instituto
de ingeniería
del conocimiento

Nº
03

20
23

Revista de Casos de Estudio en
HR Analytics

Journal of HR Analytics Case Studies

02. Explorando los factores clave del desempeño en una multinacional del sector químico.

Exploring performance-driving variables: a case study in a chemical multinational firm.

Instituto de Ingeniería del Conocimiento

ISSN: 2792-3770

La Revista de Casos de Estudio en HR Analytics

Journal of HR Analytics Case Studies

La Revista de Casos de Estudio en HR Analytics nace con la misión de facilitar el intercambio de conocimiento especializado entre profesionales y académicos en el ámbito de la **analítica de Recursos Humanos**, con el objetivo de mejorar la efectividad de las organizaciones. La entidad responsable de esta revista es la **Asociación para el Desarrollo de la Ingeniería del Conocimiento** (ADIC), siendo esta publicación on-line editada por el **Instituto de Ingeniería del Conocimiento** (IIC) con una periodicidad de un número anual.

Objetivo

La revista tiene como **objetivo** principal ser un vehículo para la reflexión y la difusión de las **buenas prácticas, últimos avances ylíneas de investigación** en el ámbito de la analítica aplicada para la toma de decisiones sobre la gestión del capital humano en las organizaciones.

La revista tiene un **carácter científico** y una **vocación divulgativa**, por ello propone artículos fundamentalmente de **carácter aplicado**. Con ellos se pretende que los profesionales de las organizaciones accedan a un conocimiento relevante acerca de cómo otras organizaciones desarrollan HRA. Y, también, acercar a los académicos el conocimiento respecto de cómo se desarrolla HRA en la práctica.

Alcance

El **enfoque de la Revista**, que pretende ser **multidisciplinar**, da cabida (entre otros) a manuscritos que: reflejen **casos prácticos** de aplicación del HRA en las organizaciones; que analicen, comparen y relacionen la utilidad de diferentes **técnicas y/o herramientas** para el abordaje de diferentes objetivos analíticos; que planteen y valoren la efectividad de diferentes **metodologías de trabajo** para el desarrollo de proyectos HRA; que ayuden a entender el **mapa de ruta** por el que transitar desde los niveles básicos del HRA hasta los niveles de excelencia; y que en general ayuden a entender cómo **mejorar la efectividad organizacional** a partir de la analítica de datos referidos a la fuerza de trabajo.

ISSN: 2792-3770

La revista está editada por el Instituto de Ingeniería del Conocimiento y tiene los siguientes órganos de gobernanza.

.02

.02

Explorando los factores clave del desempeño en una multinacional del sector químico.

Exploring performance-driving variables: a case study in a chemical multinational firm.

Laura Sanz Pulido

Correo electrónico:
laura.sanzpulido@estudiante.uam.es

Linkedin:
www.linkedin.com/in/ laurasanzpulido

Received: 31 may 2023
Received in revised form: 21 july 2023
Accepted: 8 september 2023
Available on-line: 29 november 2023

ISSN: 2792-3770

Palabras clave
gestión del desempeño, rendimiento laboral, análisis multivariable, bosques aleatorios, aprendizaje automático, modelización predictiva, sector químico, España.

Resumen

Utilizando los datos de una empresa química multinacional ubicada en España, este estudio realiza un análisis multivariable para comprender los factores que influyen en el desempeño de los empleados. En total se analizaron 18 variables organizacionales, incluyendo algunas poco comunes como el teletrabajo y la distancia del domicilio al centro de trabajo. Utilizando Python y aplicando estadísticas descriptivas, análisis de correlación, regresión lineal y modelado de bosques aleatorios, el estudio extrae conclusiones interesantes sobre el impacto de estas variables en el desempeño de los empleados y destaca la importancia de utilizar modelos complejos de aprendizaje automático para predecir resultados, especialmente para variables con poca variación. Los resultados aportan ideas relevantes no solo para la organización, sino también para los profesionales de RRHH y sugieren posibles vías de investigación en el futuro.

Keywords
employee performance, multivariable analysis, random forest modeling, machine learning, predictive modeling, chemical company, Spain.

Abstract

Based on data from a multinational chemical company located in Spain, this study conducts a multivariable analysis to understand the factors that influence employee performance. A total of 18 variables, including tele-working and distance from home to work, were analysed using Python through descriptive statistics, correlation analysis, linear regression, and random forest modelling. The study finds interesting conclusions about the impact of these variables on employee performance and highlights the importance of using complex machine learning models for predicting outcomes, particularly for variables that have low variation. The results provide important insights for practitioners and suggest potential avenues for future research.

ISSN: 2792-3770

1. Introducción

The study of factors that affect employee performance is of great interest to both scholars and practitioners in the field of human resource management. Its importance lies in the fact that high employee performance is crucial for the success of any organization. According to a study by Khan (2012), employee productivity is influenced by several factors including motivation, leadership, training, and technology. Additionally, the COVID-19 pandemic has led to an increase in telecommuting, which has further emphasized the importance of investigating how factors such as the work environment and distance to work affect employee performance. By identifying the variables that have a significant impact on employee performance, organizations can implement effective human resource management strategies to maximize employee productivity and performance, which can have a positive impact on the success and competitiveness of the company.

Previous research has shown that various individual, job-related, and organizational factors can have a significant impact on employee performance (Cerasoli, Nicklin, & Ford, 2014; Tett & Meyer, 1993). However, most studies tend to focus on a few variables, typically no more than ten, and often overlook other potentially important factors.

One such variable that has been relatively understudied in the context of employee performance is distance to work. While previous research has examined the impact of commute time on employee job satisfaction and turnover (Amponsah-Tawiah, Annor, & Arthur, 2016) the impact of distance to work on employee performance has received even less attention. This is surprising given that commuting distance has been shown to have a negative impact on physical and mental well-being (Clark, Chatterjee, Martin, & Davis, 2020) and could potentially affect employee performance through fatigue, stress, and time constraints.

Moreover, due to the growing importance of teleworking worldwide because of the COVID-19 pandemic, this variable has become increasingly important. Many employees have shifted to remote work, which may affect their distance to work and potentially impact their performance.

So, this study aims to identify factors that affect employee performance in a chemical company in Spain, with a particular focus on training and work arrangements, including distance to work and telecommuting. By examining the impact of these variables on employee performance, this study seeks to contribute to the existing literature on factors that influence employee productivity and provide insights for practitioners to optimize human resource management strategies.

Furthermore, the ability to predict employee performance is of utmost importance for the studied chemical company, especially in light of the anticipated retirements and its aging workforce. By compressively understanding the factors that influence performance and leveraging predictive analytics, the organization can proactively plan for succession, identify training needs, and strategically allocate resources to ensure a smooth transition and minimize productivity disruptions. Moreover, predictive analytics enables the company to attract and retain the right talent, bridging the generation gap and fostering a diverse and dynamic workforce. In this way, the company can adapt to the upcoming demographic changes, maintain competitiveness, and achieve long-term success with a highly capable and motivated workforce.

