

THE POWER OF MUSIC

Proyecto final de primer año

Grado en Ciencia de Datos

Autores: Diego Cotino Bolufer

Miguel Camuñas Castelló

Marcos Carrasco Panadero

Pablo Gil Martínez

Sergio Samaniego Hernández

Tutor: Jose Manuel Gil

1ºB1

18/06/2024

Index

Introduction	3
Motivation	3
Objectives	3
Methodology	4
Structure	5
State of art	5
Problem analysis	6
Proposed solution	7
Project scope	7
Deliverables (products)	7
Success criteria	8
Alignment with the Sustainable Development goals (SDG)	8
Working plan	9
Preparation and data understanding	10
Introduction	10
Data source	10
Exploratory analysis and data quality	13
Unidimensional Analysis	14
Multidimensional Analysis	16
Tools utilized	21
Results	21
Conclusions	22
Bibliography	23
Annexes	¡Error! Marcador no definido.
Annex A	¡Error! Marcador no definido.
Identification and treatment of outliers	¡Error! Marcador no definido.
Identification and treatment of missing values	¡Error! Marcador no definido.
Recodification	¡Error! Marcador no definido.
Annex B (Unidimensional analysis)	¡Error! Marcador no definido.
Annex C (Multidimensional analysis)	¡Error! Marcador no definido.

Introduction

On the street, at the public transport, at the university... People listens to music and use it as a method to escape their problems. According to Cristina Caron (2023) “music therapy has grown over the last decade.”¹ But does music really impact on mental health? Experts say yes, but how should people consum music to maximize their effects in mental wellness?

Motivation

This project is driven by the team's commitment to address the crucial issue of mental health. This topic is considered a public issue by important universities, like the university of Tulane (2021)², so contributting to this cause is an honour and a powerfull motivation. Obtaining significant results that can contributte to the topic is a pleasure and the team will be delighted to share the results with the society.

The idea about doing it about music came because it was an unexplored topic that was not talked about, outside the musical therapy. That is why, after looking for other variables, such as geographic conditions, it was finally decided by the team to work with a potential relationship with music.

Objectives

These are the goals that are aimed to fulfill:

- To find in the data analysis a series of potential relations between a certain music consumption habits and mental health.
- To develop a website that helps people to understand how music affects mental health.
- To transmit in our report how music consumption behavior could impact on the mental health of the user.
- To create a video about the project made, with its analysis, results and conclusions extracted.

Methodology

In order to make the project successfully and obtain relevant results it is used the following methodology. First of all, the project has to be planned, with the topic, the objectives and the products the project wants to have. With the objectives and the topic defined the dataset has to be found, this dataset has to be relevant and it has to have the information needed to make all the project. For this project it was found a dataset that relates mental health with music (Music & Mental Health Survey Results), it was obtained from kaggle, a web platform where the data science community share their projects and different tools like datasets.

After the dataset is chosen, the analysis of the variables to obtain results about our topics, starts. All the analysis is done with an statistical software called Statgraphics. This software was chosen because it is the best software to develop the project. The software has all the necessary tools and functions to make the project and it has a visual and attractive form to show the information and the data.

Knowing the variables the preprocess stage can be done. In this stage the individuals and the data with outliers and missing values is transformed or deleted with the goal to have an easier source to analyse. The outliers are found with the *outlier identifier* tool that the software has, and are treated being deleted or being changed by the media depending on the case. The missing values are found by the *data viewer* and are treated in the same way as the outliers. In the preprocessing stage some variables are also recodified because it is easier to make correlations with intervals instead of different values.

When the data is ready to be analysed, the unidimensional analysis begins. In this stage each variable is going to be analysed to find patterns in its behavior. This analysis is done in statgraphics too, using the unidimensional analysis tools. Depending on if the variable is categorical or numerical the analysis is done different, using different visuals like *Box and Whiskers* or the *probability graph* and using different parameters.

With the unidimensional analysis finished the multidimensional analysis is carried out. The multidimensional analysis consists in compare different variables in order to find relations. For this part it is used the multidimensional tools that the software has, the *Two factors* functionality and the *Subgroup analysis*. The variables with relations and correlations are taken for making the conclusions because are the principal variables that are interesting.

After that, the next thing that has to be done is to formulate conclusions that emerge from the work. This conclusions are obtained from the results of the analysis, with the knowledge gained from the state of the art and other parts of the project such as the problem analysis. All the conclusions have to make sense and have to be relevant to the principal objectives of the project.

In order to transmit the results there are some products. To create these products it is key to start from the objectives and link it to what has been learned. The products are thinked as a form to transmit, so they should be understandable and visually attractive. They are a website, a report and a video that provide more depth to the project, and they are the principal form to communicate the project's results.

Structure

The project is structured in several important parts that give it a cohesive and orderly form and a cohesion that makes the project more understandable to everyone.

The first part of the project is the introduction, where the project's topic and its problem is raised. This first section includes the motivation, the objectives, the methodology and the structure of the project. After the introduction, in the State of Art the problem on which the project focuses is analyzed in-depth, and it is proposed a possible solution to the problem after. After that, it is planteated the project scope where it is shown what the project will try to be. This part includes the products, the success criteria that is ,when the project is considered as a success, and the alignment with the sustainable development goals. Following the project scope is the working plan that is used during the project.

Regarding the data analysis, the data source and from where it comes is explained. With the data source understood the analysis starts with the unidimensional analysis where each variable is analysed individually. After the unidimensional analysis the multidimensional analysis is explained, in this part the relations between the variables is shown in order to find evidences. After the analysis it is explained what tools were used during the process.

After that the results of the analysis and the conclusions are shown. Finally, it is displayed the bibliography and the annexes with additional information and more explanations about what has been done in the project.

State of art

Mental health has been a recurrent topic of debate in recent years for several reasons. Firstly, it has gone from being a taboo subject in certain areas and social groups to being something on everyone's lips, in addition in recent times there has been an increase in problems in this area mainly among young people, which is a compelling reason to carry out a project focused on this issue. This project is going to analyse whether there is a relationship between the music listened to and the state of mental health, making use of a large database with different variables related to the amount of music consumed, as well as the specific styles and the state of mental health, focusing on different related problems.

According to Amanda MacMillan (2023) “more people in the US are living with mental and emotional distress”³. This may extend globally, as Andrew Moose and Ruma Bhargava (2024) point out “Since 2006, levels of reported youth happiness have declined in North America, South America, Europe, South Asia, the Middle East and North Africa.”⁴ Areas which

concentrate the vast majority of the world's population, so it can be assumed that this is a global problem.

