Cairo University

Cairo University
Faculty Of Economics and Political Science

COVID-19 Analytics: Multivariate Approach

Presented By:

Name	Code
Shahd Hesham ElSayed	5200384
Mariam Amir Hassan	5200948
Nada Ahmed Amr	5200695

Under the supervision of:

Dr. Sara Osama

Table of Contents

Introduction

The COVID-19 pandemic has not only presented a global health crisis but has also brought about significant challenges to mental well-being due to prolonged periods of lockdown and isolation.

A study was conducted to uncover the protective influence of character strengths on mental health and self-efficacy during lockdowns. Drawing data from 944 Italian respondents one month into lockdown, the research employs principal component analysis to identify key character strengths. The findings reveal four major strengths - transcendence, interpersonal, openness, and restraint.

In this report, we aim to build on this study. Beginning with a random sample of 850 participants, we will conduct factor analysis to compare results with the initial study. Additionally, cluster analysis will unveil patterns among individuals with different strengths, while multivariate regression will delve into the relationships between character strengths and mental health outcomes, offering a nuanced understanding of their interplay during the challenges posed by the pandemic.

Noting that the previous study included only **23** explanatory variables and **3** response variables namely, DASS21 (Depression Anxiety and Stress Scale), GHQ12 (General Health Questionnaire), and SEC (Self-efficacy for Covid-19).

Data Preparation

In the data preparation step, it was detected that there exists only 1 missing value that was removed before starting the analysis.

Factor Analysis

Factor analysis is a statistical technique primarily designed to unveil the latent structure within a dataset, serving the dual purpose of revealing underlying patterns and acting as a data reduction method. In this report, we will adhere to the three main steps inherent to factor analysis to gain deeper insights into the structure of the data. These steps include determining the number of factors to retain, estimating the factor loadings, and factor rotation. Factor analysis mainly works with **standardized data**, therefore, we standardized the original variables.

I. Determining the number of factors to retain:

Considering the scree plot as the best method to determine the number of factors that would be retained in the model, it was shown that we should include only four factors.

To further validate this conclusion, we employed the parallel analysis method. This technique involves comparing the observed eigenvalues with eigenvalues obtained from a randomly generated dataset of the same size. Factors with eigenvalues higher than the corresponding random eigenvalues are retained. Notably, the results of the parallel analysis were the same as those obtained from the scree plot, affirming the consistency of our decision to retain **four** factors in the factor analysis model.

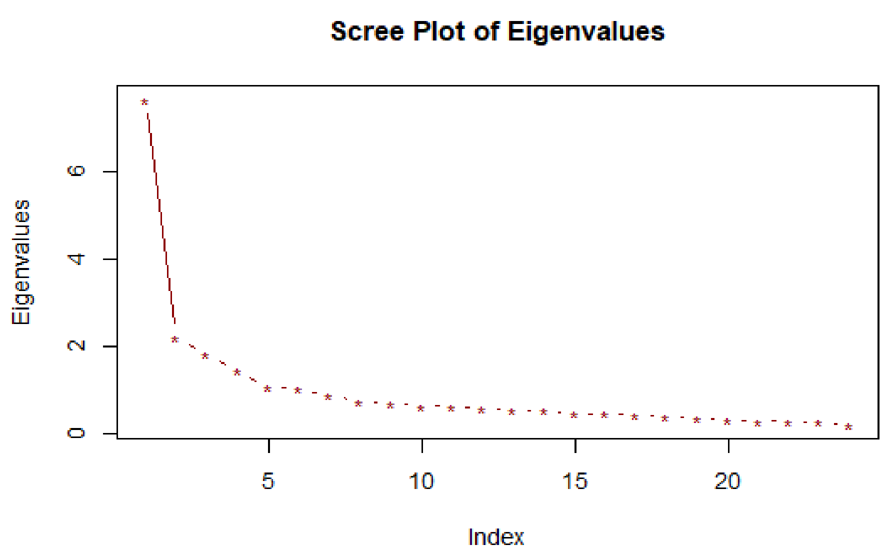

Figure (1): Scree-plot for the eignevalues

II. Estimating the factor loadings:

While estimating the factor loadings, we explored both the principal component and principal factor methods. Given our large sample size and the relatively small number of factors retained, the outcomes of both methods were close. However, we chose to proceed with the principal factor method over the principal component method. This choice is rooted in the preference for the principal factor method over the principal component method, due to the acknowledged disadvantages associated with the principal component approach.

