Reviews                 : chr  "" "" "I am not sure how how doctor stays in business. With manners lik
be a truck driver. No offense to the truckers." "Stay away from this doctor (and the term is used loosely. He per
y on my dear mother and he did t"| __truncated__ ...
 $ Positive_Reviews                 : chr  "Dr. Abdulmasshi is the best doctor out there. Amazed. I have been goin
ssih for years ever since "| __truncated__ "Dr shaher is an awesome human being and a doctor. He is good teacher
tor. I have learnt so much "| __truncated__ "" "" ...
 $ No_of_Negative_Reviews           : num  NA NA 1 2 NA 7 NA 1 4 8 ...
 $ No_of_Positive_Reviews           : num  3 2 NA NA 4 2 1 NA 14 18 ...
 $ Total_Reviews                    : int  3 2 1 1 3 9 1 2 18 13 ...
 $ PPNegative_1                     : chr  "" "nn" "sure doctor stay business manner ike truck iver offense trucke
octor term used loosely performed surgery dear mother incorrect surgery life miserable death answer "| __truncate
 $ PPPositive_1                     : chr  " best doctor amazed going   year ever since partner passed away inheri
now polite friendly import"| __truncated__ "wesome humn doctor good techer ptient doctor lernt much truly dess ev
ient strives provide best t"| __truncated__ "" "" ...
 $ n_score_bed_side                 : num  1.55e-05 1.55e-05 1.55e-05 2.30e-04 1.55e-05 ...
 $ n_score_Knowledge                : num  6.11e-06 6.11e-06 6.11e-06 1.51e-04 6.11e-06 ...
 $ n_score_communication            : num  1.47e-05 1.47e-05 3.22e-04 3.52e-03 1.47e-05 ...
 $ n_score_environment              : num  2.48e-05 2.48e-05 2.48e-05 2.64e-05 2.48e-05 ...
 $ n_score_scheduling               : num  0.000112 0.000112 0.000112 0.000112 0.000112 ...
 $ n_score_wait_time                : num  0.000162 0.000162 0.00188 0.006773 0.000162 ...
 $ n_score_costs                    : num  0.000001 0.002465 0.002466 0.002466 0.000001 ...
 $ avg_neg_score                    : num  0.000048 0.0004 0.000689 0.0019 0.000048 1 0.000048 0.00567 0.145 1 ...
 $ p_score_bed_side                 : num  1.0e-06 1.2e-06 1.0e-06 1.0e-06 1.0e-06 ...
 $ p_score_Knowledge                : num  2.88e-05 1.74e-06 1.00e-06 1.00e-06 2.04e-03 ...
 $ p_score_communication            : num  0.720607 0.000001 0.000001 0.000001 0.000143 ...
 $ p_score_environment              : num  9.50e-05 6.99e-03 1.00e-06 1.00e-06 3.01e-06 ...
 $ p_score_scheduling               : num  2.81e-02 8.47e-06 1.00e-06 1.00e-06 2.22e-04 ...
 $ p_score_wait_time                : num  7.40e-04 4.54e-04 2.73e-06 2.73e-06 3.10e-04 ...
 $ p_score_costs                    : num  1.21e-03 1.08e-04 5.77e-06 5.77e-06 4.22e-04 ...
 $ avg_pos_score                    : num  1.07e-01 1.08e-03 1.93e-06 1.93e-06 4.49e-04 7.19e-01 2.87e-01 1.93e-06
...
 $ CME.Continuing.medical.Education: int  1 1 1 1 1 1 1 1 1 1 ...
 $ Improper.Treatment.Prescription : int  1 1 1 0 0 1 1 1 1 0 ...
 $ Patient.Harm                     : int  1 1 1 0 0 1 1 0 1 0 ...
 $ Unprofessional.Immoral.Contact   : int  0 0 0 0 0 0 0 0 0 0 ...
 $ Financial.Irregularities         : int  0 0 0 0 0 0 0 0 0 0 ...
 $ avg_count_pos_rev                : num  1 1 0 0 1.33 ...
 $ avg_count_neg_rev                : num  0 0 1 2 0 ...
