

ISDS 577 CAPSTONE SEMINAR

Spring 2018 session

Musical Preferences with Cognitive Characteristics & Other Variables



Team Members

Harshal Saptarshi
Malvika Kulkarni
Shashwat Khambhatta
Taranjot Singh
Vishal Vatnani

OCEAN: The Five Factors

- Openness to Experience
- Conscientiousness
- Extroversion
- Agreeableness
- Neuroticism (Emotional Stability)



Brief Mood Introspection scale

- Fed up
- Gloomy
- Jittery
- Nervous
- Sad
- Calm



INTRODUCTION

- Research on personality types, Intelligence (GPA), musical genres and behavioral traits.
- Personalized survey form generated using Qualtrics
- The collection of data from different data sources:
 - Mechanical Turk - Online
 - Social Media Online
 - Reddit – Online
 - Physical Survey – Campus
 - In Class Survey – Campus
- 28 variables and 978 rows of data



qualtrics





VARIABLES/ATTRIBUTES

Variables	Variable Type	Number of Variables in the category
Big-5 Personality Dimensions (Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism)	Interval	5
Mood (Negativity rate, calmness rate)	Interval	2
Grade Point Average (GPA)	Range data (Categorical or Continuous) - depends on the distribution of the variable	1
College Year	Ordinal	1
Music_Rating_Clips	Interval - Rating across six audio clips	6

Music_Personal_Preferences	Interval	6
Musical_Choice	Nominal	1
Time_Spent_On_Music	Range data (Categorical or Continuous) - depends on the distribution of the variable	1
Race	Nominal	1
Age	Range data (Categorical or Continuous) - depends on the distribution of the variable	1
Gender	Binary (symmetric)	1
Commuter_Student	Binary (asymmetric)	1

Let's explore geographical locations and musical preferences

https://public.tableau.com/profile/harshal.saptarshi#/vizhome/plot_world/Sheet1

DATA EXPLORATION

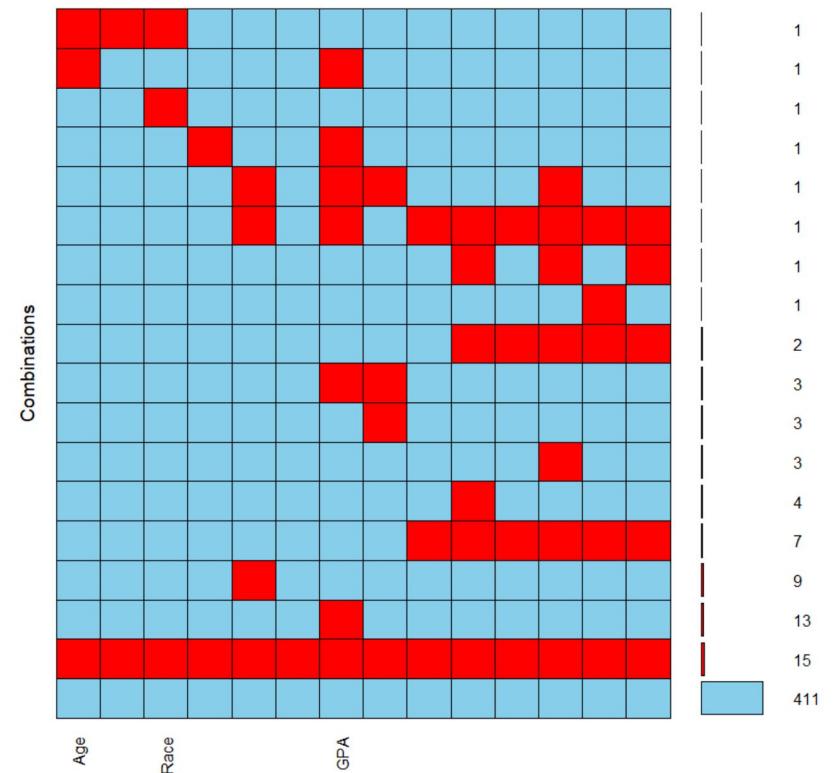
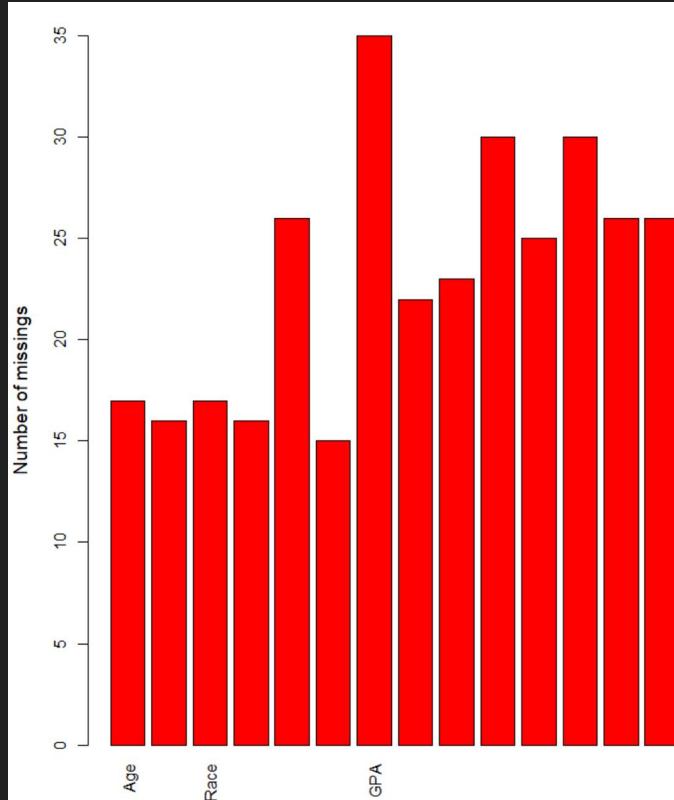
Missing Values