It's worth mentioning that the principal factor method works with the reduced correlation matrix which corresponds to the communalities.

Table (1): Factor loadings with no rotation:

Variables	F1	F2	F3	F4
Appreciation of beauty	0.5492960	0.044323415	-0.06257157	0.113165989
Bravery	0.5006385	-0.280569797	0.18409655	0.207451019
Creativity	0.5618065	-0.282863144	0.25890353	0.218918347
Curiosity	0.6961465	-0.375603839	0.14407735	0.008202443
Fairness	0.5411545	0.349779182	-0.32237820	0.232131151
Forgiveness	0.3610972	0.147952170	-0.28377041	-0.193543757
Gratitude	0.6915191	-0.055910320	-0.15768570	-0.278732757
Honesty	0.5739093	0.159042306	0.05436874	0.006870359
Hope	0.7288565	-0.270324942	0.02223582	-0.275658104
Humility	0.2502136	0.460751538	-0.22767659	-0.119011915
Humor	0.4224355	-0.230825687	-0.07230336	0.272637440
Judgment	0.4226085	0.386747763	0.51853098	0.106545104
Kindness	0.6277902	0.166237236	-0.36388011	0.203581441
Leadership	0.5963509	0.132752822	-0.15516943	0.264808393
Love	0.5277338	-0.030280054	-0.11446576	-0.060219067
Love of learning	0.4369608	-0.115380624	0.16493699	0.054176809
Perseverance	0.5763312	-0.034108635	0.16528466	-0.230071183

Perspective	0.4878440	0.209445371	0.36077636	0.138105670
Prudence	0.3203515	0.669660391	0.38981542	-0.178762223
Self-regulation	0.4522255	0.148600950	0.19127885	-0.273650498
Social Intelligence	0.6297666	0.009804963	0.01724543	0.221550854
Spirituality	0.5432862	-0.094276445	-0.15532833	-0.290991507
Teamwork	0.5092886	0.252148791	-0.32302437	0.050612435
Zest	0.7681635	-0.400642955	-0.04153086	-0.168891234

From the previous output, we can notice that loadings overlap heavily, where most of the loadings of the first factor are very high, and on the other hand, the loadings corresponding to other factors are very low. Therefore, we'll move forward to applying factor rotation to overcome this issue.

III. Correlation matrix for the factors:

Table(2): Correlation matrix for factors with no rotation:

	Openness	Transcendence	Interpersonal	Restraint
Openness	1	0.16287200	0.04154504	0.003833433
Transcendence	0.16287200	1	0.07286237	0.015827525
Interpersonal	0.041545045	0.07286237	1	0.062652521
Restraint	0.003833433	0.01582753	0.06265252	1

By examining the correlation matrix of the four factors, it becomes evident that there is no correlation among them. Consequently, the issue of multicollinearity is not present. Since the factor correlations are very small (less than 0.2), so this will influence us to use the orthogonal rotation (Varimax).

IV. Factor Rotation:

Through the utilization of factor rotation, we ease the constraints imposed on the initially extracted factors, enabling the derivation of new factors that are more readily interpretable. We tried different rotation techniques as a validation tool but we specifically chose to proceed with varimax rotation due to its ease of interpretation and due to the small observed values of correlation between the factors. Varimax is an orthogonal rotation technique that emphasizes maximizing the variance of the factor loadings on each factor, maintaining orthogonality among the factors for a clearer understanding. In this context, we will define factors with high loadings as those with loadings equal to or exceeding **0.5**.

