Abstract:
The research focuses on analysing social media usage during the early stages of COVID-19. This involves using advanced statistical methods and algorithms to derive meaningful information. The primary aim is to showcase patterns of user age groups and interactions across various social media platforms. This can offer practical implications for businessmen and policymakers.

The methods used in the research employ a multi-faceted approach, starting with descriptive statistics to characterize key variables such as time spent on social media platforms, engagement with video streaming applications, and other factors. The pre-processing stage involves cleaning the data, removing missing values, data transformation, and reduction. This also involves using algorithms such as PCA to remove unwanted data from the dataset. The research results identified distinct user clusters and behaviour patterns. Linear regression expanded the analytical focus.

This statement suggests that predicting outcomes during a social crisis is essential. The report aims to provide an overview of social media analytics during the COVID-19 pandemic. By integrating advanced statistical methods and algorithms, decision-makers can navigate the dynamic landscape of social media with greater ease and accuracy.

Introduction

In the early phase of 2020, the event of the COVID-19 pandemic resulted in a notable increase in the utilisation of social media platforms for various purposes such as communication, data transfer, online events etc. This surge significantly transformed the way individuals share data, gather information, and communicate (Tran Huyen Thi Thanhand Lu, 2021). Social media usage underwent a significant transformative shift in recent times, which became a prominent aspect of the digital world. The business problem revolves around understanding aspects and the impact of social media platforms during the initial phase of the COVID-19 surge. Research studies have emphasised the crucial role of social media platforms in sharing information, communicating, and providing community support and sentiment-sharing through various social resources (Cinelli et al., 2020). The research conducted by (Shahbazi et al., 2023)highlights the significant role of social media usage by the government in disseminating situational awareness during chaotic times. This study seeks to address research questions to unravel individual engagement with social media platforms. It aims to quantify the time investment of an individual across various platforms during the early stages of the pandemic and shed light on the duration of interaction. This study aims to identify distinct patterns and clusters of social media usage during COVID-19, and how they relate to the spread of information across various platforms. A crucial aspect of this research question involves researching the relationships between clusters, observable trends on different social media platforms, and the generation of information. This study has an objective to identify predictive patterns in social media behaviour to develop models using different tools such as the Statistical Package for Social Science (SPSS) etc. that can foresee future trends of usage of these platforms during subsequent events and emergencies (Parlak Sert & Başkale, 2023).

In the exploration of the usage of social media during the initial phases of COVID-19, this study incorporates a comprehensive overview of its prominent components of a dataset which contains data on the usage of different platforms in India. The features such as time provide quantitative analysis of the total time spent on different platforms in hours, as well as the types of platforms where data is shared, gathered, and communicated during a crisis. Additionally, these features capture nuanced aspects of diversified usage between different age and gender groups. For instance, age and gender are taken into consideration to understand better the pattern of how different groups use these platforms. Apart from that features like health and survey questions asked during COVID-19 provide both positive and negative analysis. Research done by (Sahni & Sharma, 2020) outlines the health benefits and consequences people had because of the usage of social media during the pandemic. The CSV format dataset contains ordinal, qualitative, and discrete types of data. Ordinal data includes age, date, etc. Qualitative data includes gender, survey Q&A, etc. Discrete data includes the type of social media platforms, consumed time, etc. (Anirudh Naik Gaunker, 2020).

This research incorporates data analysis frameworks containing visualisation tools to enhance data interpretation. Data analysis involves using proximity metrics and clustering algorithms, like k-means clustering (Amiri et al., 2022). This helps to categorize users into groups with shared characteristics, such as social media platforms for communication, and entertainment. This clustering facilitates the identification of the usage of social media platforms by distinct age groups. Furthermore, classification algorithms illustrated by Naïve Based and SVM etc leverage to predict future trends. Based on historical user trends of usage of social media during the pandemic, predict the likelihood of increased engagement during emergencies. Also, statistical modelling techniques such as regression analysis uncover the patterns between variables. E.g. a regression analysis helps explore relationship between the social media usage within age groups. By integrating these data analysis techniques and frameworks, the aim is to provide intricacies of social media usage and an understanding of user engagement during a crisis.

Data Pre-Processing

In this section, a comprehensive descriptive analysis is performed on the dataset. The use of various measures and visualisation tools are employed to extract meaningful insight from the data. The dataset includes multiple features of different data types.

The following table contains the present features and data type.

Variable	Data Type
Gender	Nominal
Age	Ordinal
Time	Discrete
Types of social media	Discrete, Categorical

Descriptive Analytics

The first step in descriptive analysis is to calculate the descriptive statistics of different features in the dataset. This data is presented using the Graph Summary feature of Minitab Statistical Software.

The data has been Recoded as:

- Time Spent on Social Media: TSM.
- Time Spent on Video Streaming Social Media: TVM.

