

Moodify

18FS Social Computing Project

Authors:

Nimra Ahmed - 16 723 934

Tobin Felder – 15 733 264

Lin Ma – 17743519

Dario David – 16707655

Areg Arakelyan - 16732935

Table of Contents

[Introduction 4](#_Toc515450255)

[Motivation 4](#_Toc515450256)

[Problem Statement 4](#_Toc515450257)

[Crowdsourcing Necessity 4](#_Toc515450258)

[Problem Definition 5](#_Toc515450259)

[People’s Perspective 5](#_Toc515450260)

[Age 5](#_Toc515450261)

[Gender 5](#_Toc515450262)

[Personality type 5](#_Toc515450263)

[Label’s Perspective 5](#_Toc515450264)

[Data’s Perspective 6](#_Toc515450265)

[Related Work 7](#_Toc515450266)

[Task description 8](#_Toc515450267)

[Input 8](#_Toc515450268)

[Output 8](#_Toc515450269)

[Task 1 8](#_Toc515450270)

[Task 2 10](#_Toc515450271)

[Experimental Set Up 11](#_Toc515450272)

[Goal of the Experiment 11](#_Toc515450273)

[Data 11](#_Toc515450274)

[Measures 13](#_Toc515450275)

[Relevant Task numbers 13](#_Toc515450276)

[Method to profile workers 13](#_Toc515450277)

[Age and Gender 14](#_Toc515450278)

[Personality 14](#_Toc515450279)

[Quality Control 14](#_Toc515450280)

[Results & Discussion 15](#_Toc515450281)

[Assumption 1: People of different ages will have a similar distribution of answers concerning a song’s mood 15](#_Toc515450282)

[Assumption 2: Females will have a more correct distribution over their mood categorizations than male participants 15](#_Toc515450283)

[Assumption 3: Personality types 'Neuroticism' will be better at recognizing the sadder moods than the other Personality types 16](#_Toc515450284)

[Assumption 4: Spotify’s ‘Good Vibes’ labelling will be correctly identified the most 17](#_Toc515450285)

[Assumption 5: With given options, the option ‘Melody’ of the music will be the greatest indicator of a songs mood 17](#_Toc515450286)

[Assumption 6: The free text input and the given options will not yield similar results 18](#_Toc515450287)

[Lessons Learned 20](#_Toc515450288)

[Resources 21](#_Toc515450289)

[Conclusions 22](#_Toc515450290)

[References 23](#_Toc515450291)

[Image Index 24](#_Toc515450292)

# Introduction

## Motivation

Music is all present in our everyday lives. People listen to it on a daily basis, be it at home, parties or on the way to work. However, with a plethora of different streaming services such as Spotify available, it can be hard to choose what to listen to – there are just too many options. As an attempted solution to this, Spotify was one of the first services to create all kinds of different playlists with a readily available selection of songs. Some of these playlists are based on moods. We now aim to find out if such mood labelling is a good option to discover new songs and if it works for people all over the board or just a specific subsection.

## Problem Statement

Our project aims to make people’s lives easier in the area of music choice as well as discovering new music. By making quantifiable statements about the accuracy of grouping songs together under different mood labels, we pave the way for the possibility of creating more accurate, customized playlists for the general population. With the gained data from these tasks, it might be possible to permanently enhance people’s listening experience by more accurately presenting them songs that fit with their mood and/or demographic. Further, by analyzing what aspects of music make a person decide on its mood we make a step towards a machine being able to more effectively, automatically categorize music.

## Crowdsourcing Necessity

This task presents itself for Social Computing since music is highly suggestive and an individual experience. A computer is not able to replicate the subtleties of human emotion, which is why people are required in order to make qualitative and quantitative statements to get to the ground of the connection between music and different moods. By letting people associate different songs with different moods and then asking them why they made the choice they did, we can

1. Analyze how different demographics and personality answer differently or if matching songs to moods is something universal for humans and
2. Give pointers to a future machine learning project as to what aspects of a song to focus on when categorizing human moods to songs.

# Problem Definition

Our main questions are how and why people associate songs with moods and if this is something universal or if it differs over different demographics and personalities. In order to fully answer these questions, we divided our assessment of the data into three perspectives with subcategories:

## People’s Perspective

### Age

Music perception changes with the listener’s age (Kathryn H. Arehart, 2014). Since this might also change interpretation, we want to be able to answer how different age groups differently categorize our songs.

**Assumption:** People of different ages will have a similar distribution of answers concerning a song’s mood.

### Gender

Not only age, but also gender has an influence on how humans perceive music – females are in general better at recognizing familiar tunes (Scott A. Miles, 2016). Since there is already evidence for females having a musical advantage in certain areas, we also want to determine if they more consistently connect songs with certain moods.

**Assumption:** Females will have a more correct distribution over their mood categorizations than male participants.

### Personality type

There is research suggesting that personality traits as defined by the OCEAN model and music is interconnected (Xavier Campaña, 2017). Based on that, we aim to find out if personality traits have an influence on connecting songs to specific moods. High Neuroticism is often associated with worse psychological well being, which is why we will test the following:

**Assumption:** Personality types 'Neuroticism' will be better at recognizing the sadder moods than the other Personality types

## Label’s Perspective

The origin of our song-mood association are predefined playlists from the streaming service Spotify. As a secondary goal of our evaluation, we want to determine how much the general population identifies with Spotify’s different labellings and which label resonates the most.

**Assumption:** Spotify’s ‘Good Vibes’ labelling will be correctly identified the most

## Data’s Perspective

In addition to the above, we wish to find out what it is specifically in a song that makes people associate it with a specific mood. We have designed this question as both one with suggested answers and one with open answers, in order to also determine if people let themselves be influenced by given options and if the answers add up.

**Assumption:** With given options, the option ‘Melody’ of the music will be the greatest indicator of a songs mood.

**Assumption:** The free text input and the given options will not yield similar results

# Related Work

There is a lot of different articles and scientific work published connecting music to different people’s perspectives, how it differs with age (Kathryn H. Arehart, 2014), gender (Scott A. Miles, 2016) or personality (Xavier Campaña, 2017). There is also research connecting music with mood in terms of how it relates to memory (Jäncke, 2008). We have not found, however, studies that exactly determine the categorization of songs concerning their mood based on the participant’s demographic and personality.

As to what components of a song determine its mood: There has been research done concerning automatic genre classification (Adam Sadovsky), however nothing was found specifically concerning human moods.

# Task description

## Input

The Input is a simple CSV file consisting of a song’s name and a link to the website where it is hosted, so it is possible to have it play on the website. The first two rows of the table look like this, then it continues in the same vein:

|  |  |
| --- | --- |
| Name | Link |
| Do Right | http://k003.kiwi6.com/hotlink/ljoqsgg0ue/GLADES\_-\_Do\_Right-30.mp3 |
| I’ll be There | http://k003.kiwi6.com/hotlink/mjtxmoctlr/King\_Henry\_-\_Ill\_Be\_There\_ft.\_Sasha\_Sloan\_-\_AudioTrimmer.com\_.mp3 |

## Output

The Output will again be a CSV file, which will serve as the foundation for our data analysis and consist of the answers to our tasks as well as specifications of the worker that solved them.