Due to the increase in mental health problems, investigations have been carried out to examine the potential factors contributing to this rise. Among the many factors under consideration, MacMillan A. (2023)⁵ identifies four key ones: Social media usage, Covid-19, Isolation and Loneliness, and Lack of access to care. Being the first one something that only increases due to modern era and causes problems mainly because over-comparison with what is seen on social media that often is not true. The other main factor to note is the third: Isolation and loneliness, which are enhanced factors due to social trends like decreased community involvement and fewer people starting a family.

In recent times, other factors that may affect mental health to a lesser extent have been studied, among which is the influence of music, whether the person suffering from mental health problems sings or plays it or simply listens to it. Even therapies based on this, known as music therapy, have been developed.

As Priyanjana Pramanik wrote in her recent article in medical.net (2023) music therapies have been used for physical recovery and in motor rehabilitation. However, the relationship between music and mental health continues to be poorly understood even though listening to music has been associated with mental health improvements especially in cognitive function.⁶ It is evident in the scientific community that there is a consensus on the need for further development of music therapy.

These music therapies are based on the neurological effect that music has on people, as Alison Pearce Stevens (2024) explains “Highly emotional music causes networks in the brain to release dopamine. This brain chemical, a type of neurotransmitter, plays a role in feelings of pleasure.”⁷ This, in addition to Fatima Reynold's explanation of what music allows individuals to express “Through music, individuals can express their unique experiences, struggles, and triumphs, forging connections with others who share similar backgrounds. Research has shown that exposure to diverse musical genres and artists can broaden perspectives, challenge stereotypes, and foster empathy among listeners especially when dancing together”(2023).⁸ So from a neurological such as from a social point of view experts outline the effect of music on mental health.

Other benefits of music for mental health have been outlined by Lynne Gilmour (2024)⁹ in a recent article. The project developed at the University of Stirling outlines the health benefits of live music. They propose to make it easier for young people with mental health problems to access live music events, as their condition is greatly improved when they are able to participate in such events.

In addition, medical studies have shown improvements in patients with various diseases such as schizophrenia, Parkinson or Alzheimer, as highlighted by Pearce Stevens (2024).¹⁰ As far as can be seen, music is beneficial not only to face mental health problems but also diseases that affect memory or motor skills as mentioned above.

Problem analysis

The state of the art has shown that experts agree that there is an increase in mental health problems, such as depression, in society (Goodwin et al. 2022)¹¹. And that is why it is considered a problem to be addressed.

In addition, the state of the art has shown several relationships between music and mental health, but all of them are focused from the point of view of treatments and not on how music consumption affects mental health.

It is well known that people use music to improve their mood. Is this measure sufficient? In short, the problem is that society is not aware of the extent to which music affects their mental health and society don't know how to use the music properly to obtain all the benefits that brings to.

Proposed solution

A bad mental health, like any other problem, can be remedied in several ways. The State of Art has shown that music therapy contributtes positively to mood and other diseases like Parkinson or Alzheimer. People know the power of music but again, it is not known in which grade consumption habits affect wellness. If the *How* is not explained, how would people act in consequence to improve their lives? Said this, the solution wouldn't be to find evidence whether music affects mental health. Instead, it would be to find correlations between the characteristics of music listeners and their habits with factors like anxiety, insomnia or depression. To do so, a database which cointains info about the listeners stats according to music and mental health is being analysed.

An important part of this solution will be the way it is presented. If people has not an accesible way to get the conclusions the efforts are done for nothing. For this reason, a web page will be made to show all parts of the project in an eye-catching way.

Project scope

Deliverables (products)

The project is formed by the following three products:

-The most representative is the **official web page** of the project as it contains the rest of the products. It will be done with *HTML* and *CSS*. It is meant to be an aesthetic way to deliver the correlations that the team has found. It is a public domain web because its goal is to reach as many people as possible. In order to let the user know how to flow through the website, it contains a user guide to indicate what to see in order. The website also tells about the team, its motivation and where they come from.

-The website offers a downloadable **report**. It is made by the team and it is the heart that explains the project's fundamentals. All the processes, methodology, analysis and conclusions about the project are explained from the beginning of the project, together with a presentation. It is thought for people who are used to statistic terms and knowledge.

-One of the website's pages shows the promotional **video**. It is a short 16:9 video edited with free Software. This video's goal is to help the website reach as many people as possible as the video is what is being sent when the project launches. It explains the general aspects and serves like a tutorial to teach users how to use the web properly.

Success criteria

It can be said that success has been achieved if the proposed objectives are achieved or fulfilled. To confirm that they have been achieved, they have been more measurable and quantifiable, and are the following:

- To find in the data analysis at least three potential correlations that may occur between variables.
- To develop a website that helps people to understand how music affects mental health that provides value and information to the user, having at least 250 visits by the due date.

Alignment with the Sustainable Development goals (SDG)

The project is aligned with the Sustainable Development Goal number 3 for the UPV: "Health and Wellbeing". This SDG (Sustainable Development Goal) has the objective to guarantee a healthy life and to promote the wellbeing at all the ages. One of the problems that this SDG wants to reduce is the suffering of people from mental illnesses. In this project this kind of diseases will be studied.

The project is aligned with this goal because it is treated the mental health as the principal topic of the project and it is tried to find some evidences that can contribute the topic. In this case

the music is the medicine and it will be studied how can be used to improve the wellbeing and to guarantee a healthy life.

The goal of the team, at the end of the day, is to contribute to improving all people’s mental health, in one way or another, fixating on underrated factors such as the music listened to.

Working plan

When the project started, a plan was defined with the objective to serve as a guide to be followed upon, that uses a breakdown structure and a Gantt chart. This planning has changed over time redefining the times and the Gantt chart with it. The following graphs are the last Gantt chart and the last breakdown structure.

This figure is actualized with the information about the project in the final stage. The work breakdown structure has practicaly the same parts that it had when the project started, but some changes were done. Some sections from web applications were removed because this parts were considered unnecessary and far from the principal purpose of the project, and now is only one principal section. The time of the state of art and the time of the web application were reduced and the time of all the parts of initiation and the time of reporting were expanded.