1.1. Research gaps

Although the literature on the determinants of employee performance is extensive, few studies have explored the impact of teleworking and distance from home to work on employee performance in the chemical sector in Spain. This is surprising given the increasing importance of teleworking due to the COVID-19 pandemic and the fact that the chemical industry is one of the most important sectors in Spain, contributing significantly to the country's GDP. Existing studies have mainly focused on other factors such as job satisfaction (Judge, Thoresen, Bono, & Patton, 2001; Khan, Alamdar, Nawaz, Aleem, & Hamed, 2014), leadership style (Kalambayi, Onojaefe, Kasse, & Tengeh, 2021; Tamimi & Sopiah, 2022), and training (Flegl, Depoo, & Alcázar, 2022).

Moreover, previous studies on the determinants of employee performance have mainly used traditional statistical methods, such as correlation analysis and linear regression. Few studies have utilized advanced machine learning techniques, such as random forest modeling, to predict the impact of different variables on employee performance (Motyka, 2018). Machine learning models can better capture non-linear relationships between variables and can predict outcomes more accurately, particularly for variables with low variation.

Regarding studies that use multivariate analyses to analyze employee performance, there are also limited amounts of research compared to studies that focus on a single variable or use a more limited approach. A systematic review study conducted by Motyka (2018) suggest that only 20% of studies on employee performance used multivariate analyses. This suggests that there is a need for more complex studies that analyze multiple factors that influence employee performance to have a more comprehensive understanding of predictors of performance.

Therefore, the aim of this study is to fill these research gaps by examining the effects of teleworking and distance from home to work on employee performance in the chemical sector in Spain through the use of advanced machine learning techniques. However, this study goes beyond that by analyzing a total of 18 variables obtained from HR department data (without the need for survey measurements) which is also an unusual approach. The results of this study could provide valuable insights for practitioners in the chemical industry and contribute to the development of effective human resource management strategies. Additionally, this study could pave the way for future research to explore other determinants of employee performance in the chemical industry in Spain.

1.2. Research question and objectives of the study

The preceding discussion about the research gaps on factors that influence performance steers the course of the present dissertation. Accordingly, the main research question is:

How does training with work arrangements, including distance to work and telecommuting, impact employee performance in a multinational chemical company?

In order to answer the research question and to achieve the objective, this dissertation has both theoretical and empirical objectives. These objectives are:

1. Review the literature related to possible variables that may affect performance.
2. Analyze the variables that we believe may have a relationship with performance from a quantitative standpoint and use both traditional prediction techniques such as regression and machine learning techniques like random forest to identify predictive variables. We will include variables that have been studied less, such as remote work and commute distance.

2. Literature review

Employee performance is a crucial factor in the success of any organization. Over the years, researchers have conducted numerous studies to explore the factors that influence employee performance and how it can be improved. In this section, we provide an overview of current research areas in employee performance and highlight key findings from relevant studies. Additionally, to contextualize this study within the chemical industry, which is the focus of our case study, we will review studies that have investigated the relationship between the variables we are examining and employee performance in this industry.

2.1. Mapping the research landscape of employee performance

Over the years, extensive research has been conducted on various aspects of employee performance, leading to the identification of six broad areas of research:

1. **Performance measurement and evaluation:** This area focuses on developing and evaluating metrics and methods for measuring performance in organizations. Researchers in this area aim to identify objective and reliable performance indicators that can be used to evaluate individual, team, and organizational performance (Koopmans, Bernaards, Hildebrandt, De Vet, & van der Beek, 2013; DeNisi & Murphy, 2017).
2. **Performance feedback and coaching:** This area focuses on providing feedback and coaching to employees to improve their performance. Researchers in this area examine the effectiveness of different feedback and coaching strategies and techniques, as well as the factors that influence their success (Cattell, 2006; Awaysheh, Bonet, & Ortega, 2023).
3. **Performance motivation and engagement:** This area focuses on understanding the factors that motivate and engage employees to perform at their best.

ISSN: 2792-3770

Researchers in this area investigate the role of intrinsic and extrinsic motivation, job satisfaction, organizational culture, and leadership in shaping employee performance (Riyanto, Endri, & Herlisha, 2021; Sugiarti, 2022).

4. **Performance and well-being:** This area focuses on the relationship between performance and employee well-being. Researchers in this area examine the impact of performance demands and stress on employee well-being, as well as the strategies and interventions that can enhance employee resilience and well-being (Iskamto, 2021; Wright, Cropanzano & Bonett, 2007).

5. **Performance and diversity:** This area focuses on how diversity in the workplace affects individual and organizational performance. Researchers in this area investigate the impact of diversity on team dynamics, communication, decision-making, and innovation, as well as the strategies for managing diversity to enhance performance (Joshi & Roh, 2009; Van Knippenberg, Nishii & Dwertmann (2020).
6. **Performance and technology:** This area focuses on the impact of technology on individual and organizational performance. Researchers in this area investigate the role of technology in facilitating communication, collaboration, and knowledge sharing, as well as the potential negative effects of technology on work-life balance and employee well-being (Rasool, Warraich, & Sajid, 2022; Suryadi, FoEh, & Manafe, 2022).

It is important to note that performance is a complex and multifaceted construct that is influenced by a wide range of individual, organizational, and environmental factors. As researchers continue to explore new areas of inquiry, we can deepen our understanding of performance and identify effective strategies for enhancing it in organizations.

2.2. Employee performance in chemical companies

In the context of chemical companies, numerous studies have investigated employee performance, providing valuable insights into the factors that affect performance in this industry. One key finding is the importance of safety culture in enhancing employee performance. Strong safety culture , which means a solid organizational focus on workplace safety, has been linked to higher employee engagement, job satisfaction, and improved performance (Zohar, 2010). Another significant factor that influences employee performance in chemical companies is the work environment, including job autonomy, social support, and job demands, which have been found to be important predictors of employee performance (Abdulkhaliq & Mohammadali, 2019; Karasek, 1979).

Leadership has also been highlighted as a key factor in shaping employee performance in chemical companies. Effective leadership has been linked to higher job satisfaction, better employee engagement, and improved performance (Farid, Kee, Mohamad, Hameem, & Zulkafli, 2020; Gilbreath & Benson, 2004). Furthermore, research has examined the impact of training and development programs on employee performance in chemical companies, suggesting that such programs can improve employee skills and knowledge, leading to better job performance (Al Karim, 2019).

Other organizational factors, such as job category and salary, have also been found to be related to employee performance in the chemical industry, with higher job categories and salaries being associated with better performance (Abdulkhaliq & Mohammadali, 2019). These findings suggest that various organizational factors can significantly impact employee performance in the chemical industry, highlighting the importance of considering these factors when trying to improve employee performance in this sector.

Additionally, studies have explored the relationship between employee performance and organizational performance in chemical companies, indicating a positive correlation between employee performance and organizational outcomes such as profitability, productivity, and innovation (Huselid, 1995).

(1) “Safety culture” is a work environment where safety is highly prioritized. It involves shared beliefs, attitudes, and be-haviors that emphasize safety in all aspects of work. In the chemical industry, which is inherently risky, a strong safety culture is crucial. It leads to increased employee engagement, job satisfaction, and improved performance. It also helps prevent accidents, injuries, and protects the environment.

