

Multivariate Analysis of Dry Milk Samples: An Applied Approach Using Factor and Cluster Methods

Salih Sufian Munther

Department of Management Economics Banking, College of Business Economics, Al-Nahrain University, Baghdad, Iraq.

Corresponding Author: sufian.m.salih@nahrainuniv.edu.iq



Article – Peer Reviewed Received: 15 April 2025 Accepted: 5 June 2025 Published: 25 June 2025

Copyright: © 2025 RAME Publishers This is an open access article under the CC BY 4.0 International License.



https://creativecommons.org/licenses /by/4.0/

Cite this article: Salih Sufian Munther, "Multivariate Analysis of Dry Milk Samples: An Applied Approach Using Factor and Cluster Methods", *Journal of Production and Industrial Engineering*, RAME Publishers, vol. 6, issue 1, pp. 1-11, 2025.

https://doi.org/10.26706/jpie.6.1.202 50601 **Abstract:** The main objective of this study is to explore the hidden patterns and relationships among different dry milk products through the use of factor analysis. the focus shifts to determining how many factors emerge as a result of the factor analysis This leads to identifying the latent structures by determining the number of factors revealed through factor analysis of various dry milk products. The objective is to identify the number of clustering groups and interpret the associations among the elements by jointly applying both analyses.

Firstly, a total of 20 samples were randomly drawn from multiple types of dry milk and used as the basis for the analysis. Using SPSS, both factor analysis and cluster analysis were conducted to analyze and interpret the dataset. The study's findings begin with the results obtained from the cluster analysis, a total of two factors were extracted, and explaining 91.8% of the total variance observed in this study, the interaction between variables was assessed. The initial eigenvalue explained 52.3% of the overall variation in the data; In contrast, the second eigenvalue represented 38.7% of the cumulative variance. Secondly, three groups were extracted through factor analysis resulting from the cluster analysis conducted on the collected data from a total of 20 samples of diverse dry milk types, the results are as follows:

A- The initial cluster includes full cream milk types. B- Additionally, the second group consists of full cream milk. C- The third category involves samples of infant formula.

Third, the study presents the joint design of factor analysis and cluster analysis the two-dimensional framework depicted in Figure 2 is based on figure (1). By correlating the outcomes of the two analytical methods where the cluster analysis classifies the different samples through correlating the outcomes of the two methods, where cluster analysis categorizes dry milk types into three groups (as described above). Conversely, the factor analysis suggests that there are two groups, Clusters one and two exhibited the highest degree of proximity Clusters one and two exhibited high similarity, whereas the third group was markedly distant. This outcome highlights the fundamental nature of Such findings highlight the structural characteristics of these samples, Where the first and second groups are closely related regarding the nature of their composition where both the first and second groups consist of various types of full cream milks, Groups one and two are composed of full cream milks, while group three involves infant formula products.

While this research centers on classifying dry milk samples, the statistical techniques applied Factor and cluster analysis are powerful and flexible techniques, tools that can also help inform decision-making in regional planning and resource management.

Keywords: Multivariate Statistical Methods; Factor Analysis; Cluster Analysis; Regional Resource Management; Quantitative Planning Tools.

1. Introduction

Powdered milk is defined as the substance in which the solid content is concentrated by evaporating the majority of the water content from whole milk or partially/completely skimmed milk. The moisture content in the final dried product must not exceed 2–5%.

Journal of Production and Industrial Engineering

Milk is considered a naturally homogeneous mixture of several components that exist in specific quantities, types, and characteristics, which contribute to its unique physical and chemical properties, as well as its nutritional value. These components exist in

various physical states: fat appears in an emulsified form; proteins exist in colloidal or suspended states; and some salts and the sugar lactose are found in a true solution. Additionally, milk contains vitamins (in soluble form or associated with other components), enzymes, and microorganisms.

Fat is regarded as one of the most important constituents of milk, contributing significantly to its physical characteristics and helping maintain a balance among its various components. Protein also plays a crucial nutritional role for living organisms. The human body requires approximately 45 grams of protein per day, or 0.8 grams per kilogram of body weight, and milk alone supplies about 20–30% of this requirement.

