Cairo University

Faculty of Economics and Political Science

English Section

Statistics Department

4th Year

MULTIVARIATE ANALYSIS PROJECT

"Understanding the Impact of Character Strengths on Mental Health and Self-Efficacy during the Covid-19 Pandemic"

Submitted by:

Saif Essam Abdelmaboud (5200356)

Yehia Mohamed Quolquela (5200901)

Mohamed Aly Abdelrahman (5201010)

Under The Supervision of:

Dr. Sara Osama

TA: Sama Alaeddin

Table of Contents

I. Introduction:

The Covid-19 pandemic, an unprecedented global crisis, brought about profound challenges to individuals worldwide, compelling them to adapt rapidly to a new reality of prolonged lockdowns and social isolation. Beyond the immediate health threats posed by the virus, the pandemic has been associated with a myriad of psychological consequences, prompting researchers to explore the protective factors that may mitigate the impact on mental health and well-being. This study delves into the complex relationship between Covid-19 and character strengths, with a focus on understanding how specific character traits contribute to sustaining mental health and self-efficacy during lockdown periods.

The dataset under examination comprises responses from 850 participants gathered through an online survey conducted one month after the initiation of lockdown measures. Drawing inspiration from a related study titled "“Andrà tutto bene”: Associations Between Character Strengths, Psychological Distress and Self-efficacy During Covid-19 Lockdown," which explored similar themes with a sample of 944 Italian respondents, our investigation employs advanced statistical techniques in multivariate analysis to shed light on the protective role of character strengths.

- **Objectives of the Study:** This academic report aims to:
  a. Conduct a comprehensive multivariate analysis using factor analysis, cluster analysis, and multivariate regression to unravel the underlying structure of character strengths and their associations with psychological distress, general mental health, and Covid-19-related self-efficacy.
  b. Evaluate and compare our findings with the results presented in the referenced study, providing insights into the consistency and generalizability of patterns observed in different samples.

II. Factor Analysis:

Factor analysis is a powerful statistical technique employed to uncover the latent structure inherent in a set of observed variables. It enables researchers to identify underlying factors that explain the patterns of correlations among variables, providing valuable insights into the complex interrelationships within a dataset. In this section, we present the outcomes of the factor analysis conducted on the dataset, focusing on character strengths and their potential impact on mental health and self-efficacy during the COVID-19 lockdown. This will be done through three main steps:

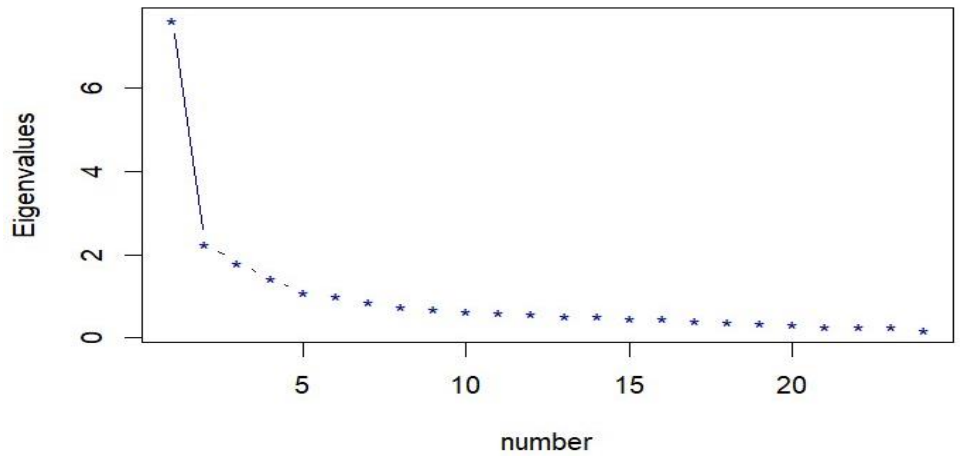
1. Determining Number of Factors:

Factor analysis commenced with the crucial task of ascertaining the appropriate number of factors to retain.