Descriptive Statistics

Variable	**N**	Mean	SE Mean	StDev	Skewness	Kurtosis
TSM	586	3.4300	0.0646	1.5633	0.29	-1.01
TVM	586	3.5700	0.0806	1.9505	0.42	-0.32

Variable	Minimum	Q1	Median	Q3	Maximum	IQR	Mode
TSM	1.0000	2.0000	3.0000	5.0000	6.0000	3.0000	3

TVM	1.0000	3.0000	3.0000	5.0000	8.0000	2.0000	3

1. Total time spent on social media applications:

Table 1 presents a summary report on the variable "Total time spent on social media" in hours, based on a sample size of 586.

- The average time spent on social media is 3.43 hours, with a standard error of 0.0646, indicating the precision of this estimate.
- The standard deviation suggests a moderate level of variability in social media usage within the sample.
- The first quartile(Q1) is 2m the median Q2 is 3 and the third quartile(Q3) is 5, illustrating that the more individuals have spent time on social media between 2 to 5 hours.
- The mode with value 3, indicated most of the individuals have spent 3 hours of duration on social media applications.

2. Total time spent on video streaming applications.

This feature displays the time spent on social media video streaming apps like Netflix, Amazon Prime, etc. in hours.

- The average video engagement is 3.57 hours, with a standard error of 0.0806, indicating precision.
- The standard deviation is 1.9505, suggesting a moderate level of variability in video streaming application usage.
- The quartile(Q1) is 3 and median(Q2) is also 3, and the quartile(Q3) is 5. This showcases that most individuals engaged in video streaming applications between 3 to 5 hours a day.
- The mode is 3, which indicates most of the individuals have spent 3 hours on video streaming applications.

Statistics of usage of social media applications by Age groups

Variable	Age (in years)	Mean	St Dev	Median	Mode
TSM	10-20	3.609	1.555	3.000	3
	21-30	3.4363	1.5499	3.0000	3
	31-40	3.316	1.662	3.000	2, 4
	41-50	2.500	1.690	2.000	2
	51-60	3.286	1.704	3.000	3
	60+	1.0000	*	1.0000	*

The table displays the statistical details for the variable "Age (in years)" and the total time spent on social media by each age group. The age group 10-20 has the highest average time of 3.609 hours. The median value is consistently around 3, indicating that the middle point of the time distribution by age is 3 hours.

Statistics of usage of different social media applications

Variable	Applications	Mean	St Dev	Median	Mode

Total time spent on social media applications.	Facebook	3.435	1.576	3.000	2
	Instagram	3.529	1.441	3.000	3
	NA	1.800	1.789	1.000	1
	Not Listed	4.118	1.654	4.000	6
	Snapchat	6.0000	*	6.0000	*
	Telegram	2.800	1.924	2.000	2
	Twitter	3.067	1.792	2.000	2
	WeChat	1.0000	*	1.0000	*
	WhatsApp	3.3887	1.5806	3.0000	2

The table displays the total time spent on various social media applications and how the mean and mode values differ across the applications.

Statistics of usage of different video streaming media applications

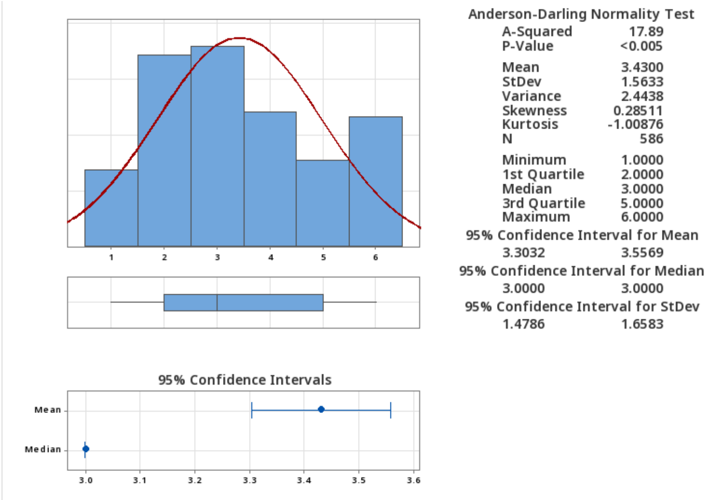
Variable	1._1	Mean	StDev	Median	Mode	N for Mode
Recoded E. Time Spent on Video	ALTBalaji	3.0000	*	3.0000	*	0
	Amazon Prime	4.060	1.914	4.000	3	18
	Hotstar	3.545	2.209	4.000	1	11
	JioTV	1.0000	0.000000	1.0000	1	2
	NA	1.429	1.158	1.000	1	12
	Netflix	4.079	1.802	4.000	3	58
	Not Listed	3.286	2.628	3.000	1	3
	SonyLiv	5.0000	*	5.0000	*	0
	Voot	2.500	1.732	2.500	1, 4	2
	YouTube	3.289	1.845	3.000	3	68

The table displays descriptive statistics for time spent on various video streaming applications. The mean and mode differ across applications, indicating a wide data spread.