However, despite the increasing popularity of telecommuting and concern about commuting distance in many industries, there is a lack of specific research on how these factors affect employee performance in the chemical industry. To the best of our knowledge, no studies have specifically explored the relationship between telecommuting or home-to-work distance and performance in this sector. Therefore, further research is needed in this area to better understand how these factors may affect employee performance in the chemical industry. Our study aims to fill this research gap and provide insights into the relationship between telework and home-to-work distance with employee performance in the chemical industry.

2.3. Approaches to data analysis in employee performance studies

Data analysis is a critical component of research on employee performance, providing insights into the relationships between various factors and their impact on performance outcomes. Researchers use a variety of data analysis techniques to analyze their data, including regression analysis, factor analysis, structural equation modeling, and cluster analysis (Zhang, Xu, Zhang, & Yang, 2021).

Regression analysis is one of the most commonly used techniques in employee performance research, allowing researchers to investigate the relationships between one or more independent variables and a dependent variable (Zhang, Xu, Zhang, & Yang, 2021). Factor analysis, on the other hand, is used to identify underlying factors that may explain the relationships between multiple variables.

Structural equation modeling (SEM) is a more advanced data analysis technique that allows researchers to examine complex relationships between multiple variables and test theoretical models (Ringle, Sarstedt, Mitchell, & Gudergan, 2020). Cluster analysis is another technique used to identify groups or clusters of individuals with similar characteristics or performance outcomes.
While these techniques are commonly used in employee performance research, it is important to note that there are other data analysis techniques available as well. For example, some researchers may use qualitative data analysis techniques, such as content analysis or thematic analysis, to analyze open-ended survey responses or interview data.

It is important to properly identify when to use traditional models such as regression, which is the most commonly used technique in performance research, and when to leverage the potential of machine learning methods for prediction and analysis. Overall, the choice of data analysis technique depends on the research question, the type of data collected, and the research design.

When it comes to studying performance in the chemical industry, researchers have employed various techniques to predict and enhance employee performance. Some of the most noteworthy studies in this field include Nguyen and Giang (2019) utilization of structural equation modeling to predict employee performance, and Lather, Malhotra, Saloni, Singh and Mittal (2019) application of machine learning techniques for the same purpose. Dauda y Akingbade (2011) used regression and classification techniques to predict performance in the industry sector, while Kariuki y Murimi (2015) focused on performance prediction of chemical enterprise employees using multiple regression. Lastly, Yuan (2022) leveraged artificial neural networks to predict employee performance in a chemical company. These studies provide valuable insights into the factors that impact performance in the chemical industry and offer useful tools for predicting and enhancing employee performance.

It's important to note that in general, the variables used in these studies tend to be more qualitative or subjective in nature, such as job satisfaction, quality of interpersonal relationships, and commitment to the company, among others.

ISSN: 2792-3770

3. Hypothesis Development: Predicting employee performance thorough an integrated model of training and work arrangements

In this section, we present a conceptual model based on the variables identified in the literature review and the specific context of our study. We also formulate hypotheses that will guide our empirical analysis. Figure 3.1 presents our conceptual model.

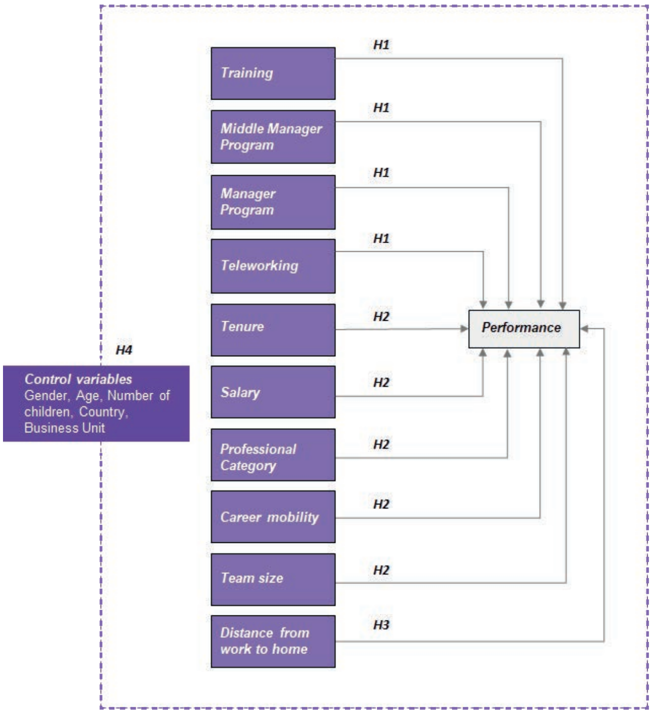

Figure 3.1 Conceptual model

Our integrated model proposes that several factors positively contribute to employee performance. First, we hypothesize that Training Middle Manager Program, Manager Program, and telecommuting/teleworking are positively related to performance, as they indicate investment in employee development and flexibility in work arrangements. Previous research has consistently shown that these variables have a positive impact on employee performance (Akter, 2016; Al Karim, 2019; Aropah, Sarma, & Sumertajaya, 2020). Employees who receive regular training and participate in development programs are likely to have better job skills and knowledge, leading to improved performance. Furthermore, employees who have the option to work remotely may have increased job satisfaction, lower stress levels, and better work-life balance, which can positively affect performance.

H1. Training, Middle Manager Program, Manager Program, and telecommuting/teleworking are positively related to performance, as they indicate investment in employee development and flexibility in work arrangements.

Secondly, we expect that tenure, salary, professional category, career mobility, and team size are positively related to performance, as they reflect experience, compensation, responsibility, opportunities for growth, and leadership skills. Numerous studies have found that these factors positively impact employee performance (Diamantidis & Chatzoglou, 2019; Cerasoli, Nicklin, & Ford, 2014; Farid, Kee, Mohamad, Hameem, & Zulkafli, 2020; Judge, Thoresen, Bono, & Patton, 2001; Khan, 2012). We hypothesize that employees who have been with the company for a longer period, have higher salaries, hold higher professional categories, have more career mobility options, and work in larger teams will have greater job satisfaction and better performance.

H2. Tenure, salary, professional category, career mobility, and team size are positively related to performance, as they reflect experience, compensation, responsibility, opportunities for growth, and leadership skills.

Thirdly, we hypothesize that distance from home to work is negatively related to performance. Previous research suggests that commuting can have a negative impact on employee well-being and job satisfaction (Page & Nilsson, 2017; Suryadi, FoEh, & Manafe, 2022). We expect that employees who have to travel longer distances to get to work may experience more stress, fatigue, and time inefficiency, leading to decreased job satisfaction and lower performance.

H3. Distance from home to work is negatively related to performance, as it may cause stress, fatigue, and time inefficiency.

Finally, we hypothesize that society within the company, business unit, and personal area are not significantly related to performance, as they may not directly affect job tasks or individual competencies. While these variables may impact employee job satisfaction and overall

  ISSN: 2792-3770

well-being, we expect that they will not have a significant impact on performance.

H4. Society, business unit, and personnel area are not significantly related to performance, as they may not directly affect job tasks or individual competencies.

