Music Enabled RunningBusiness Proposal

korte regel

**FR Corp  
  
\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**Mitchel Kuijpers  
Mailinh Luong  
Arwen Voorn  
Filip Vangelov  
  
13 september 2021

# Version History

| **Version** | **Date** | **Who** | **Change** |
| --- | --- | --- | --- |
| V0.1 | Sep 13, 2021 | Mitchel Kuijpers | * Initial setup of document and chapters |
| V0.2 | Sep 13, 2021 | Mitchel Kuijpers | * 1.Project Statement Summary   + 1.1.Client’s Background   + 1.2.Current Situation   + 1.3.Business Understanding * 5.Planning |
| V0.3 | Sep 15, 2021 | Mai Linh Luong | * 3. Scope   + 3.1 Functional Areas in Scope   + 3.2 Organizational Scope   + 3.3 Technical Scope * 7. Conclusion |
| V0.4 | Sep 15, 2021 | Mitchel Kuijpers | * 1. Project Statement Summary   + 1.1 Client’s Background   + 1.2 Current Situation   + 1.3 Business Understanding |
| V0.5 | Sep 15, 2021 | Arwen Voorn | * 2. Business Requirements   + User requirements * 6. Risk Log |
| V0.6 | Sep 15, 2021 | Filip Vangelov | * Introduction * 4. Data |
| V1.0 | Sep 16, 2021 | Mai Linh Luong | * Final check on the document * Distribution * Approval |
| V1.1 | Sep 20, 2021 | Mai Linh Luong | * Feedback processed 3. Scope |
| V1.2 | Sep 20, 2021 | Mitchel Kuijpers | * Feedback processed 2.Business Requirements |
| V2.0 | Sep 22, 2021 | Mai Linh Luong | * Feedback processed Introduction * Final check on the document * Distribution |

# 

# Distribution

This document is sent to the following:

| **Version** | **Name** | **Role** |
| --- | --- | --- |
| V1.0 | Olaf Janssen | Client |
| V2.0 | Olaf Janssen | Client |

# Approvals

This document is approved by the following:

