Analyzing Change in Weight as a Result of the COVID-19 Pandemic

Masheil Mir

June 15, 2021

Part 1

Goal

On March 12th, 2020, the Premier of Ontario declared the first drastic measure since "Corona virus" was first heard of in late 2019. Doug Ford (Ford) decided on the closure of publically funded schools across the province post March break for the duration of two weeks [1]. This was only the beginning of what has now become one of the most life altering experiences in the daily lives of everyone across the globe. As the province got through its first wave, it then entered its second in late August. This forced the Ford government to take more drastic measures and implement stronger restrictions. These measures included the closing of shopping centers, dine-in restuarants, gyms and sports/recreational centers [2]. As a result, many people were forced to make drastic lifestyle changes, espeically since they can no longer enjoy the activities they used to pre-pandemic. As a result of the gym closures due to COVID-19, a whopping 60% of those who held gym memberships cancelled or considered cancelling their subscription, and the member return rate was only a mere 31% once facilities did reopen [3]. The goal of this survey is to determine the health affects of COVID-19 on individual weight depending on their lifestyle changes.

This survey will look into the eating and exercise habits of an individual and hone in on their weight before vs. during the COVID-19 pandemic. This data can be used in other studies such as determining whether obeseity in Canadians has increased within the last 5-10 years. It will essentially allow future users to determine whether or not a correlation exists between the COVID-19 Pandemic and weight.

Procedure

To conduct this survey, it is important to take into consideration the idea of collecting data in the most time and cost efficient way as possible. Since the target audience for this survey will mainly be adult Canadians, we can collect the data from Ontario residents over the age of 18. But since obtaining the email list for Ontario residents is a bit challenging and pricey [4], we decided to take the email list for residents in the city of Markham, Ontario instead.

The frame of this survey will be a list frame taken from an address book of Markham residents [5]. This will come directly from a pre-made database that includes the information and data of Markham residents [6].

The sample frame will use the emails of the Markham residents. These emails are ones that are pre-entered by residents themselves, making them a good and reliable resource to ensure that the survey will be reaching real people. Using an email list from a city also ensures that our population is 18+, and gives us a good mix of ages, ethnicities, etc..

The procedure will be using random sampling without replacement in order to conduct the survey. This will ensure that residents of all ages and backgrounds are selected randomly and that they all have an equal chance at being included in the sample. In general, the following steps will be executed:

1) Create a survey using Google Forms
2) Send out the survey through email to Markham residents, at random.

A drawback of this process is the fact that some people may enter emails that may not be real, or they may not use these emails frequently, thus they will not see the survey in their inbox. Additionally, since we are only catering to one City, there may be a bias in regards to the number of accessible gyms in one neighbourhood vs. another. To account for the first drawback, resampling can be done on the rest of the dataset (excluding those already sampled) to guarantee survey responses. Regardless of this, there may still be a bias based on the fact that we are only considering the population of Markham residents rather than all Ontario residents, thus possibly skewing the results.

Showcasing the survey.

The survey can be found at the following link: https://forms.gle/umHYryYrnNexPLxw8

Question 1: Did you participate in physical activity before the COVID-19 Pandemic? This question is important to us because it provides us with some insight of how physically active the individual was. From this question we can deduce how cautious the individual is about their weight and physical wellbeing, and how likely they might have been to continue participating in physical activity throughout the pandemic. In other words, by asking about their activity before the pandemic we can see how they were maintaining their physical health under normal circumstances.

The drawback of asking this type of question is that in some cases (due to a medical condition) the individual might not be able to participate in physical activity and does not give the complete picture of the individual's physical health. Additionally, there is not always a guarantee that if someone is physically active at a certain moment in time, they will continue to be throughout. A better way to phrase this question would have been asking if they regularly participate in physical activity or if they only do it for a few months throughout the year.

Question 2: How many meals have you been eating on average per day since the COVID-19 Pandemic started?

When looking at how many meals a person is eating on average per day, we can analyze their eating habits in comparison to those before the pandemic. It is important to recognize that most commonly in the West, three well balanced meals are recommended to ensure you are not overeating or undereating. Assuming this is true for the average Canadian (i.e. they are eating well balanced meals), we can determine whether their food intake has increased, decreased, or stayed the same, and in return determine if there is a correlation between that and their current weight. Using this, we can assume that someone eating > 3 meals a day will perhaps gain more weight than someone who is eating < 3 meals per day.

The downside of making this assumption is exactly that; we are assuming the meals are well balanced when in reality we cannot conclude that information with the given question. The number of meals being eaten and their portion sizes, alongside age, gender, and genetics all affect metabolism in different ways. Having 3 large meals will result in a slower metabolism vs 5 smaller, spaced out meals [7]. Thus, a higher metabolism with a larger number of meals will result in burning more fat and calories, thus lowering weight.

Question 3: How has your weight changed since the COVID-19 Pandemic began? It is important to hone in on the change in weight before vs. after the pandemic as this will give us a better picture about how an individual has been taking care of themselves throughout the pandemic, if at all. Losing gyms and other facilities as perhaps their only means of getting any physical activity could have easily impacted their motivation and dedication to continuing the process at home. This question will give us an idea of how an individual's weight has fluctuated as a result of the pandemic. It allows us to compare and contrast the results in our data, such as with Question 1, to determine whether the weight change is in fact a result of the pandemic or eating habits, or if it is simply because of not being active in general, even before the pandemic.

