



Individual Coursework Submission Form

Specialist Masters Programme

Surname: Kwon	First Name: Minkyong
MSc in: Business Analytics	Student ID number: 240007680
Module Code: SMM750	
Module Title: Digital Technologies and Value Creation	
Lecturer: Philippe Blaettchen	Submission Date: 09/04/2025
<p>Declaration:</p> <p>By submitting this work, I declare that this work is entirely my own except those parts duly identified and referenced in my submission. It complies with any specified word limits and the requirements and regulations detailed in the coursework instructions and any other relevant programme and module documentation. In submitting this work, I acknowledge that I have read and understood the regulations and code regarding academic misconduct, including that relating to plagiarism, as specified in the Programme Handbook. I also acknowledge that this work will be subject to a variety of checks for academic misconduct.</p> <p>We acknowledge that work submitted late without a granted extension will be subject to penalties, as outlined in the Programme Handbook. Penalties will be applied for a maximum of five days lateness, after which a mark of zero will be awarded.</p>	
Marker's Comments (if not being marked on-line):	

Deduction for Late Submission:

Final Mark:

 %

Task 3. Defining RCC Groups

To offer RCCs effectively, it is essential to create objectively defined groups rather than targeting individuals directly. Group formation should be based on transparent and data-driven criteria to ensure fairness and consistency.

Defining Groups:

Groups can be formed using a combination of factors that significantly influence RCC acceptance. Key variables include:

- **Likelihood of Accepting RCC:** Calculated using predictive models like Logistic Regression or Gradient Boosting.
- **Employment Factors:** Including tenure (YearsAtCompany), job satisfaction, overtime status, and monthly income.
- **Cluster Analysis:** Applying clustering techniques (like K-Means) within each department to group employees based on similarities in the identified key factors.

Grouping Strategies:

1. **Homogeneous Grouping:** Employees with similar RCC acceptance probabilities and similar work-related factors (e.g., job role, department) are clustered together.
2. **Likelihood-Based Grouping:** Employees are divided into groups based on their calculated likelihood of accepting the RCC, such as high, medium, and low likelihood groups.
3. **Cluster-Based Grouping:** Using K-Means clustering to form groups within each department, considering variables that directly impact RCC acceptance.

Pros and Cons of Group Sizes:

- **Smaller Group Sizes:**
 - Pros: Better alignment with individual preferences, allowing for more targeted offers.
 - Cons: Increased complexity in group management, potentially leading to inconsistencies. Can appear discriminatory if groups are too narrowly defined.
- **Larger Group Sizes:**
 - Pros: Easier management and communication, more uniformity within each group.
 - Cons: Reduced precision in estimating RCC acceptance rates, increasing the risk of acceptance variability. Higher risk of losing essential employees from critical roles.

Ensuring Non-Discriminatory Grouping:

To ensure fairness, it is crucial to exclude variables that may directly or indirectly lead to discrimination, such as age, gender, or marital status. Instead, groups should be based solely on job-related factors. Additionally, statistical fairness tests, such as disparate impact analysis, can be applied to evaluate whether the grouping criteria disproportionately affect any demographic subgroup.

Task 4. Optimising Severance Cost Reduction Using Excel Solver

4.1 Group Definition Strategy

To form objectively defined groups, clustering techniques are applied using the top three features most positively and negatively associated with RCC acceptance. The clustering approach should use non-discriminatory and relevant features identified through data analysis and feature importance studies. Based on the data and results shared, some key factors that influence RCC acceptance are:

High Influence Factors (Positive for RCC Acceptance):

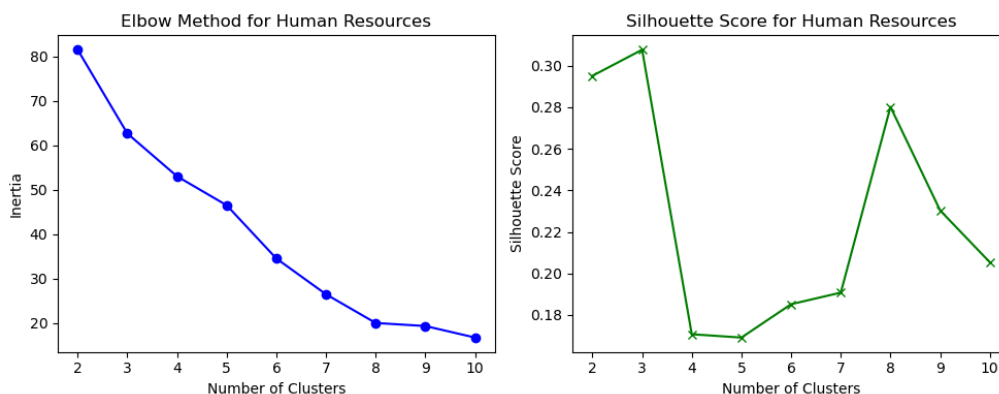
- OverTime
- NumCompaniesWorked
- YearsAtCompany

