MIS779 Portfolio AT1

ANALYSIS ON CREATIVITY AUSTRALIA





**Group 6**

**Team members**:

Anirudh Asapu:

Claire Wan: 217614144

Muqi Shao: 218613579

Saurav Lalwani: 219572734

Shruthipriya Ghanathey: 219557848

Van Hoang Nam Nguyen:

Table of Contents

[Executive Summary 3](#_Toc80697084)

[Introduction and Approach 3](#_Toc80697085)

[Assumptions 3](#_Toc80697086)

[Data Analysis 3](#_Toc80697087)

[Exploratory analysis 3](#_Toc80697088)

[Advanced analysis 3](#_Toc80697089)

[Interpretation 4](#_Toc80697090)

[Recommendations 4](#_Toc80697091)

[General Analytics Issues 4](#_Toc80697092)

[References 4](#_Toc80697093)

[Appendices 4](#_Toc80697094)

# **Executive Summary**

Creativity Australia is a charitable organization founded by notable social entrepreneur and opera singer, Tania de Jong AM (Member of the Order of Australia). It was found with the intention of creating a happier, healthier and an all-inclusive nation through their popular **With One Voice** program.

Their main objective is to help people suffering from loneliness and depression by involving them in community singing and choir. This helps people come forward and enjoy the many benefits of being part of a choir along with networking with different kinds of people and gaining mental support. Creativity Australia operates in all major territories of Australia and caters to over 6000 participants aged 9 and above belonging to all walks of life. They been doing this and more for the past 11 years through their 35+ different programs organized by hundreds of volunteers all over the nation.

# Introduction and Approach

The purpose of this report is to analyze the survey data given by Creativity Australia and then explore some useful information from people’s feedback.

As Creativity Australia’s outcomes are to improve the choirs and better support choirs, we will explore and analyze several questions in this report:

We will use some efficient tools to help our analysis such as excel, Power BI and python to make some modeling, text mining and drawing some tables.

# Assumptions

# We have made certain assumptions prior to data cleaning and pre-processing:

* Although the sample size is different for all the years survey responses, it is assumed that they represent the overall members of the choir
* The survey responses are assumed to be accurate, and appropriately represents the opinion of all the members of the choir

Data Cleaning

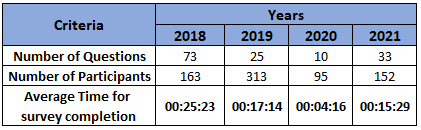
Steps undertaken to clean the spreadsheet prior to analyzing it:

* Missing values have been eliminated
* Cells with multiple values have been split
* Cells with similar values have been merged

# 

# Data Analysis

The dataset consists of survey responses of choir participants, who took the survey over the years 2018, 2019, 2020 and 2021. When analyzed the descriptive summary statistics of the dataset, the below observations were found.



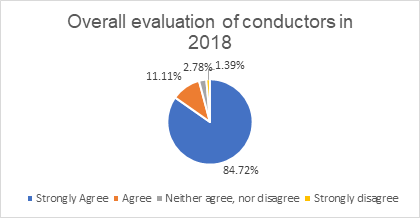
**Table 1.1 Descriptive Statistics of the Dataset**

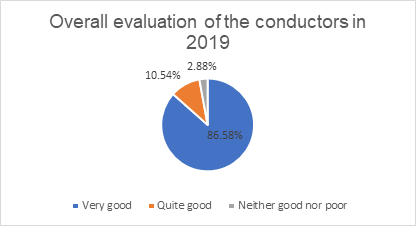
From the Table, 1.1, we can see that number of questions were asked were highest in the year 2018, and the most participants in the survey were observed in the year 313, since the year 2019 had two surveys. Moreover, looking at the average time taken for completion of the survey, it can be highlighted the year 2018 is the highest and lowest in 2020. Considering the number of questions and the time taken, it can be assumed that the survey responses are valid.