The drawback of this question is that there are many factors that may contribute to an individual's weight fluctuating. It does not solely have to be based on their level of physical activity. There may be some people who work out every day, year round, and do not have the same results as those around them. There may also be other factors, such as age, sex, sleeping patterns, and even genetics that may play a role but are not considered in this survey [8].

Part 2

Data

1. Data Collection

The survey consisted of 12 questions, with 5 categorical variables and 6 numerical variables. The way the data for this survey was collected was through a Google Form link using the procedure similar to that which was outlined in Part 1. In detail, the following steps were conducted to gather the data for this survey:

1) Create a Google Form and generate a link for public access. To ensure the entries are anonymous, details such as name, age, gender, sex, email, etc., were not asked to be included.
2) Individuals in Markham were selected and asked to spread the survey to those they know in the area with a short description about the survey's purpose.
3) The data was converted into a CSV file in order to easily analyze in R.

The drawback of the above lies within step 2 of the process. Having only sent the survey to Markham Residents that I am aware of could have created some bias, thus a better solution that will ensure the survey reaches a variety of individuals in Markham with better variability is the following:

1) Access the Markham residents database and download the CSV file. We know that these indivduals are above the age of 18 as minors are not primary owners of residency in Markham.
2) Filter only the emails of the residents in the CSV file.
3) Markham has a population of approximately 330 000 and about 68% of indivudals are of working age (16+) [9]. To include only 18+ lets take an educated assumption that this means around 60% of indivudals classify as adults. That is about 200 000 residents in Markham. Lets use 500 as our threshold for each random sample taken. After filtering the data, perform the RAND() function to randomly select 500 residents emails, without replacement.
4) Take these 500 randomly sampled emails and send the survey link out to them. If we recieve a less than 10% response rate, resample using RAND() and without replacement, and repeat the process until we have a decent number of responses.
5) Download the survey results as a CSV file in order to easily analyze in R

2. Data Cleansing

Before getting into the cleaning process of the data, note that we will be using the following acronyms throughout this study:

1. BP = before pandemic
2. DP = During Pandemic
3. HPW = Hours per week

In order to ensure that the data is clean and free of unwanted variables and values, the following steps were performed:

1) Since we are using a Google Form to generate this data, there might be unwanted columns such as a "Timestamp" that is not relevant to our study as it does not contribute to our goal. Using a built in R function, this column can be removed from the data frame.
2) It is important to check and remove any NaN or empty cells from the dataframe as this may affect/impact our data analysis later on. This can be done using a built in R function.

3) Since the data is retrieved directly through a Google Form, column names will be the questions taken from that form. This may result in a large string of text that makes it harder to see what each column is giving information for, thus the columns were renamed using the built in rename() function in R. The columns were renamed to ensure that there is still a general understanding of what the column is referring to.
4) When referring to the column names for data analysis, it might still be difficult to use the rename assignments from part 3 as they are still a bit wordy. Thus, we can name the columns for our own use again, with "before" and "after" at the end of each new name assigned to allow us to keep track of what the variable name is referencing.

Since the data cleansing is complete, we can now move onto giving a complete breakdown of the variables in the dataframe alongside important characteristics. This will allow us to see which variables are more significant and may aid us in our analysis.

3. Data Description

The data is described in the below table. This table includes the variable name, type (categorical vs. numerical), and a short description of its purpose.

Variable	Type	Description
Physical Activity BP	Categorical	This variable represents whether or not the individual participated in physical activity before the COVID-19 pandemic. It is a yes/no question, thus the response is either a yes or no.
Physical Activing DP	Categorical	This variable represents whether or not the individual participated in physical activity during the COVID-19 pandemic. It is a yes/no question, thus the response is either a yes or no.
Avg physical activity (HPW) BP	Numerical	This variable represents the average number of hours an individual spent participating in physical activity before the COVID-19 pandemic. For example, 4, which represents 4 hours.
Avg physical activity (HPW) DP	Numerical	This variable represents the average number of hours an individual spent participating in physical activity during the COVID-19 pandemic. For example, 0, which represents 0 hours.
Change in Eating Habits	Categorical	This variable represents whether or not the eating habits of the indivdual changed as a result of the COVID-19 pandemic. It is a yes/no question, thus the response is either a yes or no.
Avg Meal Count BP	Numerical	This variable represents the average number of meals an indivudal ate before the COVID-19 pandemic. For example, 3, which represents 3 meals eaten per day on average before the COVID-19 pandemic.
Avg Meal Count DP	Numerical	This variable represents the average number of meals an indivudal ate during the COVID-19 pandemic. For example, 2, which represents 2 meals eaten per day on average during the COVID-19 pandemic.
Main Types of Meals DP	Categorical	This variable describes the types of meals most frequently eaten by the individual. It consists of three options, home made meals/snaks/take out meals. The answer consists of one of the three options, for example, home made meals.
Weight BP (lbs)	Numerical	This variable represents the weight of the individual before the COVID-19 pandemic in pounds. For example, if the individual enters 156, then that means their weight before the pandemic was 156 pounds.
Change in Weight	Categorical	This variable represents the change in weight of the individual. It consists of ranges such as: gained 1-10 pounds, I stayed the same weight, I lost weight, etc.
Weight DP (lbs)	Numerical	This variable represents the weight of the individual during the COVID-19 pandemic in pounds. For example, if the individual enters 104, then that means their weight during the pandemic is 104 pounds.

4. Numerical Summaries and Plots

For all of the numerical data, we can obtain the numerical summaries that will include the Center and Spread of variable. The Ceneter and Spread includes the following values that will also be calculated as a result: mean, median, mode, standard deviation, interquartile range (IQR), minimum and maximum.