Data Visualisation

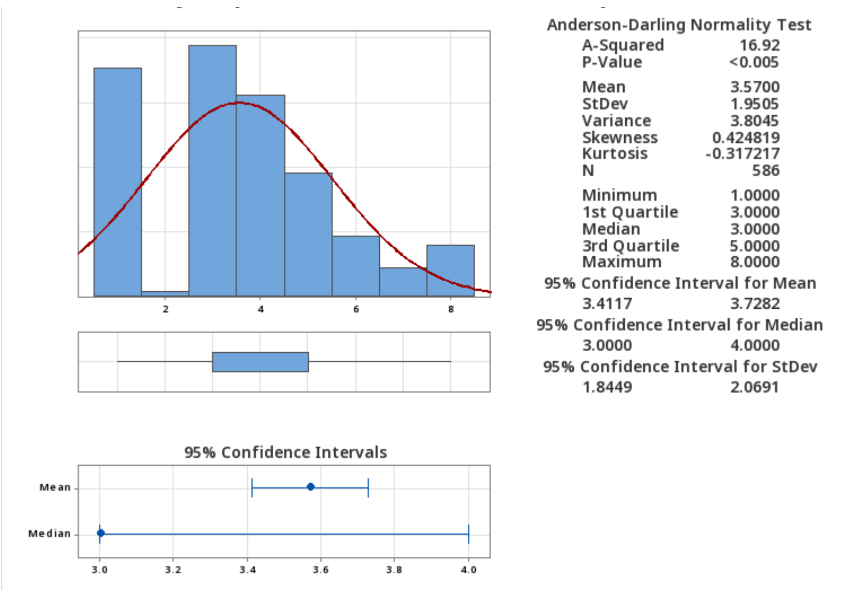
This section showcases different data visualizations, including graphs, box plots, and pie charts, to display data and its relationships with various features.

1. Summary Report for Time Spent on Social Media Applications



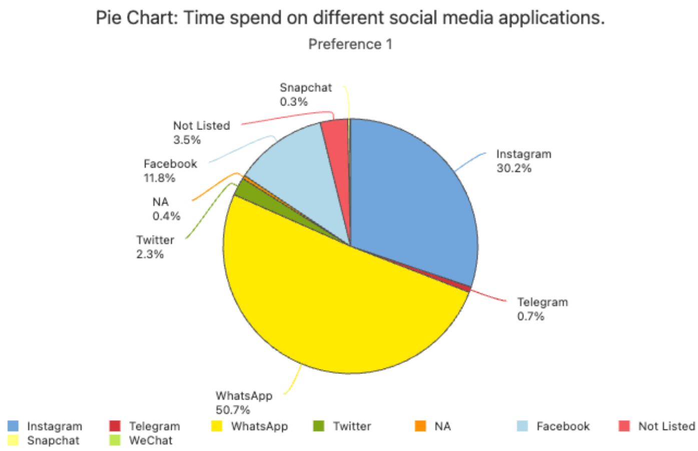
The report summarises the total time spent on social media applications and includes statistics calculated using the Anderson-Darling Normality Test. The histogram, which features a bell curve and a box plot at the bottom, indicates the skewness and spread of the data. The graph illustrates that most values fall within the range of 2 to 4 hours.

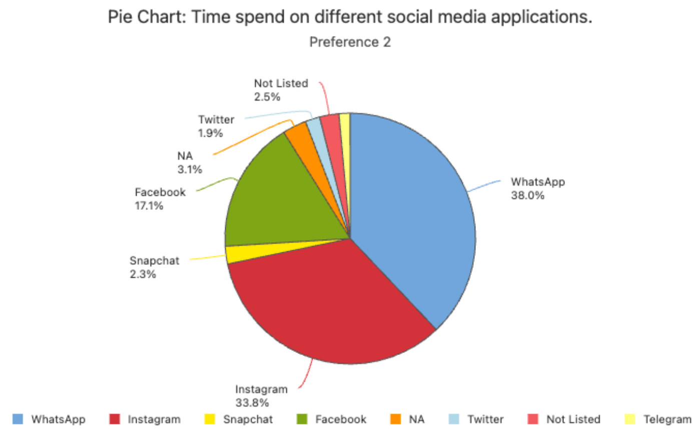
2. Summary Report for Time Spent on Video Streaming Applications



The figure displays a summary report that depicts the total time spent by individuals on video streaming applications. The bell curve is skewed to the right, and the histogram indicates that most values are concentrated between 3 to 5 hours. The box plot illustrates the spread of data within the Interquartile Range (IQR).

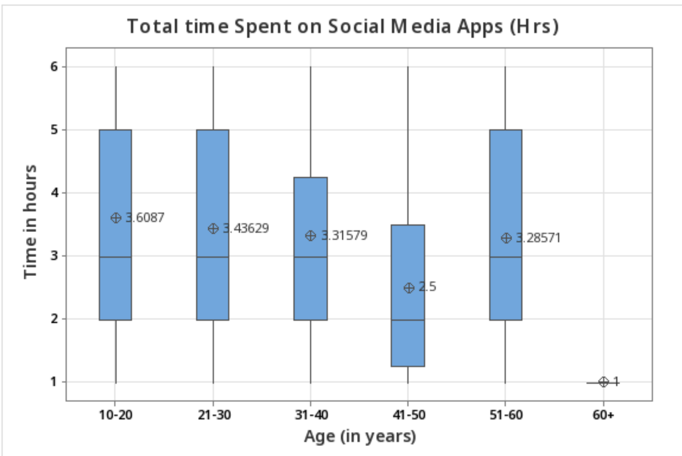
3. Pie chart of Different social media applications and their percentage of total time spent in hours.