| **Version** | **Name** | **Role** |
| --- | --- | --- |
|  |  |  |

# Table of Contents

[**Version History**](#_k4ufixjvgqe8) **1**

[**Distribution**](#_kl2vizbvbl0e) **3**

[**Approvals**](#_v9ri2d1imhrh) **3**

[**Table of Contents**](#_l7wp1t7key29) **4**

[**Introduction**](#_p3fnnqk5br96) **5**

[**1.Project Statement Summary**](#_8rg0l57vhrih) **6**

[1.1.Client’s Background](#_b2g0ej9otve7) 6

[1.2.Current Situation](#_o8ymlkb90pc) 6

[1.3.Business Understanding](#_nwdfuho5ag7h) 6

[**2.Business Requirements**](#_hminkb6p6e71) **7**

[2.1.Business requirements](#_vrm708119t38) 7

[2.2.User requirements](#_o3d0j0xmcn1m) 7

[**3.Scope**](#_rkcvysze7qau) **8**

[3.1.Functional Areas in Scope](#_334wbeskd7ys) 8

[3.2.Organizational Scope](#_49cr3ms953e5) 9

[3.3.Technical Scope](#_mb097yme7z72) 9

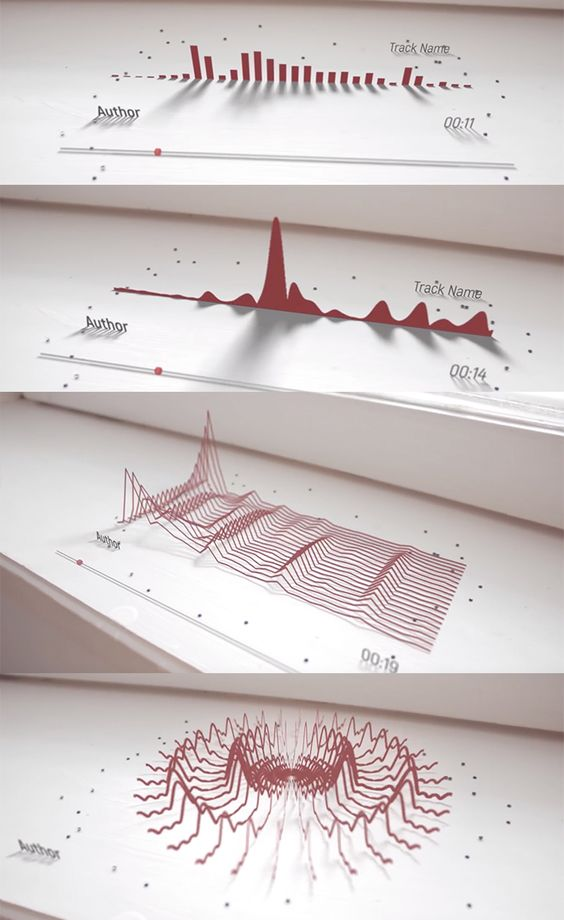
[**4.Data**](#_jpoh44wve0dz) **10**

[**5.Planning**](#_my928kghh3ne) **13**

[**6.Risk Log**](#_7c1hm39ax2qs) **14**

[**7.Conclusion**](#_kwt8u9xj9gva) **15**

# Introduction

Music as part of the human history has been around since 300,000 to 50,000 BP to be precise, for this time period we have noticed the undisputed truth that songs have impact on human behaviour, whether it is emotional, physical or purely psychological, music influences us all in a different way, it can make us happy when we are not, it can make you dance when you don't feel like it. Dancing in a silent room will feel strange, add some tones and beats to it and see how the atmosphere changes. With the progress of the music industry offering us plenty of different genres and compositions from different artists we can create a personal playlist for every human on the earth and not cross each others musical preferences, in other words different people like different songs and get influenced in a slightly different way, but the underlying fact is that there is a process of influence. 

The question we are going to give an answer to with this project is, what is the exact influence of different songs when performing tasks such as running and can we manipulate it in order to reach more efficient results. Exercising comes in different forms, some people like going to fitness, others prefer playing sports and of course running. The overall goal of exercising is to perform your tasks correctly and efficiently for your body and health. From here the question arises “*are there specific songs, genres that improve your performance*”, of course the exact results will differ from person to person, but this project will provide a proof of concept, and this concept can be applied to every other person. The original goal of the project was not about performance but health. Both health and performance can be influenced by music but it’s important to keep in mind that running at a higher performance doesn’t necessarily mean it is a healthier way of running.

# 1.Project Statement Summary

The project statement summary gives an overall summary of the project. This chapter will describe the client's background, the current situation the client is in and the business understanding, where it is explained what is being done during the project.

## 1.1.Client’s Background

The Department of ICT, Fontys University of Applied Sciences, with sponsors like Fontys IXD and Vitality Living Labs (VLL), is currently working on improving the running quality of runners with the help of recommending specific songs within a running session.

## 1.2.Current Situation

Currently there is a measurement platform that collects sensor data and music data of test runners in the field. The data that is collected is processed during the running session. Music recommendations are also generated during this session. However, data filtering currently is still limited and the music recommendations are still primitive and not yet validated.

## 1.3.Business Understanding

It is important to properly understand the structure of the business wherein the project will be made. This will allow us to have a better starting point of our data analysis and explain the different correlations and other findings we could encounter when working with the data regarding the requirements that have been set.

During this project we will be analysing data that has been gathered from a running session of one person. The idea is to find correlations between the music and running data, which could show how music affects different variables in a person's running pattern. These variables are for example the impact that a runner puts his foot down on the ground or the power in the runner's steps compared to the music that is playing.

# 

# 2.Business Requirements

Why music and running? The hypothesis is that **emotion** evoked by music can be used to steer running technique into a controlled direction (injury free).

Based on this hypothesis a (deep) machine learning technique will be used to generate song selection for optimal running performance. If this hypothesis is true and implementable, then this technique can help people worldwide improve their running technique.

Where does FR Corp come into this? **Main challenge:** design and implement a (personalized) music suggestion model.

## 2.1.Business requirements

* Filter noise from data, e.g. data during cornering, rests, climbs.
* Predicting the relation between running symmetry and musical parameters (or individual songs) is possible.
* Predict whether a song improves running quality within a session.
* Determine a proper running quality indicator (based on symmetry).