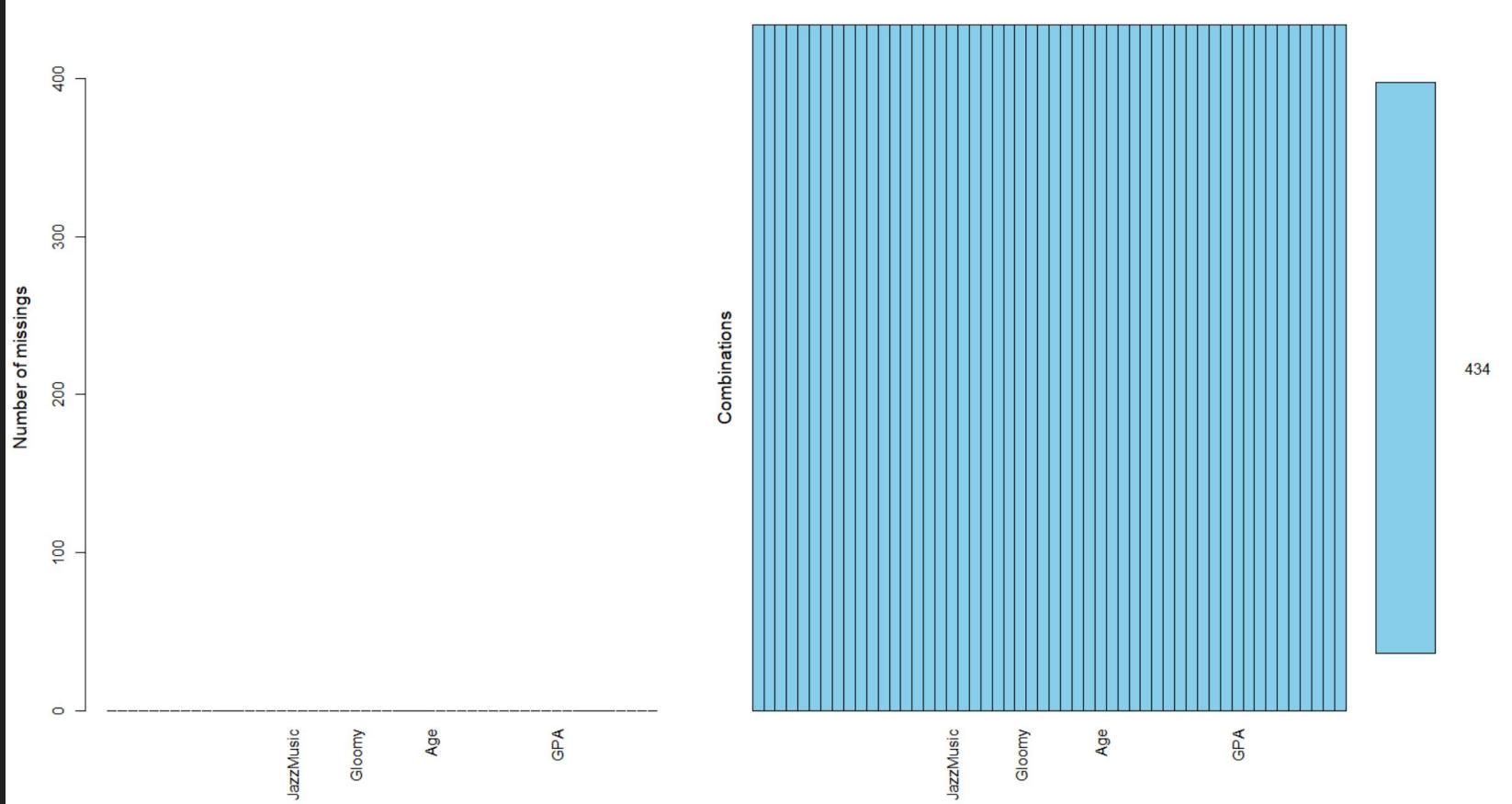
Outliers



Missing rows combinations

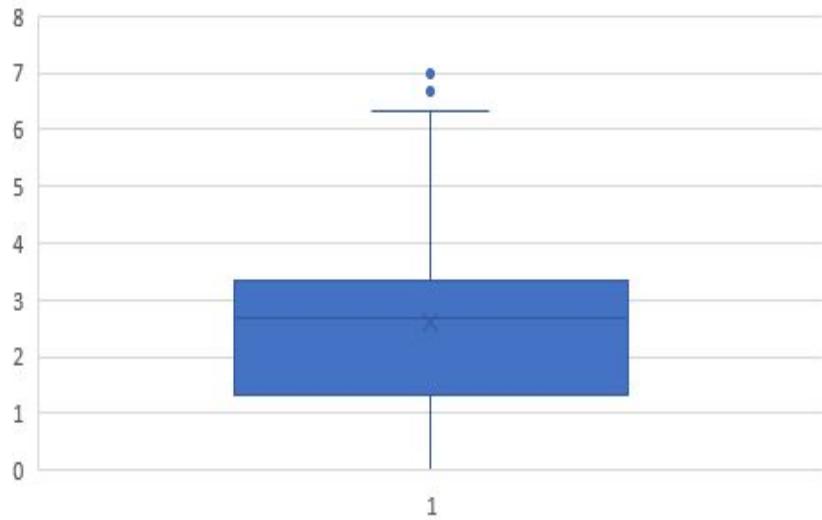


Missing rows combinations

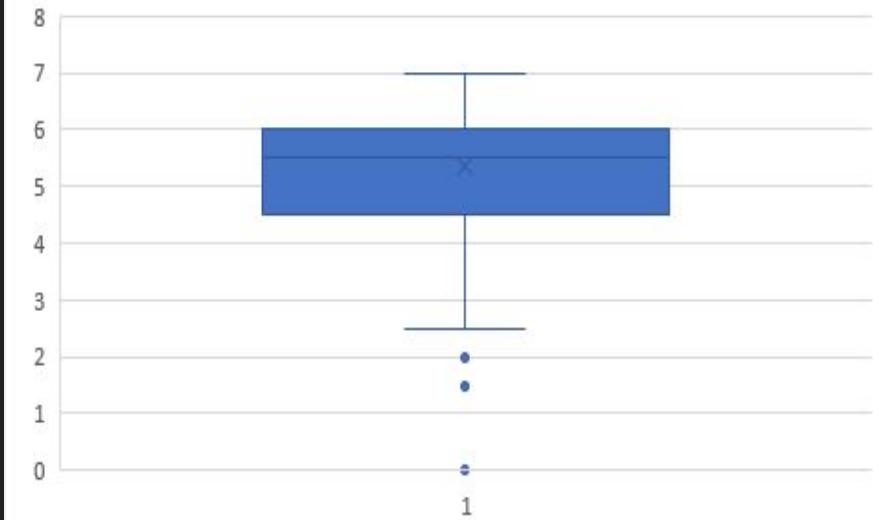


Box Plots and Outliers for unprocessed data

Scaled negative

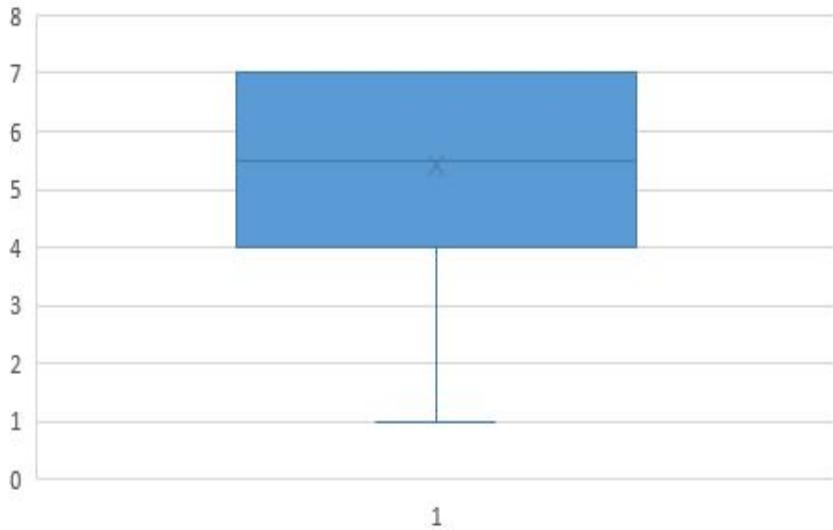


Openness to Experiences

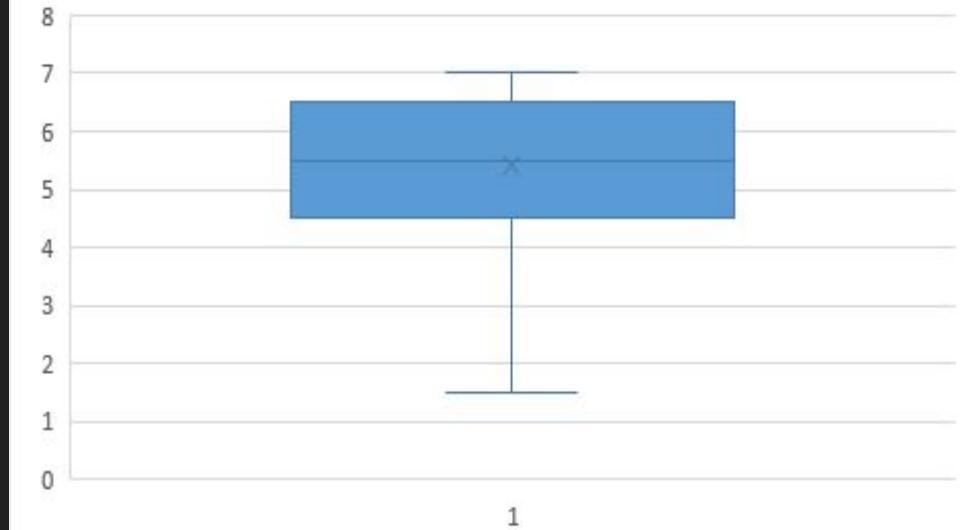


Box Plots and Outliers for cleaned data

Soft Rock/Pop



Openness to Experiences



DATA MERGING

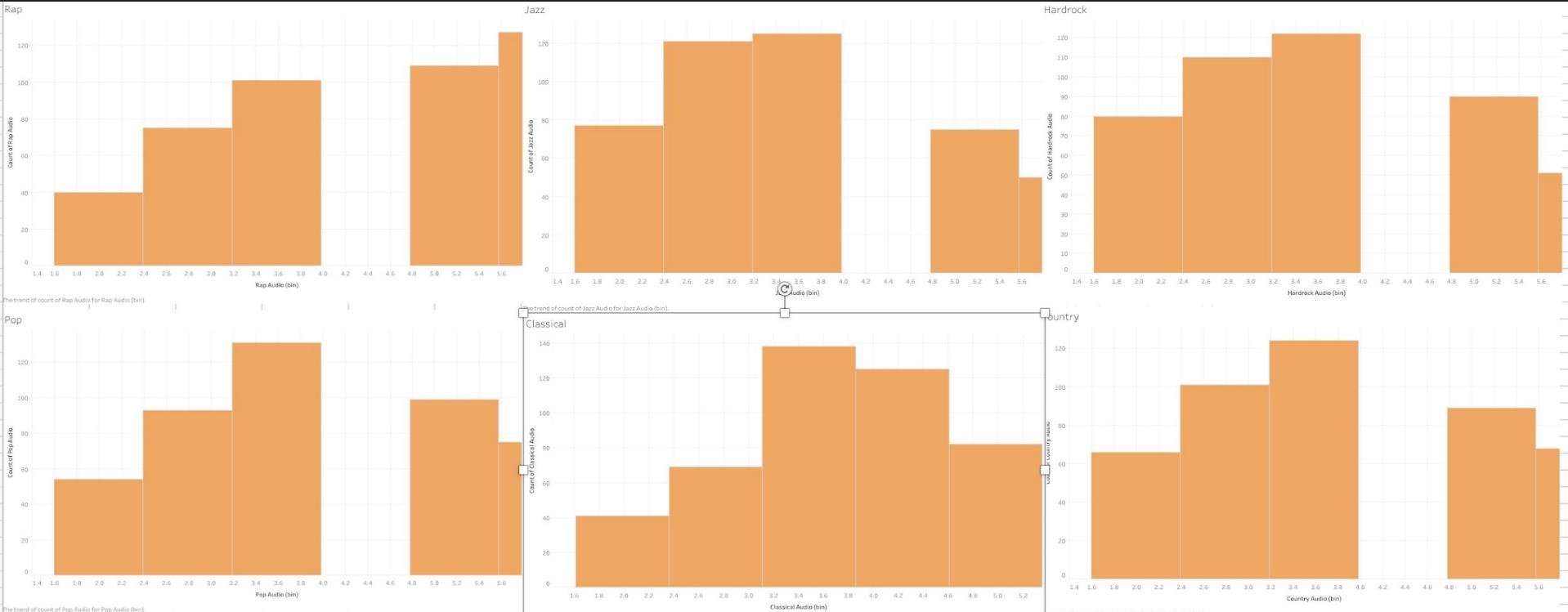
1. **Anova / t-test** - for continuous variables like Big-5 Personality Dimensions
2. **Kruskal wallis test** - for range data(ordinal variable) like GPA, Age
3. **Fisher exact test** - nominal variable like Race
4. **Chi square test** - nominal variable like Musical choice
5. **Logistic / chi squared test** - for binary variable like Gender

Example of a t-test for continuous variable

577	577_survey				
5	6				
7	5			t-Test: Two-Sample Assuming Unequal Variances	
5	7				
7	7				
				Variable 1	Variable 2
5	7			Mean	6.032967033
4	7			Variance	0.987789988
5	3			Observation	91
5	4			Hypothesized	0
7	6			df	104
5	5			t Stat	1.04907288
3	5			P(T<=t) one	0.148288183
6	6			t Critical one	1.659637437
7	7			P(T<=t) two	0.296576366
7	5			t Critical two	1.983037526
6	7				
6	6				
6	7				
5	7				