4. Methodology

4.1. Population and sample

The sample for this study was restricted to employees who were currently active in the company as of November 1, 2022, and who met the criteria of being non-intern and non-partial retiree employees with permanent contracts. This selection criteria ensured that the sample consisted of employees who had been with the company for a significant period of time. The sample size was N = 907.
However, halfway through the study, the decision was made to split the employee population into two distinct groups based on "Personal area". This division was made as these groups represented noticeably different demographics within the organization, with A1 having a higher proportion of university-educated employees and exhibiting different performance evaluations compared to A2, which consisted of employees with non-university studies. Due to these significant differences, we believed it was crucial to conduct the study for the entire population of 907 employees, as well as separately for each group. As a result of this division, the sample sizes for A1 and A2 were 393 and 514, respectively.

4.2. Variables included in the study

The variables used in this study were obtained from the organization's system and include information typically collected by the Human Resources department in medium to large-sized companies. The variables used in the study are operationalized based on their definitions in the organizational system.

The study includes several variables that are of interest.

Employee Performance. The dependent variable used in this study is employee performance, which is measured on a scale from 70 to 110, where 70 represents low performance and 110 represents high performance. For the Random Forest classification model, the variable was categorized into two, high performance starting from 95 and low performance from 94 and below. This categorization allowed for a simpler criterion to classify employ-ees based on their performance. The employee performance variable was constructed as the average of all per-formance scores for each employee over the last three years of their tenure with the company. The last three years of employee tenure were considered in the performance variable calculation due to changes in the compa-ny's performance evaluation model over time. It was determined that only recent years would be included to en-sure consistency in the evaluation process. This ensures that the measure is consistent and reliable. In addition, the performance evaluation method used is goal-setting: specific goals and objectives are established for each em-ployee, and their performance is evaluated based on whether they have achieved them or not, using a range from 70 to 100, where a score above 100 is considered exceptional.

The independent variables that may have an impact on employee performance are:

1. **Training:** the amount and variability of training received throughout their tenure with the company. It is measured by the total number of training hours completed, the number of different courses completed, and the total number of courses completed. This variable is an indicator of the investment made by the company in the development of its employees and their potential for growth within the organization.
2. **Middle Manager Program:** this is a specialized training program aimed at middle managers in the company. It is measured by the participation in this program, by a binary variable (Yes/No), indicating whether the employee has coursed the program or not.
3. **Manager Program:** this is a specialized training program aimed at managers in the company. It is measured by the participation in this program, by a binary variable (Yes/No), indicating whether the employee has completed the program or not.
4. **Telecommuting/teleworking:** whether the employee has formally requested and been granted the option to work remotely by the company. This variable is measured as a binary Yes/No variable.

5. **Tenure:** the number of years the employee has been with the company.
6. **Salary:** the employee's annual salary in dollars.
7. **Professional category:** the job position held by the employee, categorized based on their employment status (A1, A2) and their level of responsibility within the company. For employees A2, categories range from supervisor, specialist, Industrial technical staff, administrative staff, to operator. For employees A1, categories range from director, manager, middle manager, senior technical staff, technical staff, young talent.
8. **Career mobility:** the number of times an employee has changed positions or been promoted within the company during their tenure. This includes the time spent in each position, the number of promotions received, and the number of times the employee has changed roles within the company.
9. **Team size:** the number of direct reports or team members that report to the employee.
10. **Distance from home to work:** the distance in kilometers from the employee's home to the workplace.

The control variables are those that are not directly related to the research question but may have an impact on the dependent variable. In this study we have identify three control variables:

1. **Society within the company:** the specific society where the employee is working. It is measured by a categorical variable with six possible values (A, B, C, D, E, F).
2. **Business unit:** the specific business unit or department where the employee is working. It is measured by a categorical variable with seven possible values (Procurement, Sales, Technology, Strategy, HR, Operations, Finance).
3. **Personnel area:** the personnel area where the employee is classified. It is measured by a categorical variable with two possible values (A1, A2).

We also have included demographic variables such as:

1. **Gender:** the biological sex of the employee, classified as either male or female. This is a binary variable.
2. **Age:** the number of years the employee has lived since birth. This is a continuous variable.
3. **Country:** the country in which the employee is based, classified as Mexico, Spain or USA. This is a categorical variable.
4. **Number of children:** the total number of biological or adopted children that the employee has. This is a discrete variable.

All of the independent variables, including the variables of control and the demographic variables, will be taken into account when analyzing the relationship between employee performance and the independent variables, while controlling for the effects of demographic variables and variables of control. Table 4.1, located in the ap-pendix, provides a comprehensive overview of the independent variables under investigation.

4.3. Study approach and procedure

As previously stated in the chapter, the study was conducted using Python within the Visual Studio development environment. The data for the study was obtained from the company's HR management tool, which was down-loaded as five separate Excel files and combined into a single dataset.

After configuring the necessary libraries for data analysis, the dataset was cleaned by removing irrelevant col-umns and handling missing values. Missing values were evaluated for each variable, and if their proportion was deemed significant, they were either removed or replaced with a value of zero.

Exploratory data analysis was then conducted by creating visualizations to gain insights into the behavior of the variables. Numerical and categorical variables were compared separately against the target variable. Correla-tions between the target variable and numerical and categorical variables were also examined.

Next, a regression model was created to predict employee performance. Two methods were used for variable selection: firstly, selecting variables with the highest correlation, and secondly, manually selecting variables. Line-ar and Lasso regressions were performed, and one-hot encoding was used to handle categorical variables. The proportion of variables to records was evaluated and found to be valid at less than 10%. A cross-validation pro-cess was then conducted to compare the model's metrics with other alternatives.

Finally, classification models were also used for analysis. One-hot encoding was performed, and the performance variable was discretized into high (1) and low (0) categories. A random forest model was trained, and a cross-validation process was conducted to compare the model's metrics with other alternatives. Feature im-portance were also obtained and sorted in descending order.

This analysis was conducted for N = 907, and then separately for N = 514 and N = 393, using the control variable "Personal area" to divide the groups.
For the purposes of this study, we will henceforth refer to the full sample of N=907 as Group A. Within Group A, we have two subgroups: Group A1 with N=393 and Group A2 with N=514. This designation will be used throughout the analysis to clearly distinguish between the different subgroups under examination.

5. Data analysis and results

In this section, we start by describing the sample, providing details on Group A (N=907) and each of its subgroups (A1 and A2). To gain a better understanding of the relationship between the variables and performance, we began by creating visual plots to identify trends. Following this, we present the outcomes of our analysis, which encom-passes correlation analysis, regression analysis, and random forest analysis, aimed at testing our hypothesis.

5.1. Sample characteristics

The total sample size for this study is N = 907, which is divided into two subgroups A1 (n = 393) and A2 (n = 514). Most participants were from Mexico (A: 54%, A1: 58%, A2: 50%), followed by Spain (A: 44%, A1: 37%, A2: 50%) and a small number from the USA (A: 2%, A1: 4%, A2: 0%).

In terms of gender, the sample was predominantly male (A: 78%, A1: 58%, A2: 88%). A small proportion of participants had completed the Manager Program (A: 3%, A1: 15%, A2: 0%), while the majority had not (A: 97%, A1: 85%, A2: 0%). Similarly, a small proportion had completed the Middle Manager Program (A: 6%, A1: 13%, A2: 0%), while the majority had not (A: 94%, A1: 87%, A2: 0%).

