**K Means Executive Summary**

Applied Machine Learning for Analytics

Jackson Mulera Gibendi

Nexford University

Professor Nicholas Bucciarelli

August 2024

**Executive Summary**

**Executive Summary: K-Means Clustering of Employee Rating and Salary Data**

**Objective**

The primary goal of this project was to employ K-Means clustering to analyze and segment employee data based on two key metrics: **Rating** and **Salary**. This analysis aimed to identify distinct groups within the workforce to better understand employee performance relative to compensation, and to aid in making informed decisions regarding talent management, compensation strategies, and resource allocation.

**Methodology**

Using a dataset comprising employee ratings and salary information, we implemented a K-Means clustering algorithm to segment the data into coherent groups. The algorithm was executed using Pycharm a Python-based application with a graphical user interface, enabling dynamic interaction with the data, including:

Loading the employee dataset from a CSV file.

Specifying the number of clusters for analysis.

Visualizing the results through scatter plots to observe the grouping based on the standardized values of ratings and salaries.

Utilizing the Elbow Method to determine the optimal number of clusters by plotting the sum of squared distances of samples to their closest cluster center and identifying the point of inflection.

**Key Findings**

**Cluster Characteristics**: The analysis revealed distinct groups where employees with similar ratings and salary ranges were clustered together. This segmentation allowed us to visualize the relationship between how employees are rated and how they are compensated.

**Optimal Clustering**: The Elbow Method suggested a clear elbow point, indicating an optimal number of clusters that balances between too many clusters (overfitting) and too few (underfitting). This optimal number helped to maximize the interpretability and practical utility of the clustering results.

**Insights for HR Management**: By examining the clusters, we were able to identify groups of high performers either being paid at par, above, or below their peer group, which can inform targeted adjustments in compensation, training opportunities, and other HR interventions.

**Applications and Recommendations**

**Compensation Strategy**: Adjust salary structures based on clustering outcomes to ensure fair compensation correlating with employee ratings and industry standards.

**Talent Management**: Focus on clusters indicating high performance but lower salary to improve employee retention strategies.

**Performance Improvement**: Identify underperforming clusters for targeted training and development programs.

**Resource Allocation**: Allocate resources more effectively by identifying key groups of employees that may benefit from additional support or investment.

**Conclusion**

The application of K-Means clustering to employee rating and salary data provided significant insights into the workforce dynamics of the organization. It facilitated a data-driven approach to understanding complex relationships within the employee base, thereby supporting more informed decision-making in HR practices. The findings from this project are expected to contribute towards more strategic HR management, enhancing overall organizational performance and employee satisfaction.

**References**

Mulera, J. (2024). Executive Summary: *K-Means Clustering of Employee Rating and Salary Data* [*https://github.com/JacksonMulera/Module-4-Assignment-K-Means-Python-Application/new/main*](https://github.com/JacksonMulera/Module-4-Assignment-K-Means-Python-Application/new/main)