# Exploratory analysis

# 2) Conductor experience

* Overall trend





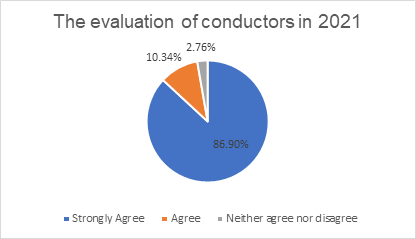
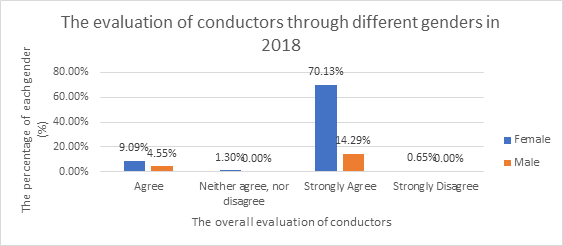


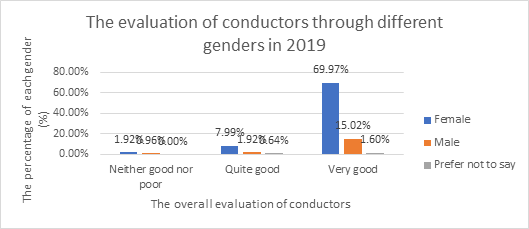
Figure: The evaluation of conductors from 2019-2021

Analysis the figures…, the respondents are more positive than negative. For example, in 2021, the percentage of people evaluate positively about the conductors than negative (86.9% > 2.76%)

Due to the large number of categorical variables in the dataset, it is difficult to visually examine all of them simultaneously. I classify variables into two categories: internal and external variables. Internal variables are evaluation-related variables that people may influence, contribute to opinions of the conductors, or have a direct connection with the conductors, such as age group, gender, and so on. External factors are those that influence the rating, such as the level of frequent participation and so on.

* Analysis of overall conductors and member’s genders





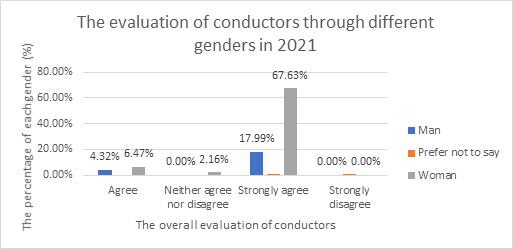
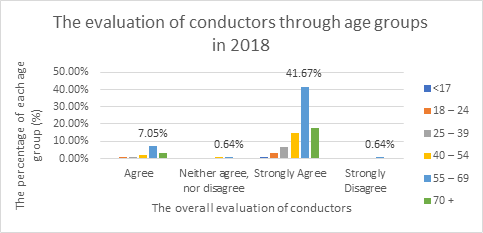
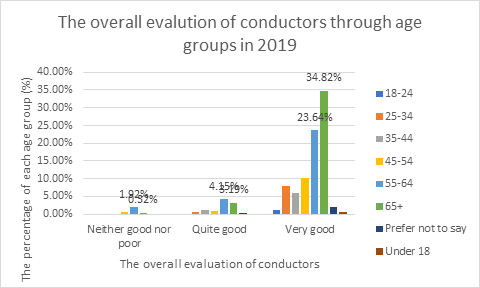


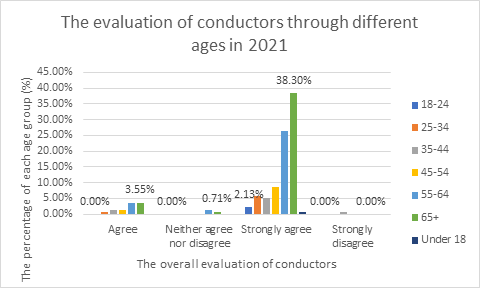
Figure: The evaluation of conductors through different genders from 2019-2021

Analysis the figures … , the female respondents are positive than the figure for males. The number of females joining the survey are more than males. The number of respondents are positive than negative.

* Analysis of overall conductors and member’s age groups



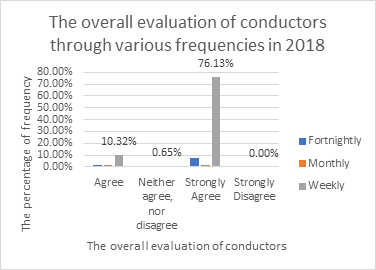




Figure…: The evaluation of conductors through different ages from 2018-2021

Analysis the figures ….., the respondents aged over 65 is more positive than the figure for less than 18 years old. More participants aged over 65+ join the survey than figure for under 18 years old.