Therefore, milk is considered one of the most valuable and comprehensive food sources available to humans. It qualifies as a complete essential food due to its inclusion of all vital nutrients in the required quantity and quality. Milk is also characterized by ease of digestion and palatability for consumers, in addition to its compatibility with most other foods. The unique nutritional qualities of milk, which are not found in any other single food, have drawn human attention and motivated researchers to delve deeply into this vital resource [1].

- To determine the number of extracted factors after applying factor analysis to various powdered milk samples.
- To identify the number of distinct cluster groups that can be visualized graphically after applying cluster analysis to the different powdered milk samples.
- To illustrate the relationship between elements in a concrete and practical way by linking the results of both factor and cluster analyses.

To achieve these objectives, the study was divided into three main sections. The first section presents a focused overview of the theoretical framework of factor and cluster analysis. The second section covers the practical application of both methods within the research, while the final section is dedicated to drawing conclusions in alignment with the study's findings.

2. Theoretical Framework

The theoretical framework is divided into two parts:

2.1 Factor Analysis

The fundamental idea of factor analysis, in its general concept, is to examine the extent to which a set of variables can be represented by a smaller number of factors. P of variables, X2...., XP X1

"in terms of a small number of factors that are used to represent the relationships among a set of interrelated variables."

"Factor analysis is based, to some extent, on a specific mathematical model used in the analysis, which is:"

Where:

$$Xi = aiF + ei \qquad \dots \dots \dots (1)$$

Xi: A standardized value of the original variable, with a mean of zero and a standard deviation of one ai Factor loading

F: The value of the factor, which has a mean of zero and a standard deviation of one across all observations."

: el¹ The portion of Xi that pertains to the variable i.

Factor analysis is applied to a wide and diverse range of operations, and the most commonly used general classification of its steps includes the following [2]:

- Correlation matrix
- Extraction of initial factors to determine the potential for data reduction
- Rotation to obtain the final solution by identifying simpler and more interpretable factors

Factor analysis data typically consists of p variables across n observations, and the process involves three main steps:

In the first step, the conditional factor loadings are determined using the Principal Component Analysis (PCA)method, by ignoring all principal components that come after the first m components. These first m principal components are then retained to serve as the m factors[3].

The second step in the analysis is known as factor rotation. In this step, the conditional factors are transformed to obtain a new set of factors that are easier to interpret. Rotation in this context involves adjusting the conditional factor loadings to achieve a simpler structure[4].

The final step involves calculating the factor scores, which are the values of the m factors for each individual observation.

When using factor analysis to derive conditional solutions, the selection of m—equal to the number of eigenvalues greater than one in the correlation matrix—is based on the rule of thumb. Factors associated with eigenvalues less than one explain less variance than those with eigenvalues greater than one, and therefore, are typically excluded[5].

In general, increasing the number of factors (m) does not affect their orientation during rotation, meaning the rotated factor structure remains consistent.

In this study, orthogonal rotation (VARIMAX) was applied within the Principal Component Analysis method. Accordingly, the new solution is represented as follows[6]:

$$F *= (GTG) - 1 GTX \qquad \dots \dots \dots (2)$$

Where*:

F: A vector consisting of M factors F1, F2, FM

X: A vector consisting of (p) Variables (X1 ... X2 ... Xp)

G : A P×M matrix of factor loadings (from variables 5, 4, 3)

2.2 Cluster Analysis

Classification is considered one of the fundamental scientific methods for data reduction, with the aim of grouping data into clusters or categories. Cluster analysis was developed to address the following problem[7]:

Given a sample consisting of n elements, each having observations on p variables, the task is to represent these elements through a graphical process that groups them into clusters, such that the elements within each cluster are similar to each other. This is a fully numerical method, and the number of clusters is not known in advance.

One of the main challenges in cluster analysis is that, despite the existence of many different methods (algorithms) for performing it, there is no universally agreed-upon best method. These various methods often produce different results for the same dataset[8].

There are numerous algorithms that have been used in cluster analysis. In this study, the focus will be on the hierarchical method. This approach begins by calculating the distances between each observation and all other observations. Then, clusters are formed through an agglomerative process, in which each element initially starts in its own individual group. Gradually, the closest groups begin to merge, until ultimately, all observations are combined into a single final cluster[9].