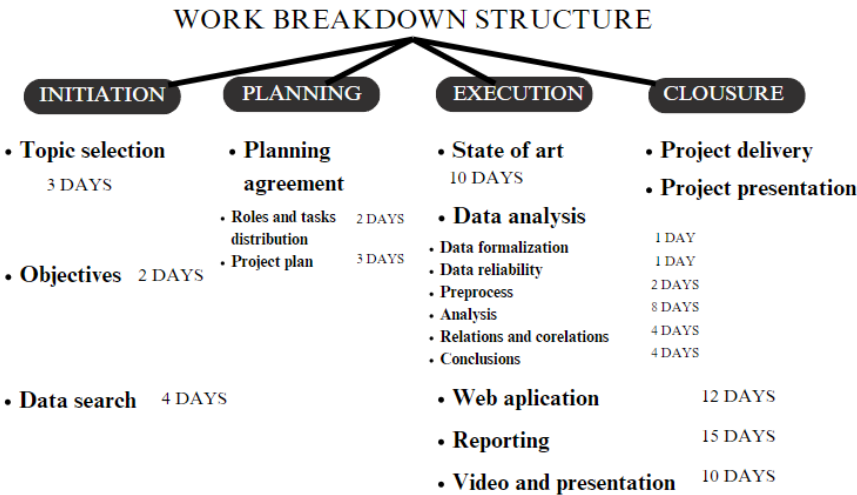


Figure 1 Work Breakdown Structure

The Gantt Chart has also been updated during the project, as it has been said before this is the last Gantt Chart that has been made which represents the project timings better than the ones made before.

As can be seen in figure 2, the project has 3 distinct phases after the choice of topic. On 28 February the first phase started, in which the objectives, data research and planning were

carried out. Then stated the second phase which is composed of the state of art, the data analysis and the report, this last started later due to it depends on the previous ones. The last phase could be call the products one because it consisted in the development of the web app, the video and the presentation.

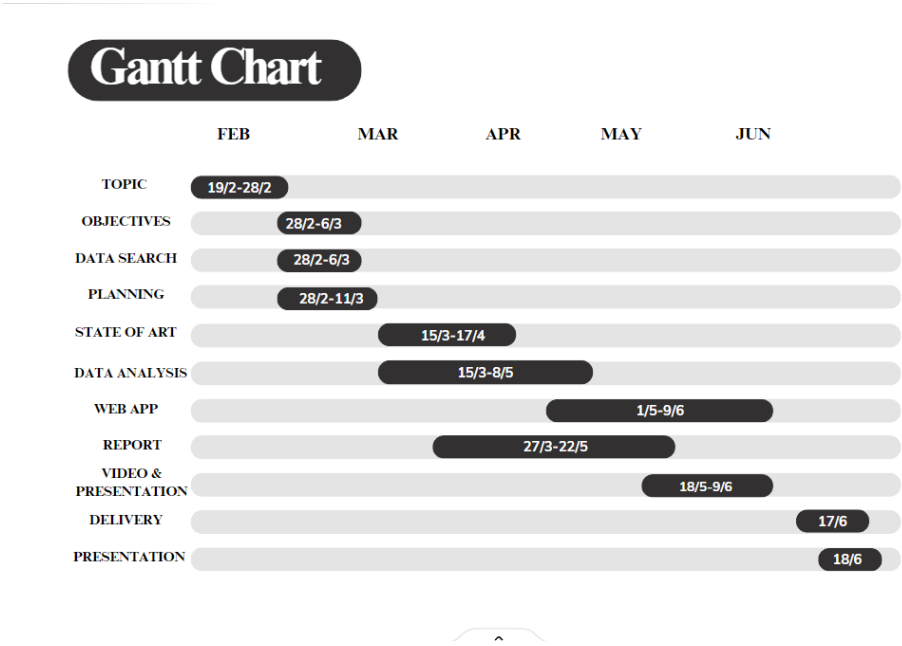


Figure 2 Gantt Chart

Preparation and data understanding

Introduction

The main part of the project is to analyze the data that the group has collected about the main topic. This analysis has the objective of finding relations and correlations in the variables to verify previously found information and to confirm the hypotesis the team had. The analysis started with the identification and reliability of the data source and its characteristics. It continued with the treatment of missing values and outliers and the analysis ended with the unidimensional and the multidimensional analysis. After this analysis the team drew conclusions.

Data source

The database was obtained from a website called “Kaggle”, and its name is “Music & Mental Health Survey Results”. From all the other databases related to the topic, this one has by far the most upvotes. It is also very well rated, with lots of users opining that the database is clean and well-documented. Another good sign is that it has been downloaded more than 20 thousand times. This is the web page’s link <https://www.kaggle.com/catherinerasgaitis/mxmh-survey-results> .

It is needed to keep in mind that some of the variables, precisely the ones about mental health disorders, are self-reported in a rating of 1 to 10, so there is a reason to watch out if aiming to a very precise analysis. There are also other details, such as the frequency of listening to a certain type of music with options such as "Sometimes" or "Never". There cannot be an exact measurement because It is not known if an answer such as “Very frequently” means 1 in every 2 songs, or 1 in every 5.

To summarize this, and taking into account everything discussed before, it has been considered to qualify this database as helpful, useful and trustable, despite the small inaccuracies that may come with this subjective way of responding. As said before, it is very likely that this database has also been helpful to other data analysts, and so it is believed that it will be useful to the team, too.

The content of the data file contains the information about 736 individuals and collects of each one, information about 33 different variables. The data file’s content is structured in columns for the variables and in rows for the individuals.

The data file contains information about the using of music, his different genres and the feelings or sensations that the different individuals have. In this information there is the individual age, the streaming service, the visualization hours, if they listen to music while working, among other interesting variables for the analysis.