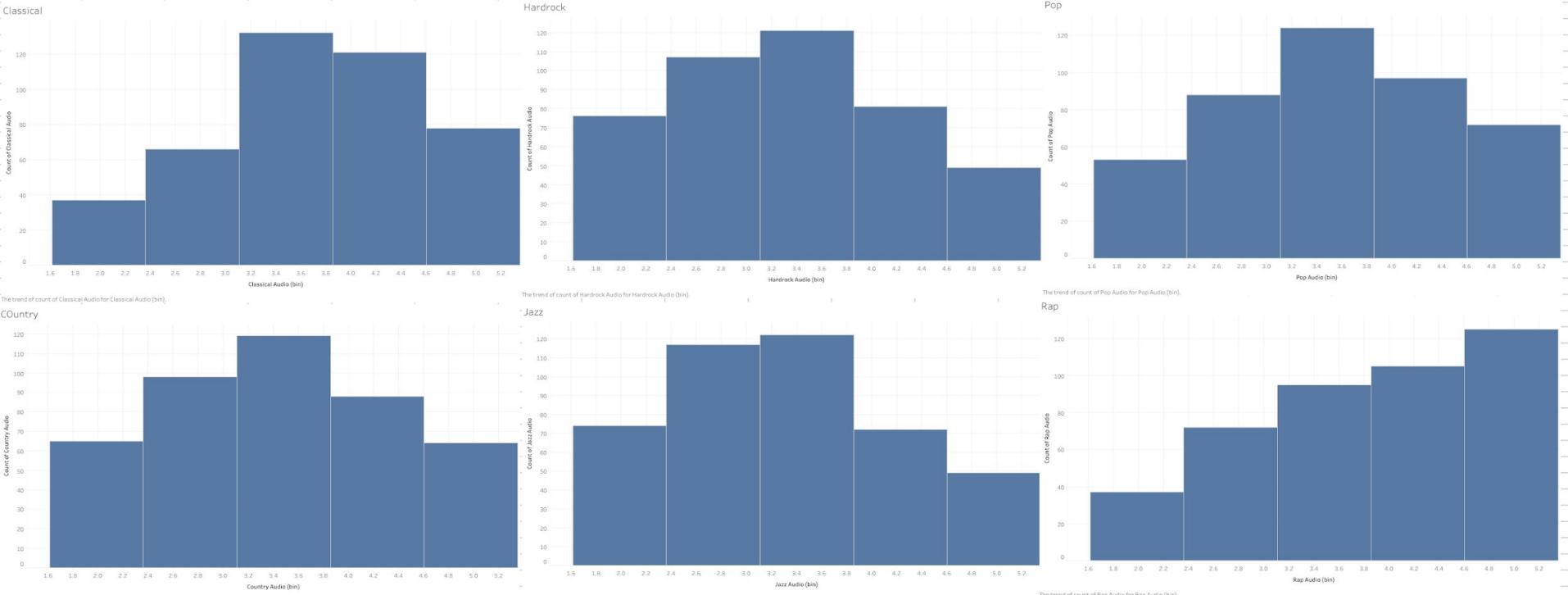
Visualisation - Before Data cleaning

Audio Ratings

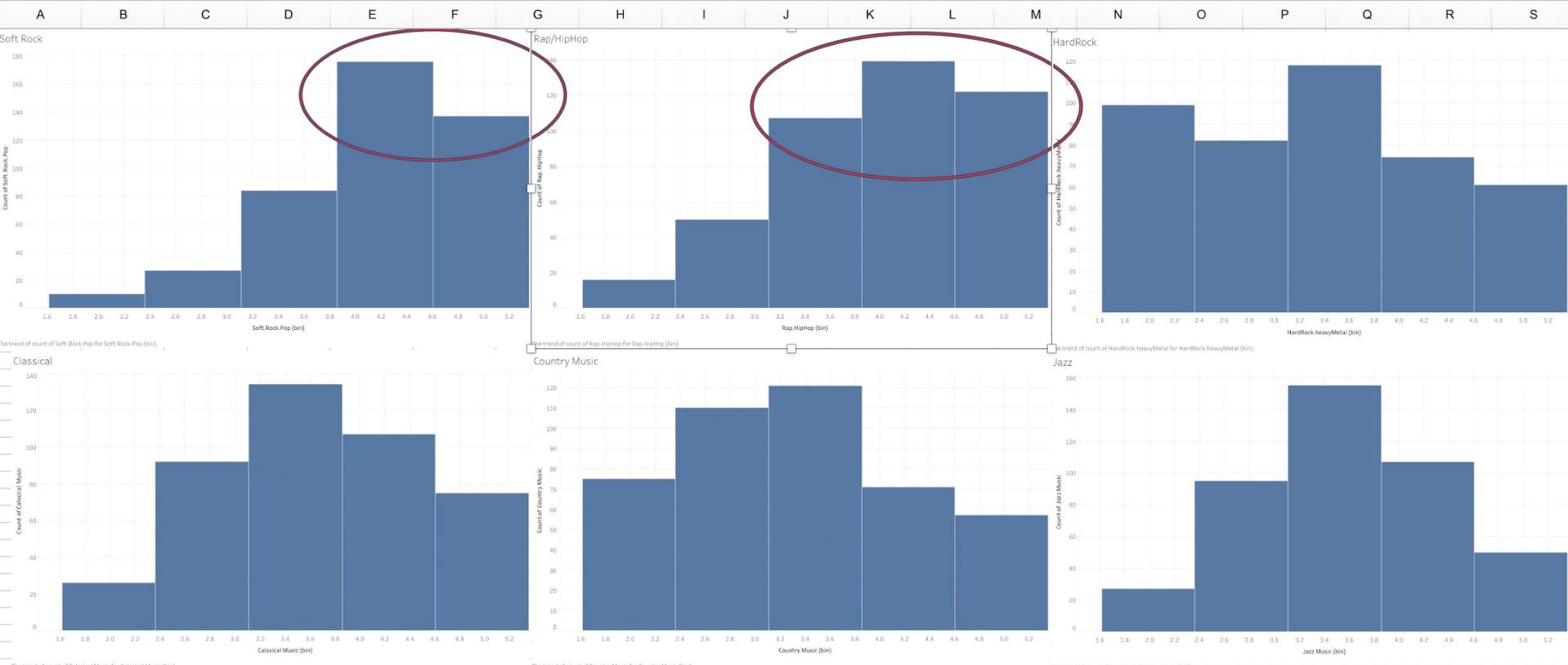


Visualisation - After Data cleaning

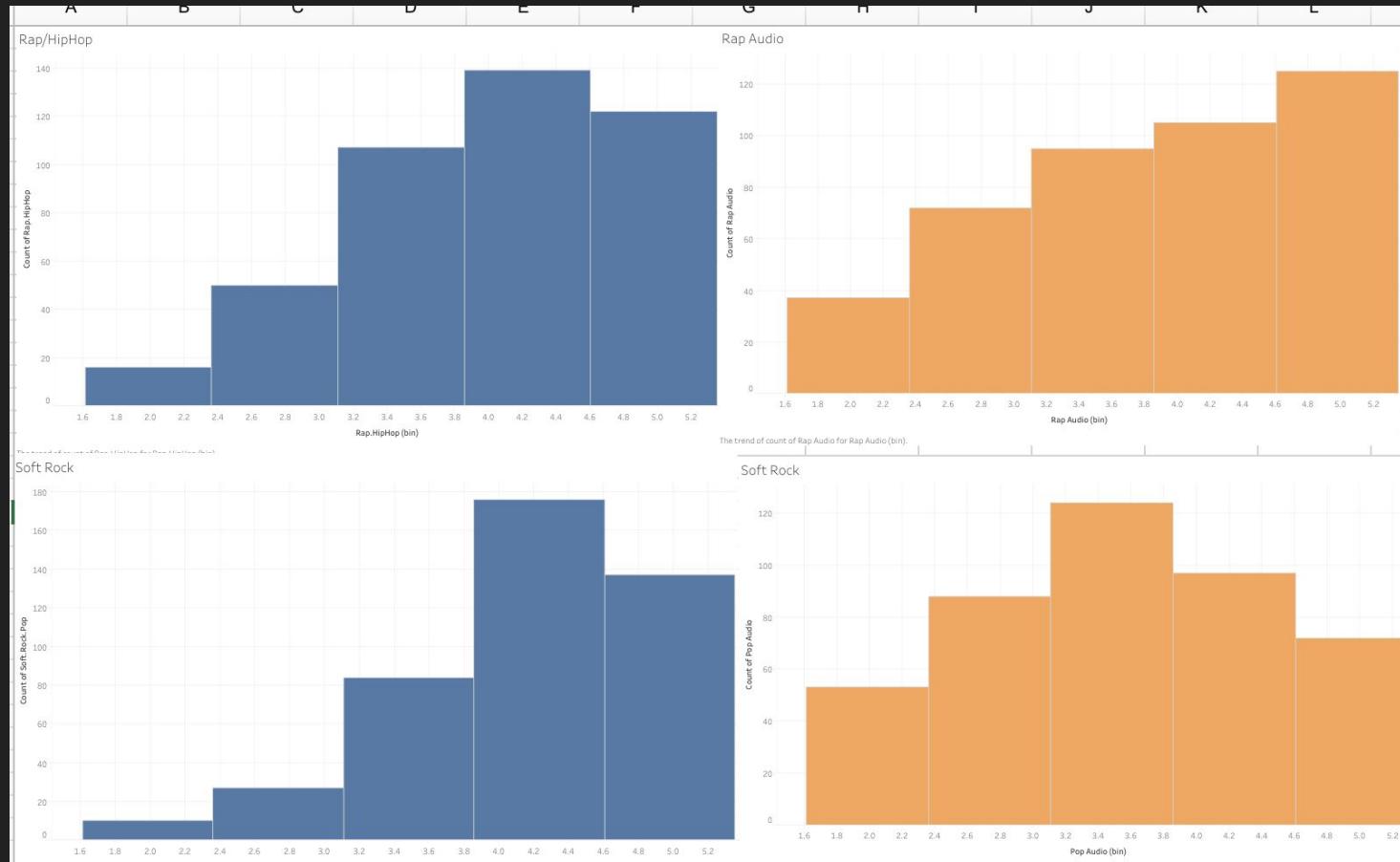
Audio Ratings



Preferred Music Genres



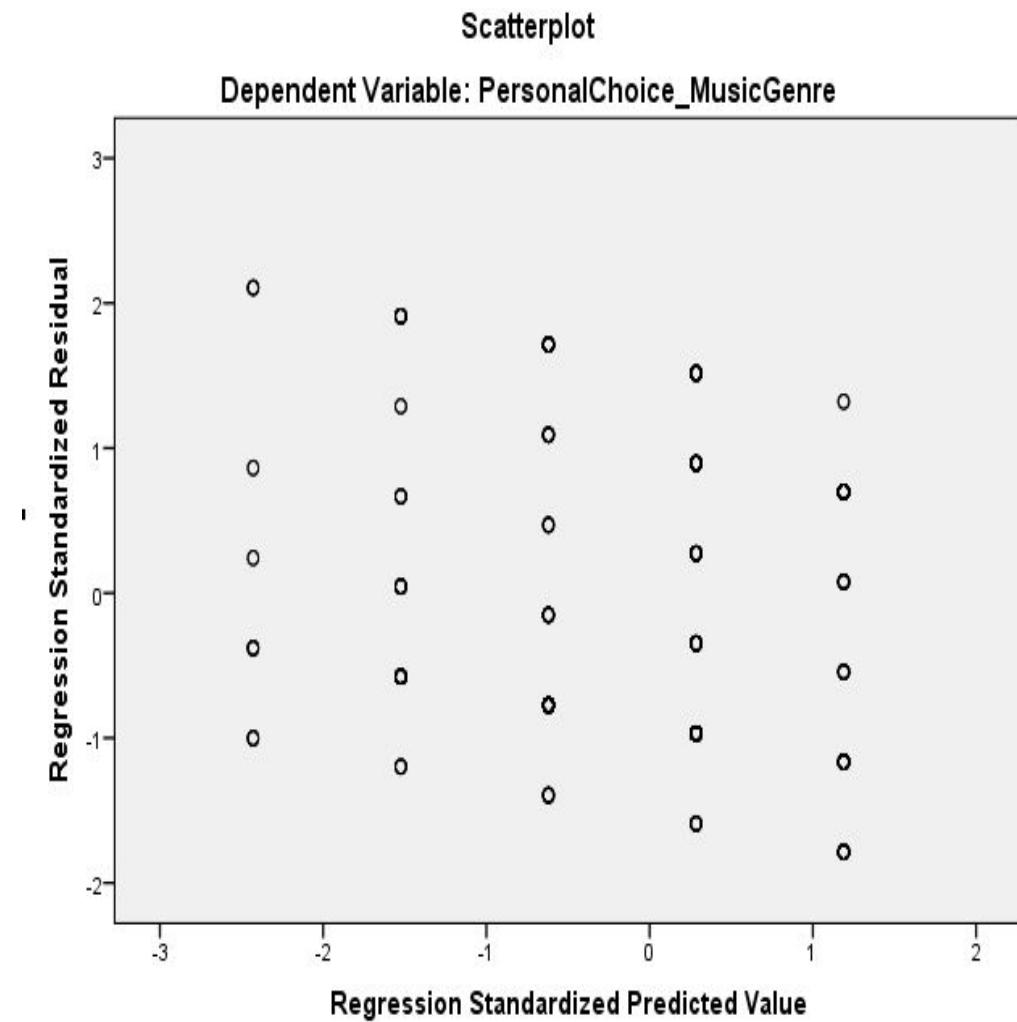
Relation between Most preferred Music genre and Highest rated Music genre



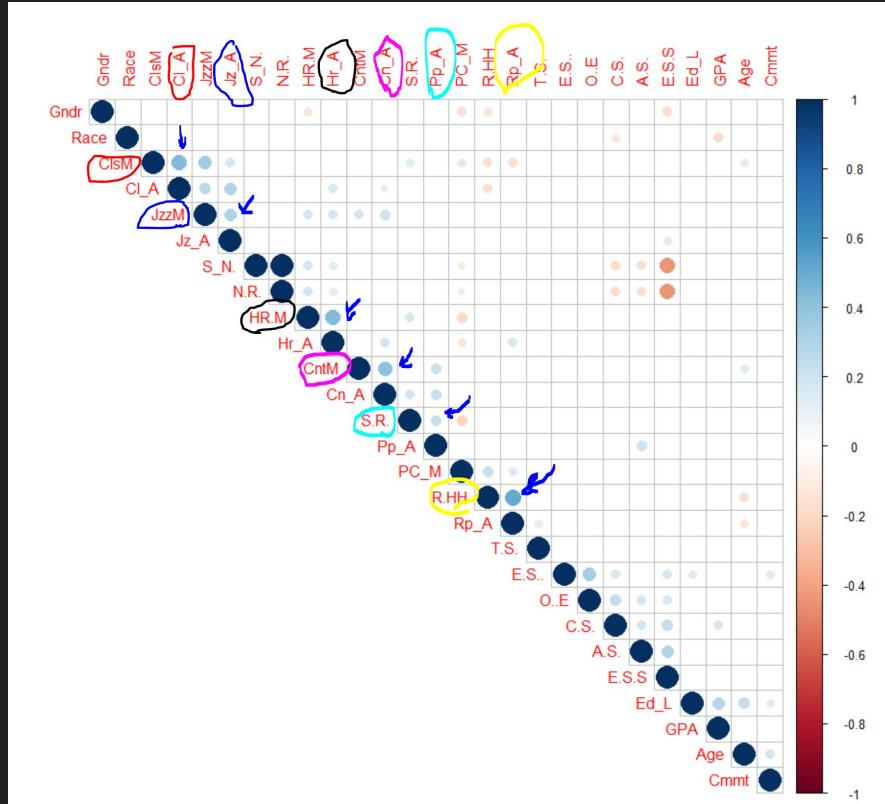
GENERAL ASSUMPTIONS

- Linearity
- Multicollinearity
- Multivariate normality
- Homoscedasticity
- Data type requirements
- Autocorrelation

Model	Collinearity Statistics		
	Tolerance	VIF	
1	Agreeableness Score (out of 7)	.773	1.294
	Conscientiousness Score (out of 7)	.775	1.290
	Emotional Stability Score (out of 7)	.619	1.616
	Openness to Experiences Score (out of 7)	.752	1.329
	CountryMusic	.708	1.413
	JazzMusic	.672	1.489
	Rap/HipHop	.619	1.616
	CalssicalMusic	.676	1.479
	Soft Rock/Pop	.848	1.179
	HardRock/heavyMetal	.701	1.427
	Fedup	.576	1.735
	Gloomy	.535	1.871
	Jittery	.602	1.660
	Nervous	.612	1.634
	Sad	.511	1.955
	Calm	.698	1.434
	Age	.808	1.237
	Gender	.824	1.214
	Race	.865	1.156
	Education_Level	.748	1.336
	Time Spent on Music	.913	1.096



Correlation between audio ratings, genre ratings and all variables



Ordinal Logistic Regression: (ref: Soft Rock/Pop)

Most Fav. Genre ~ Big 5 Personality scores

Coefficients:

	(Intercept)	Extraversion.Score..out.of.7.	Openness.to.Experiences.Score..out.of.7.
classical	-1.7312772	0.017931001	0.17280303
Country	-0.3726293	0.022095832	-0.37290380
Hard rock	-2.2288703	-0.138158352	0.26436890
Jazz	-1.1963100	0.009115382	-0.04863922