The key question this research seeks to answer is the following:

Given 20 different samples of powdered milk, each consisting of 5 variables, is there evidence indicating that these data form clustered groups? This is tested against the alternative hypothesis which states that the data indeed form clustered groupings, or whether the data instead represent an unstructured (non-hierarchical) collection.

This ambiguity—between the milk samples and the variables associated with each sample—is revealed and examined through cluster analysis.

The data used in the cluster analysis consist of a matrix of observations, where the columns represent the variables x_1 , x_2 , x_3 ..., x_p , and the rows represent n different cases.

In this study, the Hierarchical Cluster Analysis Method was employed to analyze the research data. This analysis involves the use of the following equation to calculate the distances between pairs of observations, as follows[10]:

Where:

d_{ii}: The distance between observation i and observation j

 X_i : The value of variable \boldsymbol{x}_k for observation i

 X_j : The value of variable x_k for observation j

Before calculating the distances, the original variables are transformed into standardized variables. In doing so, all p variables are treated equally in terms of their importance when determining the significance of the distances[2–4].

2.3 Research Data

The laboratory analysis data for powdered milk were obtained from the Quality Control Department / Ministry of Trade, where 20 random samples of powdered milk were collected and analyzed in the laboratory to estimate the following components for each sample : moisture content, ash content, acidity level, fat content, and total solids content, as presented in Table (1).

Models	No.	X1	X2	X3	X4	X5
NIDO s	01	3.10	5.50	1.50	29.50	96.90
Kikoze3	02	1.14	3.10	0.87	23.20	98.80
NIDOw	03	2.40	5.60	1.60	26.50	97.60
NIDOr	04	2.50	5.50	1.50	30.10	97.50
Kikoze 4	05	1.20	3.01	0.78	19.70	98.80
Full Fat Milkz	06	2.70	5.90	1.50	33.10	97.30
Blarconz	07	3.10	3.10	2.40	21.80	96.90
NIDOv	08	2.60	5.60	1.50	30.80	97.40
Carnationy	09	2.29	5.53	1.48	30.50	97.71
NIDOb	10	2.50	5.60	1.35	27.01	97.50
Sunny Boy11	11	2.20	5.70	1.39	31.70	97.80
Milak3	12	2.80	5.60	1.17	30.10	97.20
Blarcon5	13	2.22	3.15	2.70	20.70	97.78
Gloria6	14	2.40	5.70	1.44	28.20	97.60
Milak34	15	1.86	5.71	1.28	29.50	98.14
Milak22	16	1.61	5.87	1.34	27.80	98.36
NIDOo	17	2.51	5.54	1.30	30.30	97.50
Blarconp	18	2.03	3.11	2.32	23.70	97.97
Blarconm	19	2.23	3.10	2.40	23.10	97.77
NIDOj	20	1.70	5.55	1.39	30.70	98.30

Table 1. Laboratory Analysis Results of Milk Components for 20 Random Samples

 X_2 Ash content:

X₃ Acidity level:

X₄ Fat content:

X5 Total solids content:

3. Results

The following are the results obtained from the SPSS software.

Parameter	Society	Element	Latent root	Proportion of variance	Total variance explained
x11	1.001	1	2.70829	52.31	53.02
x22	1.010	2	2.03090	39.01	92.01
x33	1.002	3	0.31011	6.10	98.01
x44	1.003	4	0.10516	2.01	100.12
x55	1.004	5	0.00031	0.001	100.01

Table 2.	Initial	Output	Results	from	the	SPSS	System
							~

From the results shown above, there are two values out of the five eigenvalues that are greater than one. According to the rule of thumb, these two values are therefore taken into consideration. Accordingly, the unrotated factor model is as

follows:

X1 = (0.91545)F1 + (0.60866)F2 + el	(1.08111)	
X2 = (0.76959) F1 + (0.57140) F2 + e2	(0.91876)	
X3 = (0.02103) F1 + (0.90558) F2 + e3	(0.82052)	(4)
X4 = (0.92382) F1 + (0.56661) F2 + e4	(0.92463)	
X5 = (0.91632) F1 + (0.51720) F2 + e5	(0.95353)	
_		

The values in parentheses represent the communalities, and for example, the communality for variable X1 is :

$$(0.83795)^2 + (-0.51788)^2 = 0.97$$

It can be observed that the communalities are relatively high, indicating that most of the variance in variables X1 through

X5 is explained by two common factors. Factor loadings greater than 0.50 (regardless of sign) and underlined represent strong associations, illustrating the degree to which the variables are related to the factors.

