Mood Responses to Various Exercise Types

Sucharitha Dodamgodege[1*], Olaoluwa Ogunleye[1*], Abd Al Rahman Al[1,2]

[1]Department of Mathematics, Clarkson University, Potsdam, NY

[2]Department of Mathematics, Clarkson University, Potsdam, NY

*** Correspondence:**
Corresponding Authors
First Author: dodamgss@clarkson.edu, First Author: ogunleoa@clarkson.edu

Keywords: Vigor, Depression, Confusion, Exercise Type, Exercise Intensity

Abstract

This report presents a machine learning based approach to the prediction of a person's mood based on exercise type, exercise intensity, sleep efficiency, and sex. The dataset used during the analysis was obtained from a survey of undergraduate students during the pandemic period. The reports start with a statistical analysis and description of the variables of interest. Logistic regression model is employed to predict mood (Depression, Vigor, Confusion) level.

1 Introduction

The COVID-19 pandemic is a massive global health crisis (Bavel et al., 2020) and rapidly spreading pandemic (Bentlage et al., 2020) of recent times. As compared to the earlier pandemics the world has witnessed, the current COVID-19 pandemic is now on the top of the list in terms of worldwide coverage. This is the first time the whole world is affected simultaneously and struck strongly in a very short span of time. Initially, the death rate due to COVID-19 was around 2%, which has now increased to around 4–6% (World Health Organization [WHO], 2020). The statistics do not look so severe, but the total number of cases and the rate at which these cases are increasing day by day make the situation alarming. Exponential growth in COVID-19 cases has led to the isolation of billions of people and worldwide lockdown. COVID-19 has affected the life of nearly each person around the world. The difference between personal or professional lives has narrowed due to work-from-home instructions, and people's lives are revolving around these two due to the lockdown. People have also been pondering over a vital concern at home, i.e., the importance of their health and fitness.

Although imposing lockdown or quarantine for the population has been one of the widely used measures across the world to stop the rapid spread of COVID-19, it has severe consequences too. Recent multinational investigations have shown the negative effect of COVID-19 restrictions on social participation, life satisfaction (Ammar et al., 2020b), mental well-being, psychosocial and emotional disorders as well as on sleep quality (Xiao et al., 2020), and employment status (Ammar et al., 2020d). Announcement of a sudden lockdown of all services and activities, except few essential services, by the authorities has resulted in a radical change in the lifestyle of affected people (Jiménez-Pavón et al., 2020) and has severely impaired their mental health, which has been manifested in the form of increased anxiety, stress, and depression (Chtourou et al., 2020). The sudden changes in people's lifestyle include, but are not limited to, physical activities and exercise. Ammar et al. (2020a) have reported that COVID-19 home confinement has resulted in a decrease in all levels of physical activities and about 28% increase in daily sitting time as well as increase in unhealthy pattern of food consumption. Similar results are also reported by other researchers (Ammar et al., 2020c; de Oliveira Neto et al., 2020) as well. Although these abrupt changes have influenced every individual, many people who were regularly following their fitness activities in gyms, or in the ground, or other places before the lockdown have been affected intensely. Closure of fitness centers and public parks has forced people to stay at home, which has disturbed their daily routines and hampered their fitness activities. While compulsion to stay at home for a long period of time poses a challenge to the continuity of physical fitness, the experience of hampered physical activities, restricted social communication, uncertainty, and helplessness leads to the emergence of psychological and physical health issues (Ammar et al., 2020a,c). Varshney et al. (2020) have found that psychological problems are occurring in adults while adjusting to the current lifestyle in accordance with the fear of contracting the COVID-19 disease. However, effective coping strategies, psychological resources, and regular physical exercise can be helpful in dealing with such health-related problems during the COVID-19 pandemic (Chtourou et al., 2020).

It is important to note that physical activities (PA) and exercise not only maintain physical and psychological health but also help our body to respond to the negative consequences of several diseases such as diabetes, hypertension, cardiovascular diseases, and respiratory diseases (Owen et al., 2010; Lavie et al., 2019; Jiménez-Pavón et al., 2020). In a recent review of 31 published studies, Bentlage et al. (2020) concluded that physical inactivity due to current pandemic restrictions is a major public health issue that is a prominent risk factor for decreased life expectancy and many physical health problems (Jurak et al., 2020). Exercise is shown to keep other physical functions (respiratory, circulatory, muscular, nervous, and skeletal systems) intact and supports other systems (endocrine, digestive, immune, or renal systems) that are important in fighting any known or unknown threat to our body (Lavie et al., 2019; Jiménez-Pavón et al., 2020).

Regular physical activity, while taking other precautions, is also considered effective in dealing with the health outcomes of the COVID-19 pandemic (Chen et al., 2020). Researchers from the University of Virginia Health System (Yan and Spaulding, 2020) suggest that regular exercise might significantly reduce the risk of acute respiratory distress syndrome, which is one of the main causes of death in COVID-19 patients. Exercise and physical activities have important functions for individuals' psychological well-being as well (Stathi et al., 2002; Lehnert et al., 2012). There is sufficient literature to show that exercise can play a vital role in the promotion of positive mental health and well-being (e.g., Mazyarkin et al., 2019). However, when health promotion activities such as sports and regular gym exercises are not available in this pandemic situation, it is very difficult for individuals to meet the general WHO guidelines (150 min moderate to mild PA or 75 min intensive PA per week or combination of both) (cf. Bentlage et al., 2020). Amidst this pandemic-related restriction (home confinement and closed gyms, parks, and fitness centers), how people cope up and find ways to continue their physical fitness remains an important question.

In this paper, we are interested in analyzing how different exercise types affect the mental and physical mood of a person. We collected a survey of 155 university students over the period of 60 days during the Covid-19 lockdown in early 2020. For each of the 60days survey periods, data about a person's total sleep time, time spent exercising, specific type of exercise, and exercise intensity were collected. We have considered six categories under exercise type, namely no exercise,I- me(exercise my themself), I-you(Exercise with someone else i.e. playing badminton), I-society(exercise with a group of people i.e. playing football), I nature(outdoor activities i.e.hiking,walking) and multiple types. Exercise intensity has been considered as 0- no intensity, 1- low intensity, 2- moderate intensity and 3- high intensity. Simultaneously, data about the individual's moods like tension, depression, anger, fatigue, confusion, and vigor were recorded. Then a person's state was measured as state physical energy, state physical fatigue, state mental energy, and state mental fatigue. Also, covariate data like sex and age were recorded.

Furthermore, we performed some form of initial data cleaning on the dataset by first taking out the subject with incomplete information (like sex), then we considered a subset of the subjects who were consistent with the survey for more than 14 days.

2. Feature Description and Summary Statistics

Section 2.1: This section is for descriptive statistics for our dataset.

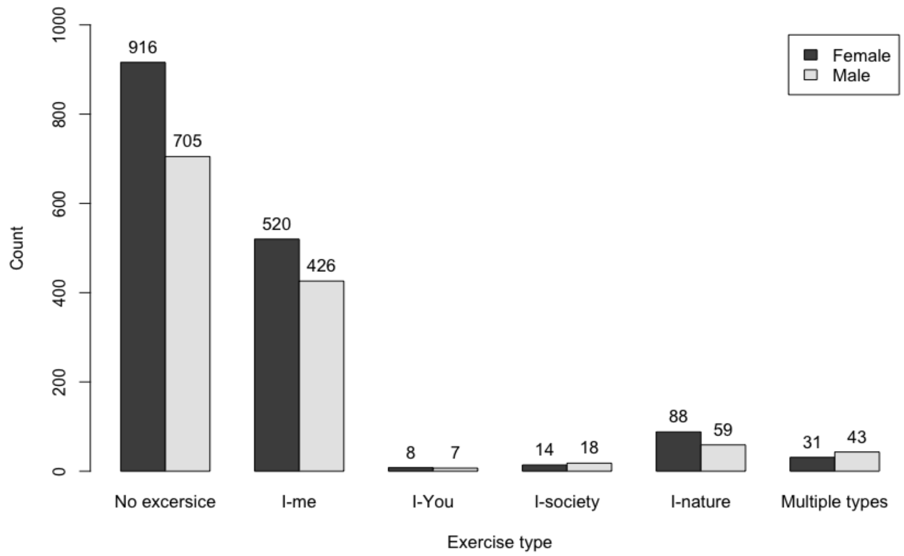

Figure 1: Gender wise count of students who have done different exercise types

The bar chart shows that many students did not exercise during this period. Among the people who exercised, I-me type was the most popular type of exercise.

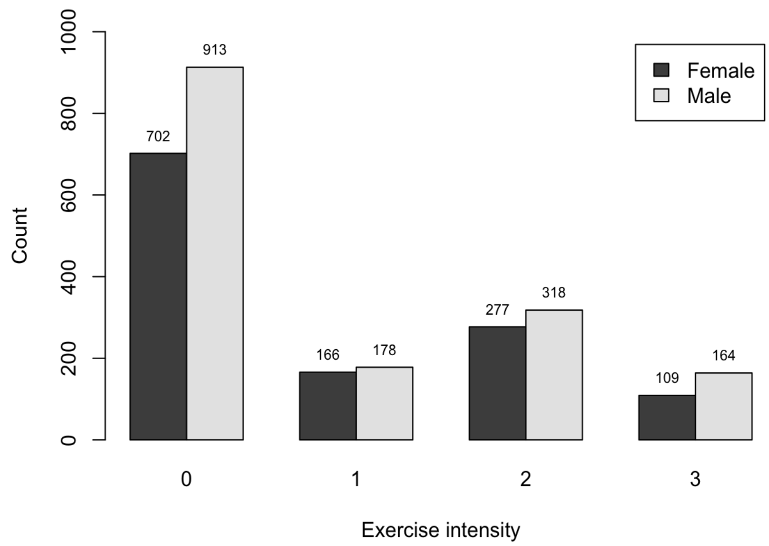

Figure 2:Gender wise count of students who have done different exercise intensities

We can see that male students have done exercise with more intensity than female students. Many female and male students have done exercise with moderate intensity.