Low Influence Factors (Negative for RCC Acceptance):

- StockOptionLevel
- JobInvolvement
- TotalWorkingYears

Using these factors, clusters can be created using **K-Means clustering**. The number of clusters can be optimized using the **Elbow Method** and **Silhouette Scores**. The analysis has shown the optimal number of clusters for each department:

- **Human Resources: 3**

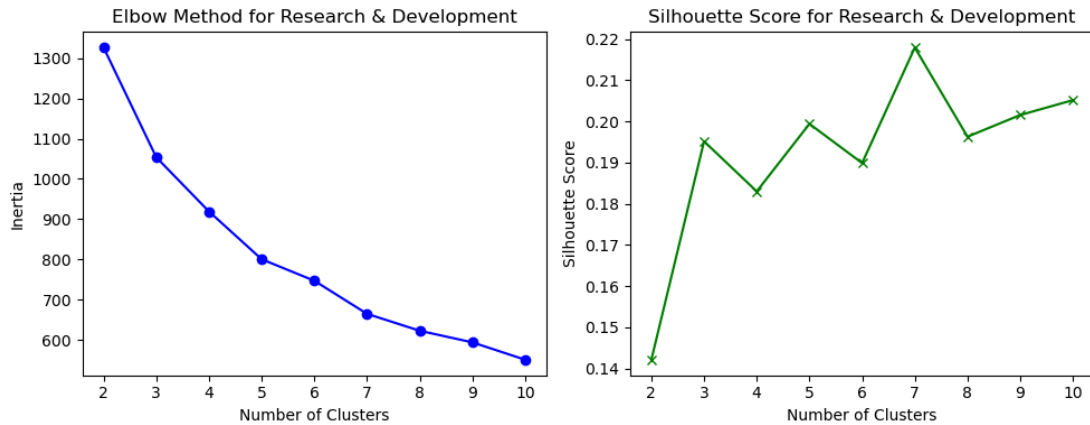


Plot 1. Elbow Method and Silhouette Score for Human Resources

Elbow Method: The inertia shows a significant drop until K=4, after which the rate of decrease slows, suggesting diminishing returns.

Silhouette Score: The score peaks at K=3, indicating the best-defined cluster structure.

- **Research & Development: 7**

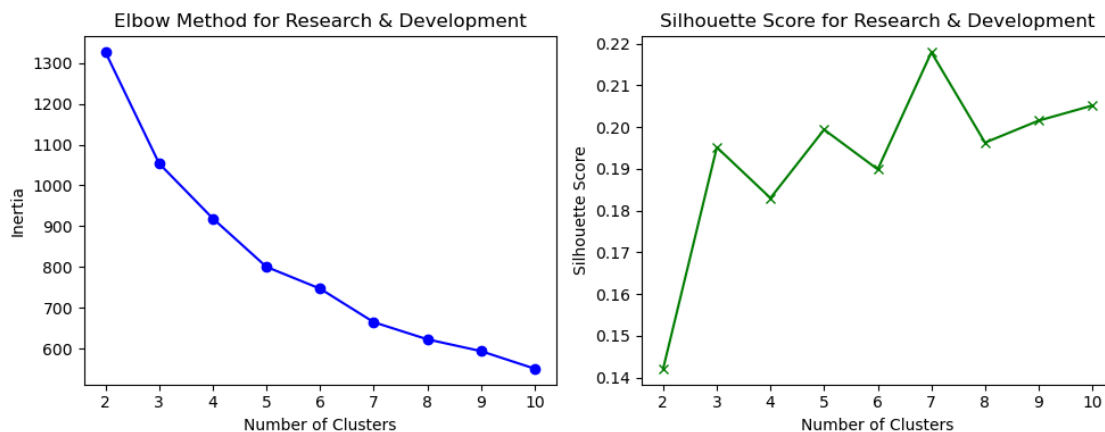


Plot 2. Elbow Method and Silhouette Score for Research & Development

Elbow Method: The elbow point appears around K=6 to K=7, where the inertia curve starts to flatten.

Silhouette Score: The highest value is at K=7, indicating this number provides better separation between clusters than others.

- **Sales: 7**



Plot 3. Elbow Method and Silhouette Score for Sales

Elbow Method: A noticeable bend occurs around K=7, beyond which inertia reduces more gradually.

Silhouette Score: The peak is at K=7, marking this as the most distinct and coherent clustering.

Validating Non-Discriminatory Clustering



Plot 4. Distributions of Gender and Marital Status in Clusters of Research and Development

Plot 4 demonstrates the distribution of gender and marital status across the six R&D clusters. The balanced and consistent distribution indicates that employees are not clustered based on personal characteristics, such as gender or marital status. This confirms that the clustering process is objective and free from demographic biases.