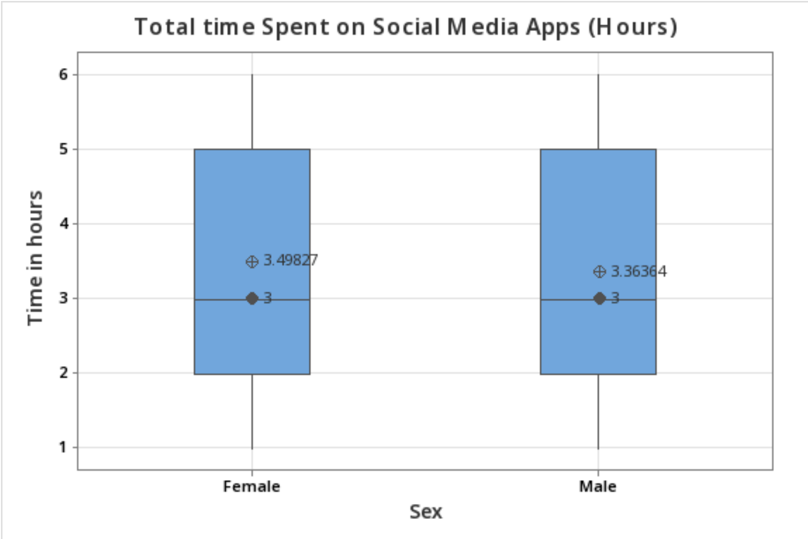


Both pie charts showcase preference 1 and preference 2 social media applications on which individuals have spent their time during COVID-19.

4. Box plot for Total Tim Spent on Social Media Application by Age and Gender groups.



The boxplot illustrates the average time spent on social media across different age groups, providing an overview of the data spread.



This boxplot displays the time spent on social media by gender. The average mean time spent by female and male users is 3.498 and 3.36 hours, respectively, with a median of 3.

Proximity Analysis:

Euclidean Distance:

The proximity analysis is done using the SPSS tool. The case processing summary states that out of 587 cases, 586(99.8%) are valid with complete data. This indicates a high percentage of valid cases in the dataset.

Proximities

Case Processing Summary

Cases					
Valid		Missing		Total	
N	Percent	N	Percent	N	Percent
586	99.8%	1	0.2%	587	100.0%

Proximity Matrix

		V6	V28	V29
Rescaled Euclidean Distance	V6	.000	.373	1.000
	V28	.373	.000	.000
	V29	1.000	.000	.000

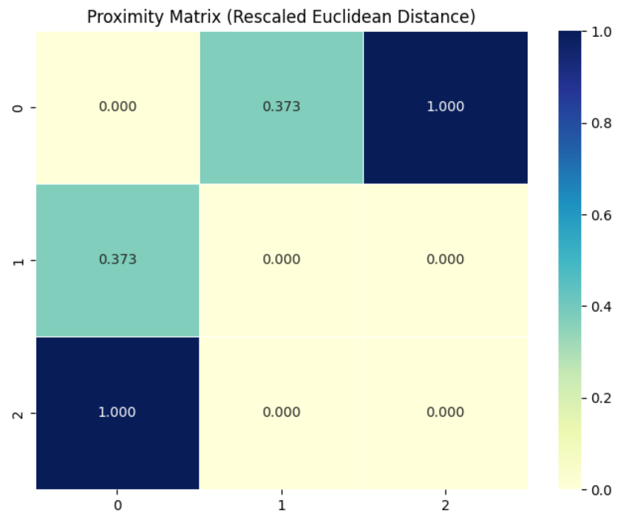
This is a dissimilarity matrix

The provided dissimilarity matrix depicts the rescaled Euclidean distance between three variables: V6, V28 and V29,

V6 = Time spent on social media applications

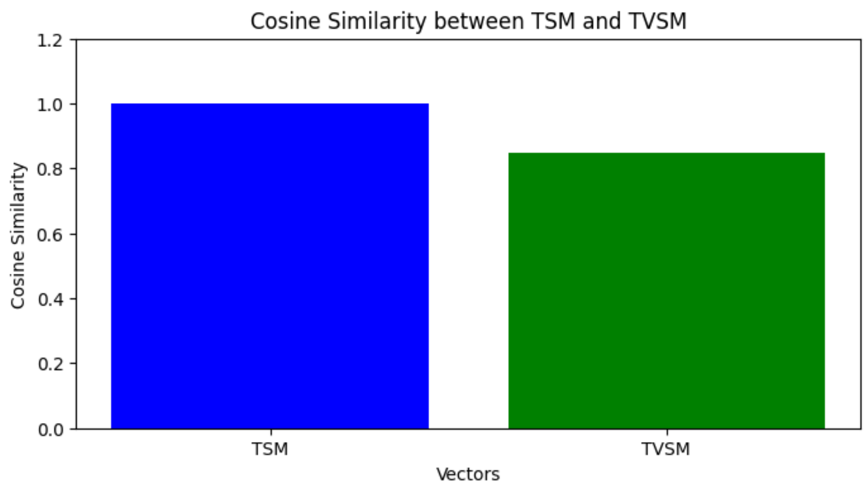
V28 = Time spent on News Applications

V29 = Time spent on Video Applications.