- **Scree Plot:**

The scree plot, a widely recognized graphical tool in factor analysis, was employed to inspect the eigenvalues and guide the decision-making process.

Figure(1): Scree Plot of Eigenvalues



- From Figure (1): The scree plot revealed a clear inflection point where the number of factors starts to level off at the fifth factor, suggesting that retaining **four factors** provides an optimal balance between explaining variance and avoiding overfitting.

- **Parallel Analysis:**

Furthermore, to bolster our decision, we employed the parallel analysis method.

- This technique involves generating a randomized dataset with the same size and structure as the actual dataset and conducting factor analysis on both datasets to obtain eigenvalues for each factor. By repeating this process multiple times, a distribution of eigenvalues is created for each factor. The key insight lies in comparing the eigenvalues from the actual dataset with the corresponding distribution from the randomized dataset. Factors with eigenvalues higher than those expected by chance are considered meaningful and retained in the final model. parallel analysis suggests that the number of factors = 4.
- Using scree plot & parallel analysis the results indicated 4 factors to retain
- The four factors explain 54.8% of the variation in the 24 variables (see table (1))

Table (1): Importance of components:

	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5	Comp.6	Comp.7	Comp.8
Standard deviation	2.760602	1.506167	1.346207	1.198377	1.04925	1.008048	0.932533	0.865853
Proportion of Variance	0.317912	0.094634	0.0756	0.059908	0.045926	0.04239	0.036277	0.031274
Cumulative Proportion	0.317912	0.412546	0.488147	0.548055	0.593981	0.636371	0.672647	0.703922

2. Extracting Factor Loadings:

The initial factor analysis was conducted using the Principal Factor Method, yielding a solution with four factors. The factor loadings before rotation are presented below in table (2):

Table (2): Factor loadings using no rotation				
Variables	V1	V2	V3	V4
Appreciation of beauty	0.5475241	0.04088	-0.06147	0.107018
Bravery	0.499023	-0.27611	0.18424	0.205392
Creativity	0.5552058	-0.28378	0.260793	0.230504
Curiosity	0.6935948	-0.38216	0.146238	0.017638
Fairness	0.5383562	0.345945	-0.33471	0.237107
Forgiveness	0.3475093	0.142839	-0.29125	-0.18108
Gratitude	0.6960424	-0.06673	-0.1526	-0.28081
Honesty	0.5722708	0.16046	0.049931	-0.00017
Hope	0.7256332	-0.27664	0.027199	-0.28217
Humility	0.254498	0.461734	-0.24124	-0.12142
Humor	0.4171633	-0.24132	-0.06711	0.271745
Judgment	0.415185	0.409172	0.510564	0.103043
Kindness	0.6298025	0.156434	-0.3555	0.200602
Leadership	0.5937686	0.140039	-0.1597	0.260526
Love	0.5287307	-0.0262	-0.105	-0.07133
Love of learning	0.4260135	-0.09788	0.15591	0.058654
Perseverance	0.5785207	-0.03295	0.161597	-0.22348
Perspective	0.4890397	0.222994	0.360948	0.141226
Prudence	0.3216354	0.681229	0.375696	-0.18127
Self-regulation	0.4537089	0.159518	0.183461	-0.28088
Social intelligence	0.6335532	0.012407	0.027163	0.206482
Spirituality	0.5369907	-0.11987	-0.15468	-0.28236
Teamwork	0.5077107	0.26009	-0.3254	0.058892
Zest	0.7683713	-0.40652	-0.02992	-0.15929

- The principal factor method operates on the reduced correlation matrix, corresponding to the communalities.
- In Table (2): it becomes apparent that loadings exhibit significant overlap. Notably, the loadings for the first factor are markedly higher compared to those for the remaining factors, while the loadings for the other factors are all low.
- This makes it difficult to interpret, hence we will have to apply a transformation to help make interpretations easier. Here necessitates the application of factor rotation to enhance the interpretability of the factor structure.