It is noticeable that all variables are jointly accounted for by both factors (Factor 1 and Factor 2), except for variable X3, which is explained only by the second factor.

This justifies proceeding with rotation, in order to obtain, as much as possible, simpler and more interpretable factors. After applying rotation (using Varimax rotation with Kaiser normalization), the following model was obtained:

$$X1 = (0.95361)F1 + (0.18372)F2 + e1$$

$$X2 = (0.18453)F1 + (0.94323)F2 + e2$$

$$X3 = (0.57656)F1 + (0.65434)F2 + e3$$

$$X4 = (0.27756)F1 + (0.93324)F2 + e4$$

$$X5 = (0.95434)F1 + (0.16245)F2 + e5$$

....(5)

The communalities remain unchanged, while the factors remain uncorrelated, making this model better than the previous one. This is because only variable

X3 is influenced by both factors (Factor 1 and Factor 2), whereas the remaining variables are explained either by Factor 1 or Factor 2 alone.

Factor 1 shows high positive loadings for variable

X1 (moisture content) and variable

X3 (acidity), and a high negative loading for variable

X5 (total solids content).

Meanwhile, Factor 2 has high positive loadings for variables

X2 (ash content) and

X4 (fat content), and a high negative loading for variable

X3 (acidity).

Using Equation (2), the factor scores are calculated based on the standardized data values (i.e., the original variables X are converted into standardized values), resulting in the following outcomes:

No.	F1	F2
01	1.30	0.51
02	-2.20	-0.74
03	0.25	014
04	0.38	0.54
05	-2.23	-0.91
06	0.74	0.99
07	1.71	-1.72
08	0.55	0.65
09	0.06	0.60
10	0.31	0.84
11	-0.19	0.36
12	0.65	0.82
13	0.52	-2.06
14	0.17	0.46
15	-0.74	0.68
16	-1.09	0.51
17	0.27	0.71
18	0.04	-1.52
19	0.39	-1.63
20	-0.1	0.66

Table 3. Factor Scores for Factors F1 and F2 for 20 Random Samples

Factor 1 primarily reflects an increase in moisture and acidity, along with a decrease in total solids content. In contrast, Factor 2 is associated with an increase in ash and fat content and a decrease in acidity.

When examining the new data matrix, we find that it contains only two factors, which replace the original variables for all samples.

In Factor 1, samples (1) and (7) have high positive scores, indicating that these samples are characterized by increased moisture and acidity levels, accompanied by a decrease in total solids content.

As for Factor 2, which ranks second in importance compared to Factor 1, samples (7), (13), (18), and (19) show high negative scores, indicating a decrease in ash and fat content, along with an increase in acidity.

When comparing, for example, sample (7) from the new data matrix (Table 2) with the original data matrix (Table 1), we observe that Factor 1 indicates an increase in both moisture and acidity, along with a decrease in total solids content. Referring to the original matrix for sample (7), there is a notable increase in moisture (X1) and a clear increase in acidity (X3) compared to the other variables, accompanied by a decrease in total solids (X5) relative to the other values.

For the same sample, Factor 2 indicates a decrease in both ash and fat content, with a corresponding increase in acidity — both factors agree on the increase in acidity.

Comparing this result with the original data matrix for the same sample confirms a decrease in ash content (X2) and a decrease in fat content (X4), along with a relative increase in acidity (X3) when compared to the other variables.

The Applied Aspect of Cluster Analysis

First: The Method Used

The Hierarchical Cluster Method was used, which involves constructing a tree-like structure or dendrogram. Within this hierarchical clustering approach, there are two main methods[11]:

- **1.** Agglomerative Method
- 2. Divisive Method

In the agglomerative method, each observation (or element) initially starts in its own individual group. In the next step, the two most similar groups are merged into a new cluster, thereby reducing the total number of groups at each stage. This process continues iteratively until all elements are combined into a single large cluster[12].