Regarding the professional category, the largest group was "Operator" (A: 36%, A1: 0%, A2: 63%), followed by "Technical staff" (A: 24%, A1: 55%, A2: 0%) and "Middle manager" (A: 9%, A1: 21%, A2: 0%). Other categories with significant representation were "Industrial technical staff" (A: 8%, A1: 0%, A2: 15%), "Supervisor" (A: 6%, A1: 0%, A2: 10%), "Senior technical staff" (A: 5%, A1: 12%, A2: 0%), "Manager" (A: 4%, A1: 10%, A2: 0%), and "Specialist" (A: 4%, A1: 0%, A2: 8%). The remaining categories ("Administrative" and "Director") had lower representation (A: 3%, A1: 2%, A2: 0% and A: 1%, A1: 2%, A2: 0%, respectively).

Concerning age, A1 was found to have a lower mean age of 43.65 years compared to A2's mean age of 47.22 years. However, A1 exhibited a higher standard deviation in age (11.09 years) than A2 (9.30 years). Additionally, A1 had a shorter mean tenure of 13.66 years compared to A2's longer mean tenure of 19.27 years.

In terms of direct collaborators, A1 had a higher mean number of 2.64, while A2 had a significantly lower mean number of 0.0058. Moreover, A1 had a lower mean performance score of 88.82, while A2 had a higher mean performance score of 91.36.

ISSN: 2792-3770

The mean time to change positions for Group A was 2.71 years, with a standard deviation of 3.12 years. The A1 subgroup had a lower mean time of 2.31 years and a standard deviation of 2.89 years, whereas the A2 subgroup had a higher mean time of 3.01 years and a standard deviation of 3.25 years.

Regarding the distance to work, the mean for Group A was 27.06 kilometers, with a standard deviation of 69.73 kilometers. The A1 subgroup had a higher mean distance of 38.67 kilometers, with a standard deviation of 90.02 kilometers, while the A2 subgroup had a lower mean distance of 18.18 kilometers, with a standard deviation of 47.04 kilometers.

The mean time in the current position for Group A was 52.68 months, with a standard deviation of 26.56 months. The A1 subgroup had a lower mean time of 47.43 months, with a standard deviation of 27.59 months, whereas the A2 subgroup had a higher mean time of 56.69 months, with a standard deviation of 25.03 months.

The mean hours of training for Group A was 0.81, with a standard deviation of 3.73. The A1 subgroup had a higher mean of 1.64 hours, with a standard deviation of 5.51 hours, while the A2 subgroup had a lower mean of 0.18 hours, with a standard deviation of 0.73 hours. Furthermore, the mean number of training courses for Group A was 13.62, with a standard deviation of 14.51. The A1 subgroup had a slightly lower mean of 13.14, with a standard deviation of 13.68, while the A2 subgroup had a higher mean of 14.33, with a standard deviation of 15.43.

Finally, regarding our target variable, the results for Group A indicate that the mean performance number is 90.26 with a standard deviation of 17.48. These values suggest that the performance variation within the group is moderate. Further analysis of the subgroups, A1 and A2, reveals interesting differences.

For A1, the mean performance number is 88.82 with a standard deviation of 25.91. The performance variation is higher than that of the overall group. Additionally, the minimum and maximum values are the same as those of the overall group. This suggests that the subgroup has lower performance than the overall group, with a higher degree of variability.

In contrast, A2 has a higher mean performance number of 91.36 with a lower standard deviation of 4.87. The performance variation within this subgroup is lower than that of the overall group and A1. Additionally, the minimum and maximum values of A2 are within the range of the overall group. This suggests that the subgroup has higher performance than the overall group, with a lower degree of variability.

In summary, Group A is primarily an aging, male-dominated population with a high percentage of operating personnel, limited movement within the company, a flat management structure, and a long-term commitment to the organization. However, comparing the two subgroups within Group A highlights significant differences. A1 consists of a younger age group with a wider range of ages, shorter tenure, and more direct collaborators. A2, on the other hand, is an older age group with a narrower age range, longer tenure, and fewer direct collaborators. Notably, A2's mean performance score is higher than A1, indicating that older employees may have a higher level of performance. Furthermore, A1 tends to have a longer commute, less time to change positions, and fewer hours of training and courses compared to A2, which has a shorter commute, more time to change positions, and more training hours and courses.

5.2. Examining variable trends

We have visually explored the relationships between the independent variables and the performance variable through the use of scatterplots and bar graphs. Our analysis was conducted with great care to ensure that the sample we used was representative, enabling us to draw meaningful conclusions. Scatterplots were used for continuous variables or those with a large number of data points, while bar graphs were employed for categorical variables or those with relatively fewer data points. The graphics for each variable are included in the appendix, from figures 5.1 to 5.12.

As for the results, we observed the following:

- Employees who change positions at least once have better performance.
- Performance decreases at around 4.1 years and then increases again.
- Employees who get promoted have better performance.
- Courses do not seem to have an impact on performance. In fact, more courses seem to lead to lower performance.
- 60% of individuals in Mexico have high performance, while only 36% do in Spain. Mexico seems to have higher performance overall.

ISSN: 2792-3770

- A1 has significantly higher performance scores than A2, with a huge jump in scores from 96 to 110. This group has high performance scores compared to A2.
- In terms of professional category and its relationship with performance, only 6% of operators have excellent performance, which is generally low. Higher positions such as managers and directors have better performance.
- Of the total number of people who telework, 56.5% have high performance.
- Apparently, the closer an employee lives to their workplace, the better their performance.
- In addition, some additional conclusions were found in subgroup A1:
- Fewer employees who are far from work have less performance than those who are close.
- For low and high performances, Spain and Mexico are equal, but for average performance, there is a significant imbalance.
- Managers and directors are either very good or very bad, while technical staff are average.
- No remarkable trends have been found when comparing the remaining variables with performance

5.3. Correlation Analysis: Pearson's and Cramer's Methods

Figures 5.13 and 5.16 present the correlation coefficients between the independent variables (categorical and numerical) for Group A. The correlation coefficients for the subgroups within Group A can be found in the appen-dix, specifically Figures 5.14, 5.15, 5.17 and 5.18.