3. Factor Correlations:

Table (3): correlation matrix between the four factors				
	MR1	**MR2**	**MR3**	**MR4**
MR1	1	-0.00943	0.002964	-0.01276
MR2	-0.00943	1	0.033614	0.013701
MR3	0.002964	0.033614	1	-0.02134
MR4	-0.01276	0.013701	-0.02134	1

- From Table (3), the correlation coefficients are close to zero, indicating that the factors (Openness, Transcendence, Interpersonal, and Restraint) are uncorrelated. This lack of correlation alleviates concerns related to multicollinearity, ensuring the independence of the factors in our analysis.
- The correlation matrix indicates that the factors are very weakly correlated. Therefore, we will use an **orthogonal rotation** to try to interpret the factors.

4. Factors Rotation:

Varimax is an orthogonal rotation method that emphasizes maximizing the variance of the factor loadings on each factor while maintaining orthogonality among the factors.

Using varimax transformation (orthogonal transformation) seems to have made interpretation easier. Variables whose loadings greater than **0.5** were chosen to determine the names of the factors.

Table (4): Factor loadings using Varimax rotation					
Variables	Openness	Transcendence	Interpersonal	Restraint	Communalities
Appreciation of beauty	0.33403	0.22263	0.36424	0.15125	0.3166854
Bravery	0.59687	0.18396	0.04956	0.09399	0.40139
Creativity	0.66872	0.19193	0.03024	0.15808	0.5099278
Curiosity	0.64504	0.47155	0.06951	0.07455	0.6488204
Fairness	0.16717	0.05781	0.72438	0.14743	0.5777545
Forgiveness	-0.06197	0.32365	0.38727	0.01467	0.2587823
Gratitude	0.22769	0.64354	0.34111	0.09340	0.5910669
Honesty	0.25603	0.27727	0.33233	0.32073	0.3557346
Hope	0.40374	0.70284	0.13434	0.09163	0.6834284
Humility	-0.23292	0.13375	0.47945	0.22113	0.3509084
Humor	0.49926	0.09860	0.20954	-0.08788	0.3106104
Judgment	0.22563	0.00080	0.09865	0.74192	0.6110947
Kindness	0.29126	0.19124	0.68157	0.04253	0.5877432
Leadership	0.36861	0.11197	0.54241	0.15145	0.4655494
Love	0.24428	0.37164	0.30240	0.08437	0.2963555
Love of learning	0.37078	0.20979	0.07882	0.17639	0.2188161
Perseverance	0.26438	0.49608	0.10665	0.29062	0.4118309
Perspective	0.34431	0.07685	0.15402	0.53938	0.4391147
Prudence	-0.16951	0.10491	0.19644	0.81437	0.7415289
Self-regulation	0.06650	0.41529	0.10797	0.39409	0.3438476
Social intelligence	0.47952	0.19476	0.36194	0.21459	0.4449163
Spirituality	0.16339	0.56901	0.23635	0.00701	0.4063823
Teamwork	0.09540	0.20050	0.61257	0.10112	0.4347735
Zest	0.54975	0.67252	0.16346	-0.02657	0.7819248

- From the output in Table 4, it is evident that the factors are ordered in terms of importance. The fourth factor emerges as the most influential, while the third factor takes a comparatively lesser role.
- Notably, Bravery, Creativity, and Curiosity exhibit high loadings on the fourth factor, suggesting that this factor can be identified as the **"Openness"** factor.
- Similarly, Gratitude, Hope, Spirituality, and Zest demonstrate high loadings on the second factor, indicating that this factor represents **"Transcendence"**.
- The **"Interpersonal"** factor is characterized by high loadings from Fairness, Kindness, Leadership and Teamwork.
- Lastly, the third factor, denoted as the **"Restraint"** factor, is shaped by heavy loadings from Judgment, Perspective, and Prudence.
- We will use the new four factor variables as a replacement for the 24 variables.

✓ Communalities were calculated to gauge the proportion of variance in each variable explained by the factors.

✓ Communalities are how much variation of each variable that is due to the factors. For the example, 61.1% of variation in variable judgment is due to the four factors.