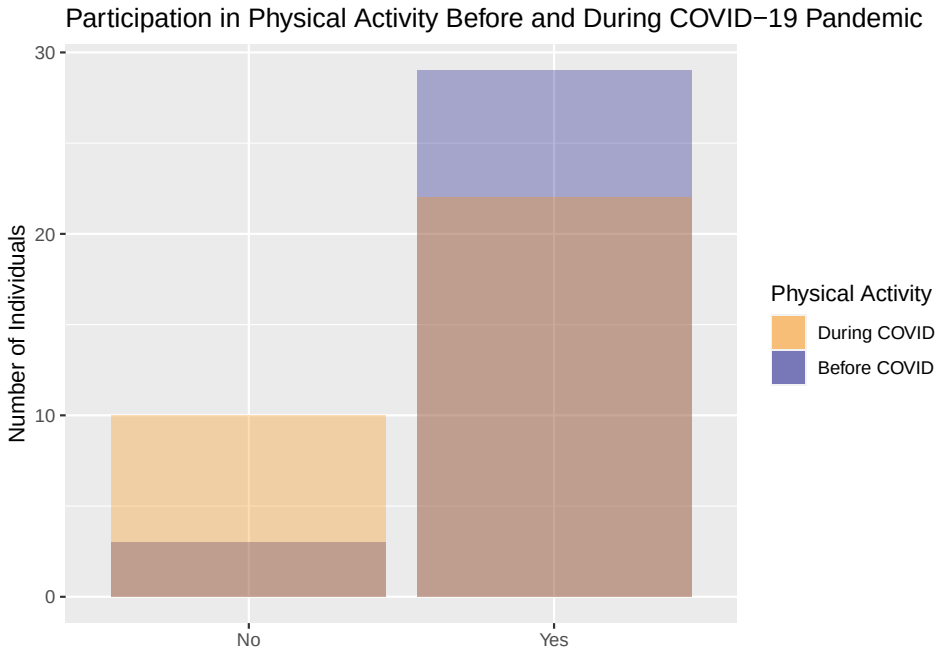
In the case of the categorical variables, plots will be presented to further showcase the before and after affects.

a. Participation in Physical Activity

Numerical Measures	Physical Activity BEFORE the COVID-19 Pandemic
Mean	0.9062500
Min	0.0000000
Max	1.0000000
Median	1.0000000
Standard Deviation	0.2961446
Interquartile Range	0.0000000

Numerical Measures	Physical Activity DURING the COVID-19 Pandemic
Mean	0.6875000
Min	0.0000000
Max	1.0000000
Median	1.0000000
Standard Deviation	0.4709291
Interquartile Range	1.0000000

The table above shows the categorical results of the participation in physical activity before and during the COVID-19 Pandemic. We will only be focussing on the mean for this variable since the rest of the values do not provide relevant information for the study. Notice that before COVID the mean was **0.90625** compared to a mean of **0.6875** during the COVID-19 pandemic, meaning that more individuals were physically active before the pandemic. Looking at the second table, notice that the overall mean is lower, meaning that there were more individuals who did not participate in physical activity during the pandemic, thus we can conclude that overall physical activity decreased during the pandemic vs. before.

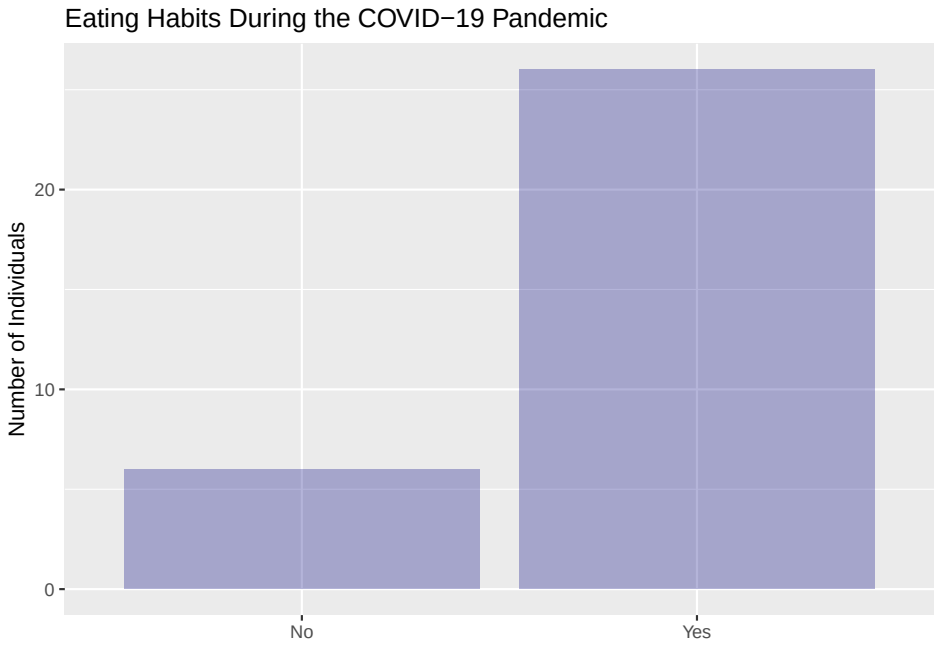


To provide some graphical insight on the physical activity of participants before and during COVID-19 pandemic we can see that the number of individuals that did participate in physical activity dropped from approximately 28 to 22 and those that did not increased from approximately 5 to 10. This means that there was a significant enough impact of those who did not participate in physical activity during the pandemic to change the mean, as shown in the tables above.

b. Change in Eating Habits During COVID-19 Pandemic

Numerical Measures	Chane in Eating Habits During COVID-19 Pandemic
Mean	0.8125000
Min	0.0000000
Max	1.0000000
Median	1.0000000
Standard Deviation	0.3965578
Interquartile Range	0.0000000

The table above shows the change of eating habits of the participants as a result of the COVID-19 Pandemic. The mean value is **0.8125** which tells us that the majority of participants experienced a change in their eating habits. Since this is a binary question with two possible answers, 0 meaning no and 1 meaning yes, we can conclude that there was a significant amount of individuals that experienced a change in their eating habits.