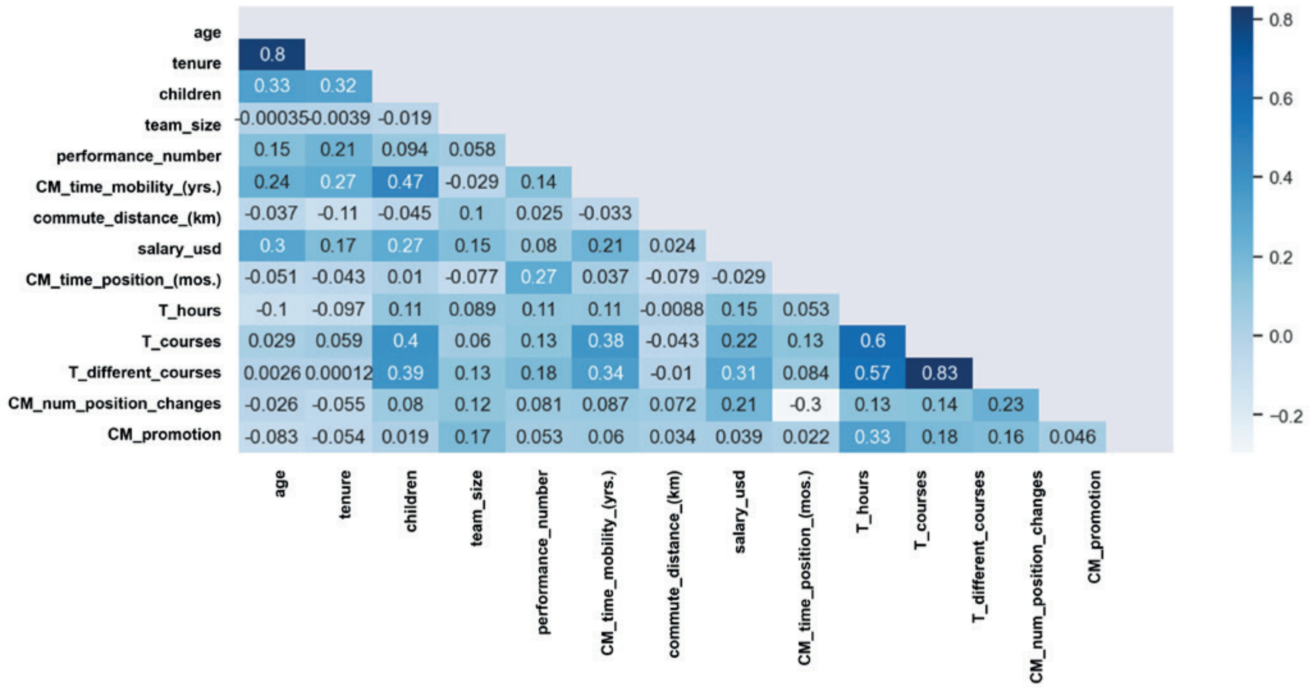

Figure 5.13 Correlation matrix A NV

ISSN: 2792-3770

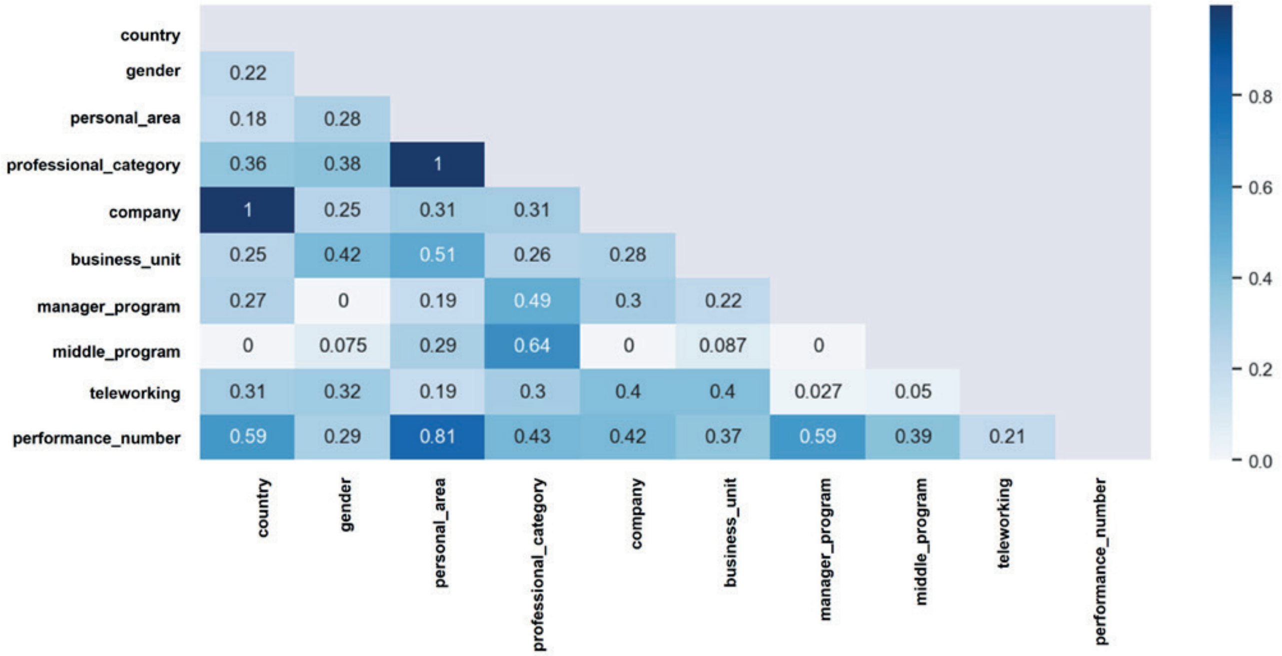

Figure 5.16 Correlation matrix A CV

It's worth noting that in A2, two variables (Manager Program and Middle Manager Program) were removed from the analysis as potential independent variables due to a lack of available data. This was done in order to adapt the selection of independent variables to the specific characteristics of this subgroup and ensure accurate and mean-ingful results.

The results suggest that variables such as "Career mobility ", "Tenure", "Training", "Salary" and "Age" are posi-tively correlated with performance across all groups. It also appears that "Manager Program" and "Middle Man-ager Program" may also be correlated with performance in certain groups. However, some variables such as "Distance from home to work", "Teleworking" and "Team size" may not have a strong relationship with perfor-mance.

In general, the correlations are not very strong, which means that the variables are not highly related to each other. However, some variables show moderate correlation with performance, such as "Carer mobility" and "Tenure". This suggests that seniority in the position and time since the last position change may be related to better performance.

5.4. Hypothesis testing: regression analysis

In order to further investigate the relationship between our independent variables and performance, we conducted both linear regression analysis and Lasso regression analysis. Unfortunately, the results of both analyses were not encouraging, as we were not able to obtain significant coefficients for any of our independent variables.

In fact, the coefficients we obtained were negative for all three subgroups. For Group A, the values were -15.32 and -0.64 in linear regression and -15.67 and -0.23 in Lasso regression. These results are similar in the subgroups.

While it may appear that the included independent variables are not good predictors of performance based on these results, it's important to consider that the issue may not lie with the variables themselves but rather with the limitations of the methods used.

The moderate variation in performance, as demonstrated by the standard deviation (std = 17.47), justifies the need for more sophisticated analysis methods to predict this variable. Linear regression and Lasso regression, while useful for analyzing relationships between variables, may not be optimal for predicting performance with such moderate variation. Random forest models, with their ability to capture complex relationships between variables and account for non-linear effects, may provide a more robust and accurate prediction of performance for Group A and its subgroups. Therefore, we will explore the use of random forest analysis in the next section to predict performance and determine the key factors that contribute to its variability.

5.5. Hypothesis testing: random forest analysis

In this section, we used a series of steps to prepare our data and train a random forest classifier. First, we used one-hot encoding to transform our categorical variables into binary features. Then, we discretized the "performance" variable into two categories: high performance and low performance based on a threshold value of 95. Next, we split our data into predictor variables (X) and target variable (y), and trained a random forest classifier with 100 estimators using the Scikit-Learn library. Finally, we used k-fold cross-validation to evaluate the performance of the model and compared it with other alternatives by calculating the F1 score. The F1 score is a measure of the classifier's accuracy that considers both the precision and recall of the model. Overall, these steps allowed us to build a powerful machine learning model to predict the performance of our target variable based on the selected predictor variables.