Rap	-0.1375152	0.107981789	-0.07002090

Ordinal Logistic Regression: (ref: Soft Rock/Pop)

Most Fav. Genre ~ Big 5 Personality scores

	Agreeableness.Score..out.of.7.	Emotional.stability.Score..out.of.7.
classical	-0.2566871	0.1003680
Country	-0.2261086	0.2494814
Hard rock	-0.1607237	0.1080100
Jazz	-0.3207600	0.2699459
Rap	<u>-0.3318626</u>	<u>0.2990021</u>
	Conscientiousness.Score..out.of.7.	
classical	0.057063375	
Country	0.129676583	
Hard rock	-0.004419483	
Jazz	0.005059337	
Rap	0.005012801	

Ordinal Logistic Regression: (p-values)

Most Fav. Genre ~ Big 5 Personality scores

```
> print(p, cutoff = 0.05)
      (Intercept) Extraversion.Score..out.of.7. openness.to.Experiences.Score..out.of.7.
classical    0.1270737                      0.8775806                      0.30190815
Country      0.7548012                      0.8662386                      0.03097859
Hard rock    0.1052122                      0.3167130                      0.19325297
Jazz          0.3684537                      0.9492056                      0.80405698
Rap           0.8637322                      0.2096376                      0.55203944
      Agreeableness.Score..out.of.7. Emotional.stability.Score..out.of.7.
classical     0.091087185                   0.431971621
Country       0.189751022                   0.100657171
Hard rock     0.379548418                   0.480196807
Jazz          0.083750317                   0.097299762
Rap           0.003291264                   0.002073778
```

Q1: Are an individual's Big-5 personality characteristics predictors of their music genre preference (most favorite)?