The following table contains the variable the data file has, their description, a comment and the variable type (if the variable is numerical or categorical). The variables in colour purple are the variables that were used by the team for the analysis:

Variables	Description	Comments	Variable type
Age	This variable offers the age o the individual.		Numerical
Primary streaming service	This variable offers the name of the principal streaming service the individual use to listen music.		Categorical
Hours per day	This variable offers the amount of hours the individual listens to music in one day.	(24 hours)	Numerical
While working	This variable define if the individual listens music while working.	(YES or NO)	Categorical
Instrumentalist	This variable define if the individual plays an instrument or not.	(YES or NO)	Categorical

Composer	This variable define if the individual compose music or not.	<i>(YES or NO)</i>	Categorical
Fav genre	This variable offers the favourite genre of the individual.		Categorical
Exploratory	This variable define if the individual actively explores new genres or artists.	<i>(YES or NO)</i>	Categorical
Foreign language	This variable define if the individual listens to music in other languages that he can't speak fluid or understand.	<i>(YES or NO)</i>	Categorical
BPM	This variable offers the beats per minute of the favourite genre.		Numerical
Frequency [Classical]	This variable define how frequently the individual listens to this genre.	<i>(never, rarely, sometimes, very frequently).</i>	Categorical
Frequency [Country]	This variable define how frequently the individual listens to this genre.	<i>(never, rarely, sometimes, very frequently).</i>	Categorical
Frequency [Edm]	This variable define how frequently the individual listens to this genre.	<i>(never, rarely, sometimes, very frequently).</i>	Categorical
Frequency [Folk]	This variable define how frequently the individual listens to this genre.	<i>(never, rarely, sometimes, very frequently).</i>	Categorical
Frequency [Gospel]	This variable define how frequently the individual listens to this genre.	<i>(never, rarely, sometimes, very frequently).</i>	Categorical
Frequency [Hip hop]	This variable define how frequently the individual listens to this genre.	<i>(never, rarely, sometimes, very frequently).</i>	Categorical
Frequency [Jazz]	This variable define how frequently the individual listens to this genre.	<i>(never, rarely, sometimes, very frequently).</i>	Categorical
Frequency [Kpop]	This variable define how frequently the individual listens to this genre.	<i>(never, rarely, sometimes, very frequently).</i>	Categorical
Frequency [Latin]	This variable define how frequently the individual listens to this genre.	<i>(never, rarely, sometimes, very frequently).</i>	Categorical
Frequency [Lofi]	This variable define how frequently the individual listens to this genre.	<i>(never, rarely, sometimes, very frequently).</i>	Categorical
Frequency [Metal]	This variable define how frequently the individual listens to this genre.	<i>(never, rarely, sometimes, very frequently).</i>	Categorical
Frequency [Pop]	This variable define how frequently the individual listens to this genre.	<i>(never, rarely, sometimes, very frequently).</i>	Categorical
Frequency [R&B]	This variable define how frequently the individual listens to this genre.	<i>(never, rarely, sometimes, very frequently).</i>	Categorical
Frequency [Rap]	This variable define how frequently the individual listens to this genre.	<i>(never, rarely, sometimes, very frequently).</i>	Categorical

Frequency [Rock]	This variable define how frequently the individual listens to this genre.	<i>(never, rarely, sometimes, very frequently).</i>	Categorical
Frequency [Video games music]	This variable define how frequently the individual listens to this genre.	<i>(never, rarely, sometimes, very frequently).</i>	Categorical
Anxiety	This variable offers the level of anxiety of the individual.	<i>(0 – 10)</i>	Numerical
Depression	This variable offers the level of depression the individual thinks he has.	<i>(0 – 10)</i>	Numerical
Insomnia	This variable offers the level of insomnia of the individual.	<i>(0 – 10)</i>	Numerical
Ocd (Obsessive compulsive disorder)	This variable offers the level of Obsessive compulsive disorder the individual has.	<i>(0 – 10)</i>	Numerical
Music effects	This variable offers the level of music affects the individual.	<i>(worsen, no effect, improve)</i>	Categorical
Permission			Categorical
Timestamp			Categorical

Figure 3 Table of variables

The description specifies what the variable means. The comments were added with the goal to specify what values can the variables have. The variable type says if the variable is categorical or numerical.

Exploratory analysis and data quality

With the data source analysed the team knows how to work with the data obtained. In the preprocess the outliers and the missing values were found and treated with the objective of having a data file analysable to search the necessary information to the project. In this stage, some variables were recoded with the finality of providing different types of variables and different forms of the same variable.

After the identification of outliers, some of the individuals with outliers were deleted because the data didn't make sense, for example individuals that say that listen to music more than 24 hours per day. Other outliers that made sense but were no representative and were so far from the media were changed for this parameter.

There were some missing values. The individuals with this kind of values were deleted because individuals with no information were no representative to the project. After these processes the data file was left with 586 individuals compared to 736 at the beginning.

The variables **Age** and **BPM** (beats per minut of favourite genre) were discretized because in the analysis could be needed to be used these variables like qualitative ones to find relations. These new variables were called **Age_recod** (age recodificated) and **BPM_recod** (beats per minute of favourite genre recodificated).

With all the preprocesses done the data file has 586 individuals (rows) and 35 variables (columns) and the data analysis can be started. Of all these variables only 13 are going to be analysed in order to find the relations interesting to the project (the variables can be useless).

For more information about the preprocess, see the annex A.

Unidimensional Analysis

The variables that have been used in this project are related to two different themes, the first one is the music, and the second one is the mental health. Regarding the music, the most relevant variables are the hours per day hearing music, the quantity of people who are more likely to explore new type of music and the effect that music has in the mood of the individuals.

The individuals of the sample listen to music between zero and ten hours each day. The variable has a great positive asymmetry, which means that the data is concentrated in low values, being the median three. Most people of the sample listen among one and five hours of music per day (Figure 47). Besides, a 73,21% of the individuals are willing to search and explore new types and genres of music (Figure 54).

A 76,11% of people in the sample think that music has an improvement effect on their mood, 2,22% that it has a negative effect, and 21,67% that music does not have an effect in them. That is perfectly appreciable in the next pie chart:

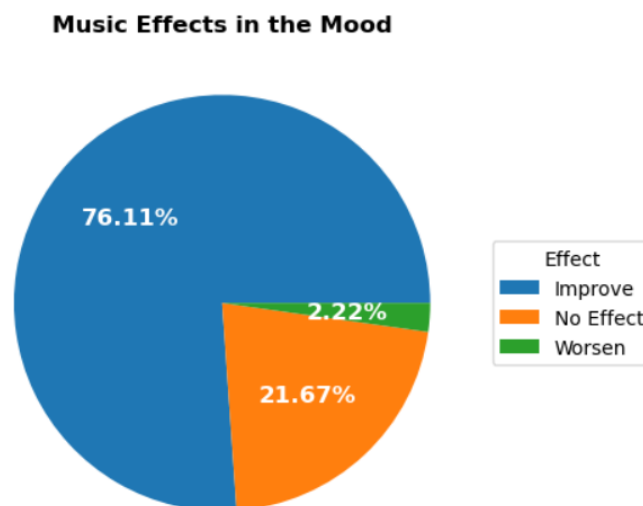


Figure 4 Music Effects in Mood (Pie chart)