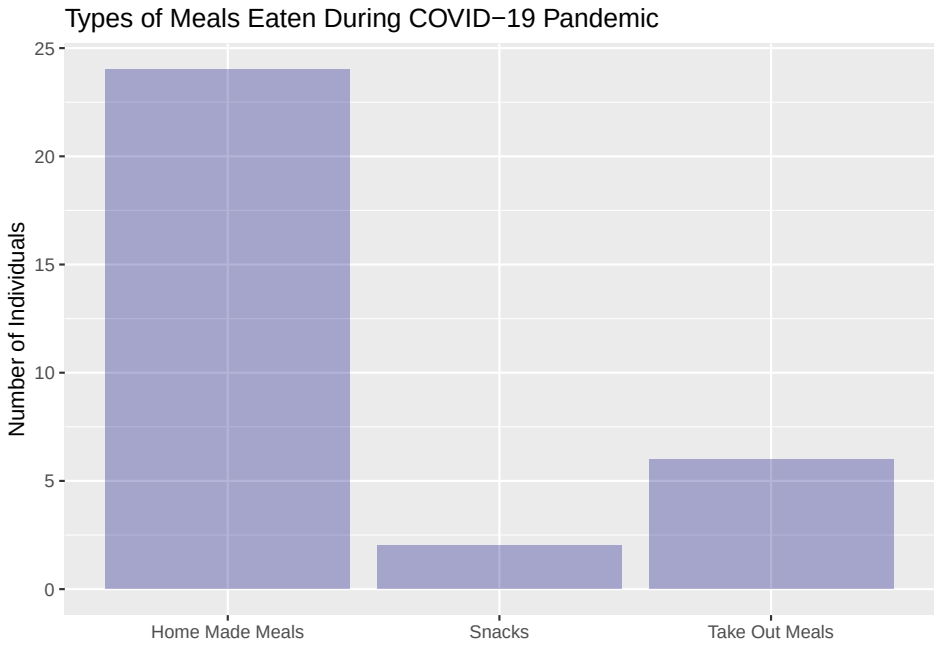


To provide some graphical insight on the change in eating habits during the COVID-19 Pandemic we can see that the number of participants that experienced some change is significantly higher than those that did not with approximately 26 agreeing and approximately 6 disagreeing.

c. Types of Meals Eaten During COVID-19 Pandemic

Numerical Measures	Types of Meals Eaten During COVID-19 Pandemic
Mean	1.4375000
Min	1.0000000
Max	3.0000000
Median	1.0000000
Standard Deviation	0.8007053
Interquartile Range	0.2500000

The table above shows the mean of the types of meals that were eaten throughout the COVID-19 pandemic. Since these numbers are referencing 1 of 3 options, 1 being home made meals, 2 being snacks and 3 being take out meals, we can see that the average is about **1.4375**, rounded down to 1. This means that the average individual ate homemade meals throughout the pandemic. This data alone does not offer much significance but when paired with the rest of our data, it may provide important insight.

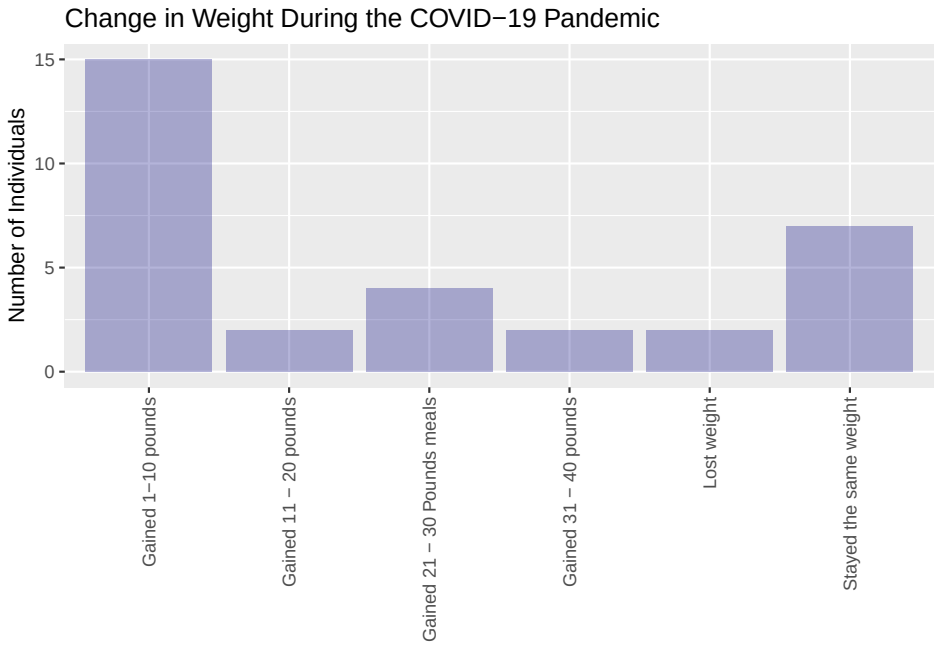


Looking at the graph above, we notice the same results as our table. Majority of individuals ate home cooked meals, and not many individuals ate snacks or take out meals throughout the pandemic. Again, this alone does not offer much significance but when paired with the rest of our data, it may provide important insight.