	Softrack/Pop	Country	Rap
Openness		-0.327	
Agreeableness			-0.346
Emotional Stability			0.305
Extroversion	-0.323		

Q1.1: Personality types with musical audio ratings

	Softrock/Pop	Country	Rap	Jazz	Classical	Hardrock
Openness						
Agreeableness	0.333					
Emotional Stability				0.225		
Extroversion					-0.155	
Conscientiousness				-0.227		

Q3: If mood of an individual influences musical preferences and ratings

	Country	Rap	Jazz	Classical	Softrock/P op	Hardrock
Mood	0.177	-1.29	NA	NA	0.129	0.289
Mood with audio ratings	NA	NA	-0.159	NA	NA	0.204

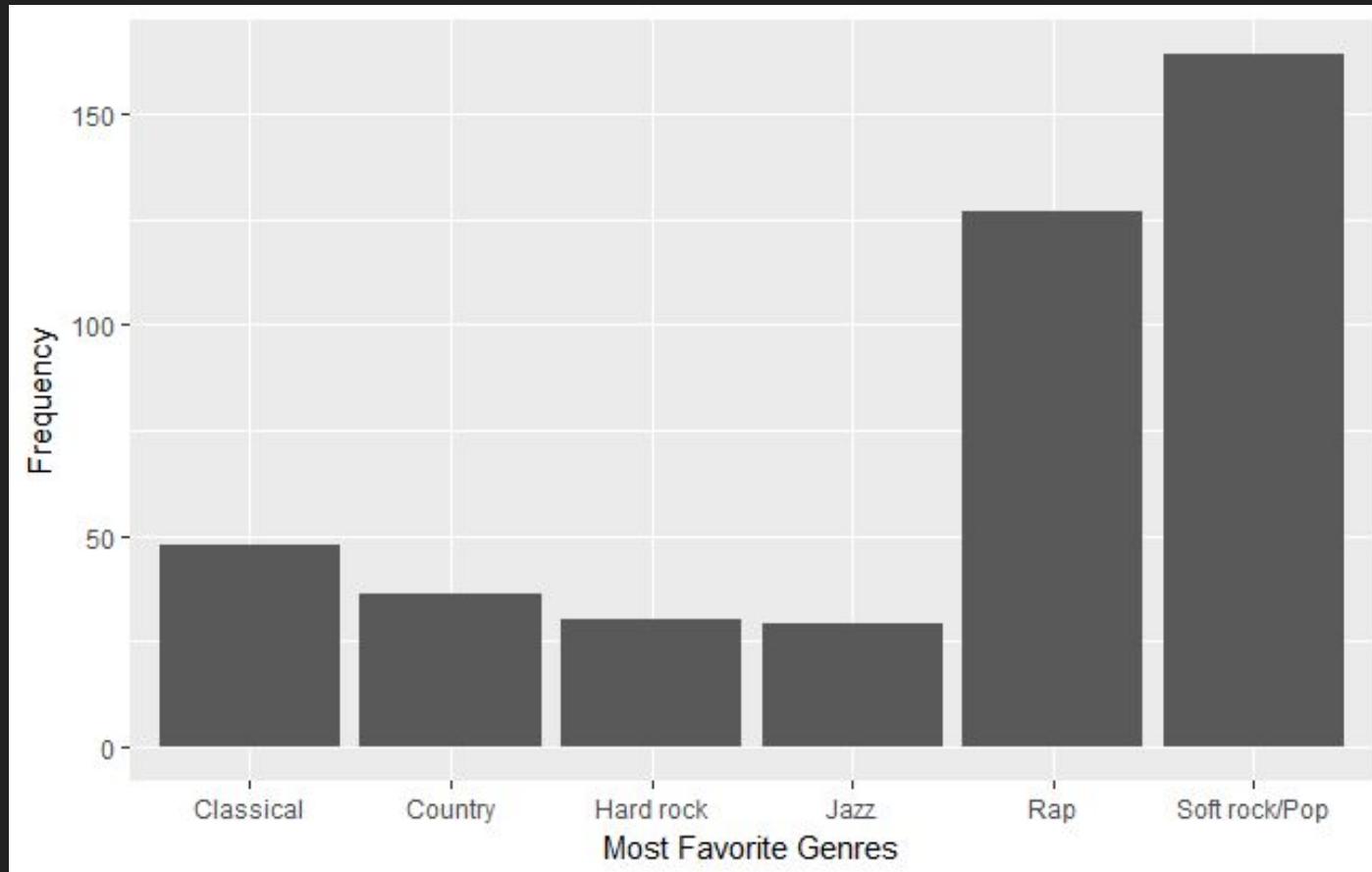
The mood scale is rated as a scale of negative feelings

The highest significance - Negative mood with Hardrock!

Q8 & 4: Demographics with musical audio ratings

	Softrock/Pop	Country	Rap	Jazz	Classical	Hardrock
Age		0.257	-0.304		0.190	
Education		0.302				
GPA						
TimeSpent OnMusic			0.136			
Race						

Most Favorite Genre



Factor Analysis

Factor Analysis on Mood Scale

Loadings:

	Factor1	Factor2	Factor3
Fedup	0.714		
Gloomy			0.989
Jittery			
Nervous		1.016	
Sad	0.620		
calm	-0.655		

Factor Analysis on Big 5 items

Loadings:

	Factor1	Factor2	Factor3	Factor4	Factor5
Extraverted_Enthusiastic			0.797		
Critical_Quarrelsome			0.322		
Dependable_selfDisciplined					-0.316
Anxious_Upset	1.054				
Open.to.new.experiences_complex			0.403		
Reserved_Quiet		-0.475			
Sympathetic_Warm			0.657		
Disorganized_careless				0.872	
calm_Emotionallystable			0.390		
Conventional_Uncreative	1.019				

Factor Analysis on genre ratings

Loadings:

	Factor1	Factor2	Factor3
CountryMusic			
JazzMusic	0.929		
Rap.HipHop			1.001
CalssicalMusic	0.377		
Soft.Rock.Pop		1.014	
HardRock.heavyMetal			

Factor Analysis on audio ratings

Loadings:

	Factor1	Factor2	Factor3
classical_Audio	1.010		
Jazz_Audio	0.308		
Hardrock_Audio			
Country_Audio		1.007	
Pop_Audio			
Rap_Audio			0.642

Q2: If the relationship in (1) holds true, then is the relationship mediated by their mood (negative-calm) or demographics?

Q2: Part 1: Relationship between Mood and Personal Choice (p-value)

Most Fav. Genre ~ Mood

	(Intercept)	scaled_Negative.Relaxed_Mood_seven_point
classical	7.222434e-02	0.11373566
Country	3.886218e-06	0.17039272
Hard rock	2.546202e-04	0.98277345
Jazz	1.750229e-04	0.94386037
Rap	2.392432e-01	0.02031162

Q2: Part 1: Relationship between Mood and Personal Choice (coefficients)

Most Fav. Genre ~ Mood

Coefficients:

	(Intercept)	scaled_Negative.Relaxed_Mood_seven_point
Classical	-0.6747847	-0.218509153
Country	-2.0509878	0.186465484
Hard rock	-1.6896477	-0.003334952
Jazz	-1.7623700	0.010980568
Rap	0.3211505	-0.228201897

Q2: Part 2 Relationship between Music Choice and Agree., Open., and Emot. (p-value)

Most Fav. Genre ~ Agree. + Emot. + Open.