Based on the results of the random forest analysis, the model's overall performance on the entire dataset was moderate, with an F1-score of 0.54, indicating that the model has limited predictive power. However, when we consider the subgroups, the results reveal interesting variations. The A1 subgroup had the highest performance, with an F1-score of 0.73, while the A2 subgroup had the lowest performance, with an F1-score of 0.29. These findings indicate that the model has a better ability to predict high performance for Group A1, while it is less effective for Group A2. These insights can be valuable in identifying the factors that contribute to high performance in each subgroup and enhancing their performance accordingly.

The F1-scores highlight that the model may have different predictive capabilities for A1 and A2, indicating the need for further exploration to comprehend the underlying factors contributing to these discrepancies. The findings suggest that the random forest model is a promising approach to predicting performance in this dataset, and additional investigations are necessary to uncover its full potential.

ISSN: 2792-3770

Variable (Group A)	F-Importance (top 10)	Variable (Group A1)	F-Importance (top 10)	Variable (Group A2)	F-Importance (top 10)
salary_usd	*0.12*	*salary_usd*	*0.10*	*age*	*0.14*
T_hours	*0.10*	*tenure*	*0.09*	*salary_usd*	*0.12*
age	*0.09*	*T_hours*	*0.08*	*T_courses*	*0.11*
T_courses	*0.09*	*age*	*0.06*	*tenure*	*0.11*
tenure	*0.09*	*T_courses*	*0.06*	*T_hours*	*0.11*
commute_distance_(km.)	*0.05*	*CM_time_position_(mos.)*	*0.06*	*T_different_courses*	*0.07*
team_size	*0.05*	*team_size*	*0.05*	*commute_distance_(km.)*	*0.06*
T_different_courses	*0.05*	*commute_distance_(km.)*	*0.05*	*CM_time_mobility_(yrs.)*	*0.04*
CM_time_position_(mos.)	*0.05*	*T_different_courses*	*0.04*	*CM_time_position_(mos.)*	*0.04*
CM_time_mobility_(yrs.)	*0.03*	*middle_program*	*0.03*	*children*	*0.01*

Table 5.2 Random Forest Results: Group A and Subgroups Comparison.Source: own elaboration by using Python

The feature importance results obtained from the random forest analysis provide insights into the factors that are most strongly associated with employee performance in each subgroup. In Group A, the most important factor is “salary” with a relatively high feature importance score (0.12). This suggests that higher salaries tend to be associated with better performance. Other important predictors include “training” (T_hours, T_courses) (0.10, 0.09) and “age” (0.09). These variables indicate that employees who have more experience and engage in continuous learning through courses tend to exhibit better performance.

Additional factors such as tenure (0.09), distance to work (0.05), and team size (0.05) also contribute to performance prediction, albeit to a slightly lesser extent. These findings suggest that employees with longer tenure, shorter commuting distances, and a larger network of collaborators may demonstrate better performance.

Upon examining the predictor importance in A1 and A2, it appears that there are no significant differences between the two subgroups. The variables that strongly influence performance in Group A are consistent across both groups. This suggests that these predictors have a significant impact on performance regardless of the subgroup.

The similarities in predictor importance between A1 and A2 indicate that implementing HR strategies focused on these common factors could benefit both subgroups. By emphasizing areas such as salary adjustments, career development opportunities, and training programs, the company can create an environment conducive to improved performance for employees.

Finally, these results demonstrate the usefulness of the Random Forest technique in identifying important factors for job success in different employee groups. Combining the feature importance information with cross-validation results can provide a more complete understanding of the factors contributing to employee performance.

ISSN: 2792-3770

6. Summary and conclusions

The following section begins with a summary discussion of the key results of this study. The last section refers to the limitations of the study, and the research areas that can be explored in future research.

6.1. Discussion of results

The findings from the feature importance analysis shed light on the potential factors that can influence employee performance. The results suggest that variables with higher importance in the model are positively related to employees' performance, supporting the idea that these characteristics can enhance employee performance. However, it is important to note that correlation does not necessarily imply causation, and further investigation is necessary to establish the true relationship between these variables.

Regarding H1, the analysis has shown that "training" has emerged as a significant predictor for job performance in the studied group. This finding aligns with previous research, particularly in the field of the chemical industry, where training and skill development have been found to have a positive impact on employee performance (Akter, 2016; Al Karim, 2019).

However, the "Middle Manager Program" and "Manager Program" did not emerge as significant predictors of performance in this analysis. This could be attributed to the fact that these programs are specifically implemented for employees identified as high potential or possessing special skills. As these programs may target a selective subset of employees, their limited influence on the analyzed dataset may explain their lack of significance as predictors in this study.

Regarding telecommuting/teleworking, although it was not identified as a significant predictor of performance in this particular sample, it is important to note that the sample size may have been insufficient to fully evaluate its impact. Telecommuting has been the subject of numerous studies highlighting both its benefits and challenges in terms of productivity and job performance. However, compared to the total sample and considering the studied groups, the percentage analyzed may have been small, limiting the ability to draw definitive conclusions about its relationship with performance in this specific study.

In relation to H2, the analysis revealed that "tenure" emerged as one of the most influential factors in predicting employee performance, along with "salary" and "team size." Additionally, "career mobility" was found to have a significant impact. However, while these variables do contribute to employee performance, their effects may not be as pronounced as those of other factors included in the model.

Conversely, the variable "professional category" did not emerge as a significant predictor of employee performance. This lack of significance suggests that other variables may have a stronger influence on performance than the professional category alone.
As for H3, "distance from home to work" was found to be an important factor overall, suggesting that the proximity of employees' residences to their workplaces significantly influences their performance.

Finally, the results provided partial support for H4, as "personnel area", “business unit” and "society" did not emerge as top factors in any group, indicating that their influence on performance may be limited.

Overall, results provide empirical support for the notion that employee development programs and flexible work arrangements can enhance performance. However, it is important to acknowledge that this approach is not exhaustive. The model's inability to predict all aspects of performance may be due to its focus on organizational rather than individual characteristics. Future research could explore this issue in greater depth. Furthermore, our study highlights the importance of using sophisticated analytical techniques, particularly when dealing with low-variability variables such as "performance".

Turning to the question of why model A1 outperforms model A2 in terms of prediction accuracy within a random forest framework, despite both models employing the same predictor variables, several factors come into play. It is plausible that there are unobserved or unmeasured variables that have varying impacts on performance outcomes across the two groups. These hidden factors, including individual motivations, personal circumstances, and team dynamics, could contribute to the divergence in predictive accuracy between A1 and A2, despite sharing the same set of predictors.

Additionally, the discrepancy may be attributed to the sample size differences between A1 and A2. With A1 having a considerably smaller sample, the populations represented are likely to be significantly distinct, resulting in greater differentiation and potentially enhancing prediction accuracy. Conversely, in A2, despite having a larger sample, the population may exhibit greater homogeneity, posing challenges for accurate prediction.

