How Do Personality Traits Affect Movie Ratings?

Stat 5740 Final Report

Team 13: Alex Cluff, Allison Nimmo, Haoming Wang

Introduction

As we searched on Kaggle for datasets, there was an interesting research studying the correlation between several personalities and movie ratings for twelve movies measured quantitatively. Since watching movies is a daily hobby for each group member, it did not take long for us to determine our topic for the project. Subsequently, we came up with our general question, that is “how do certain personality scores affect how certain people rate certain movies?” In this project report, we will specifically discuss our data sources and an overview of our data. Meanwhile, we will also split each of our specific questions related to the general one into an isolated section. In the end, we will provide a sum-up for all of our results.

Data

The data was found on Kaggle and contains both a participant’s personality scores from a Myers-Briggs test and their ratings for different movies they were assigned based on a predetermined metric.

Sources

The data was collected via online survey. It was split into two datasets - ratings and personality. Each observation in the ratings dataset contained a study participant, a specific movie id, and the rating that participant gave that movie. Each observation in the personality dataset contained a specific participant, their personality scores, predicted movie ratings, and various other metrics assigned to that specific participant.

Since the ratings dataset was so large and some participants rated upwards of 1,000 movies, we decided to compress it into a dataset containing the unique user id, and various summary statistics for that participant. Then, we joined the datasets on the userid variable.

Variables

The resulting merged dataset consisted of various personality scores on a scale of 1-7 and included: openness, agreeableness, emotional stability, conscientiousness, and extraversion. Based on those scores, twelve movies were selected and the observers predicted the participants’ ratings for those movies. We also included a variable called enjoy_watching, which is an expected enjoyment rating given to a list of 12 recommended movies by the participant. For the ratings dataset, as mentioned in the previous section, we compressed the thousands of

observations into summary statistics to make the data more manageable. In the merged dataset, we included a count of how many movies each participant rated, their mean rating, and the standard deviation/variance of their ratings.

Summary

With a data set having such a large sample size we expect to see many notable qualities. Most of which will be small numerically, but have a high power. In figure 1 we have a table of means for each of the traits with the mean rating for each of the possible values.

	value																							
	1		1.5		2		2.5		3		3.5		4		4.5		5		5.5		6		6.5		7	
	rating		rating		rating		rating		rating		rating		rating		rating		rating		rating		rating		rating	
	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N
agreeableness	3.00	4730	2.95	16677	3.35	24990	3.18	64489	3.37	92662	3.38	162895	3.36	214410	3.38	146422	3.46	87971	3.43	91287	3.39	65031	3.46	44360	3.54	12827
conscientiousness	3.00	2502	3.48	6651	3.44	11928	3.37	55663	3.37	84321	3.39	83093	3.45	126542	3.34	143436	3.29	134770	3.36	136917	3.43	126249	3.38	55552	3.29	61127
emotional_stability	3.47	8032	3.34	14598	3.29	24560	3.34	52972	3.38	86469	3.34	101150	3.24	151684	3.43	107757	3.39	113517	3.40	117551	3.39	118625	3.46	73692	3.43	58144
extraversion	3.38	53264	3.30	89142	3.33	122768	3.38	122587	3.25	129602	3.39	100721	3.40	132896	3.39	84336	3.43	68292	3.44	54638	3.58	38161	3.50	20688	3.51	11656
openness	2.89	1585	3.85	1363	3.31	6284	3.22	10635	3.39	13878	3.27	49730	3.46	73372	3.35	118577	3.42	162696	3.34	189252	3.33	193795	3.42	121654	3.39	85930

Figure 1

Specific Questions

Question 1 - Allison Nimmo

Is there a relationship between agreeableness/openness scores and whether or not the users enjoyed the movies watched?

