

Assignment Project Proposal

Advanced Data Management Project (55-706949-AF-20189)

**February 2019**

**Group 2: "Peak Analytics"**

Claudia Johnson - b8036424

Livia Hull - b7039217

Michael Lee - b7035141

Tim Bridgstock - b7042166

Olympia Stoli - b8003682

Contents

[Introduction 2](#_Toc1647350)

[Areas of Analysis and Datasets 2](#_Toc1647351)

[**Content** 5](#_Toc1647352)

[Business Questions 5](#_Toc1647353)

[Team Roles 5](#_Toc1647354)

[Division of Labour 6](#_Toc1647355)

# 

# Introduction

As a collective, this group, "Peak Analytics" has selected the "Music Industry and Event Management" scenario. We will present a business case identifying an artist with recent success on social media, a suitable venue, and a date and time for the gig. This gig is set to be the first of its kind and Peak Analytics will ensure maximum impact in terms of reputation and profit.

# Areas of Analysis and Datasets

In line with the event management company's stipulation for a "first of its kind" event with "a lot of impact", we will seek to identify an artist with recent success, and a considerable following, that has either never performed in Sheffield, or not performed in Sheffield recently, by procuring, collating and analysing data on chart performance, social media presence and subscription figures, and past and future concert activity. The analysis should include criteria to rule out unrealistic targets.

The potential event's dates will be selected from the datasets of future and past events, which include information on dates, artists, venues and ticket prices. The availability of the artist should also be taken into account, and whether other planned events could affect the volume of tickets to be sold.

After a date has been identified, we will then compile a list of venues available on this day and use mapping software to overlay these on a map of Sheffield. We will then create a heat map of Sheffield transport links. A list of venues which are located close to a sufficient number of transport links will be compiled and then compared by capacity.

After the choice of artist, date and venue, an assessment of profit is needed. Peak Analytics will estimate all the costs involved, such as the artist's fee and venue rental costs. From datasets on previous events a prediction of the number of tickets to be sold and ticket prices will be carried out so Peak Analytics will estimate the revenue from tickets and merchandise and calculate the expected net profit.

A final recommendation will be made when all the data has been collected and the chosen artist suitable for the venue available on a selected day is known. An additional benefit related to social impact of the event will be done, estimated through the number of followers on Instagram and the number of tweets within a week after the event.

For this case study, data will be collected from multiple different sources:

1. Selection of artist  
   A dataset will be built combining data retrieved from MusicBrainz, YouTube, Spotify, Songkick, Facebook and Twitter via APIs and/or web-scraping techniques. MusicBrainz provides a list of URLs, including social media links, for each artist, so can be used to produce a table collating each artist's profile names/IDs on different platforms.

* <https://musicbrainz.org/doc/Developer_Resources>
* <https://spotifycharts.com/regional>
* <https://developer.spotify.com/documentation/web-api/quick-start/>
* <https://developer.twitter.com/en/docs>
* <https://www.songkick.com/developer> (API key obtained)
* <https://developers.google.com/youtube/reporting/>
* <https://developers.facebook.com/docs/graph-api/>

1. Selection of venue  
   We will need data on venues for hire (pubs, clubs, etc.) along with information for bus stops, tram stops, car parks and taxi ranks. Unfortunately Sheffield City Council do not publish this as part of their open data so Peak Analytics have submitted a Freedom Of Information request, which will be completed by 18th March 2019. Some of the data is already freely available.

* <https://data.gov.uk/dataset/ff93ffc1-6656-47d8-9155-85ea0b8f2251/national-public-transport-access-nodes-naptan>
* <https://www.sheffield.gov.uk/content/sheffield/home/parking/city-centre-car-parking.html>
* <https://data-sheffieldcc.opendata.arcgis.com/datasets/5223281c5402427dbf1d52c7b9c2f9d9_3>
* <https://www.theguardian.com/travel/2008/may/01/uk.bars>
* <https://www.creativetourist.com/sheffield/music-venues/>
* <https://www.skiddle.com/gigs/Sheffield/>
* <https://www.yelp.co.uk/search?cflt=musicvenues&find_loc=Sheffield%2C+South+Yorkshire>
* <https://www.tickx.co.uk/article/576/sheffield-gig-guide-these-are-the-best-live-music-venues-of-sheffield/>