On the other hand, the divisive method works in the opposite direction. It starts with a single large group that contains all observations (or elements) [13]. In each subsequent stage, the most dissimilar observations are split off into smaller groups. This process continues until each observation stands alone in its own individual group [5], [6].

Second: Software Used

The prepackaged software SPSS under Windows, version 20, was used on a personal computer compatible with IBM systems.

Third: Research Data

The same dataset used in the factor analysis section was also employed in this part of the analysis.

Fourth: Research Results

After running the program, the following results were obtained:

Phase	Merged cluster(s)		
	Group 1	Group 2	
1	4	16	
2	5	8	
3	5	9	
4	4	10	
5	19	18	
6	5	13	
7	15	15	
8	2	3	
9	14	21	
10	4	11	
11	4	13	

Table 4. Clustering Results from the Cluster Analysis of 20 Different Powdered Milk Samples

12	2	14
13	7	13
14	2	18
15	2	5
16	1	6
17	2	4
18	1	4
19	2	1

Table (2) illustrates the clustering of different powdered milk samples. The table consists of three columns:

• The first column (from the left) represents the stage number, ranging from 1 to 19, corresponding to the stages at which the 20 powdered milk samples were merged.

• The second and third columns show the clusters combined at each respective stage.

From this table, the following results can be interpreted:

- Sample (4) clusters with Sample (17) at stage 1, with Sample (9) at stage 2, with Sample (8) at stage 3, and with Sample (12) at stage 6.
- Sample (3) clusters with Sample (11) at stage 4, and with Sample (14) at stage 11.
- Sample (2) clusters with Sample (18) at stage 14, and Sample (14) clusters with Sample (16) at stage 7.
- Sample (1) clusters with Sample (4) at stage 8, with Sample (15) at stage 12, with Sample (6) at stage 15, with Sample (3) at stage 17, and with Sample (2) at the final stage 19.
- Sample (7) clusters with Sample (13) at stage 13.

Sample (2) clusters with Sample (18) at Stage 14, then with Sample (7) at Stage 16, and finally with Sample (5) at Stage 18.

There is a strong correlation between Table (2) and Table (3), as Table (2) helps in interpreting Table (3).

Table (3) consists of rows and columns. The rows represent the stages (a total of 19 stages), while the columns total 39. There are two types of columns:

- The first type includes the individual samples (1–20).
- The second type consists of columns filled with one or more X marks, which appear along the column according to the stage at which clustering occurs. The length of each X-filled column depends on the stage in which the clustering takes place—the earlier the stage, the longer the column, and vice versa.

To interpret Table (3), one must refer to Table (2), as explained below:

- Sample (4) clusters with Sample (17) at Stage 1. In Table (3), this is represented by a column between the two samples that is fully filled with X marks, indicating this clustering event.
- According to Table (2), Sample (4) then clusters with Sample (9) at Stage 2. In Table (3), the column between Samples (9) and (17) contains X marks in all cells except the last, indicating the second stage of clustering.
- Sample (4) also clusters with Sample (8) at Stage 3, which is reflected in Table (3) by the column between Samples (8) and (9), signifying that Sample (4) joins with Sample (8) at this stage.
- As shown in Table (2), at Stage 4, Sample (3) clusters with Sample (11). In Table (3), this is represented by the column between Samples (3) and (11), confirming the clustering between these two samples.
- Similarly, for Samples (18) and (19), Table (2) shows that they cluster together at Stage 5. This is represented in Table (3) by the column located between Sample (18) and Sample (19), which illustrates the clustering process between these two samples.
- Sample (4) clusters with Sample (12) at Stage 6, as shown in Table (2). This is reflected in Table (3) by the column between Sample (8) and Sample (12), indicating the clustering process between Samples (4) and (12) at this stage.
- Sample (14) clusters with Sample (16) at Stage 7, as indicated in Table (2). This is represented in Table (3) by the column between Samples (14) and (16), which shows the clustering that occurred between them at this stage.