Furthermore, the variables that talk about mental health are the anxiety, depression, insomnia and obsessive-compulsive disorder that individuals think that suffer, being all of them a self-diagnosis. Every value of these four variables ranges from zero to ten. Anxiety variable has a negative asymmetry, that means its data tends to have a high value (Figure 31), also its median is six. The values of depression are more equilibrated in the sample (Figure 35), with a median of five. Besides, insomnia variable has a median of three and obsessive-compulsive disorder has a

median of two. Both have a positive asymmetry, due to the fact that most of their data has low values. In order to compare visually, this bar chart shows the medians of the four mental health issues:

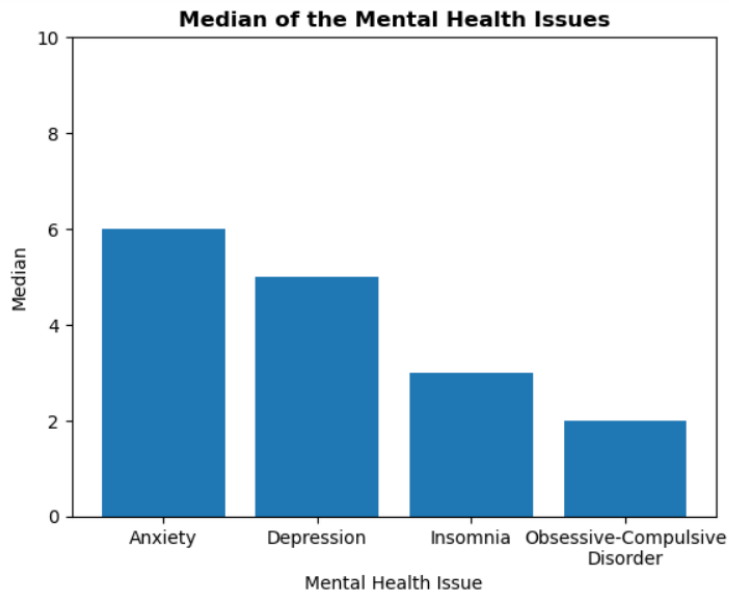


Figure 5 Median of the Mental Health Issues (Histogram)

Moreover, the individuals of the sample have been divided in ranges of age (“ “ this specific age is not included in the interval, “]” this specific age is included in the interval.): (0-16], (16-30], (30-45], (45-60], (60-75] and (75-90]. The age variable has a median of 21 and its quartiles are 18 and 27, being its interquartile range 9. The 68,6% of the individuals are included in the interval that comprise the ages between 17 and 30, the (16-30] one. That is seen clearly in this bar chart:

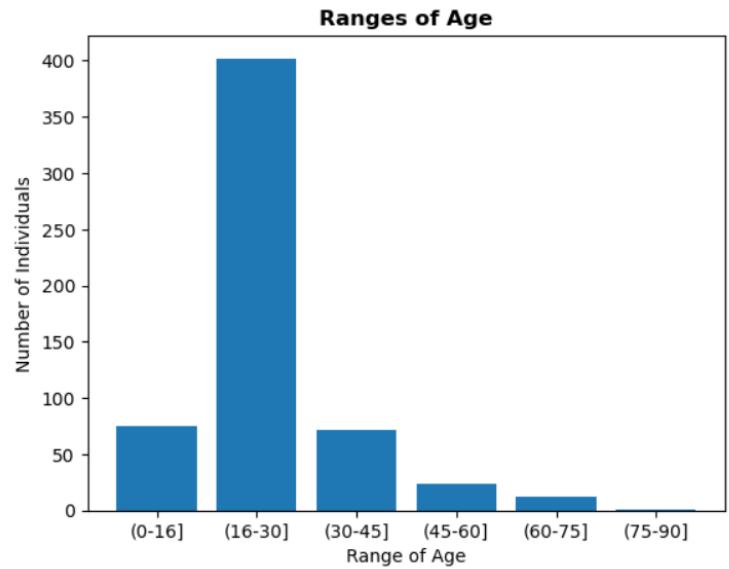


Figure 6 Ranges of Ages (Histogram)

For more information and the complete unidimensional analysis of the variables check the Annex B.

Multidimensional Analysis

A big part of the project lays on the Multidimensional Analysis. It gives answer to questions like, does listening to music more hours influence in insomnia levels?, how much does believing that music is positive has to do in your current mental state?

Standardization

The Standardization is the procedure by which the comparison among values in different units is possible. In this descriptive analysis, the quantitative variables have no unit as they are collected in a scale from 0 to 10 (**Anxiety**, **Depression**, **Insomnia**, **OCD**). In the other hand, the variables **BPM**, **Hours per day** and **Age** have every value in beat per minute, hours and years respectively. Everything said, a standardization is not necessary.

Identification, study and description of relationships

The goal of the multidimensional analysis is to relate two or more variables in order to look for an explanation of a phenomenon. To put it in other words, is the same as investigating the cause and showing how intense the consequences are.

Multidimensional Analysis of Quantitative Variables

Age (without recodification), **BPM**, **Hours per day**, **Anxiety**, **Depression**, **Insomnia** and **OCD** are quantitative variables. These variables will be related in order to find some relations. The relations are obtained from the covariance matrix, in this graphic each variable has the number of correlation with the other variables. The boxes with a value close to 0 have no correlation between them (boxes in green). It can be seen that there is a relation between **Depression** and **Anxiety** (box yellow), This relation shows that people who have more **Anxiety** have higher levels of **Depression**. Besides, there are moderate relationships in the couples Depression-Insomnia and OCD-Anxiety. Important to mention the **inexistence** of **correlation** between **Age** and **Hours per day** with **Depression**, **Insomnia**, **Anxiety** or **OCD**.