A moderate dissimilarity of 0.373 exists between V6 and V28, while the dissimilarity of 1.000 between V6 and V29 is more pronounced. The value 0.000 suggests that V26 and V28 are dissimilar.

Cosine Similarity:
The Cosine similarity is calculated using the "python Sklearn" library with the normalization of the two vector values into variables.



The cosine similarity of the two vectors shows that TSM and TVSM are very similar in the dataset., with the Cosine Similarity of 0.848165162134791.

Data Cleaning:

The pre-processing stage is a critical phase in preparing a dataset for analysis, enhancing the quality and relevance of data.

Missing values:

Result Variables					
Result Variable	N of Replaced Missing Values	Case Number of Non-Missing Values		N of Valid Cases	Creating Function
		First	Last		

V29_1	1	1	587	587	SMEAN(V29)

Result Variables					
Result Variable	N of Replaced Missing Values	Case Number of Non-Missing Values		N of Valid Cases	Creating Function
		First	Last		
V6_1	1	1	587	587	SMEAN(V6)

The missing values are replaced using the SPSS tool, replacing values by calculating the SMEAN of the feature data.

Data Transformation:

```
V5 into ZoneCode
Old Value                          New Value  Value Label

A. Lockdown zone category (As o          1  A. Lockdown zone category (As o
Containment                              2  Containment
Green                                    3  Green
Orange                                   4  Orange
Red                                      5  Red
```

Age Into AgeCoded	
Old Value	New Value
10-20	1
21-30	2
31-40	3
41-50	4
51-60	5
60+	6

The data has been standardised and transformed to ensure that variables are comparable and do not influence the analysis inappropriately.
Features such as Age and lockdown zone in covid 19 are recorded as 1,2,3 numeric values.
The Gender Male and Female are recoded as 0 and 1 respectively.

Data Reduction:

The reduction of data has been analysed using Principal Component Analysis (PCA) in Minitab.

Eigenvectors

Variable	**PC1**	PC2	PC3	PC4	PC5	PC6	PC7
V3	-0.224	-0.588	-0.184	0.299	0.402	0.461	0.326
Social	0.530	-0.292	-0.091	0.066	0.456	-0.148	-0.626
News	0.148	0.599	0.205	-0.141	0.493	0.560	0.028
Video	0.617	-0.048	0.165	0.029	0.127	-0.312	0.689
ZoneCode	0.372	0.191	-0.192	0.693	-0.439	0.332	-0.072
AgeCoded	0.193	0.077	-0.849	-0.449	-0.071	0.120	0.123
WorkoutTime	-0.302	0.406	-0.362	0.451	0.417	-0.480	0.072

Variables with higher eigenvector values have a greater impact on their respective principal component.

- Video and Social consistently show high influence across multiple principal components,
- News, V3 which is Gender and Workout Time exhibits lower influence.

After performing PCA analysis, the News and Workout Time columns can be removed from the analysis.

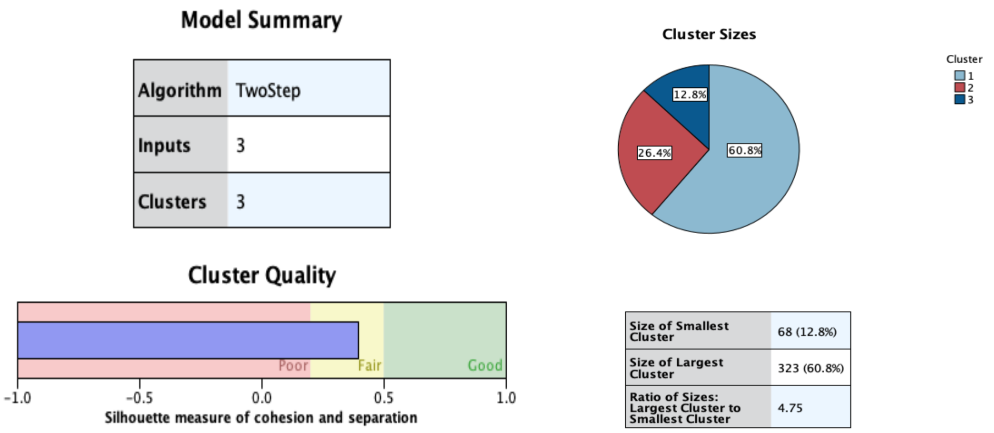
Data Processing:

In this section, advanced data processing techniques, including clustering and classification algorithms, will be applied to the finalised dataset. Clustering methods such as K-means clustering and Two-Step clustering will be implemented to analyse different patterns associated with the dataset. The classification algorithms, including Naïve Bayes and Support Vector Machines, will be employed to assess the effectiveness of the algorithm.