## Task 1

The first the worker sees of the tasks are the instructions we provide:

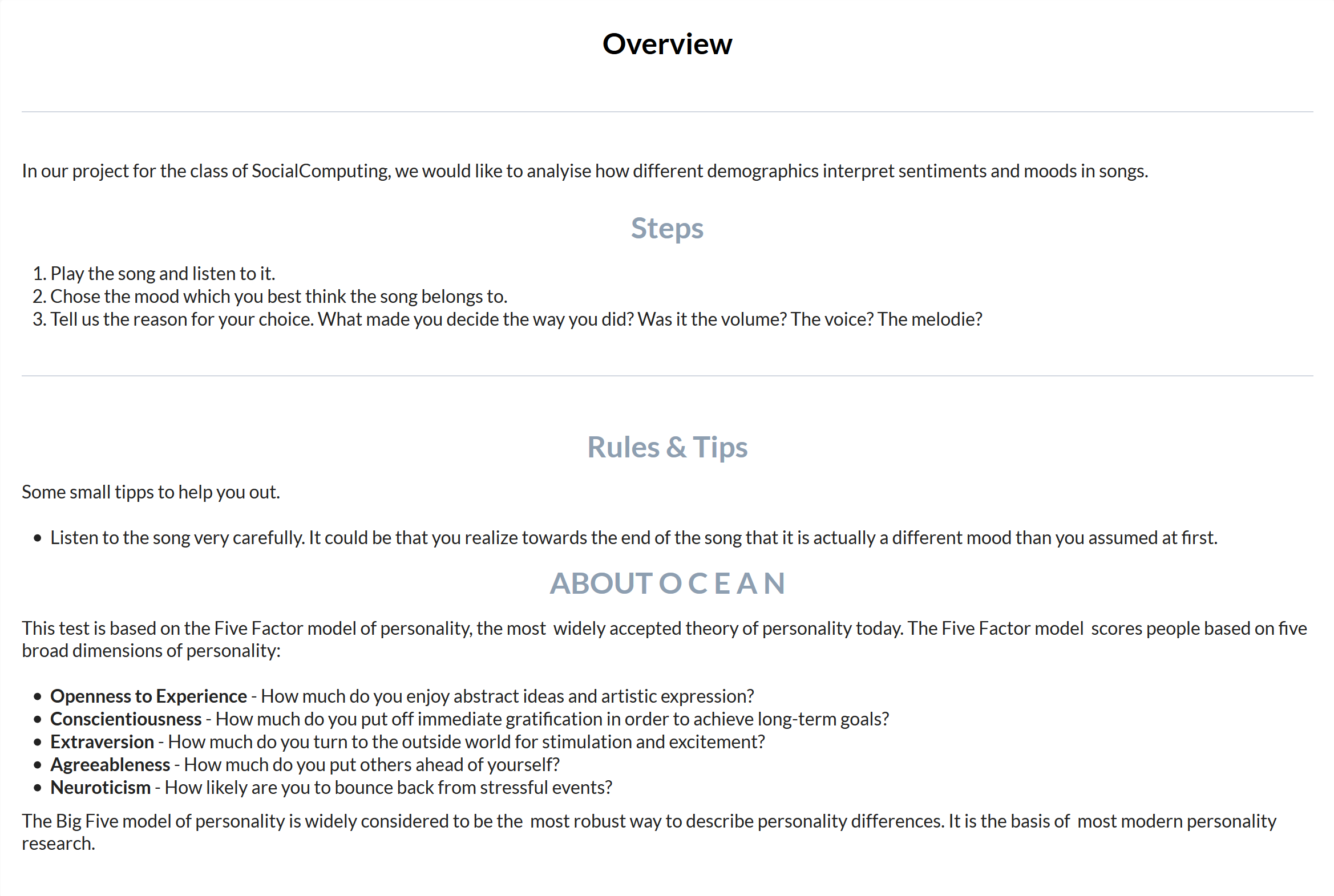


Abbildung 1: Instructions

Following that, we ask questions in order to profile the worker. These questions only have to be filled out once for the task.

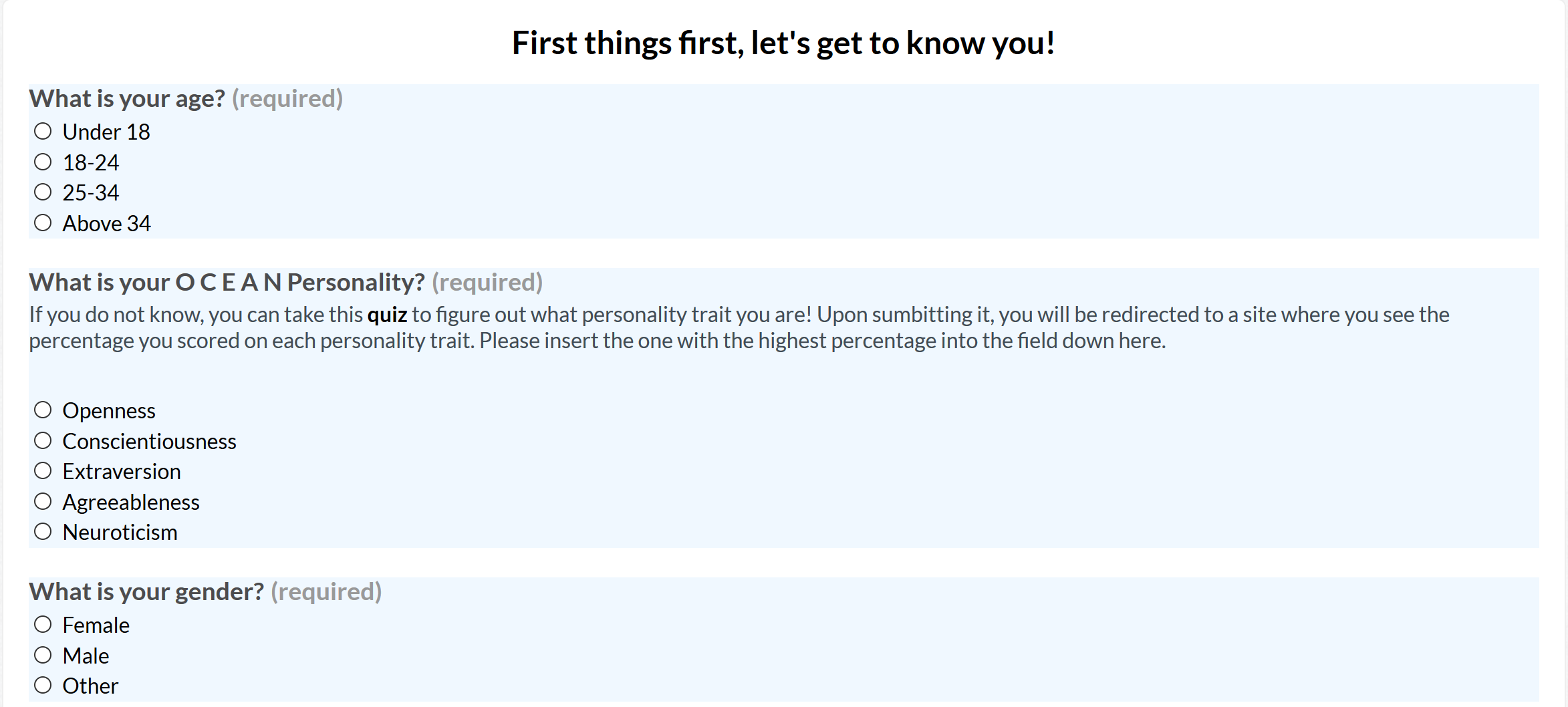


Abbildung 2: Worker Profiling

After these personal details are filled out, the worker is prompted to analyze the songs we provide.