d. Change in Weight During COVID-19 Pandemic

Numerical Measures	Change in Weight During COVID-19 Pandemic
Mean	2.843750
Min	1.000000
Max	6.000000
Median	2.000000
Standard Deviation	2.080778
Interquartile Range	4.000000

The table above shows the mean of how much weight individuals gained throughout the COVID-19 pandemic. Since these numbers are referencing 1 of 6 options (1 being an individual gained 1-10 pounds, 2 being an individual gained 11-20 pounds, 3 being an individual gained 21-30 pounds,4 being an individual gained 31-40 pounds, losing weight, staying the same weight), we can see that the average is about **2.84375**, rounded up to 3. This means that the average individual gained between 21-30 pounds throughout the pandemic.



The graph above shows that majority of individuals gained between 1-10 pounds during the COVID-19 pandemic. Notice that the number of individuals gaining between 11-40 pounds or staying the same weight seems to be steady as well. This data can be paired with other variables to conclude whether or not this could be a direct result of not having access to facilities or if there may be other reasons for weight gain.

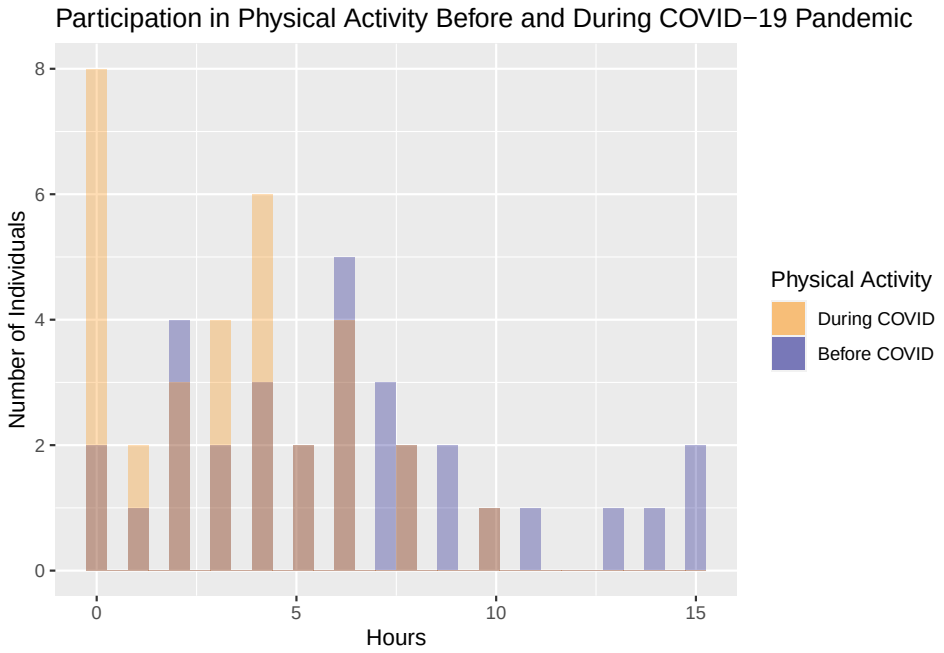
e. Average Participation in Physical Activity

Numerical Measures	Physical Activity BEFORE the COVID-19 Pandemic
Mean	6.250000
Min	0.000000
Max	15.000000
Median	6.000000
Standard Deviation	4.158163
Interquartile Range	5.250000

Numerical Measures	Physical Activity DURING the COVID-19 Pandemic
Mean	3.250000
Min	0.000000
Max	10.000000
Median	3.000000
Standard Deviation	2.723849
Interquartile Range	4.250000

The tables above show the mean, min, max, median, standard deviation and interquartile range for Physical Activity before and during the COVID-19 Pandemic. Notice that the mean hours of physical activity before the COVID-19 pandemic is **6.25**, whereas the mean hours of physical activity during the pandemic is almost half, **3.25**. These values are the sum of all hours spent participating in physical activity before and after the COVID-19 pandemic, respectively, divided by the number of data points for this variable. The min value for both tables is **0**, meaning that there are individuals in our data that did not participate in physical activity before and during the COVID-19 pandemic. The standard deviation explains how the data is spread out

from the mean value. Since the hours of physical activity range from 0 to 15 (min to max) for the first table, the standard deviation is **4.1581634**. Since the hours of physical activity range from 0 to 15 (min to max) for the first table, standard deviation is **2.7238493**. The Interquartile Range, **5.25** (before the COVID-19 pandemic) and **4.25** (during the COVID-19 pandemic), represents how spread out the middle 50% of the data points are.



A histogram can be constructed to compare the relationship between the physical activity before and after the COVID-19 pandemic. Notice that the majority of individuals spent around 7-8 hours per week participating in physical activity before the COVID-19 pandemic, however that value decreased during the COVID-19 pandemic as the histogram is slightly skewed to the right, meaning that more of the data lies closer to the left side of the plot. Thus, it can be seen that the number of hours an individual participating in physical activity decreased as a result of the COVID-19 pandemic.