Clustering Analysis:

1. Two-Step Clustering Analysis:

Age, TSM and TVM Two-Step Clustering Analysis

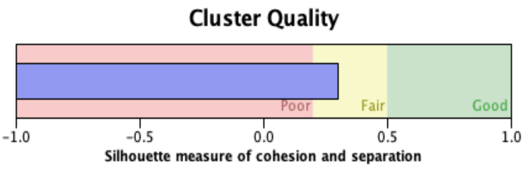


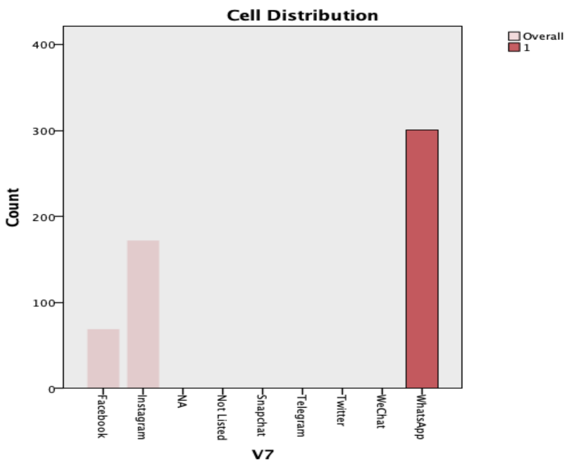
The figure illustrates a two-step clustering analysis that has been performed on the variables Age, TSM, and TVM. The Silhouette measure has been used to evaluate the cohesion and separation of the clusters. The resulting value of 0.3948 indicates moderate cohesion and separation, suggesting that the clusters are well-defined with a fair index.

In terms of cluster distribution, Cluster 1 consists of 323 instances, accounting for 60.83% of the total values. Cluster 2 and Cluster 3 contain 140 and 68 instances respectively, making up 26.37% and 12.81% of the total values.

Model Summary

Algorithm	TwoStep
Inputs	2
Clusters	2

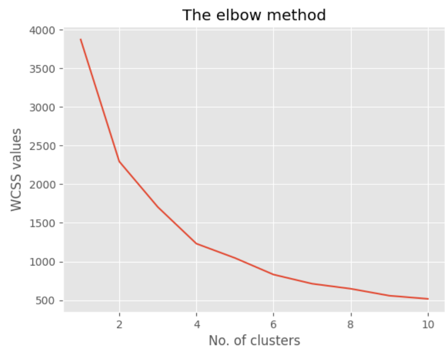




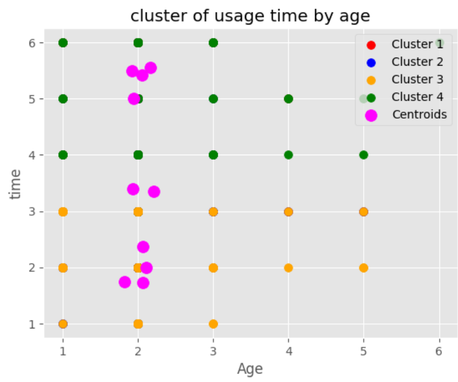
The graph presents the results of a step clustering analysis carried out on two variables, TSM and Social Media Types. The Silhouette measure for category 1 is 0.2997, indicating a reasonably good fit of clusters. The accompanying bar chart provides a breakdown of the different social media applications within each category. Among the applications, WhatsApp is the most popular with a count of 310, followed by Instagram with 172, and Facebook with 69 counts.

2. K-means Clustering Analysis

The k-means clustering analysis is performed on Age, TSM and TVM variables. To find the value of K, the elbow method is used.



The figure shows elbow method output, where the value of K = 4, an optimal cluster value for k-means clustering on the above variables.



The figure shows the K-means clustering result with four clusters. Clusters 3 and 4 have a wide area with centroids centred on age group 2, where most of the centroids lie around 2 and 3.

ANOVA						
	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
AgeCoded	.552	2	.372	583	1.485	.227
Social	244.370	2	1.614	583	151.418	<.001
Video	795.068	2	1.090	583	729.391	<.001

Based on the Age Coded p-value of 0.227, there is no significant difference among clusters. However, the Social and Video clusters show significant variation.

The clustering analysis above provides answers to the research question of identifying distinct usage patterns of social media platforms among different age groups. The K-means clusters reveal diverse groups of clusters forming a pattern that highlights age group 2 having the highest number of centroids.

Classification Analysis:

1. Naïve Bayes Classification Analysis

Case Processing Summary

		N	Percent
AgeCoded	1	69	11.8%
	2	463	79.0%
	3	38	6.5%
	4	8	1.4%
	5	7	1.2%
	6	1	0.2%
Valid		586	100.0%
Excluded		0	
Total		586	

Subset Summary

Subset	Predictor Added	Rank	Pseudo-BIC	Average Log-Likelihood
0	(Initial Subset)[a]			
1	Video	2	.718	-.713
2	Social	1	.698	-.687

a. The initial subset is empty.