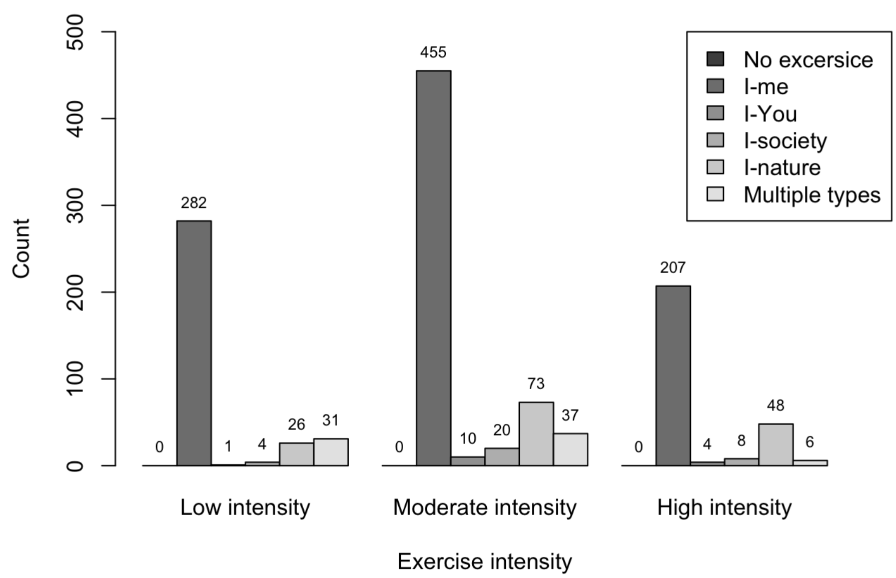

Figure 3:Exercise type wise count of students who have done different exercise intensities

This graph shows how many students have done different exercise types in different intensities.

Section 2.2: We have used boxplots to understand variation of moods like tension, depression, anger, fatigue, confusion, vigor, state physical energy, state physical fatigue, state mental energy, and state mental fatigue with exercise type and gender.

Anger

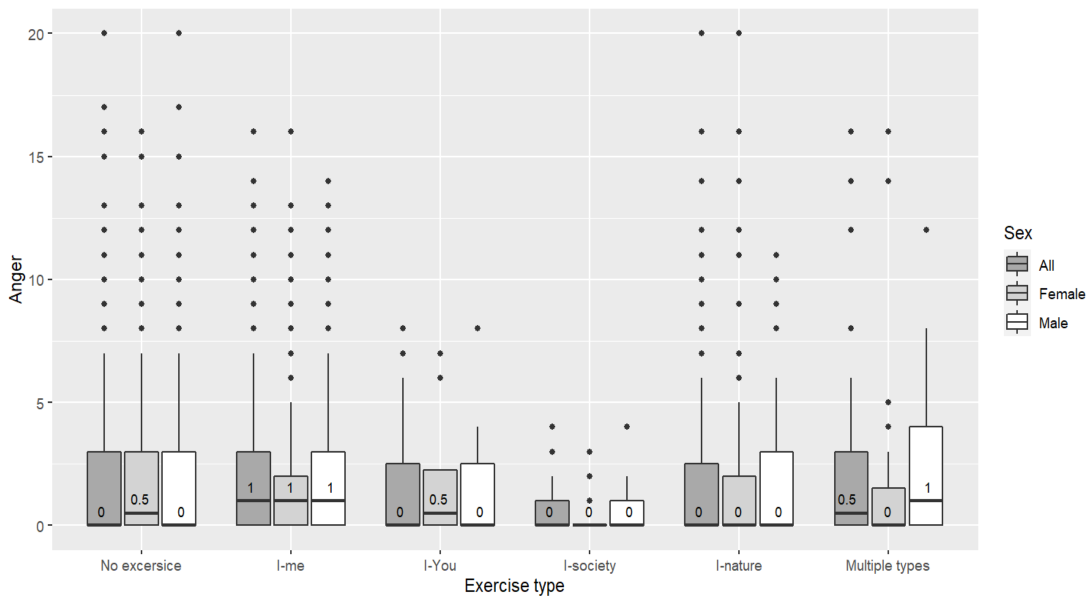

Figure 4: Variation of anger with exercise type and sex.

Confusion

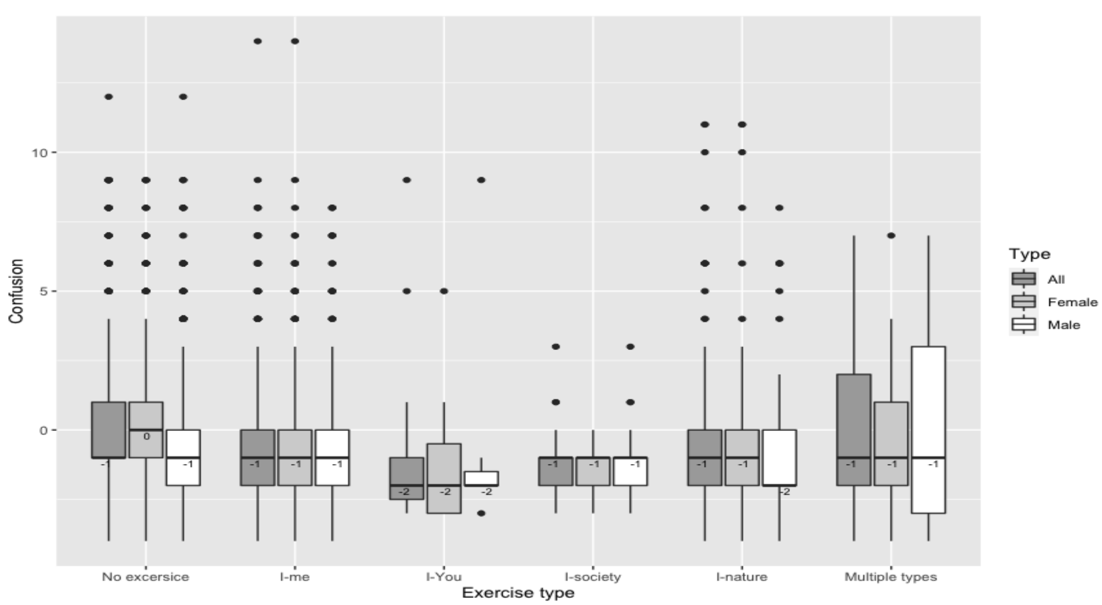

Figure 5: Variation of confusion with exercise type and sex.

Depression

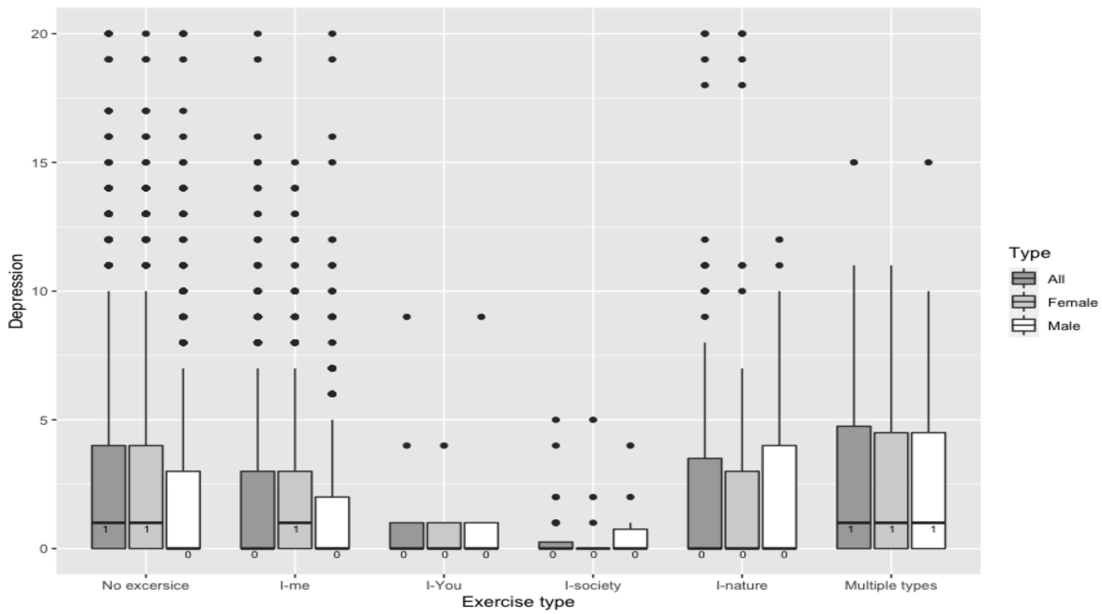

Figure 6: Variation of deperssion with exercise type and sex.

State mental energy

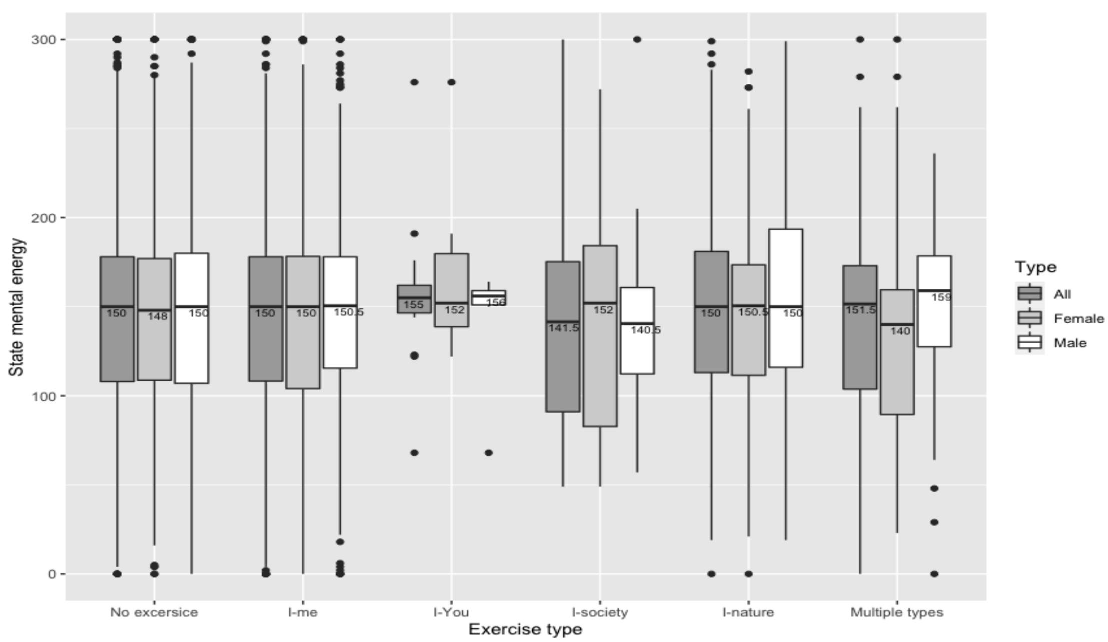

Figure 7: Variation of deperssion with exercise type and sex.