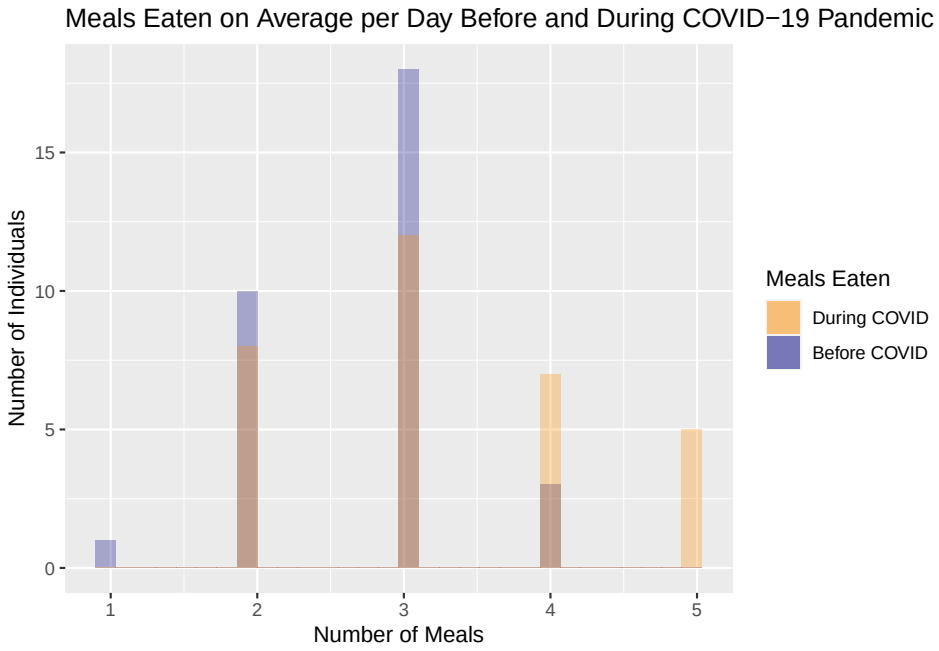
f. Average Number of Meals Eaten per Day

Numerical Measures	Meals Eaten Daily Before COVID-19 Pandemic
Mean	2.7187500
Min	1.0000000
Max	4.0000000
Median	3.0000000
Standard Deviation	0.6831792
Interquartile Range	1.0000000

Numerical Measures	Meals Eaten Daily During COVID-19 Pandemic
Mean	3.281250
Min	2.000000
Max	5.000000
Median	3.000000
Standard Deviation	1.023415

Numerical Measures	Meals Eaten Daily During COVID-19 Pandemic
Interquartile Range	1.250000

The tables above show the mean, min, and max for the average number of meals eaten per day before and during the COVID-19 Pandemic. Notice that the mean number of meals eaten per day before the COVID-19 pandemic is **2.71875**, whereas the number of meals eaten per day during the pandemic is **3.28125**. We can see that the means are relatively close to 3.0 which tells us that there is no significant difference in terms of the number of meals that were eaten before vs. during the pandemic. Looking at the min values before and during the COVID-19 pandemic, we can see that there was an increase from 1 meal minimum to 2 meals minimum and the the max number of meals per day on average went from to 4 to 5, leaving us with a medium of 3. Although this data alone does not offer much significance, it can be paired with the rest of our data, and possibly provide important insight..



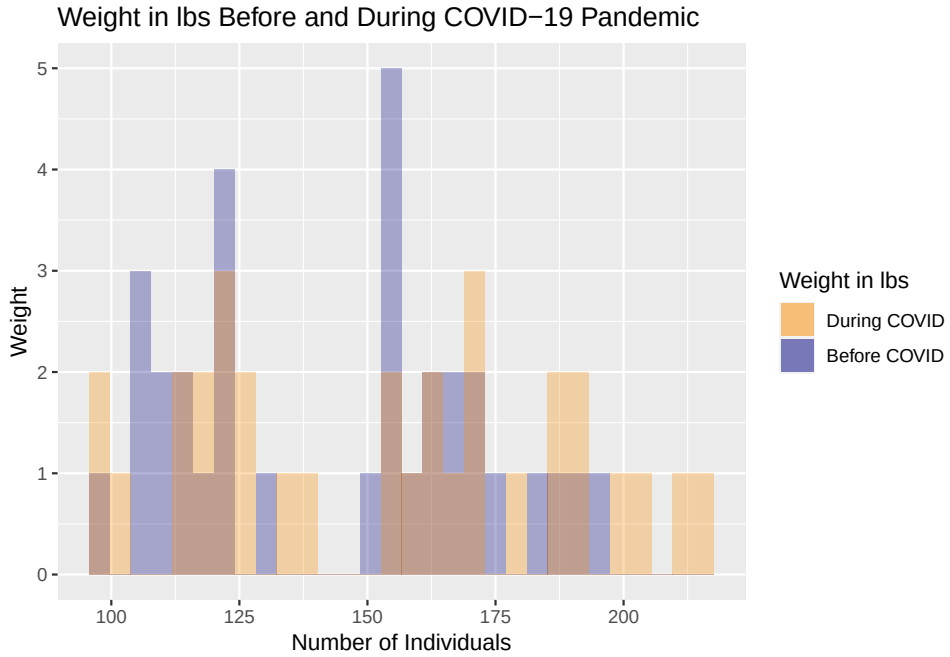
A histogram can be constructed to compare the relationship between the average number of meals eaten per day before and after the COVID-19 pandemic. Right away we can see that the histogram is shifted to the right by 1 unit when making the comparison between meals eaten on average per day before vs during COVID-19. The graph does not provide significant data other than what has already been mentioned and seen from the table, but it can be paired with the rest of our data to possibly provide important insight.