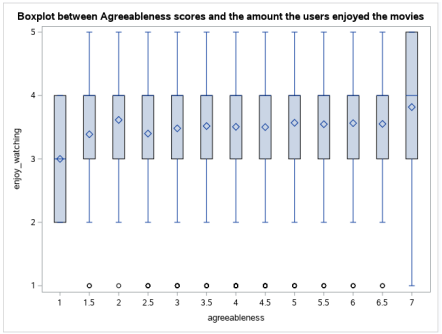


Figure 1.1

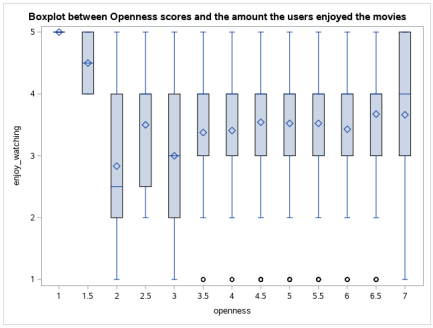


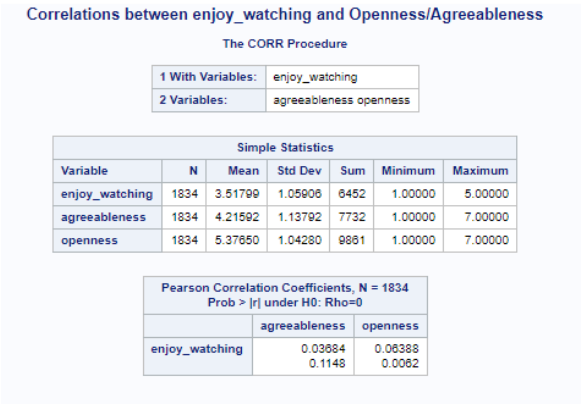
Figure 1.2

Methods

First, I created summary plots of what the distribution of enjoy_watching throughout Agreeableness and Openness was (shows in Figures 1.1 and 1.2, respectively). Then, I performed a simple correlation test using PROC CORR on both agreeableness against the variable enjoy_watching. This assumed that there is a linear relationship between the variables and there are no significant outliers. Next, I did an ANOVA test for both agreeableness and

openness against enjoy_watching. This compared the means of enjoy_watching across the different factors of the personality scores.

Results



Correlations between enjoy_watching and Openness/Agreeableness

The CORR Procedure

1 With Variables:	enjoy_watching
2 Variables:	agreeableness openness

Simple Statistics

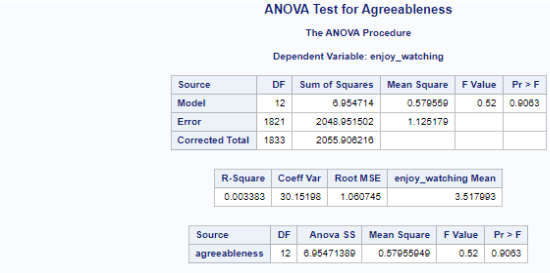
Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
enjoy_watching	1834	3.51799	1.05906	6452	1.00000	5.00000
agreeableness	1834	4.21592	1.13792	7732	1.00000	7.00000
openness	1834	5.37650	1.04280	9861	1.00000	7.00000

Pearson Correlation Coefficients, N = 1834
Prob > |r| under H0: Rho=0

	agreeableness	openness
enjoy_watching	0.03684 0.1148	0.06388 0.0062

Figure 1.3

Based on the results of the correlations table (shown in Figure 1.3), with a p-value of 0.1148, using an α-level of 0.05, there is not sufficient evidence that the correlation between enjoy_watching and agreeableness is not 0. Therefore, we can conclude that there is no correlation between agreeableness and enjoy_watching. For openness and enjoy_watching, we have a correlation coefficient of 0.063 and a p-value of 0.0062. Therefore, we can reject the null hypothesis and conclude that there is sufficient evidence that there is a correlation between openness and enjoy_watching.