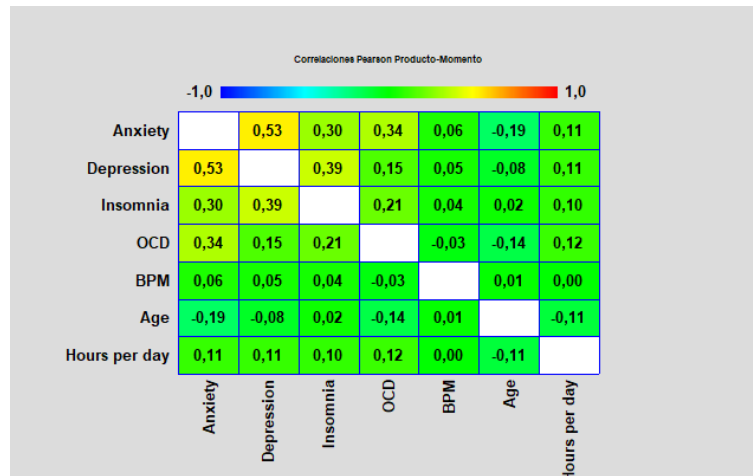


Figure 7 Correlation pearson (all the quantitative variables)

Multidimensional Analysis of Qualitative Variables

The distribution of **Age_recod** was seen in the unidimensional analysis, now this variable is compared with other variables in order to find correlations. The finality of finding correlations is to know if the different ranges of age should be treated separately. In order to see if these correlations exist it is used the “Mosaic graph” relating **Age_recod** with the rest of variables. It can be seen that the variable **Age_recod** has no correlation with any variable because the colored areas has the same size for each correlation how in the following example:

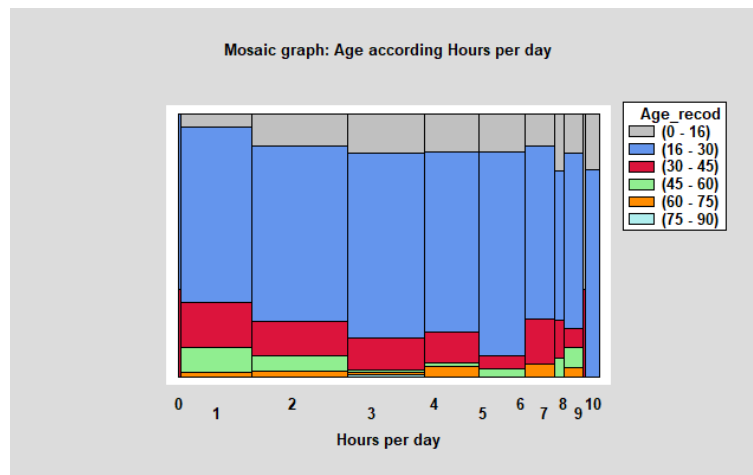


Figure 8 Age acording Hours per day (Mosaic graph)

As the variable **Age_recod** has not a relation with other variables we can use all the sample instead of different age ranges in the rest of the analysis.

After that, it will be seen if the hours of listening to music affect to mental status. The variables **Hours per day**, **Anxiety**, **Depression** and **Insomnia** are quantitative variables but as

the variables have few values they will be treated as ordinal qualitative ones. The **Kendal number**, is an statistic that tells numerically the intensity of a correlation. In this case is **0.092**, which is closer to 0 rather than 1. Due to this, variables have no correlations and it can be said that listening more hours of music per day does not contribute to lower or higher levels of **Anxiety, Depression** or **Insomnia** than listening to less hours.

Now the variables **OCD** (Obsesive compulsive disorder) and **While working** are analysed in order to find correlations. **OCD** was previously unused. It also goes from 0 to 10. It will be treated as a quantitative variable by using the **Subgroup Analysis**. The OCD average equals to 4, and the probability of listening to music while working is 79,86%. If we consider OCD values > 4 as those that are less common, the cumulative probability would be 31.06%. Knowing this it is possible to obtain the:

$$\text{Probability(While Working=Yes \& OCD > 4)} = 0.7986 * 0.3106 = 0.2480 = 24.8\%$$

$$\text{Probability(While Working =No \& OCD > 4)} = 0.2014 * 0.3106 = 0.0625 = 6.25\%.$$

In other words, you are **four times** ($24,8 / 6,25 = 3,968$) **more likely to have abnormal OCD values if you listen to music while working**. This tool makes an univariant analysis for every value of the qualitative variable (**While working**) along with every value of the quantitative one (**OCD, Anxiety, Depression, Insomnia...**).

The results are shown on the **Box&Whiskers** graph. How it can be seen the the distribution on the **Yes** is more to the right than on the **No**. So, it can be said that people with highs levels of **OCD** tend to use more music while working.

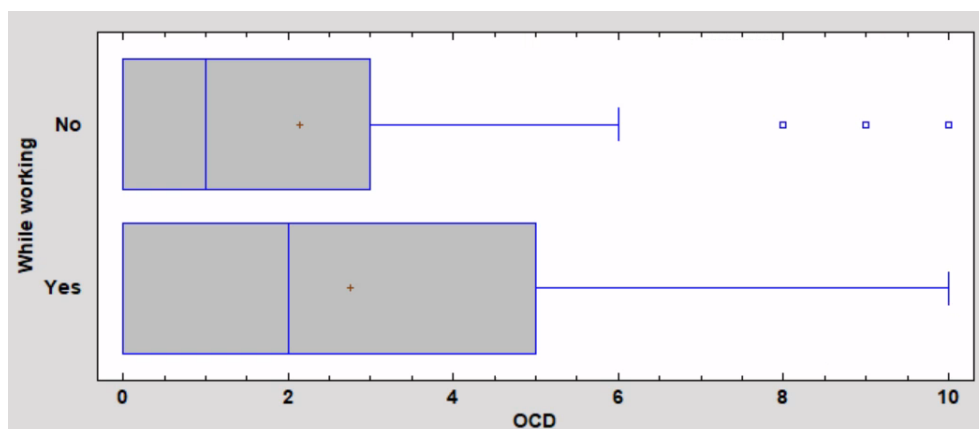


Figure 9 While working and OCD (Box&Whiskers)

After that, it will be proved if there is a relation between the **Mussic effects** and the **Hours per day** the music is listened. **Hours per day** has been used with **Anxiety, Depression** and **Insomnia** but treating them as ordinal qualitative variables. Now **Hours per day** will be treated as a discrete quantitative variable.

How it can be seen in the **Box&Whiskers** the distribution on people who think that music worsens their lives has lower values of **Hours per day** than people who think the opposite or nothing.