[[1]]

	(Intercept)	Openness.to.Experiences.Score..out.of.7.	Agreeableness.Score..out.of.7.
Classical	0.14554417	0.21165053	0.091735369
Country	0.98598802	<u>0.03759428</u>	0.207931404
Hard rock	0.05701621	0.29036285	0.427870753
Jazz	0.34874520	0.81432834	0.079486142
Rap	0.90685214	0.84275837	<u>0.002084356</u>
	Emotional.stability.Score..out.of.7.		
Classical	0.3664952847		
Country	0.0647217305		
Hard rock	0.5618008084		
Jazz	0.0876197571		
Rap	<u>0.0008870731</u>		

Q2: Part 2 Relationship between Music Choice and Agree., Open., and Emot. (coefficients)

Most Fav. Genre ~ Agree. + Emot. + Open.

Coefficients:

	(Intercept)	Openness.to.Experiences.Score..out.of.7.	Agreeableness.Score..out.of.7.
Classical	-1.54659564	0.19551673	-0.2543398
Country	0.01922471	-0.33288308	-0.2151655
Hard rock	-2.47235357	0.20136410	-0.1450952
Jazz	-1.16291936	-0.04293192	-0.3231542
Rap	0.08723938	-0.02170001	-0.3431090
	Emotional.stability.Score..out.of.7.		
Classical		0.1127151	
Country		0.2750013	
Hard rock		0.0868882	
Jazz		0.2728999	
Rap		0.3158340	

Q2: Part 3: Relationship between Emotional Stability and Mood

Mood ~ Emot.

coefficients:

	Estimate	std. Error	t value	Pr(> t)	
(Intercept)	4.4889	0.1919	23.39	<2e-16	***
Emotional.stability.Score..out.of.7.	-0.4042	0.0393	-10.28	<2e-16	***

signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.136 on 432 degrees of freedom

Multiple R-squared: 0.1967, Adjusted R-squared: 0.1948

F-statistic: 105.8 on 1 and 432 DF, p-value: < 2.2e-16

Q2: Part 4: Relationship between Music Choice and Emot. (coefficients)

Most Fav. Genre ~ Emot.

Coefficients:

	(Intercept)	Emotional.stability.score..out.of.7.
Classical	-1.553497	0.07129077
Country	-2.254626	0.15910065
Hard rock	-2.043605	0.07564257
Jazz	-2.563269	0.17831429
Rap	-1.279978	<u>0.21811426</u>

Q2: Part 4: Relationship between Music Choice and Emot. (p-value)

Most Fav. Genre ~ Emot.

	(Intercept)	Emotional.stability.Score..out.of.7.
classical	0.0063179456	0.54761893
Country	0.0006643587	0.23601637
Hard rock	0.0031002179	0.59816202
Jazz	0.0004658448	0.22623254
Rap	0.0028024147	0.01221671

Q2: Part 4: Relationship between Music Choice and Emot. and Mood (p-value)

Most Fav. Genre ~ Emot. + Mood

	(Intercept)	Emotional.stability.Score..out.of..7.	scaled_Negative.Relaxed_Mood_seven_point
classical	0.501458571	0.89985636	0.1368307
Country	0.000186644	0.04538003	0.0318782
Hard rock	0.034732731	0.56290055	0.8086395
Jazz	0.004828820	0.16515970	0.4841263
Rap	0.324578775	0.09910067	0.1721930

Q2: Part 4: Relationship between Music Choice and Emot. and Mood (coefficients)

Most Fav. Genre ~ Emot. + Mood

Coefficients:

	(Intercept)	Emotional.stability.Score..out.of..7.	scaled_Negative.Relaxed_Mood_seven_point
classical	-0.5742174	-0.01659637	-0.22833071
Country	-3.8843906	<u>0.30535469</u>	0.33159520
Hard rock	-2.2362944	0.09298969	<u>0.04198518</u>
Jazz	-3.1350810	0.22962857	0.12256256
Rap	-0.6284182	0.15971937	-0.14931489
,			

Q2: Part 5: Relationship between Music Choice and Race (p-value)

Most Fav. Genre ~ Race

	(Intercept)	Race_textAmerican Indian or Alaska Native	Race_textBlack or African American	Race_textNative Hawaiian or Pacific Islander	Race_textOther Asian	Race_textWhite
Classical	8.195851e-08	0	0.98413330			
Country	2.161502e-10	0	0.79982495			
Hard rock	9.840795e-11	0	0.71776099			
Jazz	7.029466e-11	0	0.98565001			
Rap	1.471853e-05	0	0.03488202			

	Race_textOther	Race_textAsian	Race_textWhite
classical	0.7589866e-01	0.7232716	0.590245423
Country	0.6383667e-01	0.4684076	0.023895953
Hard rock	0.5720301e-01	0.1683538	0.094028060
Jazz	0.7858117e-01	0.9999934	0.255765845
Rap	0.4616102e-05	0.3632423	0.003037448

Q2: Part 5: Relationship between Music Choice and Race (coefficients)

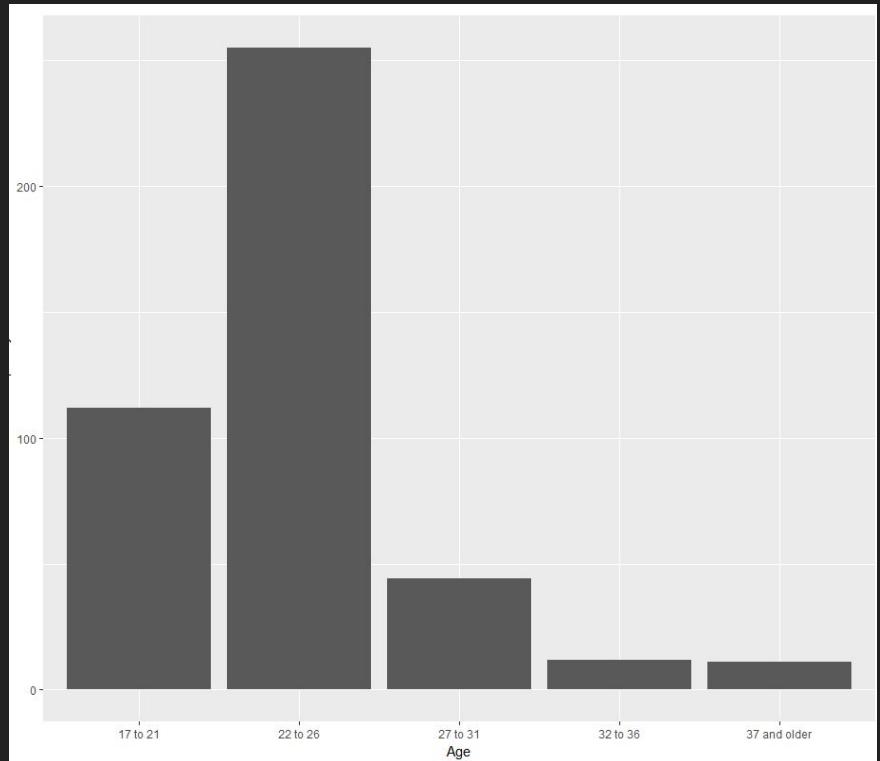
Most Fav. Genre ~ Race

Coefficients:

	(Intercept)	Race_textAmerican	Indian or Alaska Native	Race_textBlack or African American	Race_textNative Hawaiian or Pacific Islander	Race_textOther	Race_textOther Asian	Race_textwhite
Classical	-1.1435618		-4.108017					-13.2453945
Country	-1.6776430		17.387890					0.2913153
Hard rock	-1.8028033		-2.841515					0.4164820
Jazz	-1.8717980		-2.664448					-12.6216586
Rap	-0.8219765		16.532227					<u>1.3815776</u>
	Race_textNative Hawaiian or Pacific Islander	Race_textOther	Race_textOther Asian	Race_textwhite				
classical	17.470245	-0.1557180	-1.686265e-01	-0.2915215				
Country	-3.097163	-0.3147835	-4.818352e-01	<u>1.0310136</u>				
Hard rock	-2.902774	0.3211982	-1.455293e+00	0.8377205				
Jazz	18.198490	0.1670503	-5.061384e-06	0.6190329				
Rap	-4.845890	1.3418522	3.364677e-01	<u>1.0355486</u>				