Table (3): Factor loadings with varimax rotation:

Variables	F4 (Openness)	F1 (Transcendence)	F2 (Interpersonal)	F3 (Restraint)
Appreciation of beauty	0.33558139	0.216456511	0.37183824	0.15060035
Bravery	0.60287422	0.176022782	0.05328771	0.09489398
Creativity	0.66642455	0.194969128	0.03773153	0.16442642
Curiosity	0.64298918	0.469832243	0.07477584	0.08218835
Fairness	0.16133920	0.065274213	0.72084158	0.15199552
Forgiveness	-0.06017100	0.343392235	0.38495820	0.02311404
Gratitude	0.22523324	0.636738408	0.34521136	0.09243721
Honesty	0.26205241	0.271967035	0.33503917	0.32059428
Hope	0.41315588	0.693590280	0.14223333	0.09377446
Humility	-0.23308552	0.136327165	0.46814555	0.22096401
Humor	0.49601113	0.093111183	0.22318040	-0.08237080
Judgment	0.24001080	0.005108355	0.09804207	0.73563202
Kindness	0.28117068	0.187966007	0.69278332	0.03567373
Leadership	0.37377961	0.111082947	0.54220019	0.14637977
Love	0.25189096	0.363525939	0.30881516	0.07198370
Love of learning	0.38761242	0.218348264	0.07190260	0.17691091

Perseverance	0.26559259	0.496985332	0.10195071	0.29265167
Perspective	0.34514528	0.078601554	0.15467910	0.53090749
Prudence	-0.16440097	0.108452447	0.18915893	0.81265701
Self-regulation	0.08026323	0.410975937	0.10037520	0.39069652
Social Intelligence	0.48275048	0.178774716	0.37422150	0.20256893
Spirituality	0.15486463	0.572357927	0.24641165	0.02356907
Teamwork	0.09390208	0.213867256	0.60568276	0.09193403
Zest	0.54670511	0.672675938	0.16922733	-0.02866612

After rotation, the overlap issue is solved. From the above output, we can notice that there exists a change in the order of factors post-rotation which signifies an attempt to make the factor structure more interpretable and meaningful. Furthermore, it is evident that bravery, creativity, and curiosity exhibit high loadings on the fourth factor, indicating that this factor can be identified as the "Openness" factor. Similarly, gratitude, hope, spirituality, perseverance, and zest demonstrate high loadings on the second factor, implying that this factor represents "Transcendence." The "Interpersonal" factor is characterized by high loadings from fairness, kindness, and teamwork. Lastly, the third factor, denoted as the "Restraint" factor, is shaped by heavy loadings from judgment, perspective, and prudence.

V. Communalities

Table (4): communalities for variables with varimax rotation:

Variables	Communality	Uniqueness
Appreciation of beauty	0.3204124	0.6795876
Bravery	0.4062858	0.5937142
Creativity	0.5105944	0.4894056
Curiosity	0.6465238	0.3534762
Fairness	0.5730063	0.4269937
Forgiveness	0.2702658	0.7297342
Gratitude	0.5838813	0.4161187
Honesty	0.3576695	0.6423305
Hope	0.6807892	0.3192108
Humility	0.3408993	0.6591007
Humor	0.3112912	0.6887088
Judgment	0.6083980	0.391602
Kindness	0.5956095	0.4043905
Leadership	0.4674587	0.5325413
Love	0.2961486	0.7038514
Love of learning	0.2343868	0.7656132
Perseverance	0.4135728	0.5864272
Perspective	0.4310919	0.5689081
Prudence	0.7349821	0.2650179
Self-regulation	0.3380624	0.6619376
Social Intelligence	0.4460843	0.5539157
Spirituality	0.4128509	0.5871491
Teamwork	0.4298603	0.5701397
Zest	0.7808390	0.219161

Interpretation of some of the above communalities:

- The fraction of variance in Zest that is due to the common four factors is 0.7808.
- The fraction of variance in Kindness that is due to the common four factors is 0.595.
- The fraction of variance in Hope that is due to the common four factors is 0.6807.

On the other hand, uniqueness is defined as the fraction that's due to the unique effect of each variable. But, as our main concern is the communalities then there is no need to interpret uniqueness.

Cluster Analysis

Moving forward with our analysis, we applied cluster analysis as a classification tool to classify our 849 observations into clusters, using the four obtained factors.

It's worth mentioning that we applied k-means clustering, so we tried to determine the number of clusters before we moved forward to the next steps in cluster analysis. Using the histogram of the first principal component we're able to see that the data can be classified into either two or three clusters.