✓ Higher communalities suggest that a larger portion of the variance in these variables can be attributed to the common factors.

5. Comparison with the study:

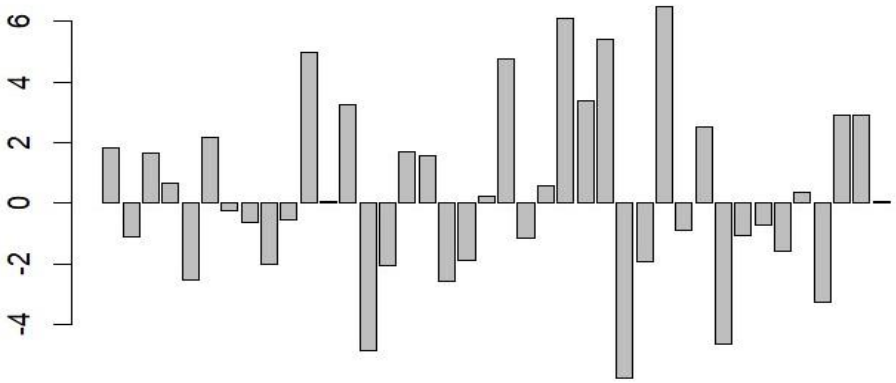
The study used the same method for determining the number of factors. For estimating the loadings, the study used an oblique transformation for the factors assuming that they are correlated. They used promax transformation. The results of the study were almost the same as our findings, as we obtained the same four factors.

III. Cluster Analysis:

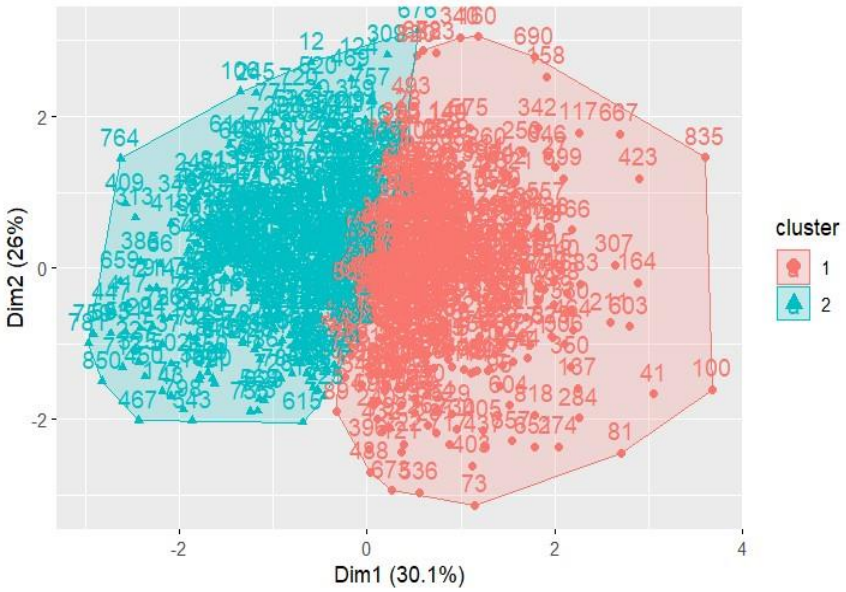
1. Cluster Evaluation:

- We will use the four factors obtained from cluster analysis namely transcendence, interpersonal, openness and restraint to classify our sample. From figure (2) we can see that we have 2 clusters (positive & negative) or 4 clusters (small positive, large positive, small negative & large negative). we used k-means clustering to perform cluster analysis. Both 2 and 4 clusters were considered but the results of 2 clusters were more appropriate for our sample.

Figure(2): histogram of the first principle component



Figure(3): Clusterplot for 2 Clusters



- The observations were classified into two clusters where the first cluster had 441 observations (51.88%) & the second cluster had 409 observations (48.12%).