Selected Predictors

	Predictors
Categorical	Social Video

Classification

Observed	Predicted 1	2	3	4	5	6	Percent Correct
1	0	69	0	0	0	0	0.0%
2	0	463	0	0	0	0	100.0%
3	0	38	0	0	0	0	0.0%
4	0	8	0	0	0	0	0.0%
5	0	7	0	0	0	0	0.0%
6	0	1	0	0	0	0	0.0%
Overall Percent	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	79.0%

Dependent Variable: AgeCoded

The provided figure displays a summary of the social and video-related variables. The lower pseudo-BIC and higher Average Log-Likelihood values indicate that the model's performance has improved. Subset 2, which includes the social predictor, is considered the final model subset as it has a Pseudo-BIC of 0.698 and a higher Average Log-Likelihood of -0.687 compared to Subset 1.

The figure also depicts the model's accuracy in predicting Age Coded Category 2, which is quite high. However, the model fails to predict other categories correctly due to their smaller sample sizes.

Regarding the accuracy of the model, it covers a significant portion of the dataset for Age Coded Category 2, with an accuracy rate of 79%. Unfortunately, the low sample sizes of the minor classes result in the model's inability to predict them accurately.

Therefore, based on the model's performance, it can be concluded that the time spent on social media and video streaming applications can be used to predict Age category 2 accurately.

2. Support Vector Machine Classification Analysis

To perform a Support Vector Machine (SVM), one needs to select a subset of the dataset that includes TSM, TVM, and Age. TSM and TVM will be independent variables, while Age will be the dependent variable. Once the subset is selected, it should be divided into train and test datasets. Finally, a classification report can be generated by comparing the test and prediction data.

```
from sklearn.metrics import classification_report

print(classification_report(y_test, y_predict))
```

```
              precision    recall  f1-score   support

           1       0.00      0.00      0.00        16
           2       0.78      1.00      0.88        92
           3       0.00      0.00      0.00         7
           4       0.00      0.00      0.00         2
           6       0.00      0.00      0.00         1

    accuracy                           0.78       118
   macro avg       0.16      0.20      0.18       118
weighted avg       0.61      0.78      0.68       118
```

Sr.no	Precision	Recall	F1-score	Support
1	0.00	0.00	0.00	16
2	0.78	1.00	0.88	92
3	0.00	0.00	0.00	7
4	0.00	0.00	0.00	2

6	0.00	0.00	0.00	1

Validation:

The given output showcases that, the model performs well in classifying class 2 instances, but it struggles with other classes due to low portion and recall values.

- Precision: For class 2, the precision is 0.78, indicating that 78% of the instances predicted as class 2 were correctly satisfied.
- Recall: For class 2, Recall is 1.00, indicating all instances of class 2 were correctly satisfied
- F1-Score: For class 2, F1-score is 0.88.
- Accuracy: Overall accuracy is 78%, a portion of correctly classified instances out of the total.

Statistical Modelling and Analysis

The statistical modelling involves linear regression analysis to predict Social Media Usage time by age groups. The models are created using the SPSS tool.

Correlations

		Social	AgeCoded
Pearson Correlation	Social	1.000	.068
	AgeCoded	.068	1.000
Sig. (1-tailed)	Social	.	.051
	AgeCoded	.051	.
N	Social	586	586
	AgeCoded	586	586

The Correlations tables show a weak positive correlation (0.068, p=0.051) between social and age-coded variables in a dataset of 586 observations.

Model Summary[b]

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	.068[a]	.005	.003	1.561	.005	2.678	1	584	.102

a. Predictors: (Constant), AgeCoded
b. Dependent Variable: Social

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6.525	1	6.525	2.678	.102[b]
	Residual	1423.107	584	2.437		
	Total	1429.631	585			

a. Dependent Variable: Social
b. Predictors: (Constant), AgeCoded

The model summary indicates a weak fit with an R-squared value of 0.0005. Although there is a slight improvement in the change statistics, the overall model performance is still limited with low R-squared and F change.

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions (Constant)	AgeCoded
1	1	1.957	1.000	.02	.02
	2	.043	6.765	.98	.98

a. Dependent Variable: Social

Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	3.25	4.12	3.43	.106	586
Residual	-2.600	2.746	.000	1.560	586
Std. Predicted Value	-1.667	6.527	.000	1.000	586
Std. Residual	-1.666	1.759	.000	.999	586

a. Dependent Variable: Social

The residual statistics display the predicted values of the dependent variable "Social," which is the total time spent on social media in hours. This analysis aims to answer the research question of predicting the value of social media usage on different platforms. The minimum predicted value is 3.25 hours, the maximum is 4.12, and the mean is 3.43. This indicates that the average use of social media during a crisis for a specific age group is 3.43 hours per day on different social media applications.