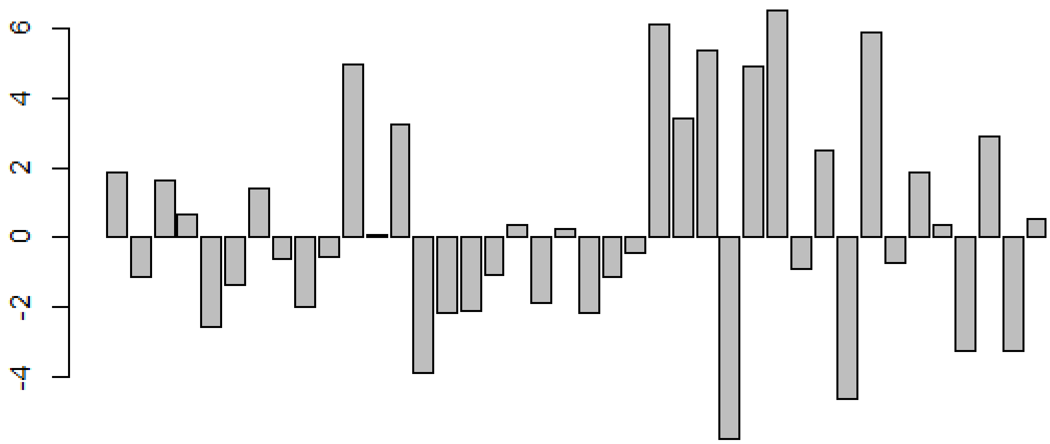

Figure (2): Histogram of the first principal component

By setting the k-means algorithm to initiate with 25 random starts, we experimented with both 2 and 3 clusters. The results indicated that utilizing 2 clusters was more appropriate as it'll be shown.

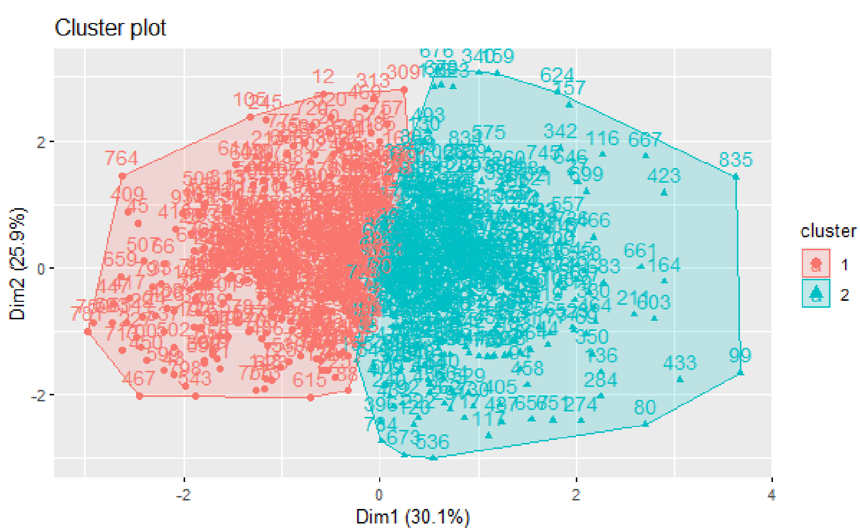

Figure (3): Cluster Plot

By creating a table of the clusters to determine the count of observations in each cluster, it was identified that the first cluster encompasses 419 observations, accounting for 49.35% of the dataset, whereas the second cluster comprises 430 observations, constituting 50.65% of the data.