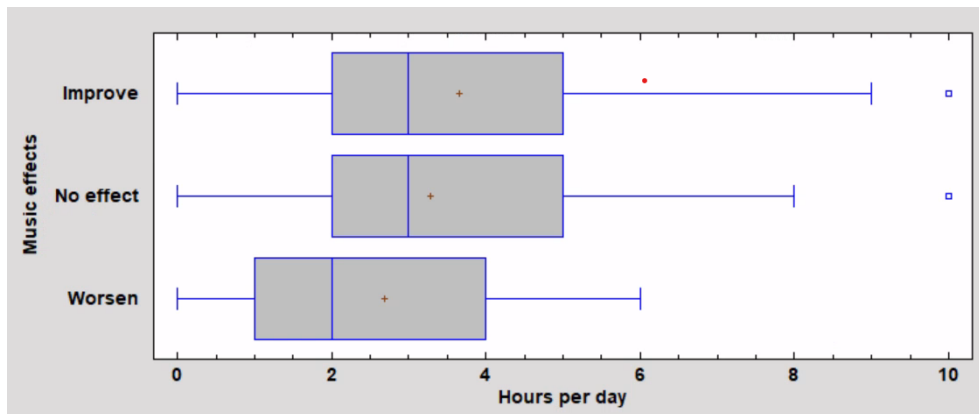


Figure 10 Music effects and Hours per day (Box&Whiskers)

Now, do people who listen to music **while working** tend to be more **exploratory**? Or are they more likely to be less **exploratory while working**?

Although being more people who listen to music while working, in both scenarios of the **exploratory** variable there is a significant difference in the frequencies. There are many more who are exploratory and listen to music while working, 85%, than people who are not exploratory and listen to music while working 68% .

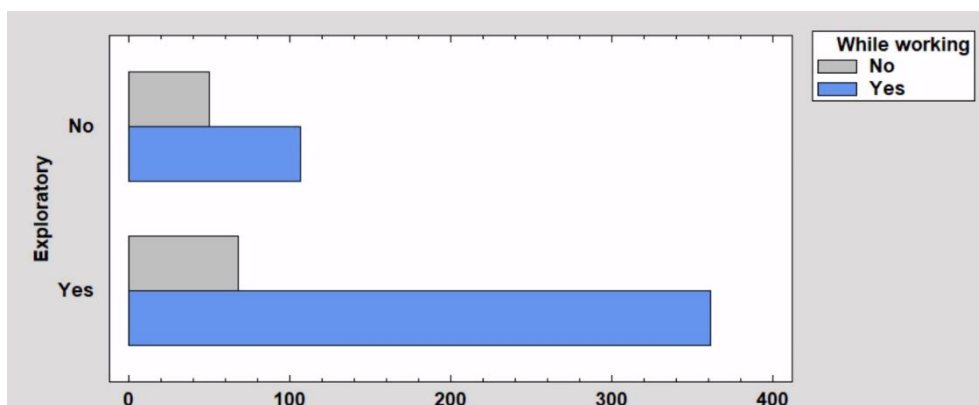


Figure 11 Exploratory and While working (Histogram)

Finally, it will be checked how the mental status of the individuals is related with the believed effect that music does in them. For this it will be used **Music effects**, a qualitative variable with 3 possible values (Improve, No effect, Worsen) and the three principal mental health problems in this project (**Anxiety, Depression and Insomnia**). Will the fact of considering music to be positive be reflected in the values of **Anxiety, Depression and Insomnia**?

There are clearly higher levels of **depression** and **anxiety** in people who believe that music worsens their lives than people who believe the opposite. Distribution of *Improve* and *No effect* is quite the same in every scenario.

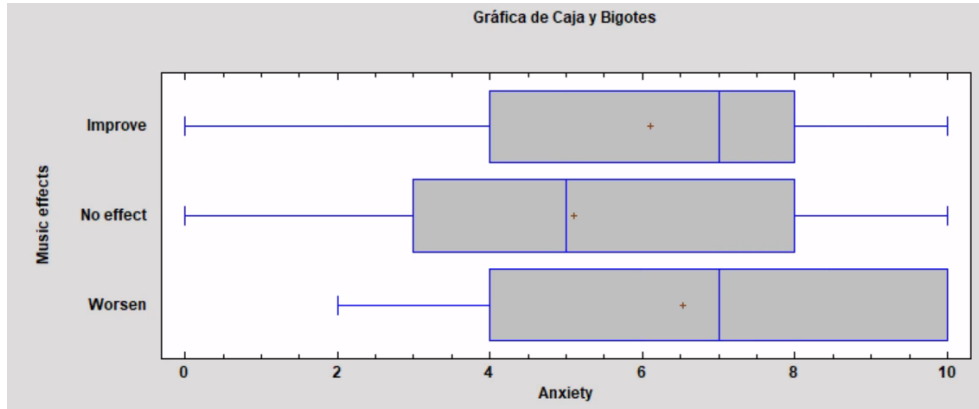


Figure 12 Music effects and Anxiety (Box&Whiskers)

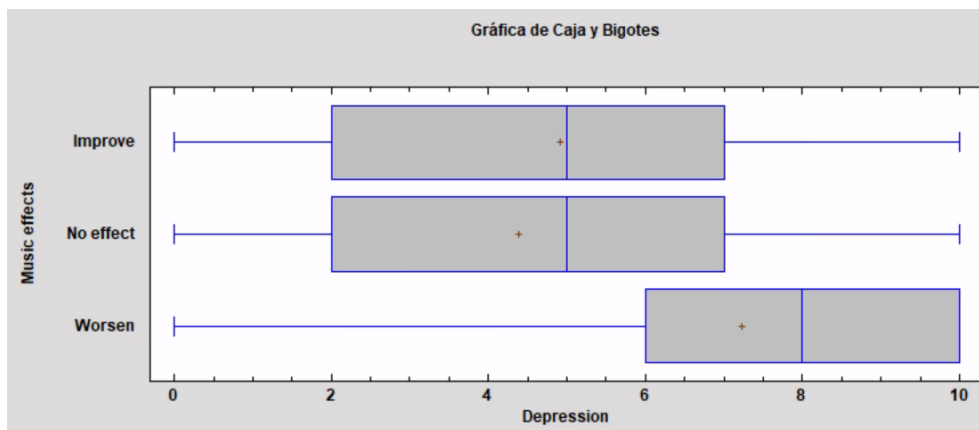


Figure 13 Music effects and Depression (Box&Whiskers)

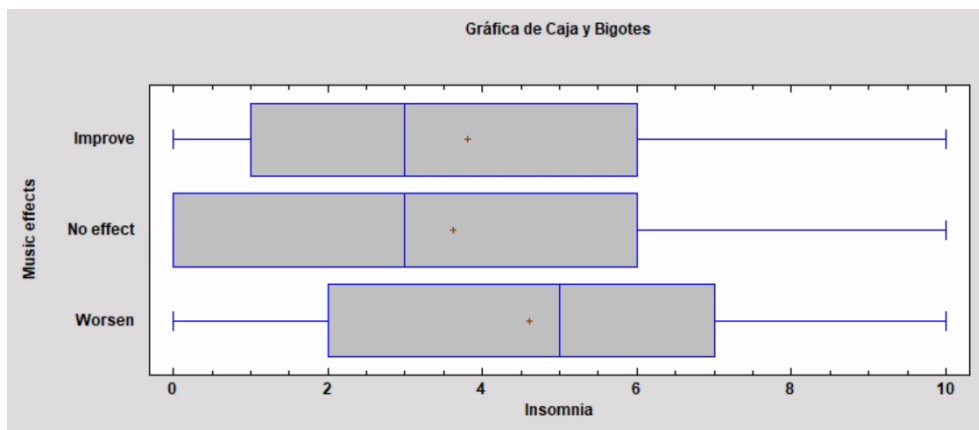


Figure 14 Music effects and Insomnia (Box&Whiskers)

For more details about the multidimensional analysis check the annex C.