State mental fatigue

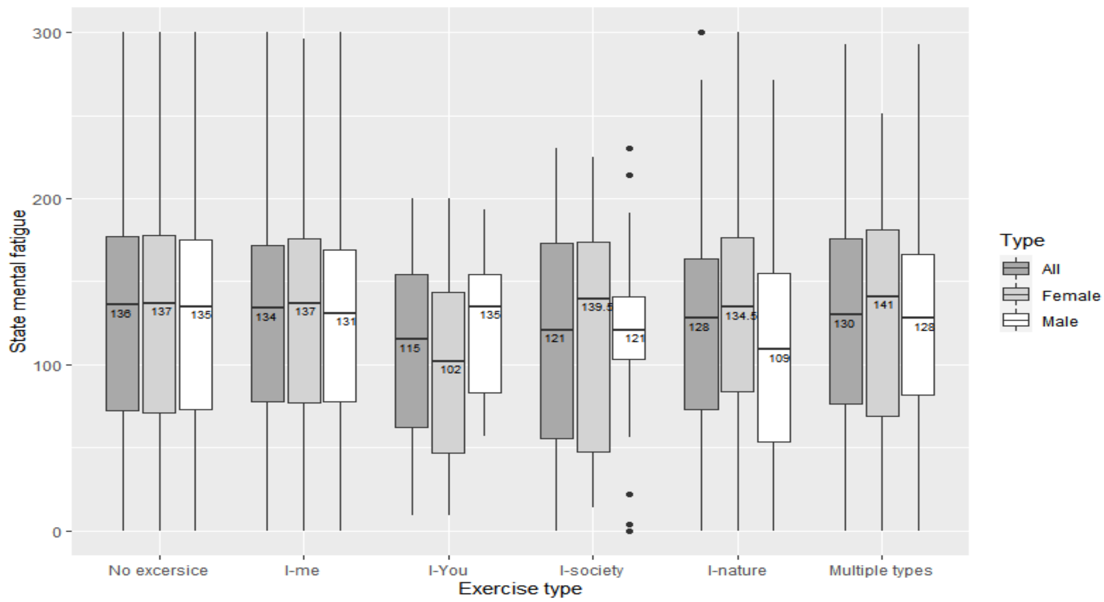

Figure 8: Variation of State mental fatigue with exercise type and sex.

State physical energy

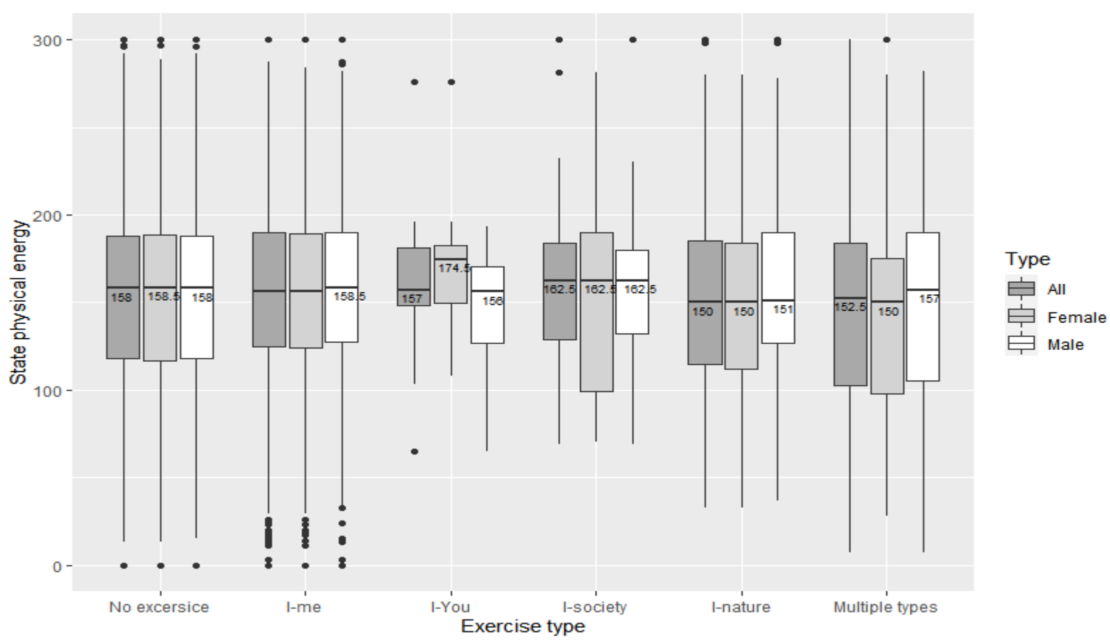

Figure 9: Variation of State mental energy with exercise type and sex.

State physical fatigue

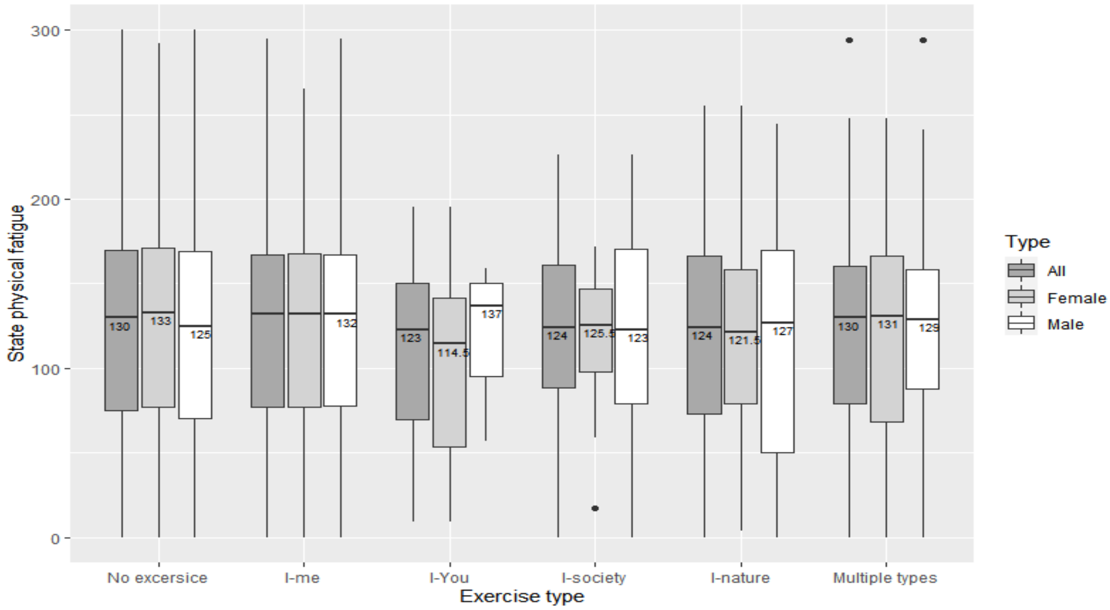

Figure 10: Variation of State mental fatigue with exercise type and sex.

Fatigue

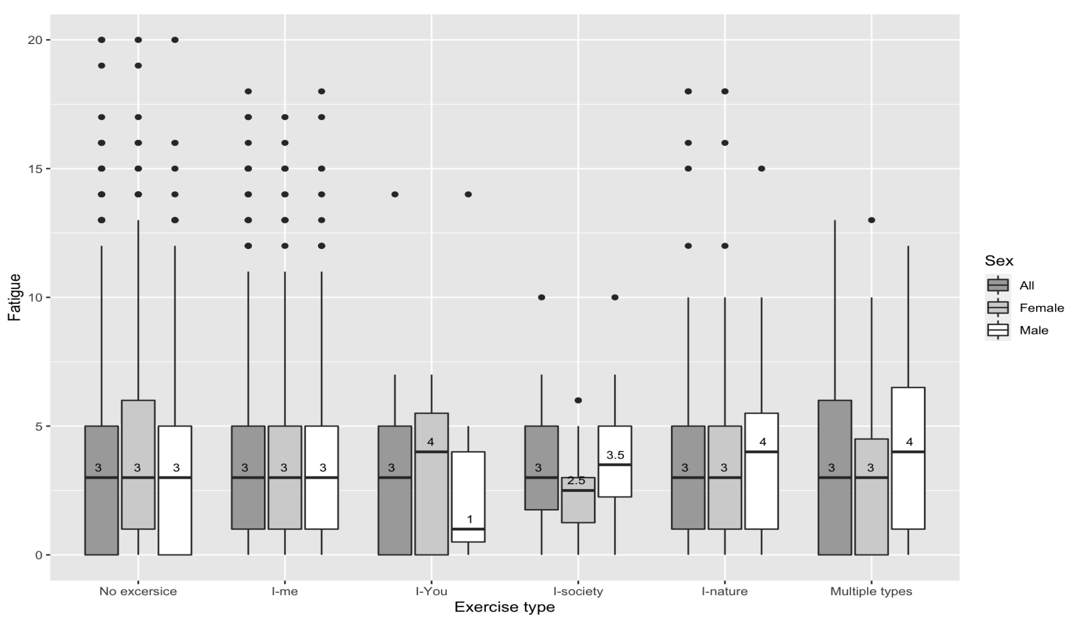

Figure 11: Variation of fatigue with exercise type and sex.

Tension

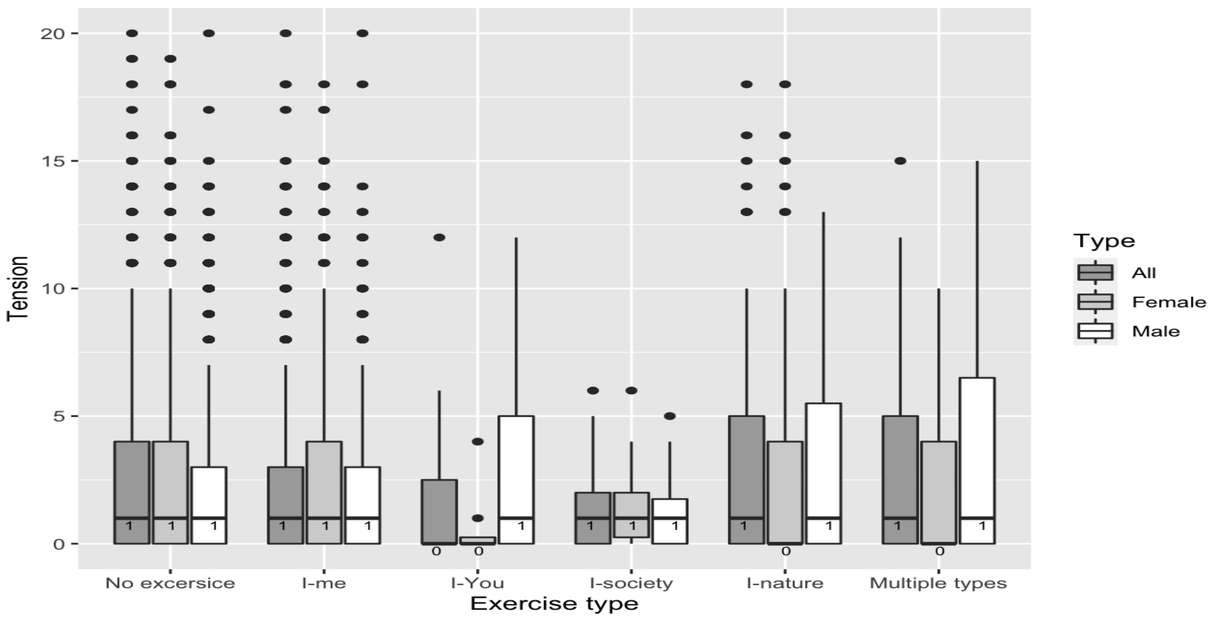

Figure 12: Variation of State mental tension with exercise type and sex.