* The frequent level of participation



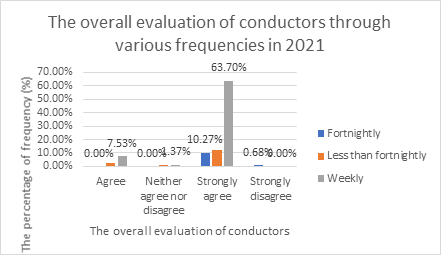


Figure … : The overall evaluation of conductors through various frequencies 2018- 2021

Analysis the figures …., the respondents participate weekly are more positive than other groups about the evaluation of the conductors (over 63%). The weekly participants join the survey than other groups.

Based on the above exploratory analysis performed, the patterns below are being observed to be consistent over the years:

* Respondents are more likely positive than negative
* Members attending choir weekly are positive about the conductors
* Women have expressed positively of the conductors comparing to Males.
* Members aged over 65 have a positive evaluation of the conductor compared to people under 18

The next step, I will make some hypothesis test if pattern is consistent over the year or not

* Two-sample tests proportion:

\_ Positive opinions are included strongly agree and agree, very good and quite good

\_ 5% level of significances (alpha = 0.05) is appropriate for any hypothesis test

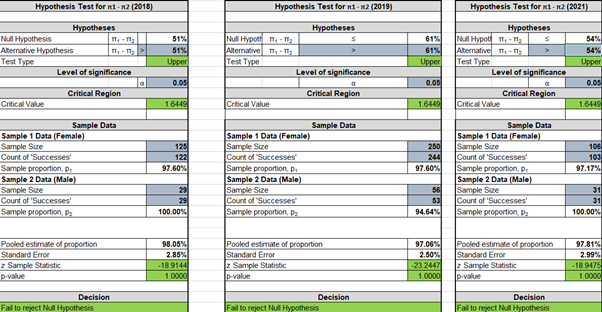


Figure: Hypothesis about overall rating through genders

As can be seen from the figure above, we can conclude that 5% level of significances, there are not enough evidence females are more positive than males giving highly overall-rating.

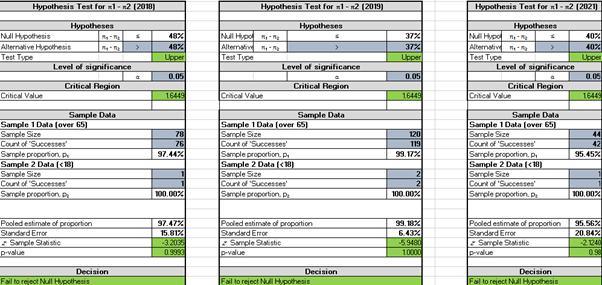


Figure: Hypothesis about ‘positive’ opinions through age group

* As can be seen from the figure above, we can conclude that with 5% level of significances, there are not enough evidence to say that members aged over 65 have a positive evaluation of the conductor consistently.

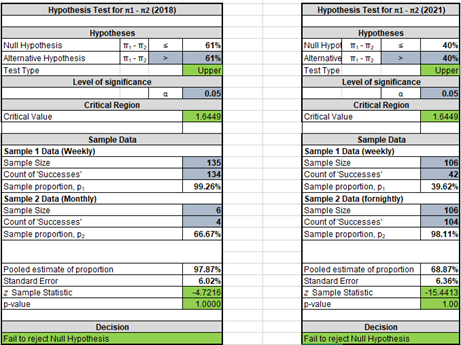


Figure: Hypothesis about ‘positive’ opinions through age groups

As can be seen from the figure above, we can conclude that with 5% level of significances, there are not enough evidence to say that members attending choir, weekly are positive about the conductors

* Sentiment analysis

We will classify whether a block of text is positive or negative through the process of sentiment analysis. It gains to analyze people’s opinion in a way that it can help improve the quality of Creativity Australia.