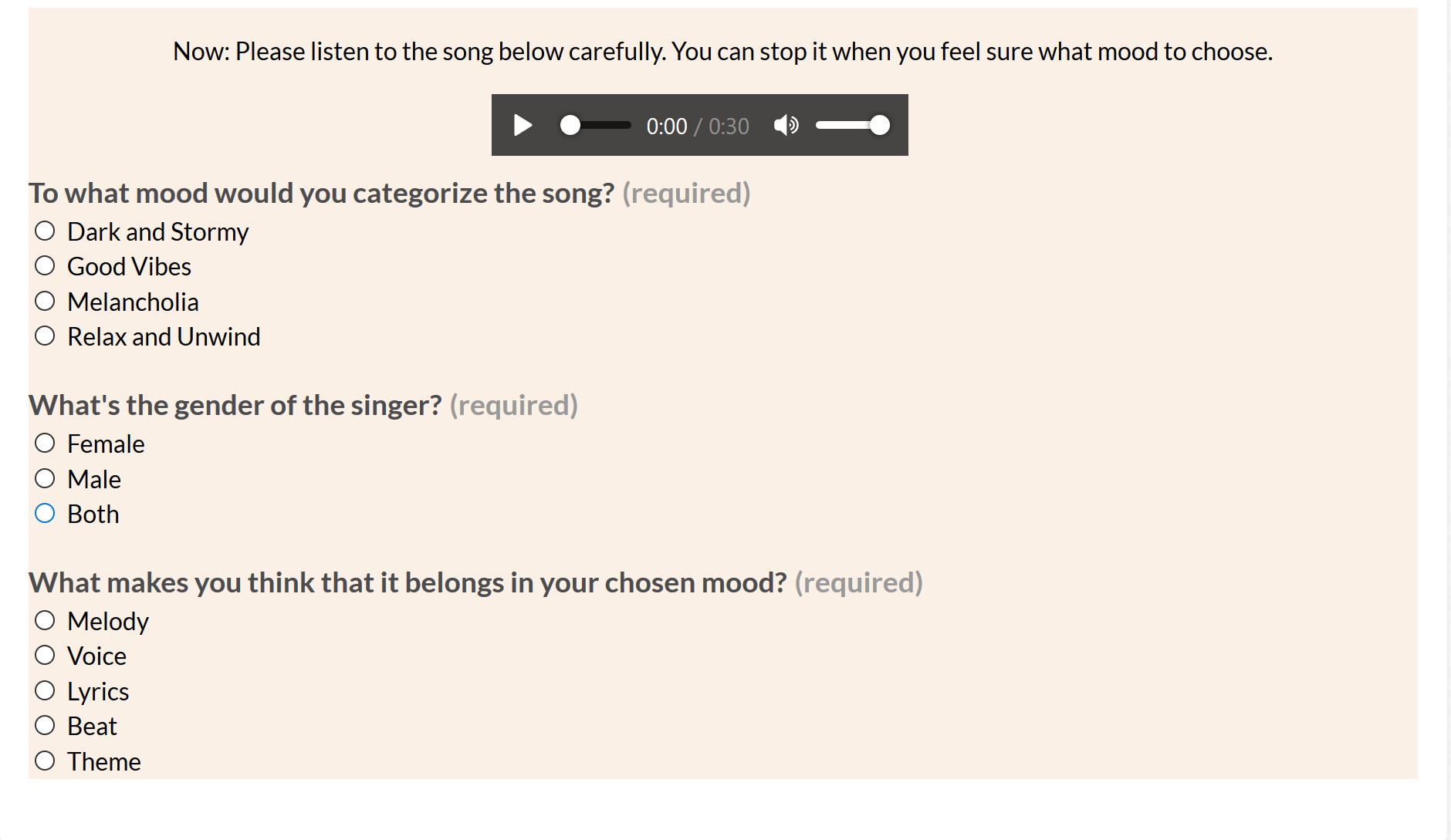


Abbildung 3: Task1 Song Evaluation

## Task 2

The Instructions and Worker profiling are the exact same in Task 1 and Task 2. Where it differs is the Song evaluation: While, in Task 1, we provide options for the worker why they put the song to their chosen mood, we gather open answers in this task.

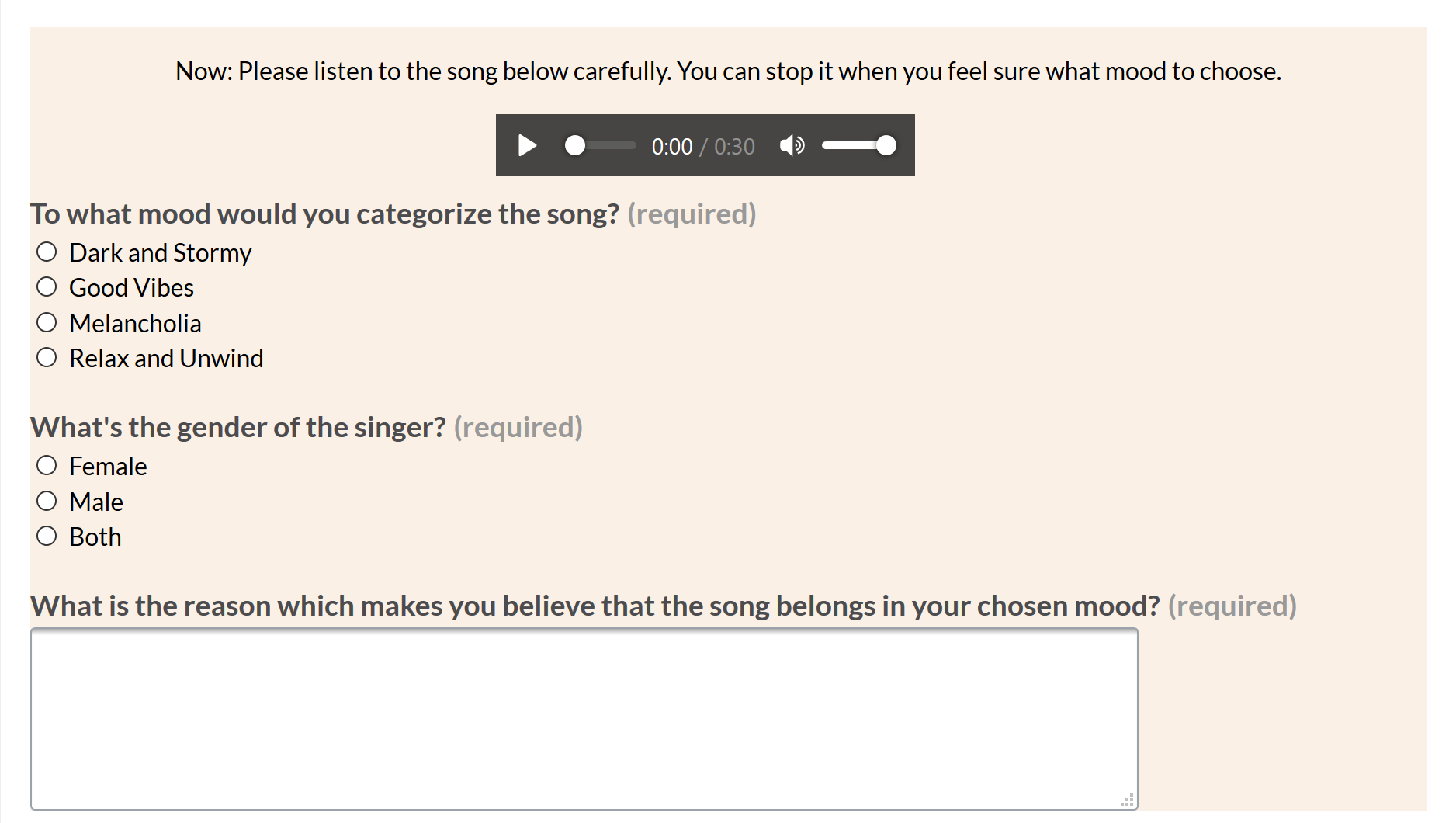


Abbildung 4: Task2 Song Evaluation

# Experimental Set Up

## Goal of the Experiment

The Goal of our Experiment is to validate or invalidate the Assumptions made under the section ‘Problem Definition’.