Tools utilized

In order to carry out the analysis different tools have been used. The main ones are Statgraphics and Excel. *Statgraphics* was used instead of other tools such as *Python* (with jupyter notebook) since it is a previously used program and it saved time learning a new tool. *Excel* was used to prepare the data and make the data file usable for *Statgraphics*.

In the preprocess stage were used the following tools: The tool used for the identification of outliers was the *outliers identifier* from analysis of one variable and the tool used for treat the data was the *data book*. The missing values were found by the *data viewer* and they were treated by the *data book*. The new variables were recoded, by the recodification tool that statgraphics provides, after copying the original variables.

In the unidimensional analysis of the variables it was used the *Frequency table* to determinate the frequencies. In order to find atypical data it was used the *Box and Whiskers* chart. Also some parameters were studied in order to understand better the form of the data, this parameters were found by the Unidimensional analysis the software provides. The visual graphs used for the unidimensional analysis were the *Pie chart* and the *Histogram*.

In the multivariant analysis was used *Statgraphics* and its functionalities: for categorical variables the functionality of *Two Factors*, for numerical variables the functionality of *Multivariable* and when both kinds were worked at the same time, the *Subgroup Analysis*. The visual graphs used for the multidimensional analysis were *Box&Whiskers*, the *Mosaic graph* and the *Histogram*.

Results

In resume, the results that can be extracted from the analysis are the following:

On the side of the univariant results, it can be stated that:

- The people listens 3 hours per day of music.
- The 76,11% thinks that music improve their mood, the 22,67% thinks that music don't affect and the 2,22% thinks that music worsen.
- The most of the people explore new genres or new types of music (73,21%).
- There is more anxiety than depression, and more depression than Insomnia and OCD.

On the other hand, these are the multivariant results:

- If you listen to music **while working**, it is more likely for you to be an **exploratory** person.
- Higher levels of **anxiety** and **depression** abound more in people who think that music worsens their lives.
- People who are **exploratory** with music listen more to **foreign language** music.
- **Age** does **not affect** how many **hours of music** you listen to.
- **Hours** of consumption, **age** and **favorite music gender** are **not correlated** with **anxiety, depression, insomnia** or **OCD**.
- There is a **relationship** between people with **anxiety** and obsessive-compulsive disorder (**OCD**)
- People who **think that music worsens their lives, listen to music less hours**.
- People who have higher levels of OCD are more likely to listen to music while working.
- There is a strong relation between anxiety and depression

Conclusions

The music is obviously an important part of the human lives. The people listen to music a mean of 3 hours per day independently the age they have. So, it is important to know how it affects to our mental health. Also, the people use to explore new genres of music frequently and that transform the music not only into a type of entertainment, but also a way to learn new languages or to disconnect at work.

It is scientifically known that music helps facing a bad mental health in terms of depression, insomnia, anxiety or even OCD. If music improves depression, it also heals anxiety as they are statistically correlated. The same happens with anxiety and obsessive-compulsive disorder. Another direct association includes OCD and the people who listen to music while working. Music with repetitive and predictable rhythms may provide a comfortable mental space for high OCD people.

The hours of consumption, the age or the gender of music do not affect mental health. But the people that think the music worsen their mood do not explore the huge world that the music is and they lose the opportunity to enjoy this art. This kind of person has also higher levels of anxiety or depression, so the people as close this door have worse mental health. In conclusion it is not important at all how the music is listened, it is important to enjoy the music as it is preferred and to believe in the power of music.

Bibliography

Table Index

Figure 1 Work Breakdown Structure	9
Figure 2 Gantt Chart.....	10
Figure 3 Table of variables.....	13
Figure 4 Music effects in Mood (Pie chart)	14
Figure 5 Median of the Mental Health Issues (Histogram).....	15
Figure 6 Ranges of ages (Histogram).....	15
Figure 7 Correlation pearson (all the quantitative variables)	17
Figure 8 Age according Hours per day (Mosaic graph).....	17
Figure 9 While working and OCD (Box&Whiskers).....	18
Figure 10 Music effects and Hours per day (Box&Whiskers)	19
Figure 11 Exploratory and While working (Histogram)	19
Figure 12 Music effects and Anxiety (Box&Whiskers).....	20
Figure 13 Music effects and Depression (Box&Whiskers).....	20
Figure 14 Music effects and Insomnia (Box&Whiskers).....	20

References

1. Caron C. (2023) *How Music Can Be Mental Health Care*. The New York Times.
2. Tulane University (2021) *Understanding Mental Health as a Public Health Issue*. Tulane.edu
3. MacMillan A. (2023). *4 Possible Reasons Why Mental Health Is Getting Worse*. health.com
4. Moose A. and Bhargava R. *A generation adrift: Why young people are less happy and what we can do about it*. weforum.org
5. MacMillan A. (2023). *4 Possible Reasons Why Mental Health Is Getting Worse*. health.com
6. Pramanik P. (2023) *Music could hold the key to developing effective mental health interventions*. news-medical.net
7. Pearce Stevens A. (2024). *Music has the power to move us physically and emotionally. Here's why*. ScienceNewsExplores

8. Reynolds F. (2023). *The transformative power of music in mental well-being*. American Psychiatric Association
9. Gilmour L. (2024). *New project supports children and young people to access the mental health benefits of live music*. University of Stirling
10. Pearce Stevens A. (2024). *Music has the power to move us physically and emotionally. Here's why*. ScienceNewsExplores
11. Goodwin, Dierker, Wu, Galea, Hoven and Weinberger (2022) *Trends in U.S. Depression Prevalence From 2015 to 2020: The Widening Treatment Gap*. American Journal of Preventive Medicine.