Table (5): Frequency table for clusters:

Clusters	Frequency
Cluster 1	419
Cluster 2	430

Table (6): Mean of clusters:

Mean of Clusters		
Factors	**1 (Character Strengths Emphasized)**	**2 (Character Strengths Deemphasized)**
Openness	0.55529931	-0.54109397
Transcendence	0.47906178	-0.46680671
Interpersonal	0.28634722	-0.27902206
Restraint	0.06104767	-0.05948599

Upon closer examination of the table, it becomes evident that the first cluster has the highest mean values for the four factors. Consequently, we propose naming this cluster "**Character Strengths Emphasized**." Conversely, the second cluster exhibits the lowest mean values for the four factors, leading to a suggestion to label it as "**Character Strengths Deemphasized.**"

The data is standardized, with a zero cut-off point distinguishing between the two groups. Group 1 exclusively displays positive values, while Group 2 has negative mean values. This differentiation suggests that Group 1 represents the "**Character Strengths Emphasized**" cluster, while Group 2 corresponds to the "**Character Strengths Deemphasized**" cluster, as ordered in the table.

Discriminant analysis

When employing discriminant analysis for validation purposes, our initial approach involved testing the three fundamental assumptions of discriminant analysis to ascertain the suitability of either quadratic or linear discriminant analysis.

- ## Assumptions:

1) Normality assumption:

Upon employing the Mardia test for multivariate normality using the (MVN) package in R, it was observed that the assumption of multivariate normality is not met. But, since we have a large sample size (849), we can say that the normality assumption is met using the Central Limit Theorem.

2) Assumption of equal covariances:

Using the (candisc) package, we obtained the following output that rejects the hypothesis of equal covariances since the p-value is very small.

```
Box's M Test
Chi-Squared Value = 69.01626 , df = 10  and p-value: 6.86e-11
```

3) Equality of means:

Since the covariances are not equal and they're unknown, we'll use the Hottelling T-test to check the assumption of equal means. The corresponding p-value of the test was **1.843973e-191** which is less than the p-value. We'll reject the assumption of equal means.

Based on these findings, it can be concluded that **quadratic** discriminant analysis is more appropriate since the assumption of equal covariance is not met.

- ## Splitting the data and validation:

The main purpose of splitting the data into two sets is to prevent the underestimation of the probability of misclassification. Following the examination of assumptions and the identification of the appropriate model (Quadratic), the dataset was partitioned into two subsets: a training set comprising 70% of the data and a test set consisting of the remaining 30%.

The quadratic model was fitted on the train data which includes **595** observations and we used the fitted model to predict the classes for the test data which includes **254** observations. Moving forward, we constructed the classification table for the test set.

Table (7): Classification table for test set:

Predicted	**Test**		**Total**
Character Strengths Emphasized	131	4	135
Character Strengths Deemphasized	2	117	119
Total	133	121	254

Based on the table, it is evident that the overall percentage of correct classification is $\frac{131+117}{254} = \mathbf{97.637}\%$ indicating a remarkably high accuracy rate. The percentage of correct classification in the **Character Strengths Emphasized** cluster is equal to **98.5%,** while the percentage of correct classification in the **Character Strengths Emphasized** cluster is **96.7%.** This outcome supports the findings of the cluster analysis, leading to the conclusion that the discriminant analysis model aligns with the previously identified clusters.

Multivariate Regression Analysis

Adapting to the fact that the ongoing study is complex by nature, and characterized by interrelated aspects that cannot be adequately explained by examining the variables individually in isolation, this study is altered to address such intricacies through the employment of a multivariate regression analysis. Hence, it was tailored to accommodate with the cases where multiple intercorrelated response variables exist.

The variables under study are all related to psychological measures which are inherently intertwined by nature, and as previously discussed, share common underlying latent factors. Therefore, so as not to disregard such joint impacts and interdependence, the multivariate approach is exploited to declare both the unique contributions of each variable whilst also taking the shared variance and covariance structure among the multiple outcomes into consideration.

As a commencement to tackle our challenge, a multivariate regression model for the three outcome variables DASS*21 (Depression Anxiety and Stress Scale),* GHQ12 *(General Health Questionnaire)* and SEC *(Self-efficacy for Covid-19)* was constructed.