ANOVA Test for Agreeableness

The ANOVA Procedure

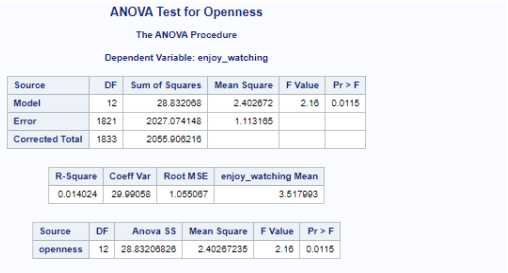
Dependent Variable: enjoy_watching

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	12	6.954714	0.579559	0.52	0.9063
Error	1821	2048.951502	1.125179		
Corrected Total	1833	2055.906216			

R-Square	Coeff Var	Root MSE	enjoy_watching Mean
0.003383	30.15198	1.060745	3.517993

Source	DF	Anova SS	Mean Square	F Value	Pr > F
agreeableness	12	6.95471389	0.57955949	0.52	0.9063

Figure 1.4



ANOVA Test for Openness

The ANOVA Procedure

Dependent Variable: enjoy_watching

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	12	28.832068	2.402672	2.16	0.0115
Error	1821	2027.074148	1.113165		
Corrected Total	1833	2055.906216			

R-Square	Coeff Var	Root MSE	enjoy_watching Mean
0.014024	29.99058	1.055067	3.517993

Source	DF	Anova SS	Mean Square	F Value	Pr > F
openness	12	28.83206826	2.40267235	2.16	0.0115

Figure 1.5

For the ANOVA test between agreeableness and enjoy_watching (shown in Figure 1.4), the F-statistic is 0.52 and the p-value is 0.9063. At an α-level of 0.05, there is not sufficient evidence to suggest that the means of enjoy_watching are different across the values of agreeableness. Therefore, we fail to reject the null. Since we failed to find sufficient evidence for both correlation and a difference in means, we conclude that there is little to no relationship between a participant's agreeableness score and whether or not they enjoyed the movies they rated.

For the ANOVA test between openness and enjoy_watching (shown in Figure 1.5), the F-statistic is 2.16 and the p-value is 0.0115. At an α-level of 0.05, we have sufficient evidence to suggest a difference in the means of enjoy_watching across the values of openness and can reject the null hypothesis. Since there was evidence for both a significant correlation and a difference in means, we can conclude there is a relationship between a participant's openness score and whether or not they enjoyed the movies they rated.

Question 2 - Alex Cluff

Are there significant differences between movie ratings given by people who are more or less extroverted?

Analysis Variable : rating		
extraversion_rd	N Obs	Mean
low	142406	3.3275073
medium	761202	3.3570655
high	125143	3.4987734

Figure 2.1

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	2459.631	1229.815	1208.22	<.0001
Error	1.03E6	1047137.395	1.018		
Corrected Total	1.03E6	1049597.026			

Figure 2.2

Methods

Assuming that the ratings are normally distributed, a single factor ANOVA test was conducted on the rating given different levels of extraversion. The original 13 levels, ranging from 1-7 counting by 0.5, have been rounded to the nearest factor of 3.5. These 3 levels at 0, 3.5, and 7 will be referred to as low, medium, and high respectively. The means for each group can be seen in figure 2.1. A multiple comparison test using Tukey's studentized range (HSD) test at significance level 0.05 was also conducted.