Vigor

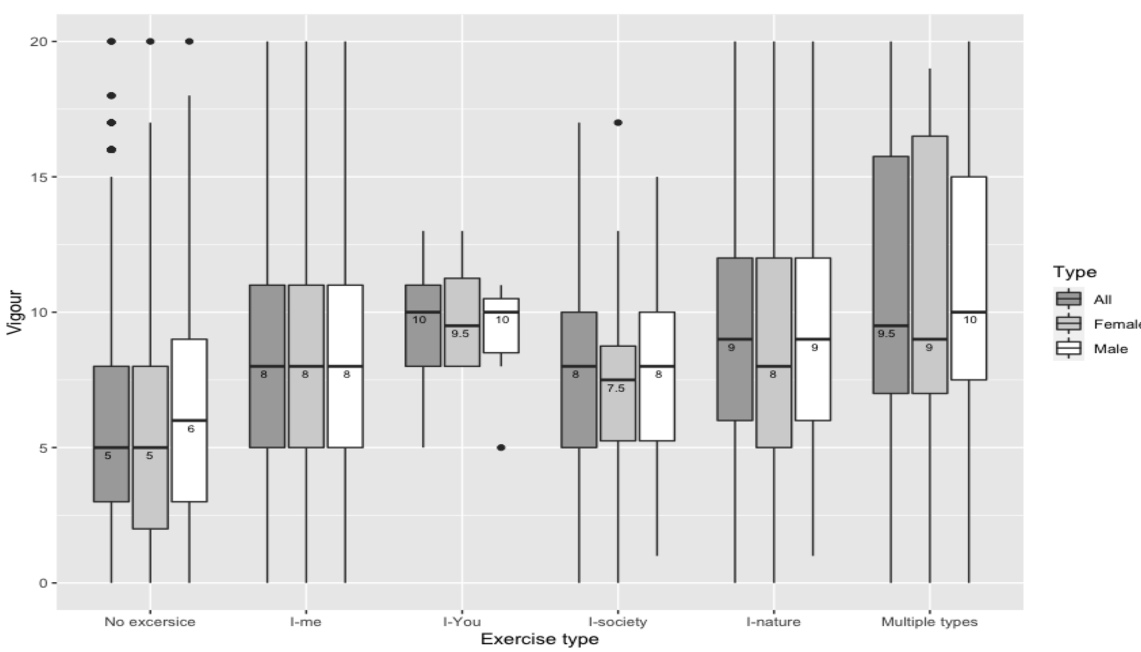

Figure 13: Variation of vigor with exercise type and sex.

Section 2.3: We have used boxplots to identify outliers for dependent variables.

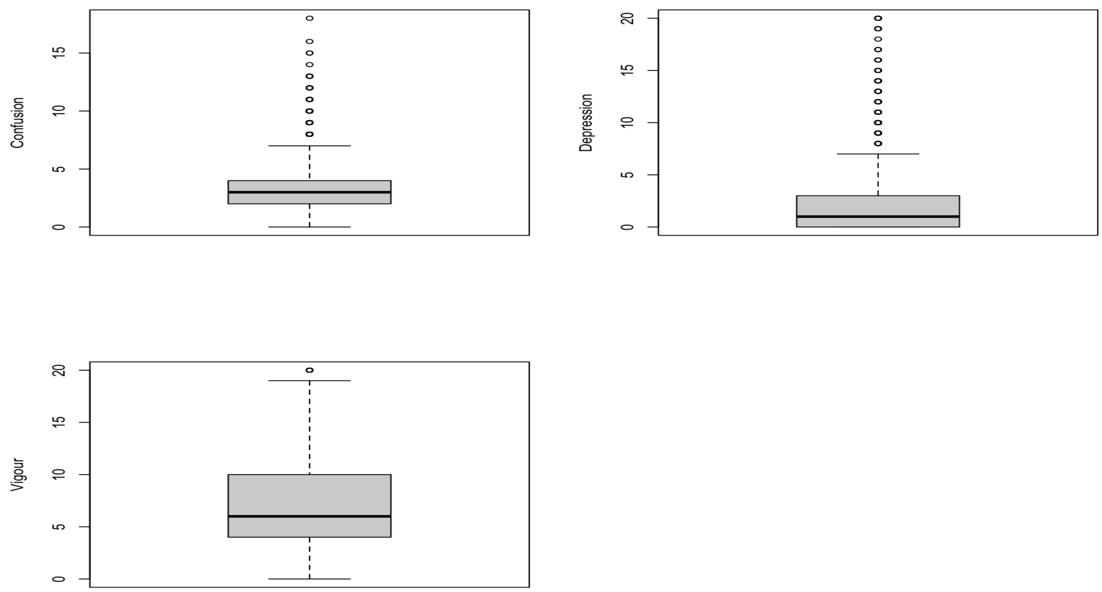

Figure 14: Boxplot for dependent variables

Then we removed outliers from each variable step by step. First we removed outliers from depression (See figure 15 row 1). Then removed outliers from confusion (See figure 15 row 2) and finally we removed outliers from vigor (See figure 15 row 3).

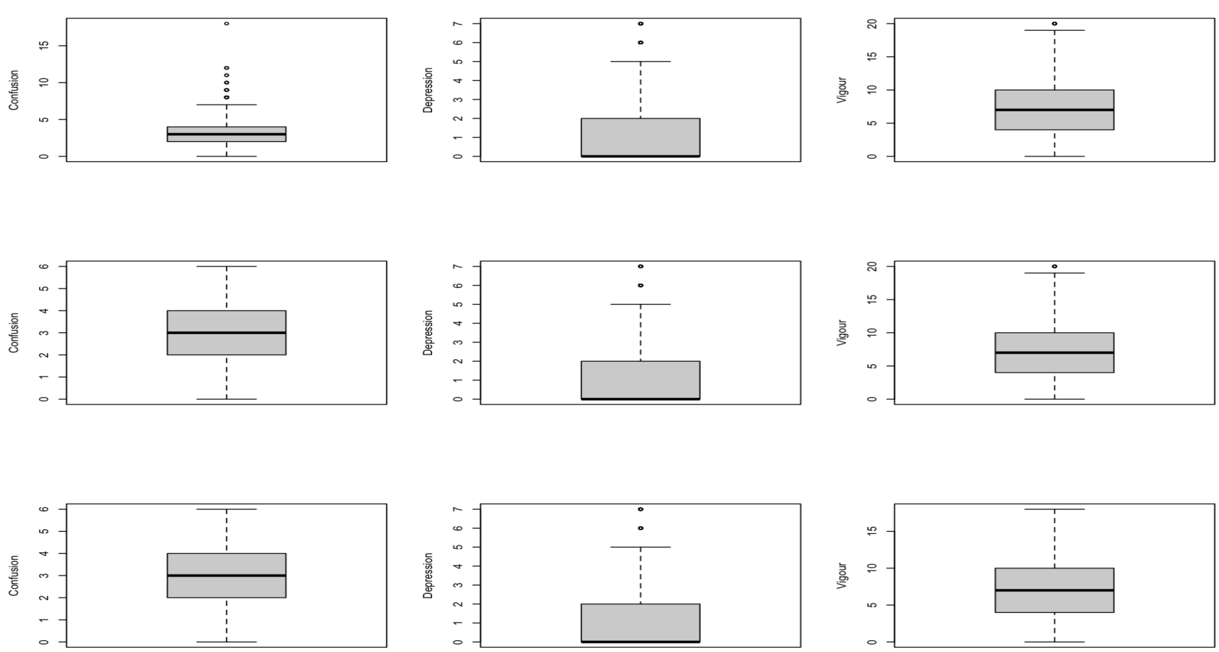

Figure 15: Removing outliers from variables.

Note: a total of 436 {possible outliers} observations were removed from 2835.

Section 2. 4: Next, we checked summary statistics for important mood types which significantly change with exercise type, exercise intensity, sleep efficiency and sex.

Table 1: Frequency table for exercise type vs vigor

Vigor level	No exercise	I-me	I-nature	I-society	I-you	Multiple types
0	154	33	1	1	0	2
1	66	14	1	1	0	0
2	56	13	3	0	0	0
3	97	23	5	2	0	2
4	169	52	6	2	0	1
5	165	76	10	2	0	2
6	108	68	13	3	0	0
7	83	68	6	3	0	4
8	119	98	14	6	3	6
9	91	72	11	2	2	3
10	76	67	9	2	3	6
11	42	47	7	3	3	3
12	42	30	10	0	1	0
13	25	39	11	1	1	1
14	28	28	6	0	0	3
15	51	2	2	0	5	15
16	13	28	1	0	0	4
17	5	7	4	1	0	6
18	3	7	3	0	0	1

Table 2:Frequency table for exercise intensity vs vigor

Vigor level	No intensity	Low intensity	Moderate intensity	High intensity
0	154	3	25	9
1	66	1	8	7
2	56	1	6	9
3	97	7	16	9
4	170	8	32	20
5	165	21	42	27
6	108	12	45	27
7	83	22	19	40
8	119	29	36	62
9	91	24	49	17
10	76	23	43	21
11	42	11	20	10
12	42	11	20	10
13	25	21	26	6
14	28	12	22	3
15	15	48	8	4
16	42	11	20	10
17	5	8	8	2
18	3	6	4	1

Table 3: Frequency table for sex vs vigor

Vigor	Female	Male
0	123	68
1	52	30
2	40	32
3	73	56
4	135	95
5	141	114
6	111	81
7	81	83
8	137	109
9	95	86
10	83	80
11	63	42
12	39	44
13	41	37
14	33	32
15	38	37
16	25	21
17	12	11
18	7	7

Table 4: Summary statistics for sleep efficiency vs vigor

Vigor	Mean	Standard deviation
0	0.84	0.17
1	0.84	0.22
2	0.88	0.15
3	0.90	0.11
4	0.89	0.16
5	0.88	0.13
6	0.89	0.14
7	0.91	0.09
8	0.91	0.09
9	0.93	0.08
10	0.91	0.12
11	0.93	0.07
12	0.93	0.09
13	0.89	0.15
14	0.91	0.13
15	0.94	0.06
16	0.93	0.08
17	0.93	0.05
18	0.95	0.05

3. Methods and Materials

Section 3.1: We ran statistical tests to test variability of moods with sex and exercise type.