### Data

The data consists of 40 different songs from four different playlists, chosen based on their mood categorization on the music streaming platform Spotify.

#### Playlist Choices

The first question that presented itself was, which mood playlists by Spotify we wanted to use as our mood reference. In order to do that, we consulted the Arousal Valence Model developed by James Rusell and Lisa Feldman Barett:

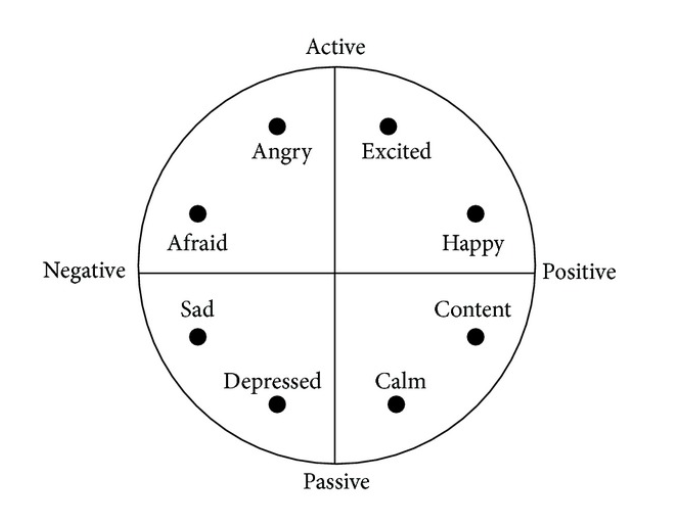


Abbildung 5: Valence Arousal Model

We aimed to find Moods that most matched the four spaces in between the labels, and after searching Spotify’s playlists we agreed on the following:

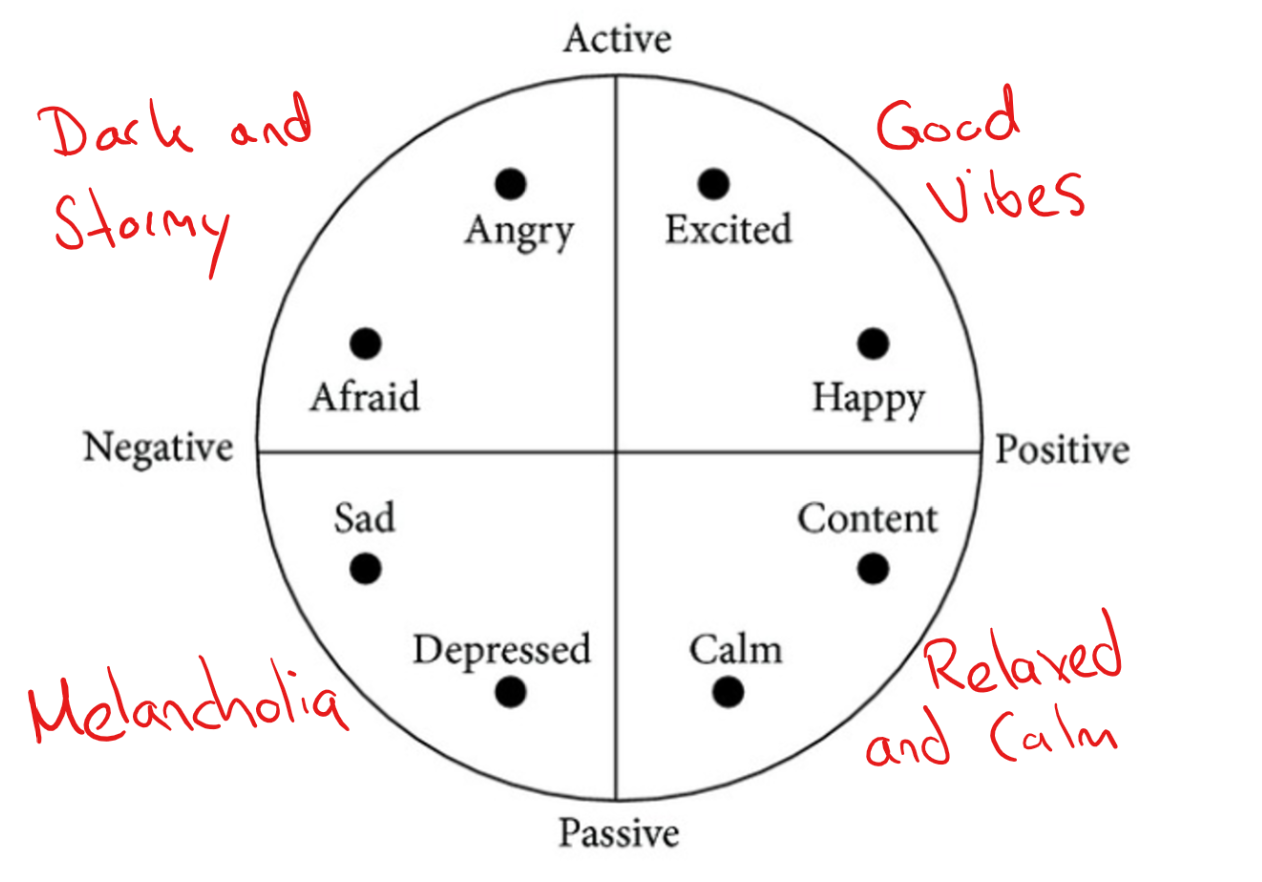
* Relaxed & Calm
  + These words are synonym to peaceful, tranquil and serene. It is a state where a person doesn’t feel any worry, anger or excitement[[1]](#footnote-1). This would correspond to the Quadrant between the axis Positive and Passive.
* Melancholia
  + Melancholia is a condition characterized by great sadness, regret, remorse and gloominess and is often connected to depression[[2]](#footnote-2). This corresponds to the Quadrant between the axis Passive and Negative.
* Good Vibes
  + Good Vibes is a colloquial term that can be translated to good energy or thoughts and prayers. It is often used in connection with other people – receiving good vibes would translate to receiving smiles, hugs and positive feelings from another person.[[3]](#footnote-3) We understand this term as an awakening of good feelings, happiness and joy. This represents the quadrant between the axis Positive and Active.
* Dark & Stormy
  + This represents the last quadrant between Active and Negative. ‘Dark’ represents negative feelings such as angst, sadness or tension while ‘Stormy’ represents action, unexpectedness.
* 

Abbildung 6: Moods&Model

### Measures

In the following we lay out how we plan to achieve the evaluations of our assumptions.