g. Weight in lbs

Numerical Measures	Weight in lbs Before COVID-19 Pandemic
Mean	144.53125
Min	97.00000
Max	195.00000
Median	155.00000
Standard Deviation	29.62996
Interquartile Range	48.75000

Numerical Measures	Weight in lbs During COVID-19 Pandemic
Mean	153.09375
Min	98.00000
Max	215.00000
Median	157.00000
Standard Deviation	35.23136
Interquartile Range	59.50000

The tables above show the mean, min, and max, median, standard deviation and interquartile range for the weight of the individuals before vs during the COVID-19 Pandemic. Notice that the mean weight of individuals before the pandemic was **144.53125**, whereas the mean weight of individuals during the pandemic is, **153.09375**. This shows that the mean weight increased overall, since the pandemic started. The min value for the table representing the data before the pandemic is **97** and the min value for the table representing the data during the pandemic is **98**. We can see that there is not a significant difference (just 1 pound), thus this is not too significant. The max value on the other hand increased by 20 pounds, with the max weight before the pandemic being **195** vs **215** during the pandemic. The standard deviation explains how the data is spread out from the mean value. Since the weight of an individual ranges from 97 to 195 (min to max) for the first table, the standard deviation is **29.6299591**. Since the weight of an individual ranges from 98 to 215 (min to max) for the second table, the standard deviation is **35.2313638**. The Interquartile Range, **48.75** (before the COVID-19 pandemic) and **59.5** (during the COVID-19 pandemic), represents how spread out the middle 50% of the data points are.



A histogram can be constructed to compare the relationship between weight in lbs before vs after the COVID-19 pandemic. Notice that during the graph seems to have shifted and become more spread out during the COVID-19 pandemic, which may be a result of many individuals gaining weight. Before the COVID-19 pandemic, the bulk of the data was between 115-160 pounds, during the pandemic it seems to be between 150-200 pounds. This data, paired with the rest of our data, can provide significant insight related to our goal for this study.

All analysis for this report was programmed using `R version 4.0.2`.

Methods

This section will conduct hyopthesis testing on the variables and will also consider the confidence intervals. The hypothesis testing done will aid in determining which variables are more or less significant.

This section will use the paired sample t-test [12], allowing the comparison between variabels considered before the COVID-19 pandemic as well as after. In order to ensure the paired t-test can be performed, there are some assumptions that much be satisfied with the data:

1. Dependent variable must be continuous (interval or ratio level)
2. The groups/samples should be focussed on the same subject/topic
3. The sample data should be randomly generated from the population
4. Paired values should have a normally distributed difference
5. Outliers should not be present in the difference between the two groups

To begin our test, we must declare our null and alternative hypothesis as follows:

Null Hypothesis: $H_0 : \mu_1 = \mu_2$ Alternative Hypothesis: $H_1 : \mu_1 > \mu_2$

This null and alternative hypothesis will be used for all of the paired sample t-tests that are conducted

1. First consider the paired sample t-test on the following variables: average physical activity before the COVID-19 pandemic and average physical activity during the COVID-19 pandemic. Note that the values for this data is considered to be in hours. The subject of both of our variables is the same, COVID-19. Random sampling is also considered in the dataset, as mentioned earlier in this report. Finally, from our histrograms, we can see that our data is approximately normally distributed, but this is acceptable as our sample size was relatively small.

Let μ_1 represent the population mean for the average physical activity before the COVID-19 Pandemic Let μ_2 represent the population mean for the average physical activity during the COVID-19 Pandemic

Hypothesis Test [12]: $t = \frac{\bar{x}_{diff} - 0}{S_x}$

Note: $\bar{x}$ represents the sample mean of the differences Note: S_x represents the estimated standard error of the mean

Confidence Interval[13]: $(a, b) = \bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2,df} * \sqrt{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})}$

2. Next, consider the paired sample t-test on the following variables: weight of an individual during the COVID-19 pandemic vs weight of an individual before the COVID-19 pandemic. Similar to our first set of variables, the subject of the two variables is the same. Random sampling was also performed on this data as mentioned earlier in this report. Finally, the histrograms for this data show that is it approximately normally distributed, as needed.

Let μ_1 represent the population mean for the weight of an individual during the COVID-19 Pandemic Let μ_2 represent the population mean for the weight of an individual before the COVID-19 Pandemic

Hypothesis Test [12]: $t = \frac{\bar{x}_{diff} - 0}{S_x}$

Note: $\bar{x}_{diff}$ represents the sample mean of the differences and S_x represents the estimated standard error of the mean

Confidence Interval [13]: $(a, b) = \bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2,df} * \sqrt{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})}$

Results

The results of the first paired sample t-test considering the average physical activity before vs during the COVID-19 Pandemic as as follows:

```
##
##  Paired t-test
##
## data:  avgPhysAct_before and avgPhysAct_during
## t = 4.4345, df = 31, p-value = 0.0001078
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.620256 4.379744
## sample estimates:
## mean of the differences
##                       3
```

Calculated Measure	Calculated Value for Event Participation
T-Test Value	4.4345000
P-Value Test	0.0001078
Confidence Interval Lower Bound	1.6202560
Confidence Interval Upper Bound	4.3797440

From the table of results above, notice that the t-test value is **4.4345** and the p-value is 1.078×10^{-4}. Since the t-value is greater than the p-value, it is safe to reject the null hypothesis and conclude that the average amount of time spent on physical activity before COVID-19 Pandemic was significantly affected when tested with the amount of time spent on physical activity during the COVID-19 Pandemic.