Kruskal–Wallis test: to determine if there are statistically significant differences between moods of each exercise group: all, female and male separately.

Null Hypothesis: Mean ranks of moods (all/female/male) for different exercise types are equal

Alternative Hypothesis: Mean ranks of moods (all/female/male) for different exercise types are different

Table 5: Kruskal Wallis' tests results

	Depression		Vigor		Confusion	
	Statistic	P - value	Statistic	P - value	Statistic	P - value
All	16.61	0.005	287.08	<0.001	16.61	<0.001
Female	9.83	0.080	174.80	<0.001	9.83	<0.001
Male	8.61	0.126	111.24	<0.001	8.61	0.126

Post – hoc analysis for depression

Null Hypothesis: There is no difference different exercise types

Alternative Hypothesis: There is difference between different exercise types

Table 6: Dunn's post – hoc analysis for depression,vigor and confusion.

Depression				
	All participants		Female	
	Statistic	p-value	Statistic	p-value
I-me - I-nature	-0.28	0.781	0.66	0.507
I-me - I-society	2.91	0.004	2.22	0.027
I-nature - I-society	2.81	0.005	1.82	0.068
I-me - I-You	1.09	0.277	1.17	0.243
I-nature - I-You	1.13	0.257	0.92	0.358
I-society - I-You	-0.77	0.441	-0.42	0.677
I-me - Multiple types	-1.37	0.172	-0.60	0.548
I-nature - Multiple types	-0.98	0.326	-0.90	0.370
I-society - Multiple types	-3.26	0.001	-2.21	0.027
I-You - Multiple types	-1.58	0.114	-1.33	0.184
I-me - No exercise	-1.87	0.061	-1.15	0.250
I-nature - No exercise	-0.60	0.546	-1.25	0.211
I-society - No exercise	-3.36	<0.001	-2.47	0.014

Depression				
	All participants		Female	
	Statistic	p-value	Statistic	p-value
I-me - I-nature	-0.28	0.781	0.66	0.507
I-me - I-society	2.91	0.004	2.22	0.027
I-nature - I-society	2.81	0.005	1.82	0.068
I-me - I-You	1.09	0.277	1.17	0.243
I-nature - I-You	1.13	0.257	0.92	0.358
I-society - I-You	-0.77	0.441	-0.42	0.677
I-me - Multiple types	-1.37	0.172	-0.60	0.548
I-nature - Multiple types	-0.98	0.326	-0.90	0.370
I-society - Multiple types	-3.26	0.001	-2.21	0.027
I-You - No exercise	-1.39	0.166	-1.35	0.177
Multiple types - No exercise	0.74	0.459	0.26	0.793

Confusion				
	All participants		Female	
	Statistic	p-value	Statistic	p-value
I-me - I-nature	-0.32	0.750	0.20	0.843
I-me - I-society	0.94	0.350	1.18	0.239
I-nature - I-society	1.01	0.312	1.03	0.303
I-me - I-You	1.14	0.255	0.	0.415
I-nature - I-You	1.20	0.231	0.72	0.469
I-society - I-You	0.41	0.684	-0.07	0.948
I-me - Multiple types	-1.26	0.209	-1.03	0.304
I-nature - Multiple types	-0.87	0.387	-1.02	0.308
I-society - Multiple types	-1.52	0.130	-1.58	0.114
I-You - Multiple types	-1.58	0.114	-1.21	0.226
I-me - No exercise	-9.62	<0.001	-7.78	<0.001
I-nature - No exercise	-4.24	<0.001	-4.03	<0.001
I-society - No exercise	-3.15	0.002	-2.77	0.006
I-You - No exercise	-2.66	<0.001	-2.02	0.043
Multiple types - No exercise	-2.04	0.008	-1.30	0.195

Vigor						
	All participants		Female		Male	
	Statistic	P Value	Statistic	P Value	Statistic	P Value
I-me - I-nature	-1.99	0.046	-1.24	<0.001	-1.72	<0.001
I-me - I-society	0.54	0.587	0.44	<0.001	0.39	<0.001
I-nature - I-society	1.41	0.159	0.91	<0.001	1.24	<0.001
I-me - I-You	-1.79	0.074	-1.55	<0.001	-0.94	<0.001
I-nature - I-You	-1.06	0.287	-1.11	<0.001	-0.30	<0.001
I-society - I-You	-1.80	0.072	-1.52	<0.001	-1.02	<0.001
I-me - Multiple types	-3.58	<0.001	-2.42	<0.001	-2.60	<0.001
I-nature - Multiple types	-1.79	0.073	-1.46	<0.001	-0.88	<0.001
I-society - Multiple types	-2.51	0.012	-1.76	<0.001	-1.81	<0.001
I-You - Multiple types	0.12	0.908	0.26	<0.001	-0.14	<0.001

<table>
<tr><td colspan="7">Depression</td></tr>
<tr><td></td><td colspan="3">All participants</td><td colspan="3">Female</td></tr>
<tr><td></td><td colspan="2">Statistic</td><td>p-value</td><td colspan="2">Statistic</td><td>p-value</td></tr>
<tr><td>I-me - I-nature</td><td colspan="2">-0.28</td><td>0.781</td><td colspan="2">0.66</td><td>0.507</td></tr>
<tr><td>I-me - I-society</td><td colspan="2">2.91</td><td>0.004</td><td colspan="2">2.22</td><td>0.027</td></tr>
<tr><td>I-nature - I-society</td><td colspan="2">2.81</td><td>0.005</td><td colspan="2">1.82</td><td>0.068</td></tr>
<tr><td>I-me - I-You</td><td colspan="2">1.09</td><td>0.277</td><td colspan="2">1.17</td><td>0.243</td></tr>
<tr><td>I-nature - I-You</td><td colspan="2">1.13</td><td>0.257</td><td colspan="2">0.92</td><td>0.358</td></tr>
<tr><td>I-society - I-You</td><td colspan="2">-0.77</td><td>0.441</td><td colspan="2">-0.42</td><td>0.677</td></tr>
<tr><td>I-me - Multiple types</td><td colspan="2">-1.37</td><td>0.172</td><td colspan="2">-0.60</td><td>0.548</td></tr>
<tr><td>I-nature - Multiple types</td><td colspan="2">-0.98</td><td>0.326</td><td colspan="2">-0.90</td><td>0.370</td></tr>
<tr><td>I-society - Multiple types</td><td colspan="2">-3.26</td><td>0.001</td><td colspan="2">-2.21</td><td>0.027</td></tr>
<tr><td>I-me - No exercise</td><td>13.95</td><td><0.001</td><td>11.15</td><td><0.001</td><td>8.35</td><td><0.001</td></tr>
<tr><td>I-nature - No exercise</td><td>8.68</td><td><0.001</td><td>6.76</td><td><0.001</td><td>5.55</td><td><0.001</td></tr>
<tr><td>I-society - No exercise</td><td>2.65</td><td>0.008</td><td>1.83</td><td><0.001</td><td>1.75</td><td><0.001</td></tr>
<tr><td>I-You - No exercise</td><td>4.00</td><td><0.001</td><td>3.28</td><td><0.001</td><td>2.29</td><td><0.001</td></tr>
<tr><td>Multiple types - No exercise</td><td>8.44</td><td><0.001</td><td>5.81</td><td><0.001</td><td>5.91</td><td><0.001</td></tr>
</table>

According to Kruskal Wallis' test results, we found that only vigor, depression and confusion change with Findings from this study suggest that I-Society activities may lead to reduced feelings of depression. This study suggests performing I-Me, I-Society, I-Nature, and I-Multiple may lead to a significant increase in feelings of vigor,depression and confusion. All exercise types reported significantly greater feelings of energy compared to days when they did not exercise.

Depression

❖ I-Society was statistically significant compared to I-Me, I-Nature, I-Multiple, and no exercise

❖ I-Society exercise activities may be associated with reduced feelings of depression due to participants socially distancing on most days during the COVID-19 pandemic

Vigor

❖ I-Me (compared to I-You, I-Society, I-Nature, and I-Multiple), I-Society (compared to I-You, I-Nature, and I- Multiple), and I-Nature (compared to I-You and I-Multiple) led to feeling significantly more vigorous

❖ I-Society exercise activities may be associated with increased feelings of vigor due to participants socially distancing on most days during the COVID-19 pandemic

❖ I-Me and I-Nature exercise activities may be associated with a significant increase in vigor due to participants being unable to go to a gym and doing exercise at home (I-Me) or outside (I-Nature) (See figure 4 - figure 14 and)

Confusion

- I-You was statistically significant compared to I-Me, I-Nature, I-Multiple, and no exercise

- This exercise type is associated with a significant decrease in confusion may be due to participants being able to talk to each other while doing this exercise.

Section 3.2: Correlation between explanatory variables

Since the explanatory variables are categorical we used the Cramer's V correlation technique at 95% confidence level to obtain the correlation between Exercise Type, Exercise Intensity and Sex.

Table 8: The Cramer's V

Variables	Exercise Type	Exercise Intensity	Sex
Exercise Type	1	0.5053	0.0057
Exercise Intensity	0.5053	1	0.0459
Sex	0.0057	0.0459	1

It is clear that exercise type has a weak positive correlation with exercise intensity.

Section 3.3: Multiple linear regression

First, we have use multiple linear regression to identify the relationship between overall mood and explanatory variables (Exrecise type, exercise intensity, sex and sleep efficiency).

We calculated mood using depression, vigor and confusion.

Total mood = $\frac{1}{3}\left(\frac{Depression}{Max\ Depression} + \frac{Confusion}{Max\ Confusion} + (1 - \frac{Vigoor}{Max\ Vigor}\right)$

Values close to 1 : Bad mood

Values close to 0 : Good mood

Exercise type, exercise intensity and sex are categorical variables. But, regression analysis requires numerical variables. Therefore, we have recorded our categorical variables using dummy variables technique.

Exercise type: exercise type has 6 levels. Therefore, we transformed that variable into 5 (6-1) variables each with two levels.

Contrast matrix:(No exercise as the reference)

- If rank = I-Me , then the column I-Me would be coded with a 1 and others with 0.
- If rank = I-You , then the column I-You would be coded with a 1 and others with 0.
- If rank = I-Society , then the column I-Society would be coded with a 1 and others with 0.
- If rank = I-Nature, then the column I-Nature would be coded with a 1 and others with 0.
- If rank = I-Nature, then the column I-Nature would be coded with a 1 and others with 0.
- If rank = Multiple types, then the column Multiple types would be coded with a 1 and others with 0.
- If rank = No exercise, then all columns would be coded with 0.

Exercise intensity: exercise type has 4 levels. Therefore, we transformed that variable into 3 (4-1) variables each with two levels.