4.2 Implementation Using Excel Solver

Model Components:

There are 17 groups in total - 3 HR, 7 R&D, 7 Sales.

1. Objective Function:

Minimize the total severance cost, calculated as the sum of the severance costs of employees selected for RCC.

- YearsAtCompany is below 10:

$$\text{Severance Cost} = \frac{1}{4} \times \text{YearsAtCompany} \times \text{Monthly Income}$$

- YearsAtCompany is more than 10:

$$\text{Severance Cost} = \frac{1}{3} \times \text{YearsAtCompany} \times \text{Monthly Income}$$

2. Decision Variables:

The number of employees who accept RCC within the group that has received RCC offer.

3. Constraints:

- Number of employees leaving

$$\sum_i \text{GroupSize}_i \times \text{RCCAcceptance Likelihood}_i \times \text{RCCOffer}_i \geq 40$$

- Salary reduction

$$\sum_i \text{Number of Employees Accepting RCC}_i \times \text{AverageMonthlyIncome}_i \times 12 \geq \text{€ 3 million}$$

- Remaining employees in each department (Human Resources, Sales, Research and Development) $\geq 80\%$ of the initial headcount.

Excel Solver Configuration:

1. **Objective Cell:** Minimize total severance cost (N25).
2. **Changing Variable Cells:** RCC offers (1: Yes, 0: No) for each group (G4:G20).
3. **Constraints:**
 - Total salary reduction constraint (B30).
 - Minimum leaving employees constraint (B33).
 - Departmental balance constraint (B36:B38).

4.3 Result

Metric	Output	Constraint	Constraint Met?
Salary Cut	€ 3,153,160	\geq € 3,000,000	Yes
Total Attrition	43.38	≥ 40	Yes
HR Attrition	0	≤ 4.2	Yes
R&D Attrition	22.25	≤ 57	Yes
Sales Attrition	21.13	≤ 27	Yes
Total Severance Cost	€ 295,907	–	–

Table 1. Optimization Result Summary

Table 1 shows the result of optimization using excel solver. All the constraints are met while minimizing the total severance cost.

Task 5. Prediction-and-Optimization Approach

Pros:

- **Data-Driven Decision Making:** The approach leverages data analysis and predictive modeling to make more informed decisions about which employees are most likely to accept RCC. By using data patterns and statistical insights, it reduces uncertainties compared to manual decision-making, making the process more systematic and evidence-based.

- **Cost Efficiency:** The model aims to minimize severance expenses by identifying employees with a higher probability of accepting RCC. This targeted approach ensures that financial resources are efficiently allocated, helping achieve salary reduction targets without incurring excessive costs.
- **Fairness and Consistency:** By defining objective, consistent groups based on relevant criteria, the approach ensures transparency and uniformity in decision-making. This consistency helps prevent subjective biases and maintains fairness when offering RCC to employees across different departments.
- **Legal and Compliance Assurance:** Implementing a data-driven, objective method helps mitigate the risk of legal and compliance issues. By relying on consistent criteria and avoiding subjective decision-making, the approach aligns with fair employment practices and can help demonstrate compliance if decisions are legally scrutinized.
- **Optimization of Resources:** The method strategically allocates RCC offers to meet the salary cut requirements while maintaining a balanced workforce within each department. By considering both financial and organizational stability, it helps preserve operational continuity while achieving cost-cutting goals.

Cons:

- **Model Accuracy Limitations:** Although the predictive model performed reasonably well, models trained on imbalanced datasets are susceptible to overfitting and may misrepresent underrepresented employee profiles. These errors can impact the precision of RCC targeting.
- **Uncertainty in Acceptance:** Despite high predictive probabilities, actual employee decisions may vary significantly due to personal, professional, or unforeseen factors. This unpredictability can result in unexpected outcomes, undermining the reliability of model-driven recommendations.
- **Individual Circumstances:** Data cannot fully reflect qualitative factors like employee morale, interpersonal dynamics, or life changes, which often play a crucial role in resignation decisions. This can lead to unexpected declines in RCC uptake or unintentional offers to employees who were not likely to leave.
- **Complexity in Group Formation:** Forming non-discriminatory and logically consistent groups while maintaining high model accuracy can be challenging. Balancing fairness and practical implementation requires careful consideration of variables to avoid group biases, which may complicate the optimization process.
- **Risk of Indirect Bias:** Even with the exclusion of explicitly discriminatory variables, bias can still creep in through correlated features. For example, 'TotalWorkingYears' is moderately correlated with 'Age' (correlation ≈ 0.65). Its use must be carefully monitored to prevent unintentional age-related discrimination.
- **Talent Loss:** The model's focus is primarily on optimization objectives — salary cuts and predicted RCC acceptance. It may overlook the strategic or institutional value of certain employees. As a result, RCCs could be offered to high performers or critical personnel, leading to loss of key talent if accepted.