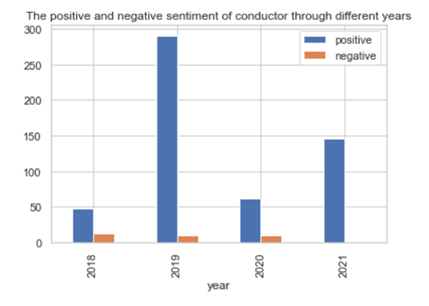


Figure …: The positive and negative sentiment though selected year

Analysis the figure …., there is enough evidence to say that we can see that the number of positive sentiments is more than negative sentiment, which mean that the respondents are really is more positive than negative. Therefore, this hypothesis is true.

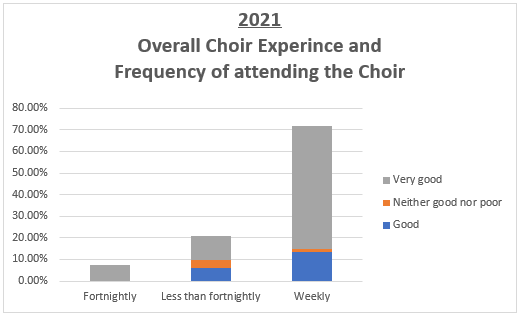
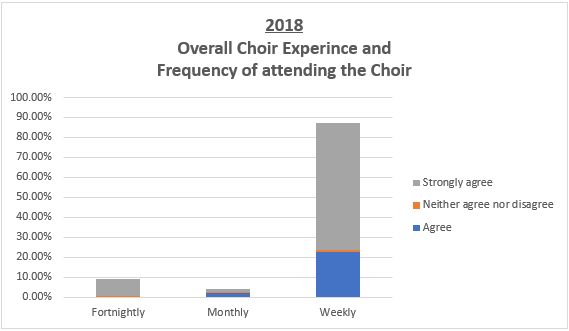
The exploratory analysis you have done so far in the project on the data will need to describe key patterns/trends/observations within the dataset. Describe only those aspects that pertain to the overall problems that you are attempting to solve. The emphasis in this section is on the understanding of how the problems you are solving, relate to the objectives (why you are solving those problems), and the analytical questions rather than advanced analytics (e.g., prediction). Include potentially useful external data/trends that are pertinent to the objectives and problems. By the end of this section, you will have a clear set of analytical problems to address. These analytical problems will mainly be addressed using your upcoming advanced analytics developments. Each analytical problem should explore a particular aspect of the overall project objective. Consider and develop each modelling effort (analytical problem) in such a way that they could inform your final recommendations. Your analytical effort should inform our current understanding, i.e., if they turn out to be correct after being tested against the data, then supporting and/or new insight should emerge.

Hypothesis to be kept here—

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

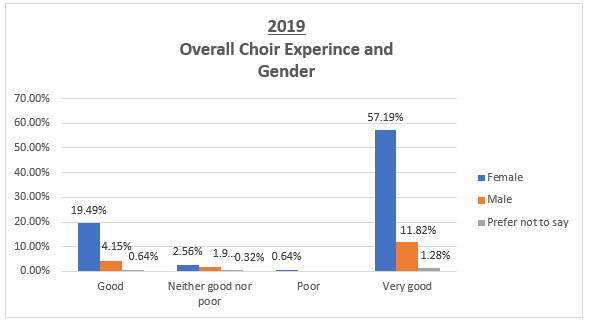
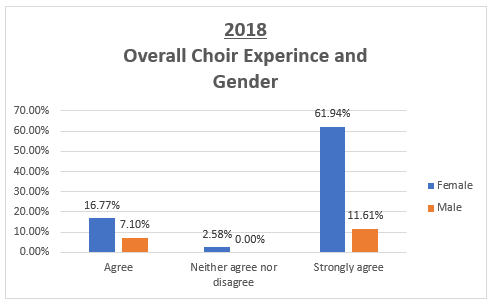
Exploratory analysis is performed on the experience of the choir, across all the years based on the Gender and the frequency of attending the choir as below:

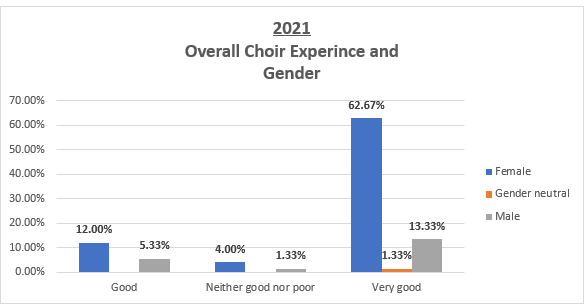
* **Frequency of attending the choir over the years**



From the figure, it can be seen that most participants like to attend the choir weekly and give a higher experience compared to the others.

