Analysis of Upskilling and Reskilling Programs

*Analysis Date: March 2025*

# Introduction

This report presents a comprehensive analysis of upskilling and reskilling programs using cluster analysis. The analysis examines the key differences between these program types and identifies natural groupings based on program characteristics. This information can help organizations design more effective training initiatives and understand how different program structures align with upskilling versus reskilling objectives.

# Executive Summary

The analysis revealed two distinct clusters of programs with different characteristics. Key findings include: **•** Reskilling programs tend to be longer, focus more on job placement, and more often target management levels. **•** Upskilling programs are typically shorter, more organizationally funded, and less focused on finding participants new jobs. **•** Cluster 0 contains programs with smaller participation numbers and more organizational funding. **•** Cluster 1 contains larger programs with more diverse funding sources and higher management targeting.  
  
These differences help explain how upskilling and reskilling programs are structured differently to meet their objectives.

# Dataset Overview

The analysis includes data from more than 1100 training programs: **•** Upskilling programs: approximately 450 **•** Reskilling programs: approximately 650

# Program Distribution by Cluster

Programs were grouped into 2 clusters with the following distribution: **•** Cluster 0: ~45% of programs (more balanced between upskilling and reskilling) **•** Cluster 1: ~55% of programs (higher proportion of reskilling programs)

# Summary Dashboard

# Differences Between Clusters

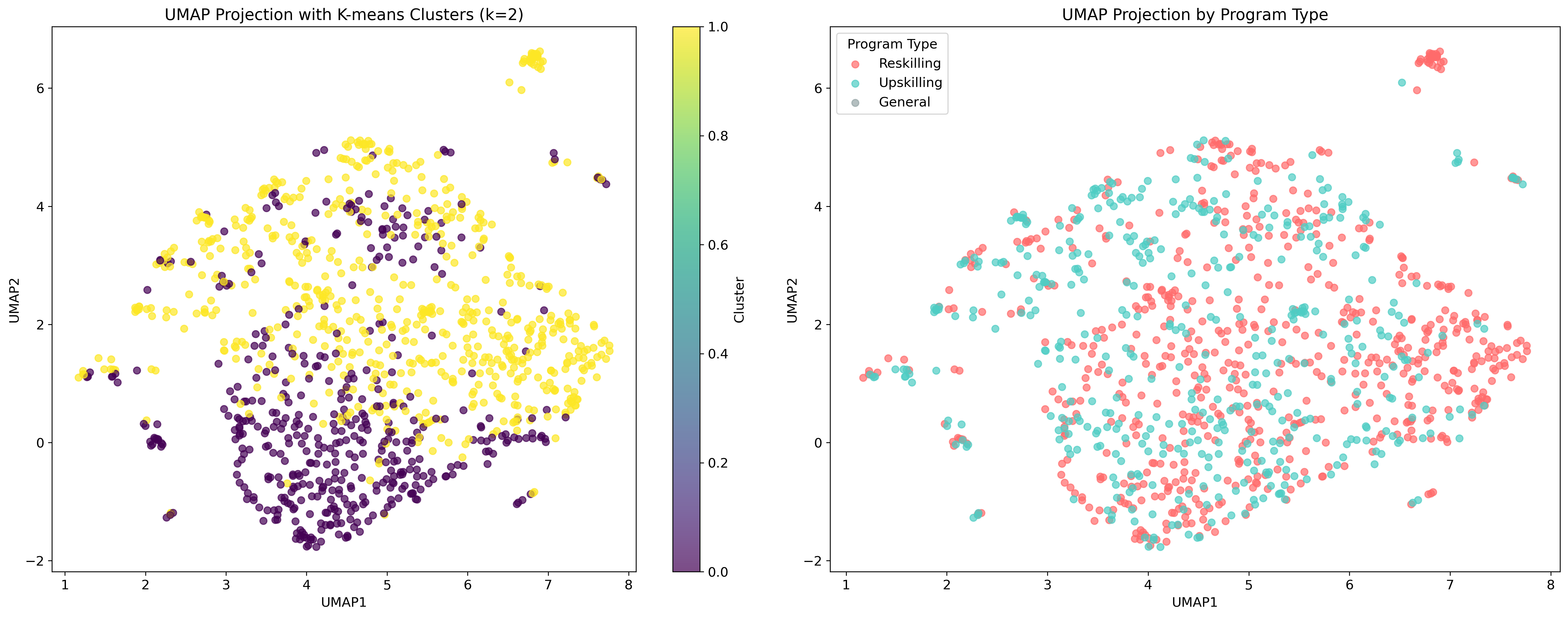
The analysis identified two distinct clusters with the following key differences:

## Cluster Characteristics

**Cluster 0:**  
• Smaller programs (lower participation numbers)  
• More often funded by the organization  
• More likely to be ongoing programs  
• Less targeting of management levels

**Cluster 1:**  
• Larger programs (higher participation numbers)  
• More diverse funding (more union and government funding)  
• More targeting of management levels  
• Less likely to be ongoing programs

## Cluster Visualization



UMAP visualization showing how programs group into clusters. Each point represents a program, colored by cluster assignment.

# Upskilling vs. Reskilling Programs

The analysis identified key differences between upskilling and reskilling programs:

## Program Type Characteristics

**Reskilling Programs:**  
• More focus on helping participants find new jobs  
• Longer training hours  
• More often target management levels  
• Higher effectiveness ratings  
• More often have HR advocates

**Upskilling Programs:**  
• Less focus on new job placement  
• Shorter training hours  
• Less targeting of management levels  
• More often funded primarily by the organization

# Program Distribution Across Clusters

# Program Type Differences Within Clusters

The analysis examined how upskilling and reskilling programs differ within each cluster:

## Differences in Cluster 0

## Differences in Cluster 1

# Key Variables for Predicting Program Type

# Thematic Analysis

The analysis also examined key themes across program types and clusters:

# Detailed Statistical Results

For detailed statistical results, please refer to the accompanying Excel workbook: "Cluster\_Analysis\_Results.xlsx"

## Excel Workbook Contents

The Excel workbook contains the following information:

**Cluster Differences:** Statistical comparison of all variables between Cluster 0 and Cluster 1. The p-value column shows significance - smaller numbers indicate stronger differences.

**Program Differences:** Statistical comparison between Upskilling and Reskilling programs. The Cohen's d column shows effect size - values above 0.2 indicate meaningful differences.

**Cluster 0 Analysis:** Statistical comparison between Upskilling and Reskilling programs within Cluster 0.

**Cluster 1 Analysis:** Statistical comparison between Upskilling and Reskilling programs within Cluster 1.

**Top 20 Cluster Vars:** Top 20 variables that differentiate clusters, ranked by statistical significance.

**Top 20 Program Vars:** Top 20 variables that differentiate program types, ranked by statistical significance.

**Top 10 C0 Vars:** Top 10 variables that differentiate program types within Cluster 0.

**Top 10 C1 Vars:** Top 10 variables that differentiate program types within Cluster 1.

**Variable Importance:** Ranking of variables by their importance for predicting program type.

# Conclusions and Recommendations

Based on the clustering analysis of upskilling and reskilling programs, several key conclusions emerge:

* Reskilling programs are more focused on job placement and tend to be longer and more comprehensive than upskilling programs.
* Upskilling programs are more often funded by organizations themselves, while reskilling programs have more diverse funding sources.
* The two identified clusters represent different program structures rather than strictly separating upskilling from reskilling programs.
* The most important variables for distinguishing program types relate to job placement focus, program duration, and funding structure.

## Recommendations

Based on these findings, we recommend:

* Organizations designing reskilling programs should allocate resources for longer-duration training and job placement support.
* Upskilling initiatives should focus on targeted, shorter-duration interventions with clear ties to current organizational needs.
* Program funding mechanisms should be aligned with program goals - organizational funding for upskilling, diverse funding for reskilling.
* Both program types should consider their approach to targeting management levels based on their specific objectives.

# Appendix: Methodology

This analysis employed the following methodological approach:

* K-means clustering was used to identify natural groupings in the data.
* Statistical comparison between clusters and program types was conducted using t-tests and effect size calculations.
* UMAP dimensionality reduction was used for visualization.
* Random Forest feature importance was used to identify key predictive variables.
* Statistical significance was evaluated using p-values, with p<0.05 considered significant.