Table (5): Number of Observations in each Cluster	
Clusters	Frequency
Cluster 1	441
Cluster 2	409

Table (6): Means of Clusters		
Factors	**1**	**2**
Openness	-0.54789	0.590757
Transcendence	-0.4492	0.484344
Interpersonal	-0.24117	0.260042
Restraint	-0.03957	0.042668

- From Table (6) we can see that cluster 1 takes low values for all the factors while cluster 2 takes high values for all the factors, therefore we will name cluster 1 “low strength factors” & cluster 2 “high strength factors”.

2. Cluster Validation (using Discriminant Analysis):

To apply Discriminant analysis, we need to check the following assumptions first,

1) Multivariate Normality Assumption:

Since we have a large sample (850 observation), the strength factors follow multivariate normal distribution using Central Limit Theorem (CLT)

2) Equality of Covariances Assumption:

Box's M Test

Chi-Squared Value = 65.11285, df = 10 and p-value: 3.86e-10

- To test the Equality of Covariances Assumption, we will use the Box M test. From the test, p-value < 0.05.
- This means that the covariances of the four factors are not equal.

3) **Equality of Means Assumption:**

- Using Hoteling T-test the p-value obtained was 6.66147e-192 which is less than 0.05. Therefore, the means of the four strength factors are not equal.
- From the aforementioned results, the quadratic discriminant analysis (does not require the equality of covariances assumption)

Table (7): validation of cluster analysis			
	Actual		
Predicted	**1**	**2**	**Total**
1	371	20	391
2	38	421	459
Total	409	441	850

- From Table (7) the percentage of correct classification for the first cluster is 90.7%, the percentage of correct classification for the second cluster is 95.46% and the overall percentage of correct classification is 93.18%. From the aforementioned results of the discriminant analysis, we can be confident of the results of cluster analysis.

IV. Multivariate Regression:

We will perform multivariate regression for the three dependent measures DASS21 (Depression Anxiety and Stress Scale) GHQ12 (General Health Questionnaire) SEC (Self-efficacy for Covid-19). We choose the explanatory variables as the four factors, Age, Gender, Work, Day, Student & Sons (the same as the study).

1) Depression Anxiety and Stress Scale (DASS_21):

Coefficients:	Response DASS_21			
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	32.62688	3.90465	8.356	2.67e-16 ***
Age	-0.05761	0.02928	-1.968	0.049434 *
factor(Gender)2	-2.78096	0.76389	-3.641	0.000289 ***
Work	0.78532	0.21938	3.58	0.000364 ***
Day	- 0.05743	0.05407	- 1.062	0.288518
factor(Student)2	0.82596	0.95975	0.861	0.389705
Sons	-1.53658	0.73923	-2.079	0.037955 *
Openness	0.09358	0.03215	2.911	0.003702 **
Restraint	0.05699	0.04871	1.17	0.242321
Transcendence	-0.27718	0.027	10.265	< 2e-16 ***
Interpersonal	0.03442	0.03226	1.067	0.286235
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1				
Residual standard error: 9.579 on 840 degrees of freedom				
Multiple R-squared: 0.2025, Adjusted R-squared: 0.193				
F-statistic: 23.55 on 9 and 840 DF, p-value: < 2.2e-16				

- The explanatory variables explain 19.3% of the variability in DASS21 (Depression Anxiety and Stress Scale). The model is statistically significant since its p-value < 0.05. the variables age, gender(male), Work, Sons, Openness & Transcendence have a significant effect on DASS21(their p-values < 0.05).

2) General Health Questionnaire (GHQ_12):

Coefficients:	Response GHQ_12			
	Estimate	**Std. Error**	**t value**	**Pr(>|t|)**
(Intercept)	23.0999	1.86371	12.395	< 2e-16 ***
Age	0.02499	0.01398	1.788	0.0741 .
factor(Gender)2	-0.92474	0.36461	-2.536	0.0114 *
Work	0.19292	0.10471	1.842	0.0658 .
Day	- 0.02124	0.02582	- 0.823	0.4109
factor(Student)2	0.7613	0.4581	1.662	0.0969 .
Sons	0.08528	0.35284	0.242	0.8091
Openness	-0.01593	0.01535	-1.038	0.2995
Restraint	0.05567	0.02325	2.395	0.0169 *
Transcendence	-0.09931	0.01289	-7.705	3.67e-14 ***
Interpersonal	0.02939	0.0154	1.909	0.0566 .
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 Residual standard error: 4.572 on 840 degrees of freedom Multiple R-squared: 0.1203, Adjusted R-squared: 0.1098 F-statistic: 12.68 on 9 and 840 DF, p-value: < 2.2e-16				

