*Fain-*ing Happiness: Analyzing the Connections between Lifestyle and Well-being during COVID-19

**Part One: Introductions and Research Questions**

**Your prototype report should begin by reintroducing your topic and restating your research question(s) as in your proposal. Your research question(s) should be (1) substantial, (2) feasible, and (3) relevant. Briefly justify each of these points as in the project proposal. You can start with the text from your proposal, but you should update your introduction and research questions to reflect changes in or refinements of the project vision. Specifically point out what has changed since the proposal. Your introduction should be sufficient to provide context for the rest of your report.**

For over a year, the COVID-19 pandemic has forced people to adopt a radically different lifestyle characterized by instability and stagnancy. Activity levels have changed significantly as people were confined to their homes due to the closure of gyms, offices, schools and places to socialize. As most activities shifted to an online platform, accessible within the comfort of their homes, people had fewer reasons to be active. Furthermore, with the rise of grocery delivery services such as Instacart and increasing popularity of food delivery services such as GrubHub, Uber Eats, etc, diet patterns undoubtedly changed as well. Along with the pandemic and these lifestyle adaptations also came unprecedented changes in well-being worldwide. Countries have reported more mental health problems during the pandemic and according to the World Happiness Report[[1]](#footnote-0), life satisfaction had a positive correlation with confidence in “their health system’s ability to respond to COVID-19”. We hypothesize that the lifestyle factors of diet and exercise could potentially be positively correlated with measures of well-being such as mental health statistics and happiness scores. In our project, we plan to analyze these connections within multiple different countries, who differ in their lifestyle patterns and the severity of COVID-19.

Our research question is: How do lifestyle factors and well-being during the COVID-19 pandemic correlate across the countries of Italy, Japan, Canada, United States, and the United Kingdom? We will conduct a substantial analysis of the correlation of various factors in our datasets, rather than just a surface-level summary of data. The lifestyle factors included in the analysis will be exercise, diet, and potentially social support, while the factors comprising well-being will be the country’s overall happiness score in the World Happiness Report and mental health as measured by Google search terms such as “depression”, “anxiety”, and “insomnia” during the past year.

Our research question is feasible for this project because we have found multiple intersecting datasets that have measures for the factors we are interested in for the countries of interest. The metrics used to measure these factors are also consistent between the datasets. Additionally, we are analyzing correlation among multiple lifestyle factors and measures of well-being, which is feasible for 4 people to investigate. In order to ensure the feasibility of the project, we are limiting our analysis to two significant lifestyle factors and two reasonable measures of well-being. Also, we are not trying to prove a causative relationship, which would be much more complex.

Lastly, this research topic is very relevant globally due to the worldwide effect of COVID-19. The majority of people around the world have had to adjust their lifestyle to the pandemic. Perhaps through understanding the diet and exercise patterns that are correlated with increased wellbeing, we can make changes in our own personal lives to proactively live happier lives during the pandemic.

Overall, our research focus has not changed significantly since our research proposal. We are still focusing on the correlation of lifestyle factors (exercise, diet, social support) and wellbeing (World Happiness Report, Google search terms). One change, however, is that we are adding South Korea as a country of study, as well as a “Worldwide” category.

**Part Two: Data Sources**

**Part 2: Data Sources After your introduction and research questions, your prototype should discuss the data you have collected and are using to answer your research questions. Be specific: name the datasets you are using and where they were collected from / how they were prepared. Briefly justify why your data are appropriate and sufficient to address your research questions. As in the introduction you can begin with the text from your proposal but be sure to update it to fit with your evolving project.**

The first set of datasets we will use is named “COVID-19 and Mental Health Search Terms” found with this link: <https://www.kaggle.com/luckybro/mental-health-search-term>. Each dataset is the breakdown of a specific country’s data: Canada, United Kingdom, United States, Italy, and Japan. Within each dataset, each column is a different search term such as “depression”, “anxiety”, and “obsessive-compulsive disorder”. Each row represents a week during the pandemic. Each value represents the search interest from 0-100, 100 representing the most popular search term at that point in time. We plan to aggregate the rows within the relevant time frame and use the mean values for each search term.

The next dataset we will use is “COVID-19 Healthy Diet Dataset” found with this link: <https://www.kaggle.com/mariaren/covid19-healthy-diet-dataset?select=Fat_Supply_Quantity_Data.csv>. Within this link, there are multiple datasets containing the percentage of total fat, protein, or caloric intake from certain food groups. We will use the dataset for average total caloric intake from different types of foods i.e. alcohol consumption, fruits, fish/seafood, or animal products within this past year. There are also columns included for COVID statistics of that country such as confirmed cases, deaths, and recovered. Each row is a different country, but we will only be looking at Canada, United Kingdom, United States, Japan, and Italy, as this is the intersection with the first dataset.

Another dataset we will use is the ”World Happiness Report” at the following link:

<https://worldhappiness.report/ed/2020/>. We will use the “Data for Figure 2.1” dataset, which consists of various factors defining overall well being. This includes the Ladder Score, a reliable measure of overall happiness factoring in GDP, life expectancy, generosity, positive and negative affect, etc.. This dataset also includes the specific countries we will be investigating, so it is a relevant source of data for assessing well-being for our project. Although we will mainly be focusing on the World Happiness Report for 2020, we may also use the World Happiness Report for 2019 to see if levels of happiness or well being have changed as a result of the pandemic.

The last dataset we will be using is “Prevalence of insufficient physical activity among adults aged 18+ years” from the World Health Organization: <https://www.who.int/data/gho/data/indicators/indicator-details/GHO/prevalence-of-insufficient-physical-activity-among-adults-aged-18-years-(age-standardized-estimate)-(-)>. This dataset shows exercise patterns for various countries. Although the data is from 2016, we will assume that a country’s overall exercise habits remain relatively constant and this data will reflect a country’s overall propensity for exercise.

**Part Three: Preliminary Results and Methods**

**The preliminary results section of your report should summarize the results obtained so far in the project. Where possible, results should be summarized using clearly labeled tables or figures and supplemented with written explanation of the significance of the results with respect to the research questions outlined in the previous section. Your results do not need to be final or conclusive for your entire project but should demonstrate substantial effort and progress and should provide a concrete proof of concept or initial analysis with respect to your research questions.**

**Your results should be specific about exactly what data were used and how the results were generated. For example, if you scraped multiple web databases, merged them, and created a visualization, then you should explain each how step was conducted in enough detail that an informed reader could reasonably be expected to reproduce your results with time and effort. Just saying “we cleaned the data and dealt with missing values” is not sufficient detail, for example. Your report itself should include an explanation of your methods, but it should also contain instructions on how to access your full implementation (that is, your code, data, and any other supplemental resources like additional charts or tables). The simplest way to do so is to include a link to the box folder, gitlab repo, or whatever other platform your group is using to house your data and code.**

**Part Four: Reflection and Next Steps**

**In this part, you should begin by reflecting on the progress of your project so far. Address the following: 1. What has been successful in the project so far or what is essentially complete and ready for the final report? 2. What has been challenging in the project so far or what is incomplete in the prototype that needs to be finished for the final report? 3. What are your next steps? These should be concrete and specific actions that your group will take to address the challenges identified in order to complete a successful final project.**

1. New Report Highlights the Impact of COVID-19 on Life Satisfaction Worldwide, 31 July 2020, worldhappiness.report/blog/new-report-highlights-the-impact-of-covid-19-on-life-satisfaction-worldwide/. [↑](#footnote-ref-0)