## 2.2.User requirements

* Looking at Spotify tracks that a user listens to and plot the danceability/energy into an exploratory data analysis.
* Document which key transitions happen in the playlist (from major to minor key) and how many.
* Figure out if the running pattern is good or bad based on symmetry, impact and efficiency.

These user requirements are based on the second meeting with the project client Olaf Janssen.

Next to the business and user requirements it is also important to note why there is an interest in the transitions between songs while running.

A large number of external factors could destroy a clear relation between a song and the way someone runs on music. However, we expect that ‘good’ music has a positive effect within a running session compared to other songs that won't have the same impact as a ‘good’ song.

One way of looking within sessions is to study transitions. Another way is to come up with an ELO-like rating system.

# 3.Scope

The scope for this project is to examine if there is a correlation between music and the performance of the running. In this chapter we will discuss the functional areas in the scope and the technical scope.

## 3.1.Functional Areas in Scope

The main scope of the project is to examine if music has an effect on the running. For that reason we are going to merge the footpods and music dataset. The dataset is provided by the client mr. Olaf Janssen. Data that has no added value to the project will be excluded in the scope of this project.

To investigate if music has an effect on the running, during the project we are going to filter out the junk dataset. Junk dataset is in this case when the runner is pausing the music or stopping for traffic lights, cars etc. So in other words, there is an outlier in the runner’s running pattern. So the ultimate goal is to filter out the junk dataset and keep the bouts.

In this project, we are also going to take a look at Spotify tracks that the user listens to and plot the danceability/energy along with other parameters into an exploratory data analysis. In this case we dive more into the music dataset and use external sources like Spotify. On top of that, there will be documentation of the effect of the running based on key transitions (e.g. from major key to minor key) in the song.

Lastly with all the analysis and research that has been made within the project, there will be an evaluation if the running pattern is good or bad based on symmetry, impact and efficiency. In short, the client is interested in rating kinds of solutions.

## 3.2.Organizational Scope

The project team consists of 4 team members who will be actively working on the project and 7 people of the company who will be co-responsible for the process of the project. Two tutors will be involved in the project as well to make sure that the project is going to reach the expectations. The client is the final decision maker in this project and will also be evaluating the work of the project.

Every Monday at 9:00 AM there will be an online stand-up on Teams with the project team where the process of the project will be discussed.

Every Tuesday at 10:00 AM there will be a company stand-up with the tutor on location TQ-5 where the process of the project will be discussed

Every Tuesday at ..:... AM there will be a client-meeting on location TQ-5 where the process of the project will be discussed.

Every Wednesday at 9:00 AM there will be a technical-tutor meeting on location TQ-5 where the process on the technical part will be discussed.

## 3.3.Technical Scope

For this project, several working environments will be used for cleaning, analysing and visualizing the data. The project team chose the following programming environments and tools:

* **Anaconda Navigator and Jupyter Notebook** - Our main working environment for the project. With the help of Python and Pandas we are going to collect, clean, munging, analyze, interpret and visualize the data
* **Github** - The working environment that we use to collect the information and data from the client that would help us with the data cleaning.

# 

# 4.Data

In this chapter you can find a full description of the data sets that are going to be used during this research. The sets were provided by Janssen,Olaf O.T.A from Fontys University of Applied Sciences. The main data sets are we are going to focus our research on are footpods.csv, footpods\_sc.csv, music.csv, phone\_activity.csv, phone\_motion.csv, phone\_location.csv and sessions.csv, these data sets are generated by a single test subject and cover around 30 sets of running sessions. Below you can see the full description of all attributes and variables of these data sets.