Contrast matrix: (No intensity as the reference)

- If rank = Low intensity , then the column Low intensity would be coded with a 1 and others with 0.
- If rank = Moderate intensity , then the column Moderate intensity would be coded with a 1 and others with 0.
- If rank = High intensity , then the column High intensity would be coded with a 1 and others with 0.
- If rank = No intensity , then all columns would be coded with 0.

Sex: The regression equation, for predicting an outcome variable (y) on the basis of a predictor variable (x), can be simply written as y = b0 + b1*x. b0 and b1 are the regression beta coefficients, representing the intercept and the slope, respectively.

Suppose that, we wish to investigate differences in moods between males and females.

Based on the gender variable, we can create a new dummy variable that takes the value:

- 1 if a person is male
- 0 if a person is female

and use this variable as a predictor in the regression equation, leading to the following the model:

- b0 + b1 if person is male
- bo if person is female

The coefficients can be interpreted as follow:

1. b0 is the average mood among females,
2. b0 + b1 is the average mood among males,
3. and b1 is the average difference in mood between males and females.

Model:

```
Call:
lm(formula = Vigour1 ~ ., data = exercFVig)

Residuals:
     Min       1Q   Median       3Q      Max 
-10.8000  -2.6548  -0.2944   2.6593  11.7538 

Coefficients:
                                     Estimate Std. Error t value Pr(>|t|)    
(Intercept)                            2.6794     0.6067   4.417 1.05e-05 ***
Exercise.TypeI-nature                  1.0619     0.3806   2.790  0.00531 ** 
Exercise.TypeI-society                -0.1580     0.7178  -0.220  0.82578    
Exercise.TypeI-You                     2.0208     1.0958   1.844  0.06530 .  
Exercise.TypeMultiple types            1.8754     0.5770   3.250  0.00117 ** 
Exercise.TypeNo excersice              2.1287     3.9141   0.544  0.58658    
Exercise.IntensityLow intensity        3.2651     0.3433   9.511  < 2e-16 ***
Exercise.IntensityModerate intensity   0.6012     0.3112   1.932  0.05351 .  
Exercise.IntensityNo intensity        -3.3241     3.9217  -0.848  0.39673    
SexMale                                0.1816     0.1655   1.097  0.27262    
Sleep.Efficiency                       4.8349     0.6465   7.479 1.05e-13 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.912 on 2383 degrees of freedom
Multiple R-squared:  0.1742,	Adjusted R-squared:  0.1707 
F-statistic: 50.27 on 10 and 2383 DF,  p-value: < 2.2e-16
```

Figure 16: Multiple linear regression output

The R^2 value is low suggesting that multiple linear regression is not the best technique to use. This is most likely due to the fact that the dependent variable is categorical.

Next we check the anova summary:

```
> anova(modelVig)
Analysis of Variance Table

Response: Vigour1
                     Df Sum Sq Mean Sq F value    Pr(>F)    
Exercise.Type         5   4699  939.77 61.4217 < 2.2e-16 ***
Exercise.Intensity    3   2012  670.68 43.8346 < 2.2e-16 ***
Sex                   1    125  124.50  8.1372  0.004374 ** 
Sleep.Efficiency      1    856  855.83 55.9354 1.047e-13 ***
Residuals          2383  36461   15.30                      
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 17: Multiple linear regression ANOVA output

Section 3.4: Logistic regression

Since our many explanatory variables are categorical we decided to run a logistic regression model to predict the mood. We built three different logistic regression models for depression, confusion and vigor. We divide the data set into the training set(80% of the data-2270 data points) and the testing set(20% of the data - 565). Fit model to the training set and testing set was used to check the accuracy.

Vigor: First, we fit the logistic regression model to the original data and our model accuracy was not good enough because of too many categories.

	0	1	10	11	12	13	14	15	16	17	18	2	3	4	5	6	7	8	9
1	0.15	0.11	0.05	0.00	0.01	0.03	0.01	0.00	0	0	0	0.07	0.02	0.10	0.19	0.12	0.07	0.04	0.01
2	0.28	0.13	0.03	0.00	0.01	0.02	0.01	0.00	0	0	0	0.06	0.04	0.12	0.16	0.07	0.03	0.03	0.01
3	0.28	0.13	0.03	0.00	0.01	0.02	0.01	0.00	0	0	0	0.06	0.04	0.12	0.16	0.07	0.03	0.03	0.01
4	0.27	0.13	0.03	0.00	0.01	0.02	0.01	0.00	0	0	0	0.06	0.04	0.12	0.16	0.07	0.03	0.03	0.01
5	0.14	0.11	0.05	0.00	0.01	0.03	0.01	0.00	0	0	0	0.07	0.03	0.11	0.19	0.12	0.07	0.05	0.01
6	0.28	0.13	0.03	0.00	0.01	0.02	0.01	0.00	0	0	0	0.06	0.04	0.12	0.16	0.07	0.03	0.03	0.01
7	0.29	0.13	0.03	0.00	0.01	0.02	0.01	0.00	0	0	0	0.06	0.04	0.12	0.16	0.07	0.03	0.03	0.01
8	0.28	0.13	0.03	0.00	0.01	0.02	0.01	0.00	0	0	0	0.06	0.04	0.12	0.16	0.07	0.03	0.03	0.01
9	0.30	0.14	0.02	0.00	0.01	0.02	0.01	0.00	0	0	0	0.06	0.04	0.11	0.15	0.07	0.03	0.02	0.01
11	0.29	0.14	0.02	0.00	0.01	0.02	0.01	0.00	0	0	0	0.06	0.04	0.11	0.16	0.07	0.03	0.03	0.01
12	0.03	0.05	0.06	0.01	0.03	0.08	0.02	0.00	0	0	0	0.10	0.04	0.09	0.20	0.18	0.05	0.06	0.02
14	0.21	0.07	0.05	0.00	0.01	0.08	0.04	0.00	0	0	0	0.03	0.02	0.09	0.16	0.11	0.06	0.05	0.02
15	0.30	0.14	0.02	0.00	0.00	0.02	0.01	0.00	0	0	0	0.06	0.04	0.11	0.15	0.07	0.03	0.02	0.01
16	0.29	0.13	0.02	0.00	0.01	0.02	0.01	0.00	0	0	0	0.06	0.04	0.11	0.16	0.07	0.03	0.03	0.01
17	0.08	0.03	0.08	0.03	0.03	0.06	0.05	0.01	0	0	0	0.02	0.03	0.08	0.12	0.11	0.09	0.10	0.07

Figure 18: Logistic regression output- Probability of categorizing inputs into vigor categories.

Interpretation: The maximum probability(0.19) of the first response is vigor = 5

Then, we categorized depression, confusion and vigor into different categories and checked for accuracy. We couldn't get good accuracy with the small intervals. Therefore, we have checked the model changing width of interval for each mood.

```
      a    b    c    d    e    f
1  0.31 0.30 0.24 0.06 0.09 0.00
2  0.49 0.32 0.13 0.03 0.03 0.00
3  0.48 0.32 0.13 0.03 0.04 0.00
4  0.47 0.33 0.13 0.03 0.04 0.00
5  0.29 0.30 0.25 0.07 0.09 0.01
6  0.48 0.32 0.13 0.03 0.04 0.00
7  0.49 0.32 0.12 0.03 0.03 0.00
8  0.49 0.32 0.13 0.03 0.04 0.00
9  0.51 0.31 0.12 0.03 0.03 0.00
11 0.50 0.32 0.12 0.03 0.03 0.00
12 0.17 0.30 0.26 0.07 0.19 0.01
14 0.30 0.30 0.24 0.06 0.09 0.01
15 0.52 0.31 0.11 0.02 0.03 0.00
16 0.50 0.32 0.12 0.03 0.03 0.00
17 0.11 0.23 0.30 0.18 0.13 0.05
```

Figure 19: Logistic regression output- Probability of categorizing inputs into vigor categories.

```
Coefficients:
         (Intercept) Exercise.TypeI-nature Exercise.TypeI-society Exercise.TypeI-You
low         2.776904            -0.5648108              0.1181346        -16.6377264
moderate    1.235210            -0.3539233              0.4266914         -0.9067452
         Exercise.TypeMultiple types Exercise.TypeNo excersice Exercise.IntensityLow intensity
low                       -0.7489626                -10.259993                       -1.498260
moderate                  -0.5893698                  3.814481                       -1.034264
         Exercise.IntensityModerate intensity Exercise.IntensityNo intensity     SexMale
low                                -0.4016269                      10.946709 -0.08723823
moderate                           -0.3187705                      -3.929381 -0.02412764
         Sleep.Efficiency
low             -2.987368
moderate        -1.315382


Std. Errors:
         (Intercept) Exercise.TypeI-nature Exercise.TypeI-society Exercise.TypeI-You
low        0.4472173             0.2587821              0.4835959       2.903431e-08
moderate   0.4999808             0.2354268              0.4313934       6.684110e-01
         Exercise.TypeMultiple types Exercise.TypeNo excersice Exercise.IntensityLow intensity
low                        0.4366509                0.09127100                       0.2400405
moderate                   0.3772912                0.09300006                       0.2176071
         Exercise.IntensityModerate intensity Exercise.IntensityNo intensity   SexMale
low                                 0.1995095                     0.09127083 0.1051326
moderate                            0.1951180                     0.09300006 0.1119996
         Sleep.Efficiency
low             0.4832866
moderate        0.5439333

Residual Deviance: 4822.751
AIC: 4866.751
```

Figure 20: Logistic regression output- for final logistic regression model for vigor.