- Sample (1) clusters with Sample (4) at Stage 8, as shown in Table (2). This is clearly represented in Table (3) by the column between Samples (1) and (4), indicating the clustering process that occurred at this stage.
- Sample (15) clusters with Sample (20) at Stage 9, as shown in Table (2). This is reflected in Table (3) by the column between Samples (15) and (20), representing the clustering process that took place at this stage.
- Sample (6) clusters with Sample (10) at Stage 10, according to Table (2). This is shown in Table (3) by the column between Samples (6) and (10), representing the clustering between the two samples.
- Sample (3) clusters with Sample (14) at Stage 11, as indicated in Table (2). This is explained in Table (3) by the column between Samples (11) and (14), which visualizes the clustering of Sample (3) with Sample (14).
- Sample (1) clusters with Sample (15) at Stage 12, as shown in Table (2). This is reflected in Table (3) by the column between Samples (12) and (15), indicating the clustering of Sample (1) with Sample (15) at this stage.
- Sample (7) clusters with Sample (13) at Stage 13, according to Table (2). This is confirmed by the column in Table (3) located between Samples (7) and (13), illustrating their clustering at this stage.
- Sample (2) clusters with Sample (18) at Stage 14, as indicated in Table (2). This is represented in Table (3) by the column between the two samples, which illustrates the clustering process at this stage.
- Sample (1) clusters with Sample (6) at Stage 15, as shown in Table (2). This is interpreted in Table (3) by the column located between Samples (20) and (6), representing the clustering that occurred between Samples (1) and (6) at this stage.
- Sample (2) clusters with Sample (7) at Stage 16, as indicated in Table (2). This is explained in Table (3) by the column between Samples (7) and (19), representing the clustering between Samples (2) and (7).
- Sample (1) clusters with Sample (3) at Stage 17, as shown in Table (2). This is reflected in Table (3) by the column between Samples (3) and (10), representing the clustering between Samples (1) and (3).
- Sample (2) clusters with Sample (5) at Stage 18, as shown in Table (2). This is represented in Table (3) by the column between Samples (5) and (13), indicating the clustering between Samples (2) and (5).
- Sample (1) clusters with Sample (2) at the final stage, Stage 19, as indicated in Table (2). This is represented in Table (3) by the column between Samples (16) and (12), illustrating the final clustering between Samples (1) and (2) at this stage.



• Figure 2. represents 20 different powdered milk samples classified using both factor analysis and cluster analysis.

Based on Tables (2) and (3), Dendrogram (1) can be constructed. This dendrogram is divided into three groups, as follows: Group One:

Includes Samples 1 to 10, from right to left: 1 (NIDO), 4 (NIDO), 17 (NIDO), 9 (Carnation), 8 (NIDO), 12 (Milak), 15 (Milak), 20 (NIDO), 6 (Full-fat milk), and 10 (NIDO).

It is observed that this group consists of full-fat milk samples.

Group Two:

Includes Samples 3 to 16, from right to left: 3 (NIDO), 11 (Sunny Boy), 14 (Gloria), and 16 (Milak).

This group also consists of full-fat milk samples.

Group Three:

Includes Samples 2 to 5, from right to left: 2 (Kikoz 1), 18 (Blarcon), 19 (Blarcon), 7 (Blarcon), 13 (Blarcon), and 5 (Kikoz 1). This group is composed of semi-skimmed milk samples.

4. Conclusions

The conclusions are divided into three parts:

First: Conclusions from Factor Analysis

Based on the results obtained in this study, it was found that two factors explained Accounting for 92.3% of the total variance among the set of variables utilized in the research. These factors are:

1. First Factor:

This factor accounts for 53.2% of the total variance, making it the primary factor in explaining the correlation matrix among the variables. The results of this factor show high loadings for the variables: moisture content, acidity, and total solids content, with loading values of +0.96, +0.59, and -0.96 respectively (see Equation 5).

The interpretation of these results reveals an inverse relationship between moisture content and total solids content in milk. As moisture increases, total solids decrease, and vice versa.

There is a direct relationship between moisture and acidity, where increased acidity leads to increased moisture due to the conversion of lactose into lactic acid Furthermore, there is an inverse relationship between acidity and total solids, as higher acidity results in higher moisture content, which in turn reduces the total solids in the milk.

2. Second Factor:

This factor is second in importance, explaining 38.7% of the total variance. The loadings for this factor show high values for the following variables: +0.94 (fat), -0.69 (acidity), +0.92 (ash) (see Equation 5).