| footpods.csv | rows: 274389 | |
| --- | --- | --- |
| **Variables** | **Description** | **Type** |
| t | Timestamp [s] of sensor event, not the exact step timestamp | datetime64[ns] |
| foot | Foot for pod 'left' or 'right' | object |
| pronation | Maximum foot pronation [deg]. Is expected to be a negative value | float64 |
| braking | Braking force on foot on impact [G] | float64 |
| impact | Downward maximal foot impact [G] | float64 |
| contact\_time | Time between initial contact and final contact of foot [ms] | int64 |
| flight\_ratio | Ratio between the time that foot is in the air to the total stride duration. | float64 |
| strike | Initial strike contact of the foot, 1=heel, 15=toe | int64 |
| power | Power of the step [W] | int64 |

| footpods\_sc.csv | rows: 197537 | |
| --- | --- | --- |
| **Variables** | **Description** | **Type** |
| t | Timestamp [s] of sensor event, not the exact step timestamp | datetime64[ns] |
| foot | Foot for pod 'left' or 'right' | object |
| cadence | Strides per minute [/min]. This is about half the step frequency | int64 |
| speed | Average speed of the foot [m/s] based on estimated user height (may not be set properly) | float64 |

| sessions.csv | rows: 39 | |
| --- | --- | --- |
| **Variables** | **Description** | **Type** |
| session\_id | id of the session, total 39 different id’s. | object |
| user\_id | id of the user (only one user\_id available) | object |
| t\_start | starting time of the session | datetime64[ns] |
| t\_end | ending time of the session | datetime64[ns] |
| duration | t\_end - t\_start , the duration of the whole session in s | float64 |

| music.csv | rows: 6945 | |
| --- | --- | --- |
| **Variables** | **Description** | **Type** |
| t | Timestamp [s] of sensor event, not the exact step timestamp | datetime64[ns] |
| track\_uri | URI of the track being played | object |
| paused | Whether the player is playing or on pause | bool |
| artist | String of the artist name | object |
| track | String of the track title | float64 |
| context\_uri | URI of the track context, usually a playlist | object |
| context | String of the track context, usually a playlist | object |
| position | Position [s] of the playhead in the track | int64 |
| repeat\_mode | State of the repeat mode, 'off', 'track', 'context' | object |
| shuffle | State of the shuffle option | bool |
| crossfade | State indicating whether crossfade is turned on | bool |

*There are more available data sets and sources for the project, these are the examples of the data sets that will be mostly used during the EDA period of the project. More information can be found here - https://github.com/olafjanssen/mergait/blob/main/data/codebook.md*