Table(8): Results for Multivariate Regression for the Response Variables (DASS21, GHQ12,SEC)

Response	DASS_21			GHQ_12			SEC		
	Estimate	Pr(>|t|)		Estimate	Pr(>|t|)		Estimate	Pr(>|t|)	
(Intercept)	32.59422	< 2e-16	***	22.97392	< 2e-16	***	3.25934	0.02271	*
Age	-0.06114	0.033767	*	0.02439	0.0798	.	0.01317	0.21833	
Gender: Male	-2.83164	0.000175	***	-0.82193	0.0239	*	0.87461	0.00179	**
Work	0.80539	0.000219	***	0.20544	0.0505	.	-0.168	0.03753	*
Student: Student	0.65237	0.4924		0.68721	0.135		-0.2674	0.449	
Sons	-1.21887	0.092792	.	0.14676	0.6753		-0.0278	0.91791	
Openness	0.10205	0.001242	**	-0.01218	0.424		0.00901	0.44167	
Restraint	0.06204	0.199844		0.05319	0.0232	*	-0.0048	0.79075	
Transcendence	-0.28545	< 2e-16	***	-0.10078	1.78E-14	***	0.11003	< 2e-16	***
Interpersonal	0.03242	0.309884		0.02921	0.0587	.	-0.0235	0.04773	*
Multiple R-squared	0.2086			0.1168			0.2412		
Adjusted R- squared	0.2001			0.1073			0.2331		
P-value	<2.2e-16			<2.2e-16			<2.2e-16		

Testing Multivariate Normal Assumption

$H_0: The\ variables\ follow\ Multivariate\ Normal\ Distribution.$

$H_1: The\ variables\ do\ not\ follow\ Multivariate\ Normal\ Distribution.$

Table(9): Mardia's Test Results for Multivariate Normality

Test	Statistic	p value	Result
Mardia Skewness	274.10411530939	4.56190205754078e-53	NO
Mardia Kurtosis	8.89457889651527	0	NO
MVN	NA	NA	NO

As shown in the table above, the p-values are all insignificant, causing the outcome result of the test to declare that the Multivariate Normality assumption is violated.
It could be worthwhile to view the tests for univariate normality to get better insights into the nature of the data under study.

Testing Univariate Normality

$H_0: The\ variables\ follow\ Normal\ Distribution.$

$H_1: The\ variables\ do\ not\ follow\ Normal\ Distribution.$

Table(10): Results for Anderson Darling Univariate Normality Tests.

Test	Variable	Statistic	p value	Normality
Anderson-Darling	DASS_21	18.9757	<0.001	NO
Anderson-Darling	GHQ_12	3.0914	<0.001	NO
Anderson-Darling	SEC	3.2274	<0.001	NO

To identify the variables that were responsible for the studied results of rejecting the previously conducted test for multivariate normality, univariate normality tests could pave the way. However, it seems that the p-values for each variable are smaller than the significance level. Hence, none of the response variables follow the normal distribution.

While the assumption of multivariate normality is desirable, its strict fulfillment may not always be necessary, especially for large sample sizes or robust statistical methods. Given that the sample size is sufficiently large, violations of normality may have minimal impact. Furthermore, our sample size is comprised of 850 observations > 20*the variables under examination which are 9 in this case. Therefore, it is not necessary to undergo transformations.

Performing MANOVA

Table(11): Type II MANOVA for the Previous Regression Model

	Df	Pillai	approx F	num Df	den Df	Pr(>F)	
Age	1	0.020215	5.756	3	837	0.00067	***
Gender: Male	1	0.020326	5.789	3	837	0.00064	***
Work	1	0.016666	4.729	3	837	0.0028	**
Student: Student	1	0.002662	0.745	3	837	0.52564	
Sons	1	0.00568	1.594	3	837	0.18941	
Openness	1	0.022217	6.339	3	837	0.0003	***
Restraint	1	0.007616	2.141	3	837	0.09359	.
Transcendence	1	0.168374	56.487	3	837	< 2.2e-16	***
Interpersonal	1	0.005943	1.668	3	837	0.17235	

Type II MANOVA is used in examining the unique contribution of each predictor variable to the variance in the dependent variables while accounting for the presence of other predictors. Moreover, the analysis uses the Pillai trace test statistic to achieve so. It measures the proportion of variance in the dependent variables explained by a predictor while accounting for other predictors.

Complying with the results of the multivariate regression model, the Type II MANOVA table shows that the variables “Student”, “Sons”, and “Interpersonal do not add a significant amount of information in terms of explaining the variability in the combination of the response variables under study. Such conclusions are obtained through the analysis of the significance of the previously discussed Pillai test statistic.