```
> p
         (Intercept) Exercise.TypeI-nature Exercise.TypeI-society Exercise.TypeI-You
low      5.322283e-10            0.02906728              0.8070110          0.0000000
moderate 1.349186e-02            0.13275503              0.3226139          0.1749184
         Exercise.TypeMultiple types Exercise.TypeNo excersice Exercise.IntensityLow intensity
low                       0.08630063                         0                    4.328429e-10
moderate                  0.11826243                         0                    2.005231e-06
         Exercise.IntensityModerate intensity Exercise.IntensityNo intensity   SexMale
low                                0.04410712                              0 0.4066561
moderate                           0.10231508                              0 0.8294350
         Sleep.Efficiency
low          6.355221e-10
moderate     1.559413e-02
```

Figure 21: p values for the logistic regression for vigor.

Logistic regression model for vigor:

$ln(\frac{P(vigor = low)}{P(vigor=high)})$ = 2.78 -0.56 (Exercise type = I-Nature) -16.64(Exercise type = I-You)-10.96(Exercise type = No exercise)-1.59(Exercise intensity = low) -0.40(Exercise intensity = moderate) -2.99 Sleep efficiency

$ln(\frac{P(vigor = moderate)}{P(vigor=high)})$ = 1.24 +3.82((Exercise type = No exercise)-1.03(Exercise intensity = low) -1.32 Sleep efficiency

Interpretation:

A unit increase of sleep efficiency is associated with the decrease in the log odds of being in low vigor vs high vigor in the amount of 2.99.

The log odds of being in low vigor vs high vigor will decrease by 0.56 if moving from exercise type= I-Me to Exercise type = I-Nature.

The log odds of being in low vigor vs high vigor will decrease by 16.64 if moving from exercise type= I-Me to Exercise type = I-You.

The log odds of being in low vigor vs high vigor will decrease by 10.96 if moving from exercise type= I-Me to Exercise type = No exercise.

The log odds of being in low vigor vs high vigor will decrease by 1.59 if moving from exercise intensity= high to exercise intensity= low.

The log odds of being in low vigor vs high vigor will decrease by 0.40 if moving from exercise intensity= high to exercise intensity= moderate.

A unit increase of sleep efficiency is associated with the decrease in the log odds of being in moderate vigor vs high vigor in the amount of 1.32.

The log odds of being in moderate vigor vs high vigor will increase by 3.82 if moving from exercise type= I-Me to Exercise type = No exercise.

The log odds of being in moderate vigor vs high vigor will decrease by 1.03 if moving from exercise intensity= high to exercise intensity=low.

Below is the result of predicting for Vigor.

Accuracy: 51.42

	high	low	moderate
high	442	379	12
low	177	774	8
moderate	242	345	15

Depression: We followed the same procedure for depression as vigor.

```
Coefficients:
                                      Values   Std. Err.
(Intercept)                      -2.82396210  0.36651559
Exercise.TypeI-nature             0.22868367  0.20298446
Exercise.TypeI-society            0.82192863  0.43034997
Exercise.TypeI-You                0.47958264  0.61103784
Exercise.TypeMultiple types       0.26761203  0.31297079
Exercise.TypeNo excersice         8.34757957 55.44729886
Exercise.IntensityLow intensity   0.16414138  0.18321126
Exercise.IntensityModerate intensity -0.14677872  0.16503569
Exercise.IntensityNo intensity   -8.39123729 55.44744959
SexMale                          -0.01454546  0.08771045
Sleep.Efficiency                  3.46607243  0.39749574

Residual Deviance: 3161.508
AIC: 3183.508
```

Figure 22: Logistic regression output- for final logistic regression model for depression.

```
> p
                    (Intercept)                  Exercise.TypeI-nature
                   1.310063e-14                           2.599088e-01
          Exercise.TypeI-society                     Exercise.TypeI-You
                   5.614513e-02                           4.325323e-01
      Exercise.TypeMultiple types            Exercise.TypeNo excersice
                   3.925122e-01                           8.803309e-01
   Exercise.IntensityLow intensity Exercise.IntensityModerate intensity
                   3.702992e-01                           3.738012e-01
     Exercise.IntensityNo intensity                            SexMale
                   8.797101e-01                           8.682868e-01
                Sleep.Efficiency
                   0.000000e+00
```

Figure 23: p values for the logistic regression for depression.

Logistic regression model for depression:

$ln(\frac{P(depression = good)}{P(depression=bad)})$ = -2.82 -0.82(Exercise type = I-Society) +3.45 Sleep efficiency

Interpretation:

A unit increase of sleep efficiency is associated with the decrease in the log odds of being in depression =good vs depression =bad in the amount of 3.45.

The log odds of being in depression=good vs depression=bad will decrease by 0.82 if moving from exercise type= I-Me to Exercise type = I-Society.

Below is the result of predicting depression.

Accuracy: 60.99

	bad	good
bad	256	775
good	159	1204

Logistic regression model for Confusion:

Frequency of different Confusion categories

0	1	2	3	4	5	6
54	244	625	664	495	198	114

Updated frequency of different Confusion categories

high (4-6)	low (0-1)	moderate (2-3)
807	298	1289

Coefficients for the logistic regression model:

```
Call:
multinom(formula = Confusion2 ~ . - Confusion1, data = exercFCon)

Coefficients:
         (Intercept) Exercise.TypeI-nature Exercise.TypeI-society Exercise.TypeI-You Exercise.TypeMultiple types Exercise.TypeNo excersice
low        -4.407860           -0.19609940              0.4083861           2.165741                   0.3708932                  9.439582
moderate   -1.278929           -0.03664938              1.0059424           1.218584                  -0.1526060                 17.438170
         Exercise.IntensityLow intensity Exercise.IntensityModerate intensity Exercise.IntensityNo intensity   SexMale Sleep.Efficiency
low                           1.66712623                           -0.3212157                     -10.26498 0.1639313         3.889744
moderate                      0.09364651                           -0.2823332                     -18.16705 0.0894299         2.466718

Std. Errors:
         (Intercept) Exercise.TypeI-nature Exercise.TypeI-society Exercise.TypeI-You Exercise.TypeMultiple types Exercise.TypeNo excersice
low        0.7452635             0.3398301              0.7973953           1.139482                   0.4416336                0.13704872
moderate   0.3465072             0.2276645              0.5565174           1.067350                   0.3796391                0.08511292
         Exercise.IntensityLow intensity Exercise.IntensityModerate intensity Exercise.IntensityNo intensity    SexMale Sleep.Efficiency
low                            0.3030826                            0.2976261                    0.13704826 0.14879943        0.8027647
moderate                       0.2307663                            0.1865868                    0.08511292 0.09573333        0.3718796

Residual Deviance: 4292.154
AIC: 4336.154
```

P - values:

```
> p
         (Intercept) Exercise.TypeI-nature Exercise.TypeI-society Exercise.TypeI-You Exercise.TypeMultiple types Exercise.TypeNo excersice
low      3.328882e-09             0.5639048             0.60854601         0.05734967                   0.4010088                         0
moderate 2.234485e-04             0.8721094             0.07067393         0.25358258                   0.6877013                         0
         Exercise.IntensityLow intensity Exercise.IntensityModerate intensity Exercise.IntensityNo intensity   SexMale Sleep.Efficiency
low                         3.785706e-08                            0.2804722                             0 0.2705952     1.263349e-06
moderate                    6.848847e-01                            0.1302423                             0 0.3502233     3.286904e-11
```

$ln(\frac{P(confusion = low)}{P(confusion = high)})$ = -4.41 +2.17 (Exercise type = I-You) +9.44(Exercise type = No exercise)+1.67(Exercise intensity = low) -10.26(Exercise intensity = No intensity) +3.89 Sleep efficiency

$ln(\frac{P(confusion = moderate)}{P(confusion = high)})$ = -1.28 +1.01((Exercise type = I-Society)+17.74(Exercise type = No exercise) -18.17(Exercise intensity = No intensity)+2.47 Sleep efficiency

Interpretation:

A unit increase of sleep efficiency is associated with the increase in the log odds of being in low confusion vs high confusion in the amount of 3.89.

The log odds of being in low confusion vs high confusion will increase by 2.17 if moving from exercise type= I-Me to Exercise type = I-You.

The log odds of being in low confusion vs high confusion will increase by 9.44 if moving from exercise type= I-Me to Exercise type = no exercise.

The log odds of being in low confusion vs high confusion will increase by 1.67 if moving from exercise intensity= high to exercise intensity= low intensity.

The log odds of being in low confusion vs high confusion will decrease by 10.26 if moving from exercise intensity= high to exercise intensity= No intensity.

A unit increase of sleep efficiency is associated with the increase in the log odds of being in moderate confusion vs high confusion in the amount of 2.47.

The log odds of being in moderate confusion vs high confusion will increase by 1.01 if moving from exercise type= I-Me to Exercise type = I-Society.

The log odds of being in moderate confusion vs high confusion will increase by 17.74 if moving from exercise type= I-Me to Exercise type = No exercise.

The log odds of being in moderate confusion vs high confusion will decrease by 18.17 if moving from exercise intensity= high to exercise intensity= No intensity.

Below is the result of predicting for Confusion.

Accuracy: 56.56

	high	low	moderate
high	144	3	660
low	10	13	275
moderate	87	5	1197

Below is the result of predicting for Confusion but biased category.

Accuracy: 74.9

	high	low	moderate
high	0	2	310
low	0	13	285
moderate	0	4	1780

P- value for Confusion logistic regression:

```
> p <- (1 - pnorm(abs(z), 0, 1)) * 2
> p
                       (Intercept)                   Exercise.TypeI-nature
                      1.310063e-14                            2.599088e-01
             Exercise.TypeI-society                     Exercise.TypeI-You
                      5.614513e-02                            4.325323e-01
         Exercise.TypeMultiple types               Exercise.TypeNo excersice
                      3.925122e-01                            8.803309e-01
      Exercise.IntensityLow intensity Exercise.IntensityModerate intensity
                      3.702992e-01                            3.738012e-01
       Exercise.IntensityNo intensity                              SexMale
                      8.797101e-01                            8.682868e-01
                   Sleep.Efficiency
                      0.000000e+00
```

P- value for Depression logistic regression:

```
> p <- (1 - pnorm(abs(z), 0, 1)) * 2
> p
                     (Intercept)              Exercise.TypeI-nature
                    1.310063e-14                       2.599088e-01
           Exercise.TypeI-society                 Exercise.TypeI-You
                    5.614513e-02                       4.325323e-01
       Exercise.TypeMultiple types            Exercise.TypeNo excersice
                    3.925122e-01                       8.803309e-01
  Exercise.IntensityLow intensity Exercise.IntensityModerate intensity
                    3.702992e-01                       3.738012e-01
    Exercise.IntensityNo intensity                              SexMale
                    8.797101e-01                       8.682868e-01
                 Sleep.Efficiency
                    0.000000e+00
```

4. Conclusions

These findings support the need for physical activity, as they suggest that performing physical activity increases feelings of energy. Specifically, individual physical activity resulted in the greatest increase in feelings of energy. These findings also suggest that performing physical activity with other people may lead to reduced feelings of depression. Lastly, findings suggest that performing physical activity individually, in nature, with other people, or doing a mix of dimensions may lead to increased feelings of vigor. Some limitations included that we were not able to collect mood data before the lockdown, therefore we do not have information on how participants' moods changed due to the lockdown from COVID-19. Also, mood was measured after each day instead of directly after exercising, therefore analysis may not have been representative of participants' mood right after their physical activity. Future research should focus on collecting mood right after completing physical activity for each exercise dimension.

2. References

1. https://www.moresteam.com/whitepapers/download/dummy-variables.pdf
2. https://socialsciences.mcmaster.ca/jfox/Courses/SPIDA/dummy-regression-notes.pdf
3. http://www.sthda.com/english/articles/40-regression-analysis/163-regression-with-categorical-variables-dummy-coding-essentials-in-r/
4. http://www.ub.edu/rfa/R/regression_with_categorical_dependent_variables.html
5. https://www.frontiersin.org/articles/10.3389/fpsyg.2020.590172/full
6. https://www.clarkson.edu/news/clarkson-university-student-research-finds-women-exercising-others-feel-less-depressed-rest
7. https://stats.oarc.ucla.edu/r/dae/multinomial-logistic-regression/
8. https://www.r-bloggers.com/2020/05/multinomial-logistic-regression-with-r/

Codes in R