1. Selection of date  
   Datasets on future and past events can be obtained from the following sources

* <https://www.programmableweb.com/category/events/apis?category=20142>
* <https://newsapi.org/>
* <http://blog.sonicbids.com/the-best-and-worst-days-of-the-week-to-play-a-show>
* <https://www.ticketmaster.co.uk/search?q=sheffield>
* <https://www.viagogo.co.uk/Sheffield/Concert-Tickets>

1. The cost-benefit analysis:  
   In addition to datasets already specified, the following sites give methodology for analysis:

* <https://www.sciencedirect.com/science/article/pii/S0261517713000034>
* <https://www.smartsheet.com/free-cost-benefit-analysis-templates>
* <https://www.youtube.com/watch?reload=9&v=7tdKkeNClPE>
* <https://www.youtube.com/watch?v=BMIXJFiA2MQ>
* <https://www.mindtools.com/pages/article/newTED_08.htm>

1. Final recommendation:  
   In addition to datasets specified above, the following sites provide datasets needed:

* <https://www.instagram.com/developer/>
* <https://www.datasciencecentral.com/profiles/blogs/the-free-big-data-sources-everyone-should-know>
* <https://toolbox.google.com/datasetsearch>

# Content

## Business Questions

1. Which artists have not performed in Sheffield since 1st January 2018, and:

* Have had at least three Spotify Top 30 songs in that time
* Meet at least two of these criteria:
* >500,000\* YouTube subscribers
* >500,000\* Facebook likes
* >500,000\* Twitter followers

*\* criteria to be honed - possibly to include an upper limit to rule out unrealistic targets*

1. What is the right date for this event? Is the artist available? Is there any competition on that date?
2. What venue has the largest capacity, is available on the date, and is within close proximity of at least three transport links?
3. What is the best ticket selling price? What is the expected profit from this event?
4. What is the final recommendation with total profit and maximum impact on social media?

# Roles and Responsibilities

Each member has been assigned a title within the project team, and has also taken on an overview responsibility for one of the business questions. It is acknowledged that work package dependencies, individual specialisms and troubleshooting will require collaboration beyond the scope of the Team Roles and Division of Labour described here.

## Team Roles

**Project Manager:** TIM  
For this project, this role is strategic and means maintaining an overview of the progress of the project, identifying bottlenecks and seeking resolutions. An understanding of the specific strengths of the team members is essential in order to identify how problematic tasks can be efficiently resolved through collaboration or reallocation. The Project Manager will give the sign-off for the final report.  
  
**Program Coordinator:** LIVIA  
The Program Coordinator's role is to monitor and report progress of the project. In order to achieve this, the Program Coordinator will set up a GANTT chart program with milestones and deadlines and also will inform the Project Manager of progress.   
  
**Technical Expert Coordinator:** MICHAEL  
The Technical Expert Coordinator's role is to provide solutions to any technical problems which arise during the project, as well as provide ongoing input into technologies used to achieve the project aims. An understanding of each task and how each team member plans to complete this will be required, along with having time scheduled to discuss tasks, carry out research and feed back to the team.

**Troubleshooting Coordinator:** OLYMPIA  
The Troubleshooting Coordinator's role is to keep a Risk Register by ensuring everybody has entered risks and mitigation actions specific to their tasks. A periodic review of these risks will be carried out.   
  
**Communication Coordinator:** CLAUDIA  
The Communication Coordinator role is to ensure that internal and external communication is in line with the project strategy. It includes organising review meetings, minute-taking, ensuring all members share their work in a common place as well as regular communication with the Course Leaders.

## Division of Labour

Tim is responsible for identifying an artist that could be hired for the event. In the role of Project Manager, he will also be reacting to, and resolving, any apparent delays in the progress of the project.

Olympia is responsible for providing the appropriate date for the event and also responsible for the diagnosis and the solution of any problem during the project.

Michael is responsible providing an appropriate venue for the event. This will consider various factors such as transport links, availability and capacity.

Livia is responsible for the cost-benefit of the event and reporting the final expected net profit from the event. Livia is also responsible for smooth running of the program and mapping out deadlines.

Claudia's main role is responsibility for combining all findings from all members into the final recommendation as a draft report. In addition, Claudia will also assess the social media impact. Claudia will ensure that communication is efficient and aligns to the project plan and strategy.

# Addendum

The group acknowledges that elements of this proposal are likely to require revision or adaptation in response to:

* Proposal feedback
* Future course content
* Unavailability/limitation of data, or legal/ethical considerations
* Workload practicalities