```

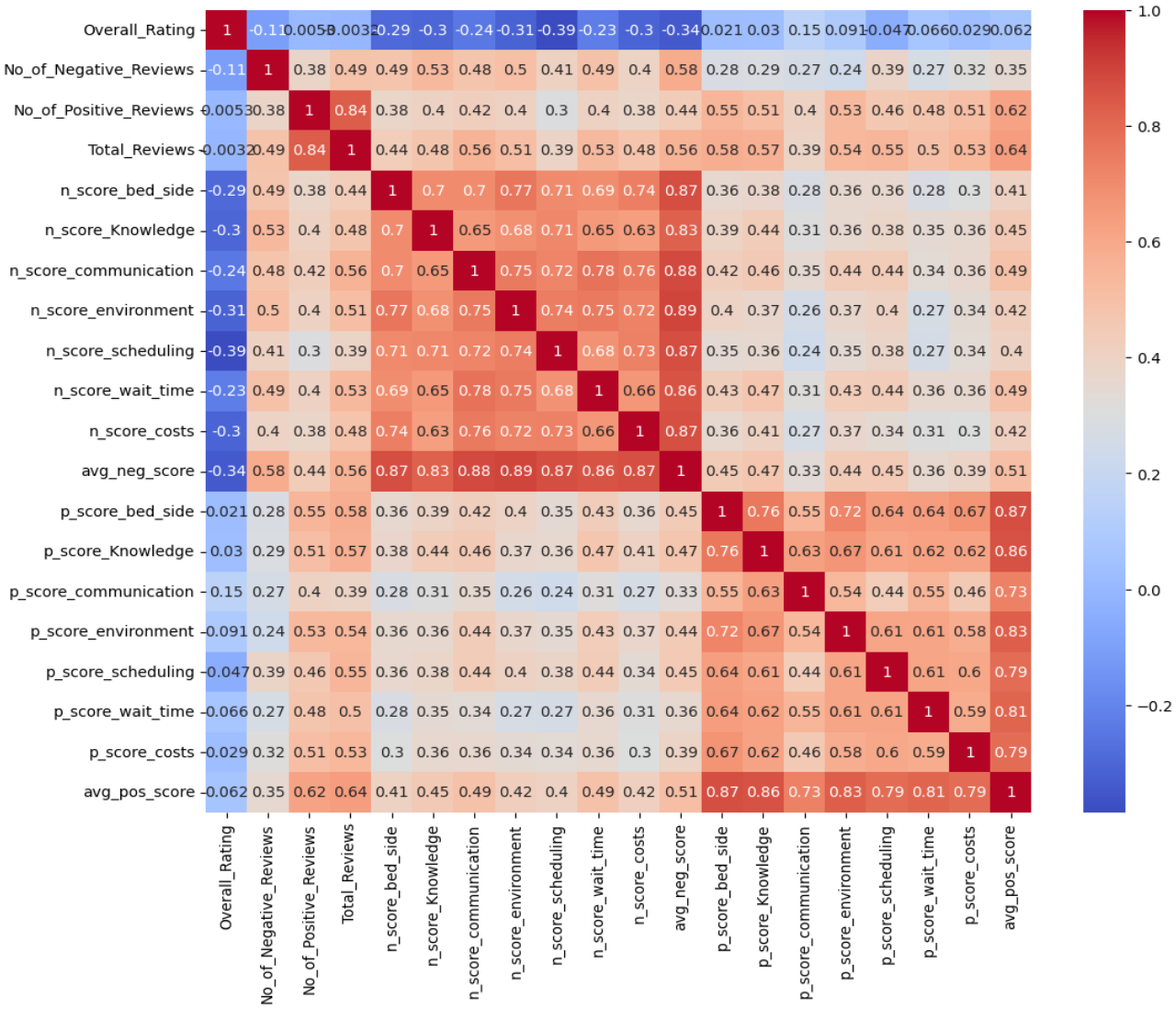
Appendix A2 - Results of the ANOVA test between Overall Rating and the following sanction reason topics

	df	Sum Square	Mean Square	F Value	Pr(>F)
Continuing Education Residuals	1 200	1.56 284.05	1.561 1.420	1.099	0.296
Improper Prescription Residuals	1 200	0.0 285.6	0.0038 1.4280	0.003	0.959
Patient Harm Residuals	1 200	0.51 285.10	0.5105 1.42I 55	0.358	0.55
Immoral Contact Residuals	1 200	0.09 285.91	0.0941 1.4276	0.066	0.798
Financial Irregularities Residuals	1 200	1.11 284.50	1.106 1.423	0.778	0.379

Appendix A3 - Results of the bar chart showing the distributions of sanctioned doctors by gender



Appendix A4 - Results of the Correlation Matrix



<u>Appendix A5</u> - Results of multiple regression models

	Residual SE	R-Squared	Adjusted R-squared	F-stat	p-value
Model 1	1.066 187 df	0.2553	0.1996	4.58	4.163e-07
Model 2	0.9048 196 df	0.4381	0.4238	30.57	2.2e-16
Model 3	1.074 122	0.1357	0.07198	2.129	0.03187
Model 4	1.071 181	0.2734	0.1931	3.405	5.8e-06
Model 5	1.178 188	0.08666	0.0235	1.372	0.1761
Model 6	1.199 196	0.01377	-0.01139	0.5473	0.7403
Model 7	1.081 180	0.264	0.1782	3.075	2.563e-05