- The explanatory variables explain 10.98% of the variability in GHQ_12 (General Health Questionnaire). The model is statistically significant since its p-value < 0.05. the variables gender(male), restraint & Transcendence have a significant effect on GHQ_12 (their p-values < 0.05).

3) Self-efficacy for Covid-19 (SEC):

Coefficients:	SEC			
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3.03453	1.423082	2.132	0.033266 *
Age	0.014	0.010671	1.312	0.189898
factor(Gender)2	0.932714	0.278408	3.35	0.000844 ***
Work	-0.15986	0.079954	-1.999	0.045887 *
Day	0.036627	0.019681	1.861	0.063082
factor(Student)2	-0.222493	0.34979	-0.636	0.5249
Sons	-0.030402	0.269418	-0.113	0.910183
Openness	0.009293	0.011717	0.793	0.427918
Restraint	-0.005755	0.017753	-0.324	0.745891
Transcendence	0.1086	0.009841	11.035	< 2e-16 ***
Interpersonal	-0.020174	0.011756	-1.716	0.086506 .
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 Residual standard error: 3.491 on 840 degrees of freedom Multiple R-squared: 0.2455, Adjusted R-squared: 0.2365 F-statistic: 29.85 on 9 and 840 DF, p-value: < 2.2e-16				

- The explanatory variables explain 23.65% of the variability in SEC (Self-efficacy for Covid-19). The model is statistically significant since its p-value < 0.05. the variables gender(male), Work & Transcendence have a significant effect on SEC (their p-values < 0.05)

➢ **Comparison with the study:**

The R^2 values obtained from our sample is very similar to the study. For DASS21, the study obtained an R^2 value of 0.22 vs our sample .2 which is greater than our sample by 0.02. For GHQ_12, the study obtained an R^2 value of 0.13 vs our sample .12 which is greater than our sample by 0.01. For SEC, the study obtained an R^2 value of 0.26 vs our sample .246 which is greater than our sample by 0.014.

V. Conclusion:

Our thorough multivariate analysis underscores the complex relationship between character strengths and mental health outcomes during the Covid-19 pandemic. Our investigation, employing factor analysis, cluster analysis, and multivariate regression, not only provides a nuanced understanding of the protective role of character strengths but also establishes parallels with a related study, reaffirming the robustness and generalizability of our findings.

The scree plot and parallel analysis converged on the identification of four latent factors representing character strengths. These factors, namely Openness, Transcendence, Interpersonal, and Restraint, collectively explained 54.8% of the variance in the 24-character strength variables. Varimax rotation facilitated a more interpretable factor structure, guiding the naming of factors based on high-loading variables.

K-means clustering based on the identified factors revealed two distinctive clusters - one characterized by low strength factors and the other by high strength factors. The cluster validation, employing discriminant analysis, demonstrated high accuracy in cluster classification, affirming the robustness of our cluster analysis results.

The multivariate regression models shed light on the influence of character strengths on psychological distress (DASS21), general mental health (GHQ12), and self-efficacy for Covid-19 (SEC). Notably, Transcendence emerged as a consistent predictor across all three models, signifying its pivotal role in promoting mental well-being and self-efficacy during challenging times.

Our examination of R^2 values in comparison to a related study revealed remarkable consistency, highlighting the reliability and generalizability of our models. Despite slight variations, the alignment of our findings with the study underscores the robustness of character strength associations across different samples.

As society grapples with ongoing challenges, our findings provide a foundation for informed interventions and contribute to the evolving landscape of positive psychology.