learn

Assignment 2

Course: COMP-3705 (277885) Unsupervised Machine Learning

Data	3 3 points	2 2 points	1 1 point	0 0 points	Criterion Score
CSV data loaded into a pandas DataFrame; qualitative and quantitative descriptions given; dataset checked and cleaned; scatter plots and histograms with comments.	Data loaded and described (both qualitative and quantitative). Dataset checked for missing values, duplicates, and incorrect types; cleaned appropriately with documentation. Correlation heatmap, scatter plots, and histograms provided with insightful comments.	Most tasks completed, but some details missing or explanations incomplete. Work may be somewhat disorganized, but understanding evident in comments.	Several steps missing or poorly documented (e.g., cleaning not explained or only superficial descriptions). Visualizations weak or lack meaningful interpretation.	No meaningful work done. Data not described, not cleaned, or visualizations and comments missing.	/ 3

Modeling	3 3 points	2 2 points	1 1 point	0 0 points	Criterion Score

Modeling	3 3 points	2 2 points	1 1 point	0 0 points	Criterion Score
Data scaled; at least 10 Agglomerative clustering models using sklearn and 10 clustering models using scipy created with various hyperparameters; silhouette scores calculated and tabulated; best model selected and justified.	Data scaled correctly. At least 10 models per clustering method created with varied hyperparameters. Silhouette scores calculated and tabulated. Best model chosen with strong justification, considering both scores and meaningful interpretation. Clear, well-organized documentation.	Most modeling steps completed, but some hyperparameter exploration or justification missing. Observations partially documented.	Only some modeling steps completed (e.g., few models, missing scores, or unclear justification).	No meaningful modeling work.	/ 3

Conclusion	3 3 points	2 2 points	1 1 point	0 0 points	Criterion Score

Conclusion	3 3 points	2 2 points	1 1 point	0 0 points	Criterion Score
Scatter plot given with good choice of features; histogram of cluster sizes; each cluster described quantitatively and in words; actionable marketing insights provided; conclusions summarized in Markdown.	Scatter plot created with clusters clearly visualized. Histogram of cluster sizes included with thoughtful interpretation. Each cluster described using descriptive statistics. Marketing recommendations clear, actionable, and well-documented in Markdown.	Most steps completed, but some cluster descriptions or insights incomplete. Recommendations partially detailed.	Several steps missing; observations or recommendations weak.	No meaningful conclusions, visualizations, or insights provided.	/ 3

Iteration & Reflection	2 2 points	1 1 point	0 0 points	Criterion Score

Iteration & Reflection	2 2 points	1 1 point	0 0 points	**Criterion Score**
Reflection on whether the chosen model produces meaningful and well-separated clusters; iterations documented.	Clear reflection on how well the chosen model produced meaningful, well-separated clusters. If initial results were not ideal, any adjustments made are documented along with observations on their effect and justification.	Some justification provided; clusters partially meaningful; reflection incomplete.	Weak or missing justification; clusters unclear or poorly evaluated.	/ 2

Organization	2 2 points	1 1 point	0 0 points	**Criterion Score**
Notebook structure, clarity of explanations, and professionalism of writing.	Well-structured, easy to follow; Markdown explanations clear; professional formatting. Correct file naming convention (Asg#_Group#_FirstnameLastname...). Includes acknowledgment of group contributions.	Mostly organized; some flow or clarity issues; Markdown explanations partially clear. Correct file naming or contributions section missing.	Poorly structured; difficult to follow; explanations unclear or missing.	/ 2

| Total | | | | / 13 |

Overall Score

A
8 points minimum

B
7 points minimum

C
6 points minimum

D
5 points minimum

F
0 points minimum