

Assignment 2

This assignment draws on lecture and lab content from Module 2: Creating dynamic reporting systems. The knowledge and skills assessed by the assignment will be covered in lectures/labs by the end of Module 2. Consequently, the order of the assignment instructions below does not necessarily reflect the order the material is covered in lectures/labs.

Overview

The data context for Assignment 2 is music listening preferences and/or habits. This means behaviours or opinions related to listening to music, for example, what people prefer to listen to, when do they listen to music, how do they listen to music, etc.

You will need to develop a form/survey about music listening preferences and/or habits, use the form to collect data, and then use this data to create a report with R Markdown.

You will need access to a computer that has [R](https://cran.r-project.org/) [↗] [\(https://cran.r-project.org/\)](https://cran.r-project.org/) installed and [RStudio](https://www.rstudio.com/products/rstudio/download/) [↗] [\(https://www.rstudio.com/products/rstudio/download/\)](https://www.rstudio.com/products/rstudio/download/) installed. **You will also need to install the R package {tidyverse}.** If installing R, RStudio, or the tidyverse onto your computer is a problem, you can use the free level of [RStudio Cloud](https://rstudio.cloud/) [↗] [\(https://rstudio.cloud/\)](https://rstudio.cloud/).

Form design and data collection

- Read the summary information about [NZ on Air's 2020 Childrens Media Use survey](https://www.nzonair.govt.nz/research/childrens-media-use-survey-2020/) [↗] [\(https://www.nzonair.govt.nz/research/childrens-media-use-survey-2020/\)](https://www.nzonair.govt.nz/research/childrens-media-use-survey-2020/). Review the contents of the the [PDF report for the survey](https://www.nzonair.govt.nz/documents/558/2020_Childrens_Media_Use_Research_Report.pdf) [↗] [\(https://www.nzonair.govt.nz/documents/558/2020_Childrens_Media_Use_Research_Report.pdf\)](https://www.nzonair.govt.nz/documents/558/2020_Childrens_Media_Use_Research_Report.pdf) to help you generate some ideas for questions you could use in your own survey.
- Develop a new Google form. Use the four guidelines for designing forms discussed in Module 2.
 - Decide on what questions you will use for your form, in order to collect data on music listening preferences and/or habits. Think carefully about what aspect(s) of music listening preferences/habits that you **specifically want to focus on**.
 - Develop between five to eight questions that align with your focus:
 - At least two questions that will generate numeric data using a *short answer* question with *response validation*.

- At least two questions that will generate categorical data using *multiple choice questions* with fixed options (i.e. no “other” option).
 - At least one question that will generate categorical data using *check boxes* with fixed options (i.e. no “other” option).
 - Ensure that responses are anonymous.
- Create a new Google form and change the settings of your Google form so that it is **NOT** restricted to only users from the University of Auckland. Create different kinds of questions and add text, visual, and structural elements as needed.
- Collect data using your Google form:
 - Link your Google form to a new Google sheet and publish your Google sheet as a CSV.
 - Send your Google form to friends and family - aim for at least 20 responses.

R Markdown report

- **NOTE:** You need to re-use at least two techniques you learned from Assignment 1 about working with R Markdown to create HTML, in addition to what is specified below..
- Within your *stats220* folder on your computer, create a new project using RStudio that is called “Assignment2”. Create a new Rmd file within the *Assignment2* project called “index.Rmd”
 - Make the title “Assignment 2” and put your name as the author.
 - Edit the YAML of the *index.Rmd* file so that:
 - the subtitle is “STATS 220 Semester One 2022”.
 - For the `r` `setup` chunk:
 - change the argument of `knitr::opts_chunk$set()` to `echo = FALSE`.
 - load the tidyverse library using `library(tidyverse)`.
- Structure the report (the *index.Rmd* file) using second-level headings as follows:
 - Introduction
 - Links to form and data
 - Analytics
 - Learning reflection
- Under the *Introduction* section of your report:
 - Explain what aspect(s) of music listening preferences/habits decided to focus on for your report.
 - Write at least 50 words.
- Under the *Links to form and data* section of your report:
 - Include the link to your Google form (the same one you shared with your respondents).
 - Include the URL for your published CSV file (the same one you used within your R code).
- Under the *Analytics* section of your report:

- Read the data from your published CSV file into a data frame called `music_data`.
- Rename the variables of your data frame and print the five rows of your data using `slice()` and `knitr::kable()`.
- Include two examples of using inline R code to apply functions to your data frame or variables (extracted as vectors).
- Include at least three bar charts produced using `{ggplot2}` with appropriate labels.
- Provide some comments that discuss what you have learned from your data.
- Under the *Learning reflection* section of your report:
 - Describe in your own words at least ONE important idea you learned from Module 2 *Creating dynamic reporting systems*.
 - Explain how you have re-used two techniques you learned from Assignment 1 (these can't be things specifically asked for in this assignment).
 - Discuss what things related to data technologies that you are more curious about exploring further.
 - Write at least 150 words for this section of your report.
- Knit your *index.Rmd* file to create a self-contained *index.html* file.

For this assignment, you will submit BOTH your *index.html* and *index.Rmd* files

Marking guide

The assignment will be marked out of 10.

- The report contains an *Introduction* section that meets the stated requirements (1 mark)
- The report contains a *Links to form and data* section that provides links to the Google form and published CSV file (1 mark)
- The Google form and data collected meet the stated requirements for the number and type of questions (1 mark)
- The report contains an *Analytics* section that contains:
 - a table of the first five rows of the data with renamed variables (1 mark)
 - two examples of using inline R code that meet the stated requirements (1 mark)
 - at least three bar charts that meet the stated requirements (1 mark)
 - comments that discuss what was learned from the data (1 mark)
- The code used in the R Markdown file (*index.Rmd*) is readable e.g. appropriate use of indenting and “white space” (1 mark)
- The report contains a *Learning reflection* section that meets the stated requirements (1 mark)

- The report demonstrates two techniques that have been re-used from Assignment 1 (1 mark)