The residual ranges from -2.600 to 2.746 with a mean of 0.000 showing that, on average, the model is performing well in predicting the usage of social media.

Overall, the residuals are centred around zero, which indicates that the model is well-behaved and performing positively for the linear regression model.

Conclusion

The pre-processing results of this study offer valuable insights into social media usage during the early stages of COVID-19. The findings provide a comprehensive understanding of user behaviour, including patterns of different age groups and interactions across various social media platforms. In conclusion, these results shed light on how people are using social media during the pandemic.

During the descriptive analysis stage, we explore important variables such as "Time spent on social media (TSM)" and "Time spent on video streaming social media (TVM)". This helps us to calculate the average time spent on these platforms, analyse user engagement across different age groups and platforms, as well as identify trends and user preferences. To gain even deeper insights, we also look at the standard deviation and skewness of the data, which helps us to understand the distribution of these features.

During the data analysis process, proximity analysis was used to identify the relationships between variables. This analysis revealed both similarities and dissimilarities within the dataset. The dissimilarity matrix helped to understand the patterns within the dataset, while the similarities assisted in identifying similar features and patterns. Additionally, it highlighted any unwanted data present in the dataset.

During the pre-processing stage, various steps were taken to refine the dataset and make it suitable for advanced analysis. These steps included data cleaning, integration, reduction, and transformation. The main objective of these steps was to ensure that the data was relevant and optimal for advanced analysis. To reduce unwanted features, the Principal Component Vector (PCA) was used to decide on the data reduction.

Clustering and classification algorithms provided comprehensive insights into the dataset. The K-means and Two Step Clustering algorithms demonstrated the suitability of the data and identified different groups and structures, offering a deeper understanding of user segmentation. On the other hand, the Naïve Bayes classification provided a predictive model for age categorization based on observed patterns. The integration of the Support Vector Machine expanded the analytical view, providing insights to explore relationships and predict social media behaviours.

The statistical model, Liner Regression gave a predictive model of predicting the average social media usage within different age groups. This gave an important insight into how individuals made use of social media concerning different patterns of social media usage across various platforms.

The study's findings were compared to existing literature, and the results confirmed the study's robustness. The practical implementation of these findings highlights their relevance for organizations, policymakers, and governments that aim to optimize and utilize social media strategies during persistent crisis scenarios.

It is important to note the limitations of this study. The dataset used in this study was limited to a specific region and a particular period during the early stages of the COVID-19 pandemic. It is essential to consider the representativeness of the dataset and the evolving nature of social media trends. To improve the accuracy of the results and gain a comprehensive understanding, it is recommended to continuously monitor social media trends across different platforms. Refining the predictive model and expanding the dataset would also help in generating more accurate results.

This study adds to the existing literature on social media analytics and provides practical insights on how to effectively use digital platforms. With the increasing usage of social media as a reliable source of information, the findings of this study can guide decision-making and strategy development. Furthermore, this study can also stimulate further research in the vast field of social media analytics.

References:

Amiri, F., Abbasi, S., & Babaie mohamadeh, M. (2022). Clustering Methods to Analyze Social Media Posts during Coronavirus Pandemic in Iran. *Journal of AI and Data Mining*, *10*(2), 159–169. https://doi.org/10.22044/jadm.2022.11270.2285

Anirudh Naik Gaunker. (2020). *Social Media Usage Data During early Days of Covid*. Kaggle. https://www.kaggle.com/datasets/anirudhnaikgaunker/social-media-usage-data-during-early-days-of-covid

Cinelli, M., Quattrociocchi, W., Galeazzi, A., Valensise, C. M., Brugnoli, E., Schmidt, A. L., Zola, P., Zollo, F., & Scala, A. (2020). The COVID-19 social media infodemic. *Scientific Reports*, *10*(1), 16598. https://doi.org/10.1038/s41598-020-73510-5

Parlak Sert, H., & Başkale, H. (2023). Students' increased time spent on social media, and their level of coronavirus anxiety during the pandemic, predict increased social media addiction. *Health Information & Libraries Journal*, *40*(3), 262–274. https://doi.org/10.1111/hir.12448

Sahni, H., & Sharma, H. (2020). Role of social media during the COVID-19 pandemic: Beneficial, destructive, or reconstructive? *International Journal of Academic Medicine*, *6*(2).

https://journals.lww.com/ijam/fulltext/2020/06020/role_of_social_media_during_the_covid_19_pandemic_.5.aspx
Shahbazi, M., Bunker, D., & Sorrell, T. C. (2023). Communicating shared situational awareness in times of chaos: Social media and the COVID-19 pandemic. *Journal of the Association for Information Science and Technology*, *74*(10), 1185–1202. https://doi.org/https://doi.org/10.1002/asi.24814
Tran Huyen Thi Thanh and Lu, S.-H. and T. H. T. T. and N. B. Van. (2021). Social Media Insights During the COVID-19 Pandemic: Infodemiology Study Using Big Data. *JMIR Med Inform*, *9*(7), e27116. https://doi.org/10.2196/27116