In addition, notice that the 95% confidence interval is (1.620256, 4.379744). Since this interval does not contain 0, we can confidently reject the null hypothesis and say, with 95% confidence that the average physical activity participation before the COVID-19 pandemic decreased by 1.620256 but increased by 4.379744

The results of the second paired sample t-test considering the weight of an individual during vs before the COVID-19 Pandemic as as follows:

```
##
##  Paired t-test
##
## data:  weight_during and weight_before
## t = 4.914, df = 31, p-value = 2.748e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   5.008731 12.116269
## sample estimates:
## mean of the differences
##                  8.5625
```

Calculated Measure	Calculated Value for Event Participation
T-Test Value	4.9140000
P-Value Test	0.0000275
Confidence Interval Lower Bound	5.0087310

Calculated Measure	Calculated Value for Event Participation
Confidence Interval Upper Bound	12.1162690

From the table of results above, notice that the t-test value is **4.914** and the p-value is 2.748×10^{-5}. Since the t-value is greater than the p-value, it is safe to reject the null hypothesis and conclude that the weight of an individual after COVID-19 Pandemic was significantly affected when tested with the weight of an individual during the COVID-19 Pandemic.

In addition, notice that the 95% confidence interval is (5.008731, 12.116269). Since this interval does not contain 0, we can confidently reject the null hypothesis and say, with 95% confidence, that the the weight of an individual during the COVID-19 pandemic decreased by 5.008731 but increased by 12.116269

Conclusion

In conclusion, it is evident that weight is being affected as a result of the COVID-19 pandemic. After conducting several types of analyses on the variables, including numerical summaries, paired t-tests, and creating plots, it was noticable that many individuals gained weight throughout the course of the COVID-19 pandemic and being at home. The goal of this study was to determine whether the restrictions and lifestyle changes as a result of the pandemic caused changes in an indivduals overall weight, and it was shown that eating habits slightly changed, many individuals stopped participating in physical activity, and most individuals gained between 1-10 pounds since the pandemic started. In addition, the average number of meals daily increased with the minimum number of meals before the pandemic being 1, and the minimum after the pandemic being 2, thus individuals were eating 1 extra meal on average since the COVID-19 pandemic began.

It is clear that many individuals are lacking the exercise and physical activity they need as a result of not having access to the facilities. This study can be further used to hone in on other significant health impacts that individuals might undergo as a result of the pandemic, especially ones that may be linked to weight gain and health risks

Bibliography

1. Nielsen, K. (2020, April 24). *A timeline of COVID-19 in Ontario. Global News; Global News.* https://globalnews.ca/news/6859636/ontario-coronavirus-timeline/.

2. DeClerq, K. (2020, September 28). *Ontario is now in the second wave of the COVID-19 pandemic. How bad will it get? Toronto; CTV News.* https://toronto.ctvnews.ca/ontario-is-now-in-the-second-wave-of-the-covid-19-pandemic-how-bad-will-it-get-1.5122765

3. Toneguzzi, M. (2020, September 8). *Gym Memberships and Attendance Plummet in Canada Amid COVID-19: Survey. Retail Insider.* https://retail-insider.com/retail-insider/2020/09/gym-memberships-and-attendance-plummet-in-canada-amid-covid-19-survey/#:~:text=The%20percentage%20of%20people%20keeping,40.44%20percent%20for%20the%20U.S

4. Buy Email Lists | Build Your Own Marketing List | Bookyourdata.com. (2021). *Bookyourdata.com.* https://www.bookyourdata.com/tool/main/wjhhZdl1s2yM01ei

5. Wikipedia Contributors. (2020, October 7). *Sampling frame. Wikipedia; Wikimedia Foundation.* https://en.wikipedia.org/wiki/Sampling_frame#Types_of_sampling_frames

6. Name Directory - Markham, A to Yogen Fruz in Canada | Canada411 People Finder. (2021). *Canada411Ca.* https://www.canada411.ca/markham-1

7. Munsters, M. J. M., & Saris, W. H. M. (2012). *Effects of Meal Frequency on Metabolic Profiles and Substrate Partitioning in Lean Healthy Males. US National Library of Medicine, 7(6), e38632.* https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3374835/#:~:text=Increasing%20eating%20frequency%20has%20been,body%20fat%20storage%20%5B4%5D

8. National Institute of Diabetes and Digestive and Kidney Diseases. (2021, May 14). *Factors Affecting Weight & Health | NIDDK. National Institute of Diabetes and Digestive and Kidney Diseases.* https://www.niddk.nih.gov/health-information/weight-management/adult-overweight-obesity/factors-affecting-weight-health

9. Markham, ON - Demographics | Townfolio. (2021). *Townfolio.* https://townfolio.co/on/markham/demographics

10. How to remove a column from an R data frame? (2020). Tutorialspoint.com. https://www.tutorialspoint.com/how-to-remove-a-column-from-an-r-data-frame

11. Riederer, E. (2020, November 23). *10.2 The kableExtra package | R Markdown Cookbook. Bookdown.org.* https://bookdown.org/yihui/rmarkdown-cookbook/kableextra.html

12. Kent.edu. (2018). *LibGuides: SPSS Tutorials: Paired Samples t Test.* https://libguides.library.kent.edu/SPSS/PairedSamplestTest

13. Yale.edu. (2021). *Confidence Intervals.* http://www.stat.yale.edu/Courses/1997-98/101/confint.htm