# 

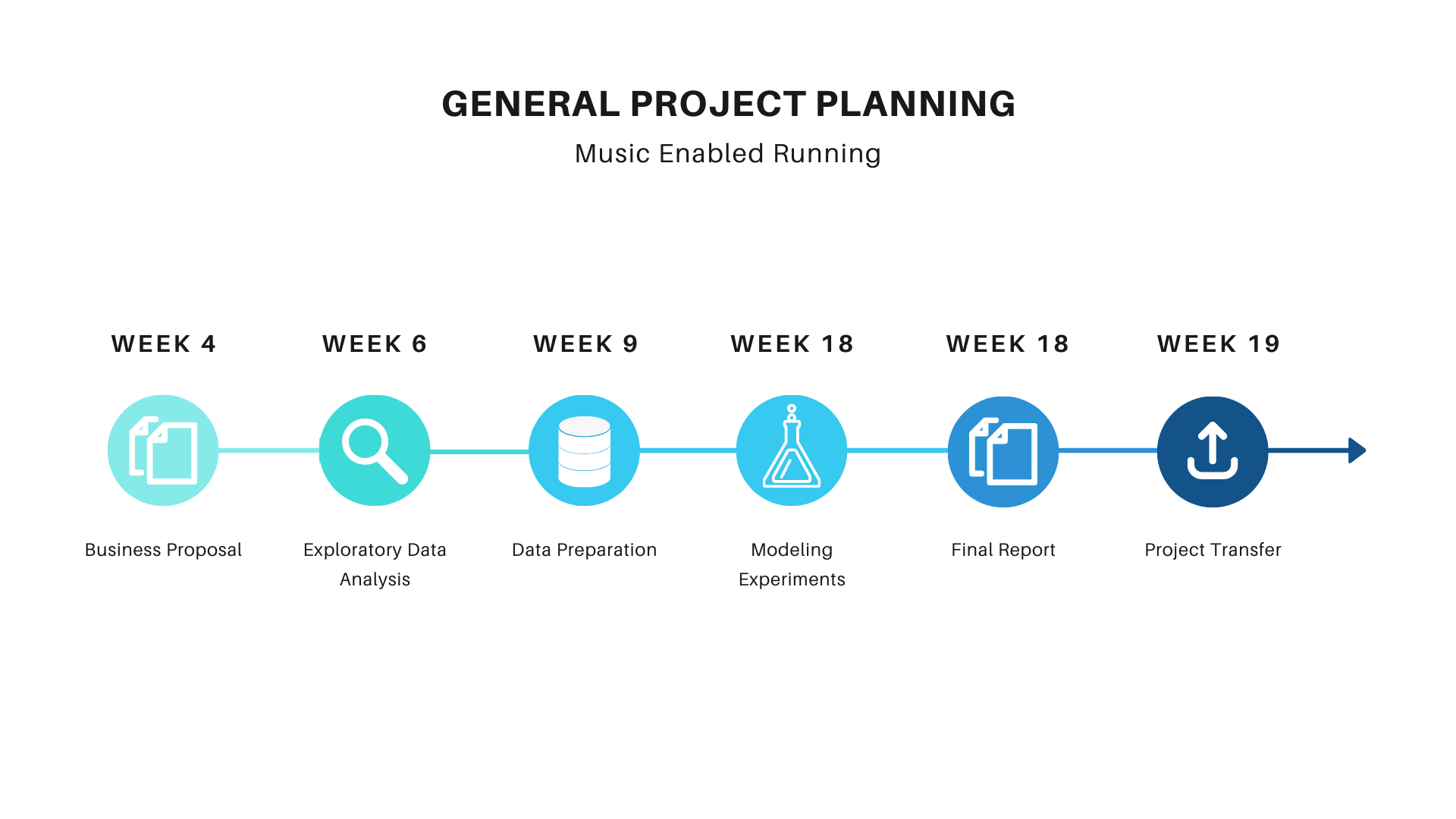
# 5.Planning

In this chapter we will show the project planning that we are going to use during this project. This will give an overview of how the project has been planned and what deadlines have been set. The project will be done in three different phases. These are the following.

* **Phase 1: Business Proposal**
  + During the first phase we will take a look at the business understanding and analytic approach for the rest of the duration of the project in phase two and three. During this phase the business proposal will also be delivered to the client.
* **Phase 2: Data Quality**
  + During the second phase the quality of the data will be analysed. We will look at the requirements of the project, understand the data and prepare the data for the third phase of the project.
* **Phase 3: Machine Learning & Reporting**
  + During the last phase of the project we will begin with the modelling of the data. During this phase we will also evaluate what has been made, gather feedback and make the final report, where our findings will be described.

The first phase incorporates a business proposal for the project. Phase two and three will be held in repeating iterations during the project duration. During weeks 2-4 phase one will be held once. Between weeks 4-10 phases two and three will be executed once and in weeks 11-18 phases two and three will be executed a second time. During the second iteration we hope to improve on the first iterations of phase two and three.

On the image below the general project planning can be seen. In this general overview we can see in what week a deadline is set for a specific deliverable, which will help us keep track of the project.



# 

# 6.Risk Log

| **Risk assessment** | | | | |
| --- | --- | --- | --- | --- |
| **ID** | **Risks** | **Gross** | **Gross** | **Gross** |
|  |  | **Likelihood** | **Impact** | **Risk** |
| R1 | A single test subject is used, results may not be representative of all users. | High | Low | Medium |
| R2 | Other factors have an impact on running performance, weather, underground, footwear, fitness and mood. | High | High | High |
| R3 | Music tastes differ per person, the machine learning algorithm needs to select music based on the music taste. | High | Medium | High |
| R4 | Outliers can impact the average of columns | Medium | Medium | Medium |

# 7.Conclusion

The primary goal for this project is to examine if there is a correlation between music and the performance of the running. By doing so, it’s important that the provided data will be cleaned and merged properly.

For this project we will also be diving more into music. So the question that we will be asking here is if music has an effect on the performance of running. To dive even more into music, we are also going to take a look at Spotify tracks that the user is listening to and plot the danceability/energy into an exploratory data analysis. This means that external source Spotify will be used for this research. In addition to this, there will be a documentation of the effect on the running mashed on key transitions in the song.

The final goal for this project is to provide insight if there is a good or bad running pattern based on symmetry, impact and efficiency. This insight can be provided with the help of all the analysis and research that has been made within the project.