#### Assumption1: People of different ages will have a similar distribution of answers concerning a song’s mood.

We will visualize how workers of different age groups have chosen moods and analyze if there are discrepancies depending on the age. In order to ensure that the results are not skewed because of the sample’s random natures, we will also factor in which mood would have been correct according to Spotify and if there is noticeable deviation.

#### Assumption2: Females will have a more correct distribution over their mood categorizations than male participants.

We will again visualize how workers of different gender have chosen moods and analyze if there are discrepancies depending on the gender. Again we will ensure quality by double-checking accuracy.

#### Assumption3: Personality types 'Neuroticism' will be better at recognizing the sadder moods than the other Personality types

We will proceed similarly as explained above and compare the answers of different personality types.

#### Assumption4: Spotify’s ‘Good Vibes’ labelling will be correctly identified the most

We will check if the workers have labelled the songs the same as Spotify and visualize the percentage of ‘correct’ answers given.

#### Assumption5: With given options, the option ‘Melody’ of the music will be the greatest indicator of a songs mood.

We visualize the different options chosen in a pie chart and analyze the results.

#### Assumption6: The free text input and the given options will not yield similar results

We compare the options chosen with the free text input given.

### Relevant Task numbers

A single task consists of the instructions, the profiling data (age, gender and personality, including having to take a quick OCEAN test) and then 10 rows of songs. Each row consists of a 30-second song segment and the questions: Which mood to categorize the song to, the gender of the singer and why this mood was chosen.

Time to complete a task is on average six minutes.

## Method to profile workers

### Age and Gender

Two out of three relevant factors for us are the age and gender, which we ask of the worker at the beginning of the task.

Age is separated into 4 categories: Under 18, 18-24, 25-34 and above 34. These categories were chosen based on Spotify’s user statistics[[4]](#footnote-4), as it is such a big streaming service that it is representative for music listeners all over the globe.

### Personality

The third factor we needed for our evaluation is the personality. First, we needed to determine a personality model to utilize. Since it was also used in other music-based research (Xavier Campaña, 2017) as well as in combination with Crowdsourcing (Gabriella Kazai, 2011), we quickly decided on the OCEAN personality model.

The description of the model, taken from truity.com, is:

“This test is based on the Five Factor model of personality, the most widely accepted theory of personality today. The Five Factor model scores people based on five broad dimensions of personality:

* **Openness to Experience** - How much do you enjoy abstract ideas and artistic expression?
* **Conscientiousness** - How much do you put off immediate gratification in order to achieve long-term goals?
* **Extraversion** - How much do you turn to the outside world for stimulation and excitement?
* **Agreeableness** - How much do you put others ahead of yourself?
* **Neuroticism** - How likely are you to bounce back from stressful events?

The Big Five model of personality is widely considered to be the most robust way to describe personality differences. It is the basis of most modern personality research. “ (truity.com, kein Datum)

## Quality Control

In order to ensure quality control, we implemented gold questions. As it is not possible to mark subjective mood placements as right or wrong, we added an additional question solely for ensuring quality: We asked the workers what the gender of the person singing the song is. In six gold questions for each task we check if the worker has marked this question correctly. If not, we assume he didn’t listen to the song and disregard his answers.

# Results & Discussion

The results are analyzed based on percentages of correct guesses. A correct guess is defined by the participant categorizing a song to the same mood as Spotify did, which we took as the ground truth.

A total of 17 workers completed our task, after quality control a total of 271 rows were worked on, meaning 271 songs were categorized.

## Assumption 1: People of different ages will have a similar distribution of answers concerning a song’s mood

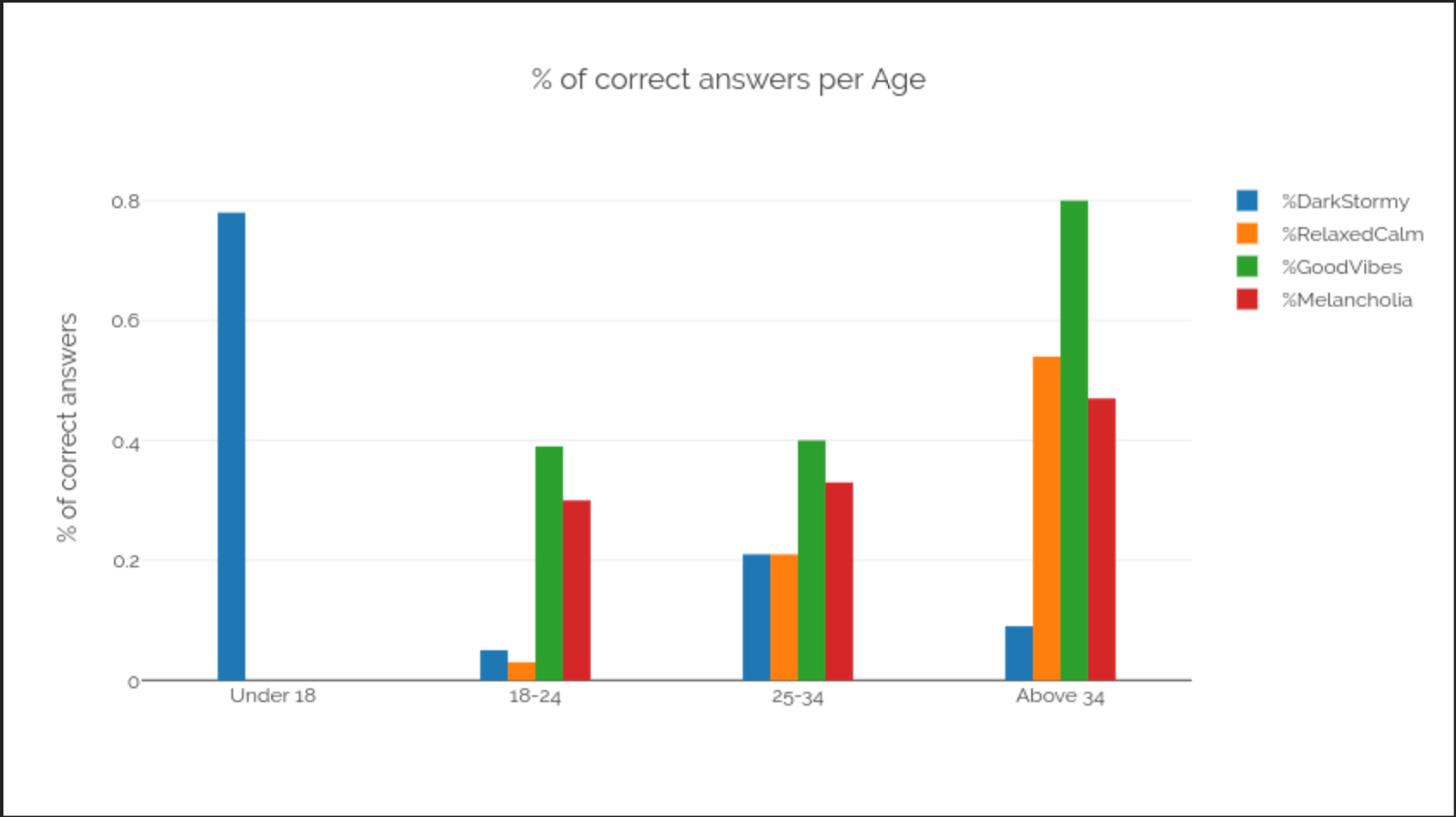


Abbildung 7: Age Distribution

Each age group categorized the moods ‘Dark and Stormy’ and ‘Relaxed and Calm’ differently – there is no consistency across the ages. The only similarity lies in the categorization of ‘Good Vibes’ and ‘Melancholia’ in the age groups ’18-24’ and ’25-34’.