```
library(plyr)
library(ggplot2)
library(dplyr)
data1 <- read.csv("Covid data_14_sex 1.csv")
data2 <- read.csv("Covid data_14_sex 2.csv")
comb <- bind_rows(data1,data2)
comb$Sex[comb$Sex=="1"]="Male"
comb$Sex[comb$Sex=="2"]="Female"

comb$Confusion = comb$Confusion + 4   # removing negative from Confusion

par(mfrow = c(2,2))
boxplot(comb$Confusion, ylab = "Confusion")
boxplot(comb$Depression, ylab = "Depression")
boxplot(comb$Vigour, ylab = "Vigour")

# removing outlier depression-wise
summary(comb$Depression)
low = 0 - 1.5*(3-0)
upp = 3 + 1.5*(3-0)
comb_clean <- subset(comb, comb$Depression<upp & comb$Depression>low)

#new boxplot
par(mfrow = c(2,2))
boxplot(comb_clean$Confusion, ylab = "Confusion")
boxplot(comb_clean$Depression, ylab = "Depression")
boxplot(comb_clean$Vigour, ylab = "Vigour")

# removing outlier confusion-wise
summary(comb$Confusion)
upp = 4 + 1.5*(4-2)
comb_clean_c <- subset(comb_clean, comb_clean$Confusion<upp &
comb_clean$Confusion>low)

#new boxplot
par(mfrow = c(2,2))
boxplot(comb_clean_c$Confusion, ylab = "Confusion")
boxplot(comb_clean_c$Depression, ylab = "Depression")
boxplot(comb_clean_c$Vigour, ylab = "Vigour")

# removing outlier confusion-wise
summary(comb$Vigour)
```

```
low = 4 - 1.5*(10-4)
upp = 10 + 1.5*(10-4)
comb_clean_v <- subset(comb_clean_c, comb_clean_c$Vigour<upp &
comb_clean_c$Vigour>low)

#new boxplot
par(mfrow = c(2,2))
boxplot(comb_clean_v$Confusion, ylab = "Confusion")
boxplot(comb_clean_v$Depression, ylab = "Depression")
boxplot(comb_clean_v$Vigour, ylab = "Vigour")

# all plot in one
par(mfrow = c(3,3))
boxplot(comb_clean$Confusion, ylab = "Confusion")
boxplot(comb_clean$Depression, ylab = "Depression")
boxplot(comb_clean$Vigour, ylab = "Vigour")
boxplot(comb_clean_c$Confusion, ylab = "Confusion")
boxplot(comb_clean_c$Depression, ylab = "Depression")
boxplot(comb_clean_c$Vigour, ylab = "Vigour")
boxplot(comb_clean_v$Confusion, ylab = "Confusion")
boxplot(comb_clean_v$Depression, ylab = "Depression")
boxplot(comb_clean_v$Vigour, ylab = "Vigour")

#  defining new  variables
exercise = comb_clean_v
plot(exercise$Confusion, exercise$Sleep.Efficiency)
plot(exercise$Depression, exercise$Sleep.Efficiency)
plot(exercise$Vigour/max(exercise$Vigour), exercise$Sleep.Efficiency)
plot(exercise$Sex1, exercise$Sleep.Efficiency)
exercise$Sex1[exercise$Sex=="Male"] = 1
exercise$Sex1[exercise$Sex=="Female"] = 0
cor.test(exercise$Sex1, exercise$Sleep.Efficiency)

# factor confusion, depression and Vigour
exercF = exercise
exercF$Confusion1[exercF$Confusion==0] = "0"
exercF$Confusion1[exercF$Confusion==1] = "1"
exercF$Confusion1[exercF$Confusion==2] = "2"
exercF$Confusion1[exercF$Confusion==3] = "3"
exercF$Confusion1[exercF$Confusion==4] = "4"
exercF$Confusion1[exercF$Confusion==5] = "5"
exercF$Confusion1[exercF$Confusion==6] = "6"
```

```
for (var in 0:18)
{
  exercF$Vigour1[exercF$Vigour==var] = as.character(var)
}

for (var in 0:7)
{
  exercF$Depression1[exercF$Depression==var] = as.character(var)
}

exercF$Exercise.Type[exercF$Exercise.Type==0]="No excersice"
exercF$Exercise.Type[exercF$Exercise.Type==1]="I-me"
exercF$Exercise.Type[exercF$Exercise.Type==2]="I-You"
exercF$Exercise.Type[exercF$Exercise.Type==3]="I-society"
exercF$Exercise.Type[exercF$Exercise.Type==4]="I-nature"
exercF$Exercise.Type[exercF$Exercise.Type==5]="Multiple types"

exercF$Exercise.Intensity[exercF$Exercise.Intensity==0]="No intensity"
exercF$Exercise.Intensity[exercF$Exercise.Intensity==1]="Low intensity"
exercF$Exercise.Intensity[exercF$Exercise.Intensity==2]="Moderate intensity"
exercF$Exercise.Intensity[exercF$Exercise.Intensity==3]="High intensity"

# define new variable for Vigour, Depression, Confusion
exercFVig  = exercF[c("Vigour1", "Exercise.Type", "Exercise.Intensity", "Sex",
"Sleep.Efficiency")]
exercFDep  = exercF[c("Depression1", "Exercise.Type", "Exercise.Intensity", "Sex",
"Sleep.Efficiency")]
exercFCon  = exercF[c("Confusion1", "Exercise.Type", "Exercise.Intensity", "Sex",
"Sleep.Efficiency")]

# removing 5 missing values
sum(is.na(exercFVig$Exercise.Intensity))
exercFVig = subset(exercFVig, is.na(exercFVig$Exercise.Intensity)==FALSE)
exercFDep = subset(exercFDep, is.na(exercFDep$Exercise.Intensity)==FALSE)
exercFCon = subset(exercFCon, is.na(exercFCon$Exercise.Intensity)==FALSE)

# Table for descriptive summary
with(exercFVig, table(Vigour1, Exercise.Type))
with(exercFVig, table(Vigour1, Exercise.Intensity))
with(exercFVig, table(Vigour1, Sex))
with(exercFVig, do.call(rbind, tapply(Sleep.Efficiency, Vigour1, function(x) c(Mean = mean(x),
StandardDeviation = sd(x)))))
```

```
# Logistic model for Vigour ~ Sex, Exercise type, Exercise internsity, Sleep efficiency.

# Split vigour into categories to improve prediction accuracy
unique(exercFVig$Vigour1)
table(exercFVig$Vigour1)

for (var in 0:18)
{
  if (var >= 0 & var <=5){
    exercFVig$Vigour2[exercFVig$Vigour1==as.character(var)] = "low"
  } else if (var >=6 & var <=8){
    exercFVig$Vigour2[exercFVig$Vigour1==as.character(var)] = "moderate"
  } else if (var >=9 & var <=18){
    exercFVig$Vigour2[exercFVig$Vigour1==as.character(var)] = "high"
  } else
    print("Zero")
}

# Model
logReg = multinom(Vigour2 ~. -Vigour1, data = exercFVig)
summary(logReg)
exp(coef(logReg))
head(round(fitted(logReg), 2), 30)

# test
test = exercFVig
test$result = predict(logReg, newdata = test, "class")
tab = table(test$Vigour2, test$result)
Tab

# Calculating accuracy - sum of diagonal elements divided by total obs
round((sum(diag(tab))/sum(tab))*100,2)

# Obtaining p - value
z <- summary(logReg)$coefficients/summary(logReg)$standard.errors
z
p <- (1 - pnorm(abs(z), 0, 1)) * 2
p
```

```
# Logistic model for Depression ~ Sex, Exercise type, Exercise internsity, Sleep efficiency.

# Split Depression into categories to improve prediction accuracy
table(exercFDep$Depression1)

for (var in 0:7)
{
  if (var == 0){
    exercFDep$Depression2[exercFDep$Depression1==as.character(var)] = "good"
  } else if (var >=1 & var <=7){
    exercFDep$Depression2[exercFDep$Depression1==as.character(var)] = "bad"
  } else
    print("Zero")
}


# Model
logReg = multinom(Depression2 ~. -Depression1, data = exercFDep)
summary(logReg)
exp(coef(logReg))
head(round(fitted(logReg),3), 30)  # Visual fitted model

# test
test = exercFDep
test$result = predict(logReg, newdata = test, "class")
tab = table(test$Depression2, test$result)
tab


# Calculating accuracy - sum of diagonal elements divided by total obs
round((sum(diag(tab))/sum(tab))*100,2)


# Obtaining p - value
z <- summary(logReg)$coefficients/summary(logReg)$standard.errors
z
p <- (1 - pnorm(abs(z), 0, 1)) * 2
p
```

```
# Logistic model for Confusion ~ Sex, Exercise type, Exercise internsity, Sleep efficiency.

# Split Confusion into categories to improve prediction accuracy
table(exercFCon$Confusion1)

for (var in 0:6)
{
  if (var >= 0 & var <=1){
    exercFCon$Confusion2[exercFCon$Confusion1==as.character(var)] = "low"
  } else if (var >=2 & var <=3){
    exercFCon$Confusion2[exercFCon$Confusion1==as.character(var)] = "moderate"
  } else if (var >=4 & var <=6){
    exercFCon$Confusion2[exercFCon$Confusion1==as.character(var)] = "high"
  } else
    print("Zero")
}


# Model
logReg = multinom(Confusion2 ~. -Confusion1 -Sex, data = exercFCon)
summary(logReg)
exp(coef(logReg))
head(round(fitted(logReg), 2), 30)      # Visual fitted model


# test
test = exercFCon
test$result = predict(logReg, newdata = test, "class")
tab = table(test$Confusion2, test$result)
tab

# Calculating accuracy - sum of diagonal elements divided by total obs
round((sum(diag(tab))/sum(tab))*100,2)

# Obtaining p - value
z <- summary(logReg)$coefficients/summary(logReg)$standard.errors
z
p <- (1 - pnorm(abs(z), 0, 1)) * 2
p
```

https://www.moresteam.com/whitepapers/download/dummy-variables.pdf

http://www.sthda.com/english/articles/40-regression-analysis/163-regression-with-categorical-variables-dummy-coding-essentials-in-r/

https://rpubs.com/julianhatwell/annr

http://uc-r.github.io/ann_regression

http://uc-r.github.io/ann_regression

https://www.youtube.com/watch?v=xOzh6PMk21I