* **Overall choir experience based on the participant’s Gender over the years**





From the figure, it can be seen that people the female participants attending the choir are higher than the male participants. Moreover, they are seen to be giving a higher rating compared to Men and others.

Therefore, based on the analysis, it can be **Hypothesized** that,

**Women attending the choir are expected to have a more positive choir experience compared to men and others.**

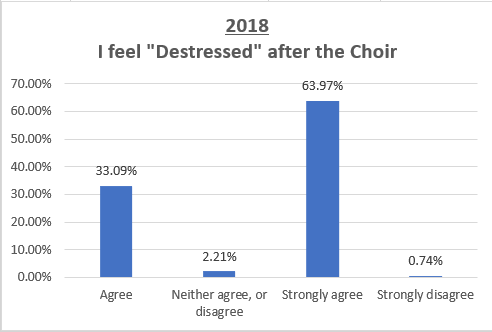
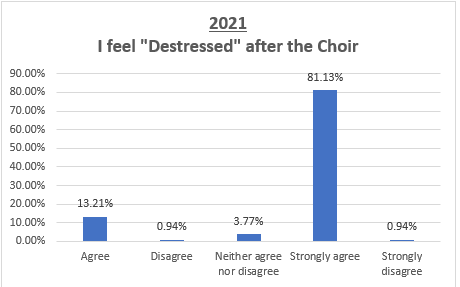
Based on research on human psychology, it is believed that women can experience a higher emotional state than men. It is proven that women have reported a high intensity in reaction than men, which aids to provided hypothesis (Poláčková Šolcová and Lačev, 2017).

To verify the claim, we use Hypothesis Test of two sample t-test.

-

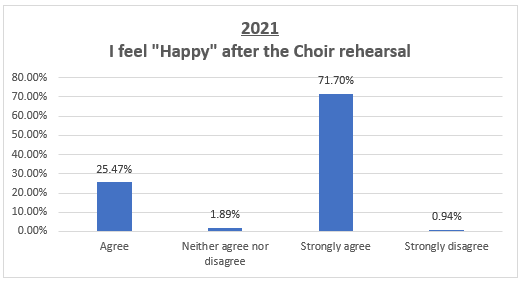
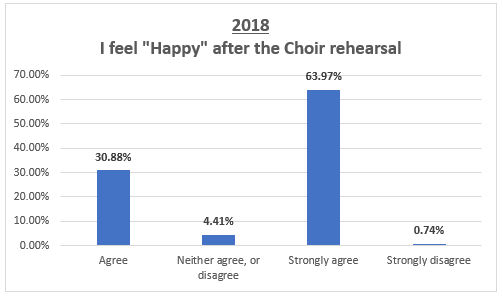
Moreover, further exploration is done to understand the participant’s weekly choir experience on feeling “Destressed”, “Happy”, “Less Anxious/Depressed” and “Improve self-confidence” as below:

* **Feeling Destressed after the Choir:**

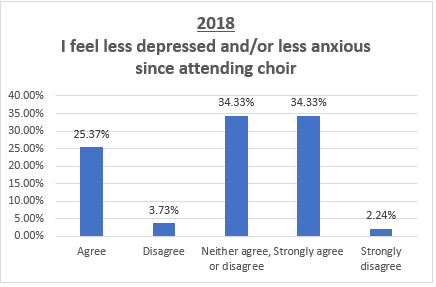
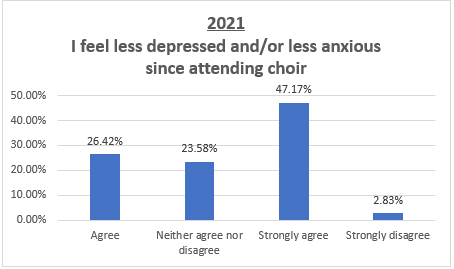
It can be seen that majority of the participants have agreed to feel destressed after attending the choir and the over the years, more people agree to it.