As a result, based on our data this Assumption is **false**.

## Assumption 2: Females will have a more correct distribution over their mood categorizations than male participants

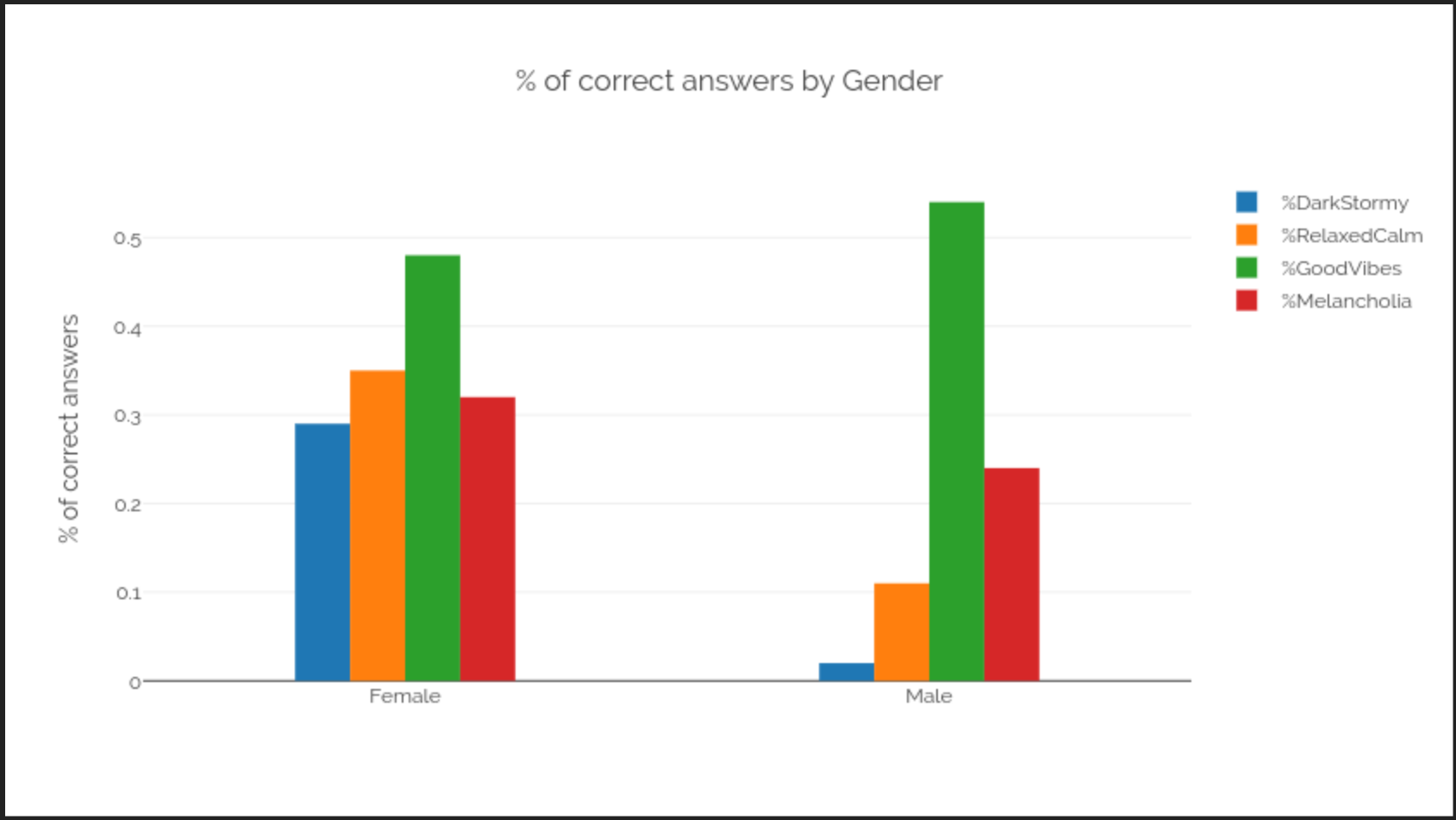


Abbildung 8: Gender Distribution

Both genders have a roughly similar percentage concerning the categorization of the moods ‘Good Vibes’ and ‘Melancholia’. However, females were clearly able to better recognize the moods ‘Relaxed and Calm’ and ‘Dark and Stormy’.

Overall, mood recognition is distinctly better in females, which is why according to our data this assumption is **true.**

## Assumption 3: Personality types 'Neuroticism' will be better at recognizing the sadder moods than the other Personality types