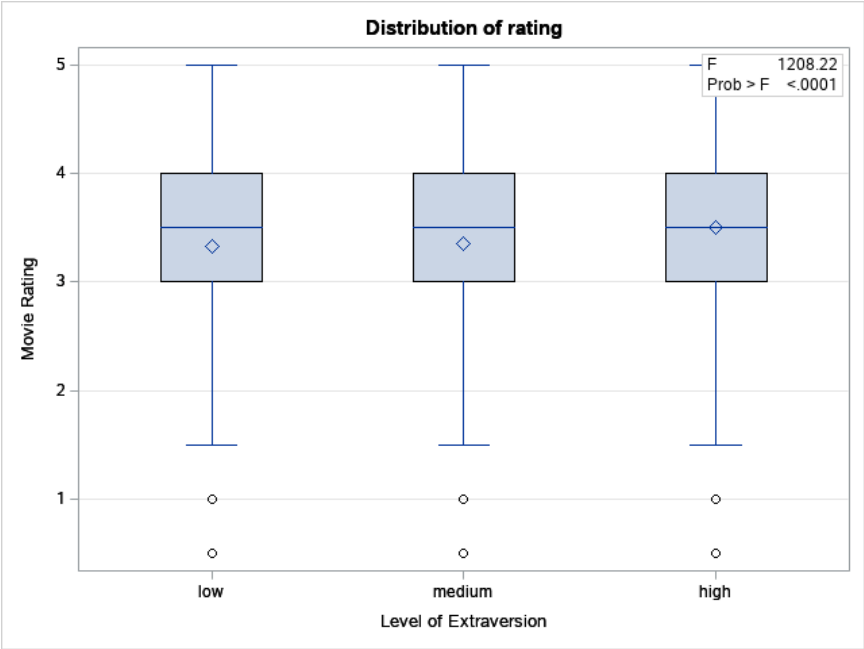


Figure 2.3

Results

It is clear from figure 2.3 that the mean rating is different for each level of extraversion. Low extraversion has the lowest average rating and high has the highest average rating (figure 2.1). The overall F-value of the ANOVA test is 1208.22 with a p-value of less than 0.0001. The HSD comparison test in figure 2.4 showed that all 3 comparisons have significant differences.

Comparisons significant at the 0.05 level are indicated by ***.				
extraversion Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
high - medium	0.141708	0.134495	0.148921	***
high - low	0.171266	0.162104	0.180428	***
medium - high	-0.141708	-0.148921	-0.134495	***
medium - low	0.029558	0.022731	0.036385	***
low - high	-0.171266	-0.180428	-0.162104	***
low - medium	-0.029558	-0.036385	-0.022731	***

Figure 2.4

Question 3 - Haoming Wang

Methods

There are two specific questions for this subsection: **a)** Is there a correlation between emotional stability scores and movie ratings? **b)** What's the relationship between the agreeableness scores and movie ratings? For both questions, I have used PROC REG

statements to construct a simple linear regression model for each movie rating related to the emotional stability and agreeableness scores respectively. In other words, we will get an expression of the fitted line, $y = \alpha + \beta x$ for some α, β, for each model. Then, based on the result tables presented in the SAS output, I have conducted an one-sample t test to test whether there is a predictive linear relationship between emotional stability or agreeableness scores and movie ratings, and H_0: $\beta = 0$ vs H_A : $\beta \neq 0$. In SAS, I first used PROC MEANS and PROC FREQ for the basic information about my variables in each question, specifically, the variable emotional_stability for emotional stability scores and the variable agreeableness for agreeableness scores. The figures 3.1, 3.2, and 3.3 below show some basic information and properties about the variables emotional_stability and agreeableness. (SAS, 2018)

Variable	N	Mean	Std Dev	Minimum	Maximum
agreeableness	1834	4.2159215	1.1379232	1.0000000	7.0000000
emotional_stability	1834	4.5616140	1.3925485	1.0000000	7.0000000

Figure 3.1

agreeableness	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	6	0.33	6	0.33
1.5	18	0.98	24	1.31
2	39	2.13	63	3.44
2.5	100	5.45	163	8.89
3	170	9.27	333	18.16
3.5	299	16.30	632	34.46
4	391	21.32	1023	55.78
4.5	273	14.89	1296	70.67
5	193	10.52	1489	81.19
5.5	146	7.96	1635	89.15
6	101	5.51	1736	94.66
6.5	76	4.14	1812	98.80
7	22	1.20	1834	100.00

Figure 3.2

emotional_stability	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	13	0.71	13	0.71
1.5	29	1.58	42	2.29
2	58	3.16	100	5.45
2.5	97	5.29	197	10.74
3	143	7.80	340	18.54
3.5	181	9.87	521	28.41
4	255	13.90	776	42.31
4.5	229	12.49	1005	54.80
5	186	10.14	1191	64.94
5.5	217	11.83	1408	76.77
6	204	11.12	1612	87.90
6.5	127	6.92	1739	94.82
7	95	5.18	1834	100.00