Henceforth, this suggests the removal of the three insignificant explanatory variables.

Table(12): Results of for Multivariate Regression for the Response Variables (DASS21, GHQ12,SEC) After the Removal of the Insignificant Variables.

Response	DASS_21			GHQ_12			SEC		
	Estimate	Pr(>|t|)		Estimate	Pr(>|t|)		Estimate	Pr(>|t|)	
(Intercept)	34.84944	< 2e-16	***	25.28725	< 2e-16	***	1.739999	0.159636	
Age	-0.08086	0.000677	***	0.01764	0.12483		0.015333	0.082169	.
GenderMale	-2.86005	0.000134	***	-0.86212	0.01717	*	0.921838	0.000916	***
Work	0.77422	0.000345	***	0.19272	0.06508	.	-0.164237	0.040573	*
Openness	0.10698	0.000678	***	-0.01149	0.44964		0.007746	0.506579	
Restraint	0.08069	0.08706	.	0.0613	0.00735	**	-0.012044	0.491691	
Transcendence	-0.28271	< 2e-16	***	-0.09344	5.19E-14	***	0.104175	< 2e-16	***
Multiple R-squared	0.2045			0.1109			0.2376		
Adjusted R- squared	0.1989			0.1046			0.2321		
P-value	<2.2e-16			<2.2e-17			<2.2e-18		

Contrary to our expectations, the R-squared values whether adjustments have been done or not have decreased, indicating that the model fit has worsened. This suggests that the removal of 3 variables at once was excessive. Hence, the removal of only the variable that was completely insignificant in the first model "Student" will be appropriate.

Table(13): Results for Multivariate Regression for the Response Variables (DASS21, GHQ12,SEC) After the Removal of the Variable "Student":

Response	DASS_21			GHQ_12			SEC		
	Estimate	Pr(>|t|)		Estimate	Pr(>|t|)		Estimate	Pr(>|t|)	
(Intercept)	33.10982	< 2e-16	***	23.73773	< 2e-16	***	2.995162	0.03121	*
Age	-0.0711	0.00429	**	0.01398	0.2452		0.017227	0.06228	.
GenderMale	-2.8064	0.000192	***	-0.78477	0.0308	*	0.861734	0.00204	**
Work	0.79093	0.000259	***	0.18679	0.074	.	-0.16123	0.0445	*
Sons1	-1.19355	0.098885	.	0.18263	0.6018		-0.04033	0.88063	
Openness	0.10236	0.001173	**	-0.01292	0.3962		0.009139	0.43425	
Restraint	0.06418	0.183186		0.05361	0.0219	*	-0.00521	0.77127	
Transcendence	-0.28649	< 2e-16	***	-0.10165	1.04E-14	***	0.110403	< 2e-16	***
Interpersonal	0.03248	0.308446		0.02897	0.0609	.	-0.02346	0.04796	*
Multiple R-squared	0.2081			0.1149			0.2411		
Adjusted R- squared	0.2006			0.1065			0.2339		
P-value	<2.2e-16			<2.2e-17			<2.2e-18		

Simultaneously comparing all the previous results for the ongoing multivariate regression analysis, we can confirm that the best fitting model for the data is the last one that only excludes the explanatory variable "Student" during the construction process.

Such a claim was reached through the comparison of the values computed for the adjusted coefficient of determination "R-Squared" across the three models. Moreover, the highest values were achieved in the last model concluding that the model is the most able to explain the highest proportion of variability in the outcome variables when it is constructed as in **table (13).**

Models' interpretations

Models' Goodness of fit:

The adjusted R-squared provides a measure of the proportion of variance in the response variable that is explained by the model, adjusted for the number of predictors.

Overall, judging by the values of the adjusted coefficient of determination, each model is statistically significant, suggesting that at least some of the predictor variables are somehow associated with the respective response variables. For example, **Transcendence:** Showed a consistent and significant negative association with DASS21, GHQ12, and a positive association with SEC. It indicated that higher levels of transcendence were linked to lower distress and better mental health.