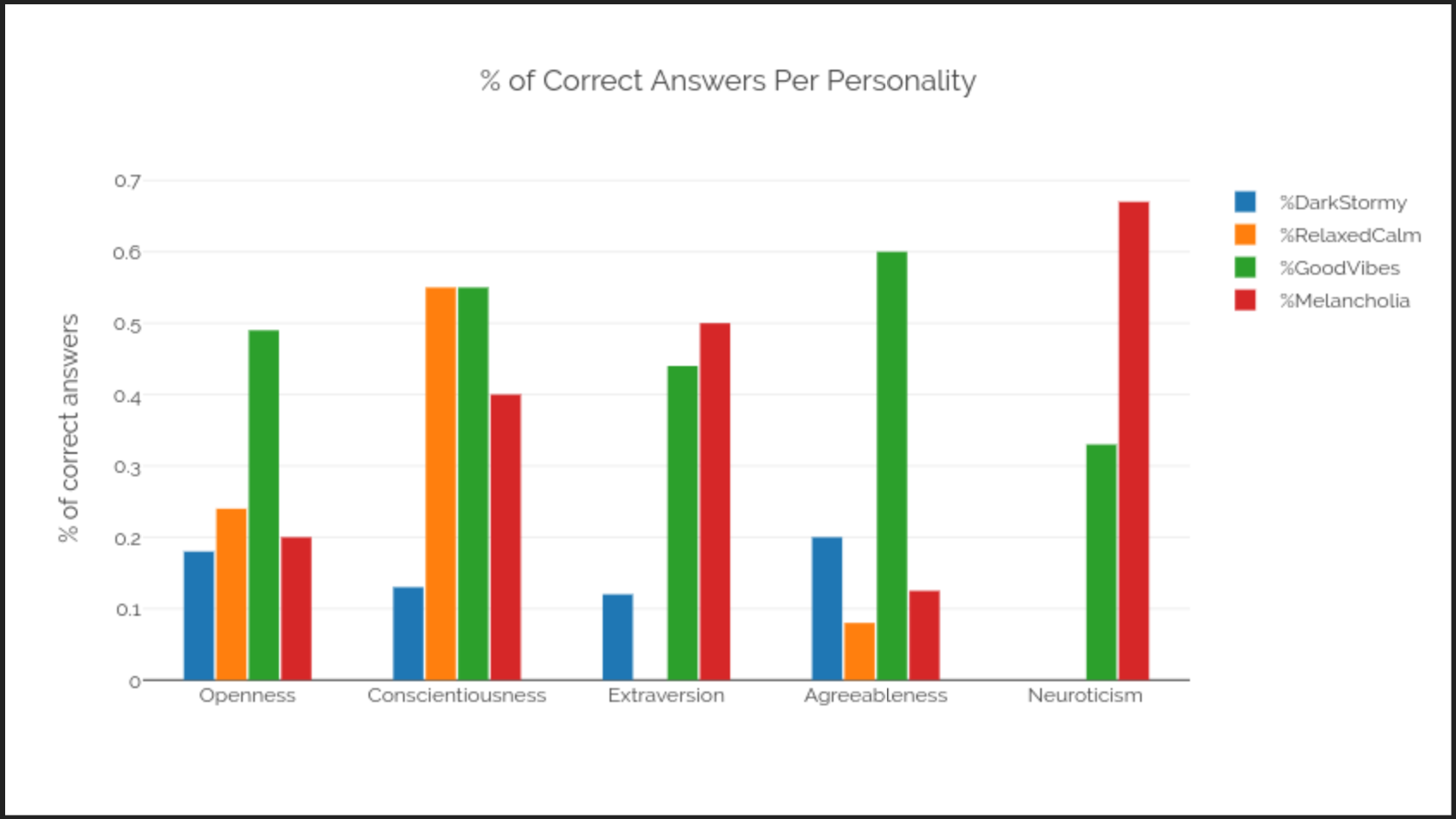


Abbildung 9: Personality Distribution

Our example for ‘Neuroticism’ only shows data for two moods. However, since one of them is happy (‘Good Vibes’) and one sad (‘Melancholia’), we can still test our assumption.

Our workers for personality type ‘Neuroticism’ has the worst overall correctness for assigning songs to the mood ‘Good Vibes’. Furthermore, they have by far the highest assignment rate for songs that belong to the category ‘Melancholia’. While there is a tendency for all other personality types to better recognize the happy moods, ‘Neuroticism’ is the opposite.

Due to a high correctness concerning ‘Melancholia’ and a low one for ‘Good Vibes’, our data suggests this answer is **true**.

## Assumption 4: Spotify’s ‘Good Vibes’ labelling will be correctly identified the most

With ‘Good Vibes’ having a correctness percentage of just over 50% and the next highest being below 30%, our data clearly suggests this Assumption to be **true.**

## Assumption 5: With given options, the option ‘Melody’ of the music will be the greatest indicator of a songs mood

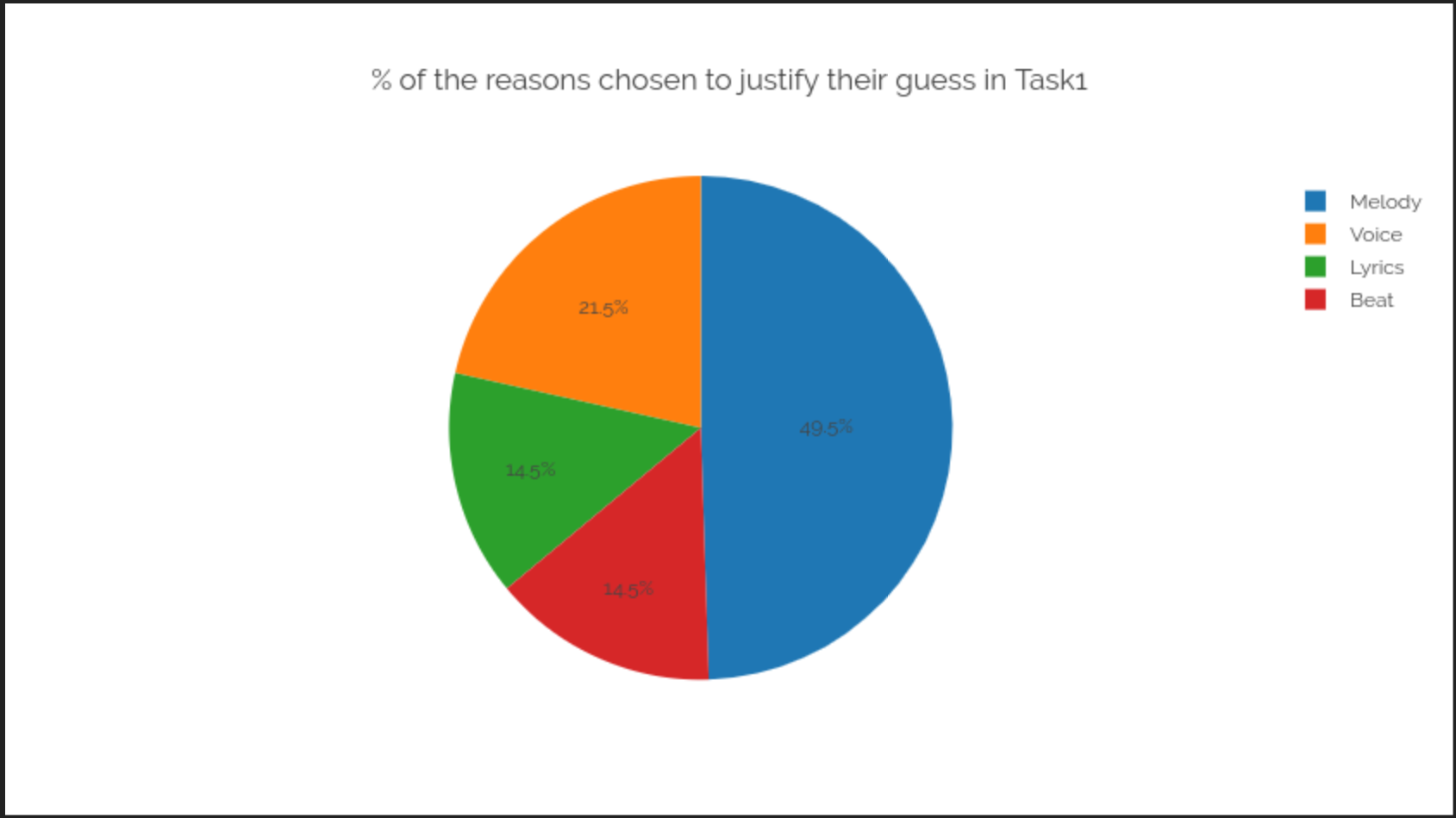
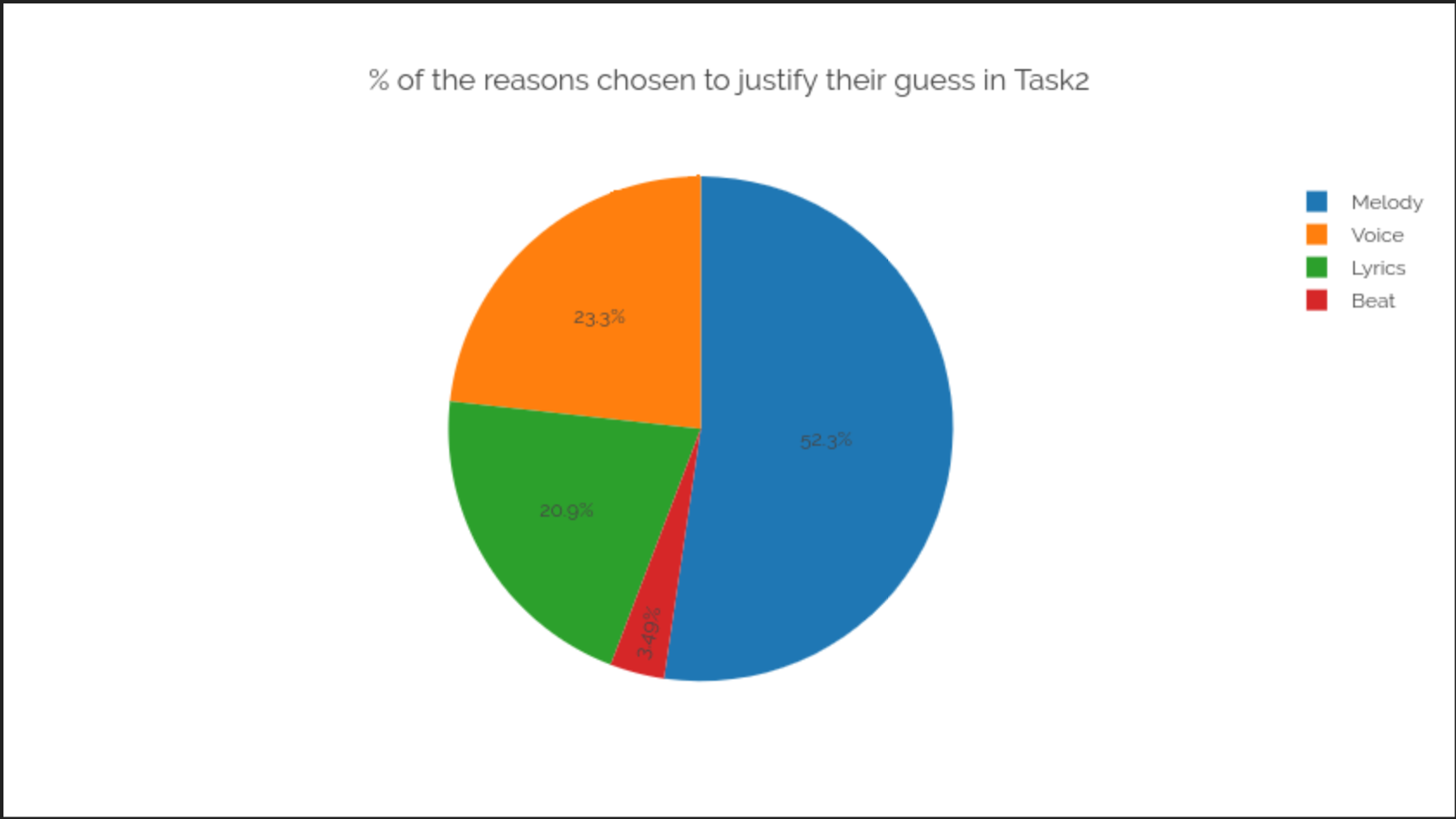