Implicit Assumptions:

- **Historical Data Reflects Future Behavior:** A central assumption is that past patterns of employee behavior are indicative of future responses to RCC offers. While historical data provides valuable insights, external changes—such as shifts in the job market, company direction, or economic climate—may limit the generalizability of these patterns. This assumption is moderately likely to hold, but regular model updates with fresh data are essential to maintain accuracy.
- **Model Accuracy is Sufficient:** The approach assumes the predictive model can reliably estimate RCC acceptance probabilities. However, limitations due to imbalanced data, overfitting, or underrepresentation of certain employee profiles can reduce the model's effectiveness. Although validation metrics suggest reasonable performance, this assumption may not always hold, particularly in edge cases.
- **Predicted Probability \approx Actual Acceptance:** It is assumed that predicted acceptance probabilities will translate into actual behavior. For example, a 0.7 probability within a group is expected to mean that 70% of employees in that group will accept the RCC offer. In reality, personal factors—such as health, family circumstances, or relationships with colleagues—can influence decisions in ways that the model cannot capture. This makes the assumption less likely to be perfectly fulfilled, introducing uncertainty into outcomes.
- **Group Formation is Objective and Fair:** The method assumes that clusters are formed using relevant, work-related features rather than personal attributes. Validation using demographic distributions (e.g., gender, marital status) in the R&D department supports the non-discriminatory nature of clustering.
- **Compliance with Legal and Ethical Standards:** By relying on objective data and applying consistent, group-based rules, the process assumes it minimizes legal and ethical risks. This assumption is likely to hold, especially as the method avoids individual targeting and enhances transparency in how RCC offers are allocated.

Overall, while most assumptions are reasonably fulfilled, the uncertainty around personal decision-making and external market conditions means results should be interpreted with caution, and processes should remain flexible to adapt to unexpected outcomes.

Conclusion:

The prediction-and-optimization approach is a practical way to make data-driven decisions, like offering RCC. It helps reduce uncertainties, cut costs, and maintain fairness by using data to guide decisions. However, it has its challenges, like the risk of inaccurate predictions, unexpected individual choices, and potential biases. Also, it assumes that past data will predict future behavior accurately, which isn't always true. While it's a useful tool, it's important to use it carefully, considering its limitations and regularly checking its accuracy.

Appendix

- Output Using Excel Solver

Data					Decision variable		Constraints				Objective
Department	Average Severance Cost	Average RCC Acceptance Likelihood	Average Monthly Income	Group Size	RCC Offer	Salary Cut	Department	Number of Employees Accepting RCC	Total Attrition in Department	Severance Cost	
Human Resources 1	6699.20	0.5763	3964.64	14	0	0	HR	0	0	0	
Human Resources 2	36086.33	0.2298	8996.00	5	0	0		0	0	0	
Human Resources 3	127889.17	0.2562	17813.00	2	0	0		0	0	0	
Research & Development 1	12913.53	0.4655	6069.76	21	0	0	R&D	0	0	0	
Research & Development 2	5555.24	0.4636	4935.81	48	1	1317897.48		22.2506		123607.3366	
Research & Development 3	9439.78	0.2862	5743.86	28	0	0		0	0	0	
Research & Development 4	11936.16	0.5075	5605.49	49	0	0		0	0	0	
Research & Development 5	106482.68	0.2403	13138.85	20	0	0		0	0	0	
Research & Development 6	9504.53	0.3893	4542.35	89	0	0		0	0	0	
Research & Development 7	17276.63	0.3403	11918.90	30	0	0		0	22.2506	0	
Sales 1	10451.27	0.3705	5854.35	17	1	442483.704	Sales	6.2985		65827.32161	
Sales 2	87048.93	0.2765	12984.40	5	0	0		0	0	0	
Sales 3	10482.42	0.3850	5486.96	51	0	0		0	0	0	
Sales 4	18139.82	0.4638	6578.50	20	0	0		0	0	0	
Sales 5	118527.52	0.3298	13198.44	9	0	0		0	0	0	
Sales 6	10082.10	0.3936	11916.50	12	1	675408.154		4.7232		47619.7944	
Sales 7	5823.77	0.4812	5915.62	21	1	717370.558		10.1056	21.1273	58852.68852	
Total Number of Employees					Total Number of Groups with RCC Offers						Total Severance Cost
441					4						295907.14
Constraints											
Salary Cut											
>= 3153159.89 3000000											
Total Attrition											
>= 43.38 40											
Attrition of Each Department Maximum 20% Each Department											
HR	0.00	<=	4.2								
R&D	22.25	<=	57								
Sales	21.13	<=	27								
Department											
HR				21							
R&D				285							
Sales				135							
Total Number of Employees											
441											