Figure 3.3

Results

Taking the result of movie 1 as an example, the figures 3.4 and 3.5 below have shown that the t statistic and scatterplot for testing whether the slope β_1 of the linear regression equation between emotional stability scores of movie ratings is equal to zero. Under the null hypothesis, I got p-value equal to 0.40 which is greater than $\alpha = 0.05$, so we failed to reject the null hypothesis. Similarly, I have conducted t-tests for the rest 11 movies and the corresponding p-values are 0.18, 0.32, 0.51, 0.38, 0.35, 0.80, 0.46, 0.48, 0.41, 0.81, 0.59 which are all greater than $\alpha = 0.05$,so we failed to reject the null hypothesis for each of these cases either.

On the other hand, the figures 3.6 and 3.7 have shown that the t statistic and scatterplot for testing whether the slope $\beta_2$of the linear regression equation between agreeableness scores and movie ratings is equal to zero. I got a p-value equal to 0.0003 which is less than $\alpha = 0.05$,

so I reject the null hypothesis. Similarly, I have conducted t-tests for the rest 11 movies and the corresponding p-values are <.0001, 0.002, 0.0027, 0.0001, 0.0008, 0.0044, <.0001, 0.,0006, 0.0037, 0.0019, <.0001. Since all of them are less than $\alpha = 0.05$, I can conclude that there exists a linear correlation between the emotional stability scores and the movie ratings and there is strong evidence that the audiences' agreeableness scores and movie ratings are not linearly correlated.

Since there are 1834 recorded participants for each test, by the Law of Large Numbers, it is reasonable to assume the samples satisfy a normal distribution and that participants' responses are independent of each other. So, it is eligible for us to conduct t tests in this case.

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	4.33833	0.03068	141.38	<.0001
emotional_stability	1	0.00539	0.00643	0.84	0.4020

Figure 3.4

Figure 3.5

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	4.24268	0.03426	123.82	<.0001
agreeableness	1	0.02852	0.00785	3.63	0.0003

Figure 3.6

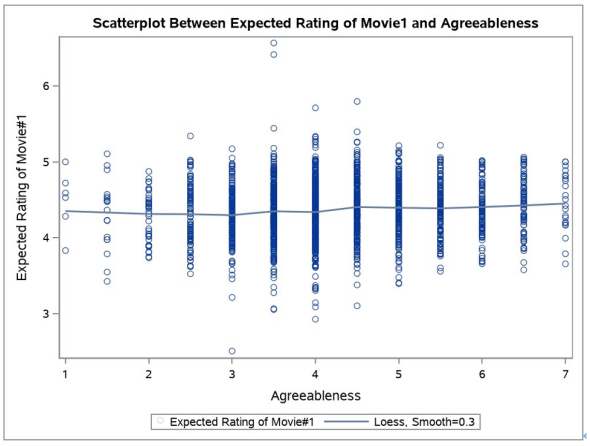


Figure 3.7

Discussion

During our analysis we have concluded that personality traits affect the rating of movies in a number of different ways. We have found through testing that a high rating is associated with high extraversion. This result agrees with our understanding of the trait. We also found that higher openness correlates to a higher expected enjoyability, while higher agreeableness does not.

We came across a few limitations when doing our analysis. It is evident that we used very few scatterplots for traits compared to ratings. The data was discrete for both the traits and the ratings. This made it difficult when trying to create graphs that represent the data in a valuable way. We have found it more valuable to use mean tables to view differences. We also found that trait did not have as large of an effect on ratings as we expected. The differences in means for question 2 were very small. We were only able to conclude significance because of the large sample size.

References

SAS Institute, Inc. (2018) SAS Studio 3.8 (Enterprise Edition) for Linux [Computer software]. Cary, NC.

"SAS Documentation." *SAS Customer Support Site*, support.sas.com/en/documentation.html.

Ali, Arslan. "Top Personality Dataset." Kaggle, September 27, 2020. https://www.kaggle.com/arslanali4343/top-personality-dataset