Q2: Part 6: Relationship between Music Choice and Age

Most Fav. Genre ~ Age



Q2: Part 6: Relationship between Music Choice and Age (p-value)

Most Fav. Genre ~ Age

[[1]]

	(Intercept)	Age_text27 to 31	Age_text17 to 21
classical	1.938992e-08	0.15308903	<u>0.05077486</u>
Country	3.323266e-10	0.70766219	0.96851200
Hard rock	1.279283e-10	0.81700413	0.77190295
Jazz	2.410960e-11	0.16370947	0.42086880
Rap	3.752007e-02	0.09179418	<u>0.03696795</u>

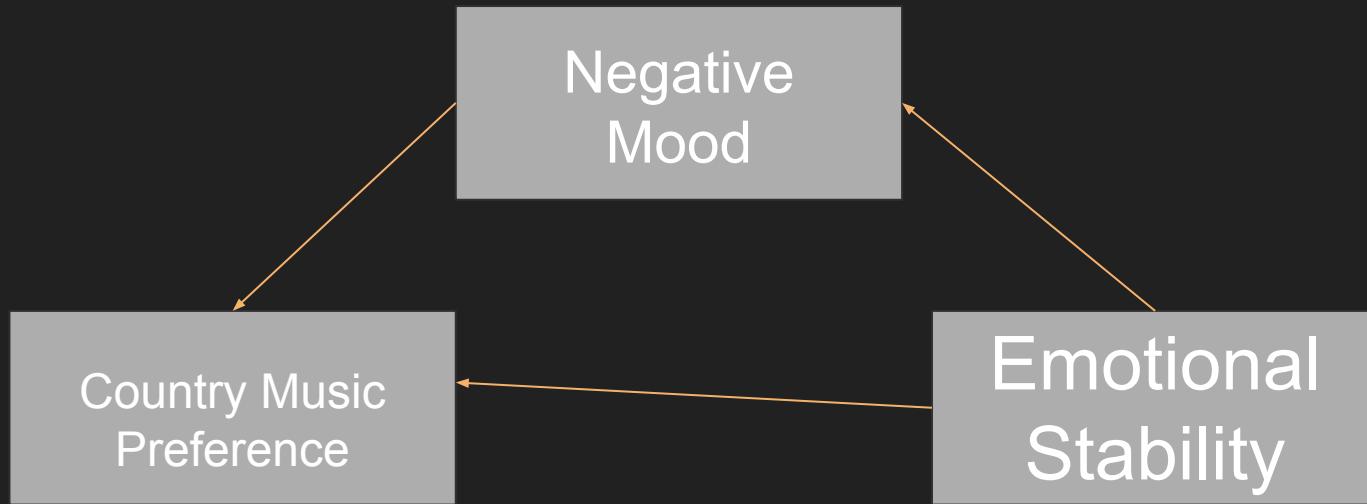
Q2: Part 6: Relationship between Music Choice and Age (coefficients)

Most Fav. Genre ~ Age

Coefficients:

	(Intercept)	Age_text27 to 31	Age_text17 to 21
classical	-1.1735115	0.6428629	<u>-1.10376166</u>
Country	-1.4836630	-0.2509761	0.01732014
Hard rock	-1.5789880	-0.1556540	-0.13866398
Jazz	-1.8021271	0.7606510	-0.47513840
Rap	-0.3262111	-0.8975803	<u>0.55446842</u>

Answer: Complete Mediation



Q5: Relationship between Time spent on music and Number of hours spent listening to music

Time spent on music ~ Big 5 traits

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	3.53608	0.59753	5.918	6.68e-09	***
Openness.to.Experiences.Score..out.of.7.	0.18548	0.08405	2.207	0.0279	*
Agreeableness.Score..out.of.7.	-0.09543	0.08295	-1.150	0.2506	
Emotional.stability.Score..out.of.7.	0.13340	0.07068	1.887	0.0598	.
Conscientiousness.Score..out.of.7.	-0.13097	0.07228	-1.812	0.0707	.

Signif. codes:	0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1				

Q6: If the relationship in (5) holds true, is it mediated by the fact that they are a commuter student?

To perform mediation, we first analyzed if the time spent on music was related with commuter student factor. We found a positive relationship as shown below

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept)    3.7778    0.1299  29.084 <2e-16 ***
factor(commuter_text)yes 0.4562    0.1833   2.489   0.0132 *  
---
signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Q6: If the relationship in (5) holds true, is it mediated by the fact that they are a commuter student?

Then we ran an analysis to see if commuter factor could be explained by any of the personality types. The relationship was found to be insignificant, as shown below. And thus we concluded that no mediation effect exists

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	1.42833	0.63860	2.237	0.0253 *
openness.to.Experiences.Score..out.of.7.	-0.04727	0.08910	-0.531	0.5957
Agreeableness.Score..out.of.7.	-0.12860	0.08825	-1.457	0.1451
Emotional.stability.Score..out.of.7.	0.05224	0.07496	0.697	0.4859
Conscientiousness.Score..out.of.7.	-0.13647	0.07703	-1.772	0.0764 .

Q4: Does the most popular music has any relation with Intelligence, Age, or personality dimension

Most popular music - Soft Rock/Pop

Only one Relation of personality dimension - Openness to New experiences(positive relation)

Model Summary ^b										
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					
					R Square Change	F Change	df1	df2	Sig. F Change	Durbin-Watson
1	.110 ^a	.012	.010	.976	.012	5.295	1	432	.022	1.884

a. Predictors: (Constant), Openness.to.Experiences.Score..out.of.7
b. Dependent Variable: Soft.Rock.Pop

Model	Coefficients ^a							Correlations				Collinearity Statistics	
	Unstandardized Coefficients		Standardized Coefficients		95.0% Confidence Interval for B		Zero-order	Partial	Part		Tolerance	VIF	
	B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound						
1	(Constant)	3.416	.228		15.007	.000	2.969	3.863					
	Openness.to.Experiences.Score..out.of.7	.095	.041	.110	2.301	.022	.014	.176	.110	.110	.110	1.000	

a. Dependent Variable: Soft.Rock.Pop

Q9: Are the ratings across music genre and preferred music genre significantly different? (paired t-test)

Country	yes	0.02986
Rap/Hip Hop	yes	0.0002714
Pop/Soft Rock	yes	2.2e-16
Jazz	yes	1.359e-07
Classical	no	0.3537
Hard Rock/Heavy Metal	no	0.889

Question 10: If the answer to (9) is positive, what factors explain the difference?

Genre	Factor	p-value	Estimate
Jazz	Openness to Experience	0.00297	0.123986
Pop/Soft Rock	Openness to Experience and Agreeableness	0.00850 and 0.04148	0.12142 and -0.09268
Rap	Age: 27 to 31	0.0401	0.31818

Summary

1. Country music preference is negatively related to Openness to New Experience Score
2. Agreeableness is negatively related to Rap music preference
3. Emotional stability is positively related to Rap music preference
4. Negative Mood fully mediates the relationship between Emotional Stability and Country music preference

Summary

5. The more open people are, the more number of hours of music they listen to
6. Most favorite among all the individuals is SoftRock/Pop
7. Extroversion is negatively related with softrock/pop & classical genres
8. People with high agreeableness are more likely to rate pop higher and emotional stability is positively related with jazz

Thank You

Q/A