* **Feeling Happy after the Choir rehearsal:**



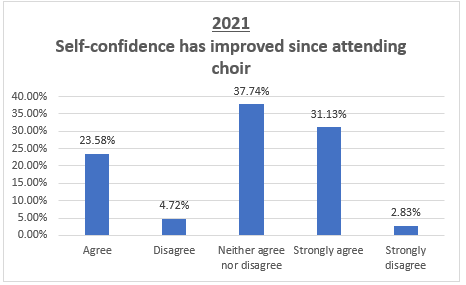
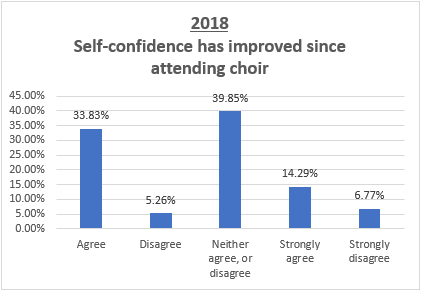
It can be seen that majority of the participants have agreed to feel happy after attending the choir rehearsal and the over the years, people agreeing has increased.

* **Feeling less depressed/less Anxious after the Choir rehearsal**

It can be seen that majority of the participants have agreed to feel less depressed or anxious since attending the choir and the over the years more people agree on this sentiment.

* **Improved self-confidence after attending the Choir**



It can be seen that over the years, majority of the participants had neutral sentiments on choir helping to improve their self-confidence, with a considerate increase in the number of people agreeing on it.

Therefore, based on the analysis, it can be **Hypothesized** that,

The experience of the participants have improved over the years with improved self-confidence, feel less destressed, less anxious, and depressed and happy after attending the choir.

**Over the years, participants experienced a higher self-confidence, and felt less destressed, less anxious/depressed, and happier after attending the choir.**

Since the global pandemic COVID-19 has impacted our daily lives, there is a change in lifestyle due to the implemented control measures like lockdowns. Moreover, the choir’s proven research improving the mental health by singing aids to the provided hypothesis(Launay and Pearce, 2021). To verify the claim, we use Hypothesis Test of two sample t-test with individual factor analysis over the years.

# Advanced analysis

Informed by the exploratory analyses, develop *an advanced analytics modelling plan* comprising of a set of models you will be developing to work on the analytical problems outlined in the previous section.

Text mining and other analysis to be kept here----

# Interpretation

Interpret the results of the exploratory analysis. What do these findings mean for to the overall objectives? What are the important observations/patterns? Were there any surprises? How could these be understood, given the methods followed in the analysis so far?

Insights of the data----

# Recommendations

List some preliminary recommendations following from the exploratory data analysis. Recommendations should emerge from the data analysis (all backed up by your results and the literature if possible) and be aligned with the overall problems and objectives. Specifically, recommendations should be driven by your exploratory data analysis (data driven) and the literature, and, they should be actionable.

**We should give recommendations----**

# General Analytics Issues

State any ethical issues, data privacy related matters, and complications or subjective consent matters. Justify the analytics methods used from a technical perspective. Assess and report any risks that may be involved in the analysis of the data as well as in the deployment and practical use of the methods developed and recommendations made in the report.

# References

1. Poláčková Šolcová, I. and Lačev, A., 2017. Differences in male and female subjective experience and physiological reactions to emotional stimuli. *International Journal of Psychophysiology*, 117, pp.75-82.
2. Launay, J. and Pearce, E., 2021. *Choir singing improves health, happiness – and is the perfect icebreaker | University of Oxford*. [online] Ox.ac.uk. Available at: <https://www.ox.ac.uk/research/choir-singing-improves-health-happiness-%E2%80%93-and-perfect-icebreaker> [Accessed 24 August 2021].

# Appendices

Include a set of appendices, clearly outlining all the workings of your summary exploratory analysis. It should be clear to an external reader how the analysis were conducted and thus how your findings were arrived at.