Abbildung 10: Choice justification

Melody was by far the most chosen option, with nearly half of all workers selecting it. So based on our data, this Assumption is **true.**

## Assumption 6: The free text input and the given options will not yield similar results



While in the free text input long explanations were given instead of just single words, a text search for keywords revealed that most of our options were also freely given by the workers. What’s more, the distribution of answers is nearly the same as when we suggested answers.

We can conclude that giving options to choose from did not influence the worker and that based on our data, this Assumption is **false.**

# Lessons Learned

Over the course of this project, we learned a great many lessons. Some of them were technical, such as plotting out data from a csv file as a source or setting up a Crowdsourcing task for people to solve.

An important part of what we learned was experimental task design. We learned how to scientifically approach a question, set up an experiment to find answers to the question and then visualize data to help with interpreting the received results. We also learned the value of quality control, as it is very possible with a large number of participants that some will not take the task seriously and just fill out anything.

Nearly more valuable, however, were lessons learned concerning group projects, be it now in university or for later in life. We realized that it is necessary for someone to step up, take control and organize the group – if this is not done, there is a danger that everyone will see someone else as responsible for successfully carrying out the task, which can lead to nothing getting done at all.

Further, we saw the value in time management – if the project is not organized early on, tasks divided equally and preparations for successful deployment made, things can get rushed as deadlines are approaching, either because things were put off for too long or not everyone is doing their part.

# Resources

Input data files as .cvs as well as our data analysis code is in an attached .zip folder.

# Conclusions

Based on our data, it is safe to say that a song’s mood is interpreted differently depending on who listens to it.

People of different ages do not perceive a song’s mood the same way: Younger people (below 35 years of age) especially have trouble with assigning a song to its mood, while the data suggests that the older one is, the more likely it is one chooses the correct mood to a song (formulated differently, more people choose the same mood to a song).

There is also a noticeable difference when it comes to the gender of the person categorizing a song. Based on our data, males are overall worse at categorizing songs, though there was a bit of a higher rate of success concerning the mood ‘Good Vibes’.

Further, our data suggests a strong correlation between personality type and mood perception. We assumed that ‘Neuroticism’, a personality type generally associated with worse mental health than others, would be better at recognizing and assigning sad moods. Our results show that this is indeed true: Workers with the personality type ‘Neuroticism’ not only assigned songs to the mood ‘Melancholia’ with greater success than the others, they in fact had the highest percentage of correct assignment out of them all.

Overall, though, the mood ‘Good Vibes’ saw the highest success rate all over the board: No matter the age, personality type or gender, ‘Good Vibes’ was assigned correctly to songs almost exclusively the most.

Why is it that such huge differences lie in perception and assignment of music? We asked our workers what it is in a song that makes them associate it with a certain mood. We posed this question both in the format of suggested options and open-end questions, which however both delivered the same results: Over 50% of asked participants noted that a songs melody is most telling what mood it belongs to.

# References

Adam Sadovsky, X. C. (n.d.). *Song Genre and Artist Classification via Supervised Learning from Lyrics.*

Gabriella Kazai, J. K.-F. (2011). *Worker Types and Personality Traits in Crowdsourcing.* Glasgow.

Jäncke, L. (2008, August 8). Music, memory and emotion. *BMC Biology*.

Kathryn H. Arehart, P. N. (2014, Mar-Apr). Effects of Age on Melody and Timbre Perception in Simulations of Electro-Acoustic and Cochlear-Implant Hearing. *HHS Author Manuscripts*. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3970813/

Scott A. Miles, R. A. (2016, March 1). Sex Differences in Music: A Female Advantage at Recognizing Familiar Melodies. *frontiers in Pychology*.

*truity.com*. (n.d.). Retrieved from https://www.truity.com/test/big-five-personality-test

Xavier Campaña, R. A. (2017). Experience in Applying Data Mining Techniques to Musical Content . *International Journal of Applied Engineering Research* .

# Image Index

[Abbildung 1: Instructions 5](#_Toc515447679)

[Abbildung 2: Worker Profiling 6](#_Toc515447680)

[Abbildung 3: Task1 Song Evaluation 6](#_Toc515447681)

[Abbildung 4: Task2 Song Evaluation 7](#_Toc515447682)

[Abbildung 5: Valence Arousal Model 8](#_Toc515447683)

[Abbildung 6: Moods&Model 9](#_Toc515447684)

[Abbildung 7: Age Distribution 12](#_Toc515447685)

[Abbildung 8: Gender Distribution 13](#_Toc515447686)

[Abbildung 9: Personality Distribution 14](#_Toc515447687)

[Abbildung 10: Choice justification 15](#_Toc515447688)

1. https://www.thefreedictionary.com/calmness [↑](#footnote-ref-1)
2. http://www.dictionary.com/browse/melancholia [↑](#footnote-ref-2)
3. http://www.yourdictionary.com/vibes [↑](#footnote-ref-3)
4. https://www.statista.com/statistics/475821/spotify-users-age-usa/ [↑](#footnote-ref-4)