From these results, a direct relationship is observed between fat content and ash content in milk. As fat content increases, ash content also increases due to the presence of various organic and inorganic substances in milk fat.

Second: Conclusions from Cluster Analysis

The results obtained from this study revealed the presence of three groups resulting from the cluster analysis of the 20 powdered milk samples, as follows:

- Group One: Included full-fat milk samples
- Group Two: Also included full-fat milk samples
- Group Three: Included semi-skimmed milk samples

These findings provide guidance to milk producers in classifying milk into full-fat and semi-skimmed types, which in turn helps consumers choose according to their preferences and health standards.

- The first and second groups together contained two-thirds of the total number of full-fat milk samples.
- The third group contained one-third of the total number of semi-skimmed milk samples.
- Group One included the largest number of samples (10 powdered milk samples), with 60% of them being NIDO milk samples, indicating a high level of homogeneity in this group, which contained half of the total samples.
- Group Three included the second largest number of samples (6), of which 66.66% were Blarcon milk samples, indicating a significant concentration of this brand in the semi-skimmed group.

Third: Combined Conclusions from Both Analyses

The results from factor analysis and cluster analysis were found to be consistent, as illustrated in Figure (1) and Figure (2). Cluster analysis categorized the samples into three distinct groups, while factor analysis revealed a strong similarity between Groups One and Two, which is consistent with reality, as both groups were composed of full-fat powdered milk samples.

Group Three, on the other hand, consisted of semi-skimmed milk samples.



References

- [1] M. M. Ali, T. S. Jawad, and S. N. A. Al-Shabeeb, *Principles of Dairy Science*, 1st ed. Mosul: Dar Al-Kutub Printing House, University of Mosul, 1980.
- [2] B. F. Manly, *Multivariate Statistical Methods: A Primer*. Chapman & Hall, Ltd., 1986.
- [3] M. Kendall, *Multivariate Analysis*. London: Charles Griffin and Company, 1975.
- [4] N. Fieller, "Further Multivariate Analysis: Working Notes," NRJF, University of Sheffield, 2001.
- [5] J. F. Hair, R. E. Anderson, and R. L. Tatham, *Multivariate Data Analysis with Readings*, 1987.
- [6] M. Sollenborn, *Clustering and Case-Based Reasoning for User Stereotypes*. Citeseer, 2004.
- [7] A. A. Ansseif, A. N. Sabti, and Z. A. Mohammed, "Determining the weights of integer linear programming model using the Analytic Hierarchical Process (AHP)," *J. Prod. Ind. Eng.*, vol. 5, no. 2, pp. 48–58, 2024.
- [8] B. F. Mohammed, "The use of quantitative techniques and their role in the administrative decision-making process in organization," *J. Prod. Ind. Eng.*, vol. 3, no. 2, pp. 21–30, 2021.
- [9] S. Mathur and S. R. Pulivarti, "Improvements in productivity and quality for ductile iron flange castings using simulation technique," *J. Prod. Ind. Eng.*, vol. 3, no. 2, pp. 16–20, 2023.
- [10] H. K. Dubey, M. P. Singh, M. Giripunje, and J. P. Modak, "Logarithmic linearization approach for determining the torque required for plastic (PVC) shredding operation through human powered flywheel motor (HPFM)," J. Prod. Ind. Eng., vol. 3, no. 2, pp. 31–41, 2022.
- [11] A. B. Rajbhoj and A. D. Diwate, "Process parameter optimization of sheet metal arc welding of AISI 1020 mild steel," *J. Prod. Ind. Eng.*, vol. 1, no. 1, pp. 1–11, 2020.
- [12] S. M. Salih and G. K. Joad, "Designing a simulation model for the milk carton filling," *J. Ilmiah Matematika*, *Kebumian dan Angkasa*, vol. 3, pp. 161–171, 2025.
- [13] A. Hamad and M. Zaid, "The effect of organizational justice on developing employee loyalty in tourism organizations: An applied study in the Iraqi Tourism Authority," *Enterprenuership J. Finance Bus.*, vol. 6, no. 2, pp. 156–170, 2025. doi: 10.56967/ejfb2025525.