Converting such numbers into percentages, we would obtain information indicating that 20.06% for DASS_21, 10.65% for GHQ_12, and 23.39% for SEC of the total variability for the respective outcome variable has been explained by the predictor variables included in the corresponding model. These values suggest that the models account for a substantial portion of the variability in the response variables, with the highest explanatory power observed in the SEC model.

Models' Overall Significance

Each model is overall statistically significant since their corresponding p-values for their respective F-statistic is <2.2e-16.

Comparison between the original study and our study

Factor Analysis:

Both our study and the original research share a common goal: uncovering latent structures in datasets using factor analysis. While the original study emphasizes discerning the number of factors and their arrangement, our study extends this exploration by delving deeper into factors, estimating loadings, and employing rotation techniques, all aimed at enriching interpretative clarity and depth.

Methodologies:

1) Determining Number of Factors:

• Original Study: Employed Parallel Analysis (PA) and Minimum Average Partial (MAP) analysis to identify the optimal number of factors, concluding four factors as the most suitable structure.

• Our Study: Utilized the scree plot and parallel analysis, consistently suggesting the retention of four factors for their analysis.

Estimating Factor Loadings:

• Original Study: Utilized a principal component analysis with an oblique rotation (promax) and selected factors based on factor loadings ≥ 0.30.

• Our Study: Employed both principal component and principal factor methods, choosing the latter due to advantages associated with larger sample sizes and rotated factors using varimax for clearer interpretation.

2) Findings:

Original Study:

• Identified four factors (transcendence, interpersonal, openness, and restraint) explaining 55% of the variance.

• Dismissed the extraction of 5 or 6 factors due to minor additional explained variance and cross-loadings.

Our Study:

- Confirmed the presence of four factors through scree plot and parallel analysis.
- Factors ordered by importance, with the fourth factor being most significant and the third least.
- Identified specific variables contributing to factors: "Openness," "Transcendence," "Interpersonal," and "Restraint."
- Emphasized communalities, attributing variance in specific variables (Zest, Kindness, Hope) to the common four factors.

Similarities:

- Both studies settled on a four-factor solution despite employing different methodologies for factor retention.
- Aimed for clarity and interpretability by using rotation techniques (Original Study: promax, Our Study: varimax).

3) Conclusion

Both studies applied rigorous methodologies to uncover latent structures within their datasets, arriving at a consensus of a four-factor solution. While employing different techniques for factor retention and loadings estimation, the studies converged on the significance of communalities and the need for clear interpretation through rotation techniques.

In essence, despite methodological differences, both studies align in their ultimate identification of a four-factor structure. Moreover, the comparison emphasizes not only the alignment in the number of factors but also highlights similarities in factor importance, strengthening the reliability of the identified factor structures.

Multivariate Regression Comparison:

Both studies delve into the impact of various factors, including character strengths, demographics, and psychological measures during the COVID-19 pandemic. Here's a comparative analysis of the key findings from both studies:

1) Similarities:

- **Transcendence:** Both studies highlight the consistent and significant positive association between transcendence as a character strength and better mental health outcomes. Higher levels of transcendence are linked to lower distress, anxiety, and better self-efficacy during quarantine in both studies.

- **Work-related Changes:** Both studies suggest a connection between work-related changes and mental health. The original study mentions that drastic work-related changes are linked to increased general distress, while our study shows a positive association of work with distress and anxiety but a negative association with self-efficacy.

2) Some Differences:

- **Demographic Variables:** The original Study focuses more on gender and work-related changes, while our study explores age, gender, and work status as significant predictors.

3) Conclusion:

Both studies underscore the importance of transcendence in predicting better mental health outcomes during the pandemic.

Conclusion

In summary, our study investigated the impact of COVID-19 on individuals' personalities and mental health in Italy, revealing four main factors through factor analysis. Varimax rotation aligned our results with the original study. Cluster analysis unveiled two clusters, validated by discriminant analysis. Multivariate regression using MANOVA confirmed consistent patterns, aligning closely with the original study. This research provides valuable insights into the enduring effects of the pandemic on psychological dimensions, highlighting the reliability of our findings across multiple analytical approaches.