6.2. Limitations of the study

As with any research study, there are limitations and assumptions that need to be considered when interpreting the results. One limitation of this study is that we have assumed that employee performance is accurately evaluated in the company, but this may not necessarily be the case. Although measures have been taken to ensure that performance is evaluated fairly, such as constructing it using the average scores of employees over the past three years, no performance evaluation model is perfect, as managers may have biases or may manipulate scores to favor certain outcomes, given that performance is often tied to incentives. This issue has been documented in prior research, highlighting the need for caution when interpreting performance data (e.g., Javidmehr & Ebrahimpour, 2015; Prendergast & Topel, 1993). Therefore, future studies could benefit from examining the validity and reliability of performance measures across different organizations and contexts.

Another potential limitation of this study is that the variables used were obtained from the organization's system, and it is possible that some variables that affect performance may not have been included. Additionally, the study assumes that the variables included in the analysis have a linear relationship with the company's performance.

Additionally, the study is based on cross-sectional data, which limits the ability to make causal inferences. While the study identifies significant associations between the variables, it is not possible to establish a cause-and-effect relationship. Future studies could employ longitudinal data to establish causality.

Overall, while this study provides valuable insights into the factors that impact employee performance in a multinational chemical company, caution should be exercised in generalizing the findings and interpreting the results.

6.3. Future research directions

This study has shed light on the factors that can predict employee performance in a multinational chemical company. However, there are still areas that warrant further investigation. Specifically, while our findings suggest that machine learning models may outperform traditional statistical methods in predicting employee performance, additional research is needed to confirm this claim. In particular, it would be worthwhile to explore whether variables with low variance are better predicted using machine learning models compared to traditional methods such as regression.

Another potential avenue for future research is to examine the relationship between employee performance and other factors such as job satisfaction and engagement. While our study focused primarily on demographic and work arrangement factors, there may be other factors that have a significant impact on employee performance that were not explored here.

Finally, it would be beneficial to replicate this study in other industries to determine if the findings are consistent across different contexts. Additionally, conducting longitudinal studies would provide insight into how these factors influence employee performance over time.

In summary, this study opens up new areas of inquiry into the predictors of employee performance in a multinational chemical company. Further research in these areas could provide valuable insights for practitioners looking to improve their workforce management strategies.

7. References

- Abdulkhaliq, S. S., & Mohammadali, Z. M. (2019). The impact of job satisfaction on employees' performance: A case study of Al Hayat Company-Pepsi Employees in Erbil, Kurdistan Region-Iraq. Management and Economics Review, 4(2), 163-176.
- Akter, N. (2016). Employee training and employee development is the predictors of employee performance; a study on garments manufacturing sector in Bangladesh. IOSR Journal of Business and Management, 18(11), 48-57.
- Allen, T. D., Golden, T. D., & Shockley, K. M. (2015). How effective is telecommuting? Assessing the status of our scientific findings. Psychological science in the public interest, 16(2), 40-68.
- Alexander III, E. R., Penley, L. E., & Jernigan, I. E. (1992). The relationship of basic decoding skills to managerial effectiveness. Management Communication Quarterly, 6(1), 58-73.
- Al Karim, R. (2019). Impact of different training and development programs on employee performance in Bangladesh perspective. International Journal of Entrepreneurial Research, 2(1), 8-14.
- Amponsah-Tawiah, K., Annor, F., & Arthur, B. (2016). Linking commuting stress to job satisfaction and turnover intention: The mediating role of burnout. Journal of Workplace Behavioral Health, 31, 104-123. https://doi.org/10.1080/15555240.2016.1159518.
- Aropah, V. D. W., Sarma, M. M., & Sumertajaya, I. (2020). Factors Affecting Employee Performance during Work from Home. International Research Journal of Business Studies, 13(2).
- Awaysheh, A., Bonet, R., & Ortega, J. (2023). Performance feedback and productivity: Evidence from a field experiment. Production and Operations Management, 32(1), 98-115.
- Cattell, A. (2006), "Improving Employee Performance through Workplace Coaching: A Practical Guide to Performance Management", Industrial and Commercial Training, Vol. 38 No. 2, pp. 106-106. https://doi.org/10.1108/00197850610653180
- Clark, B., Chatterjee, K., Martin, A., & Davis, A. (2020). How commuting affects subjective wellbeing. Transportation, 47, 2777-2805.
- Cerasoli, C. P., Nicklin, J. M., & Ford, M. T. (2014). Intrinsic motivation and extrinsic incentives jointly predict performance: A 40-year meta-analysis. Psychological Bulletin, 140(4), 980-1008.
- Dauda, Y. A., & Akingbade, W. A. (2011). Technological change and employee performance in selected manufacturing industry in Lagos state of Nigeria. Australian journal of business and management research, 1(5), 32-43.
- DeNisi, A. S., & Murphy, K. R. (2017). Performance Appraisal and Performance Management: 100 Years of Progress? Journal of Applied Psychology, 102(3), 421-433. https://doi.org/10.1037/apl0000085.
- Diamantidis, A. D., & Chatzoglou, P. (2019). Factors affecting employee performance: an empirical approach. International Journal of Productivity and Performance Management, 68(1), 171-193.
- Farid, M. A. B. M., Kee, D. M. H., Mohamad, M. S. B., Hameem, M. S. U. B., & Zulkafli, A. Z. B. (2020). The Role of Leadership and Leaders' Behavioral Characteristic on Employees: A Study of Petronas Berhad. International Journal of Applied Business and International Management (IJABIM), 5(3), 63-68.
- Flegl, M., Depoo, L., & Alcázar, M. (2022). The Impact of Employees' Training on Their Performance Improvements. Quality Innovation Prosperity, 26(1), 70–89. https://doi.org/10.12776/qip.v26i1.1665.
- Gilbreath, B., & Benson, P. G. (2004). The Contribution of Supervisor Behavior to Employee Psychological Well-being. Work & Stress, 18(3), 255-266.
- Griffin, M. A., Patterson, M. G., & West, M. A. (2001). Job satisfaction and teamwork: The role of supervisor support. Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior, 22(5), 537-550.

ISSN: 2792-3770

- Huselid, M. A. (1995). The Impact of Human Resource Management Practices on Turnover, Productivity, and Corporate Financial Performance. Academy of Management Journal, 38(3), 635-672.
- Iskamto, D. (2021). Stress and its impact on employee performance. International Journal of Social and Management Studies, 2(3), 142-148.
- Javidmehr, M., & Ebrahimpour, M. (2015). Performance appraisal bias and errors: The influences and consequences. International Journal of Organizational Leadership, 4, 286-302.
- Joshi, A., & Roh, H. (2009). The role of context in work team diversity research: A meta-analytic review. Academy of management journal, 52(3), 599-627.
- Judge, T. A., Thoresen, C. J., Bono, J. E., & Patton, G. K. (2001). The Job Satisfaction–Job Performance Relationship. Psychological Bulletin, 127(3), 376-407. https://doi.org/10.1037/0033-2909.127.3.376
- Kalambayi, J., Onojaefe, D., Kasse, S., & Tengeh, R. (2021). The influence of leadership styles on employee performance in construction firms. EUREKA: Social and Humanities, 1(1), 34-48. https://doi.org/10.21303/2504-5571.2021.002037.
- Kariuki, A., & Murimi, C. (2015). Employee empowerment and organization performance of Tata chemicals Magadi Ltd, Kenya.
- Khan, M. I. (2012). The Impact of Training and Motivation on Performance of Employees. Business Review, 7. https://doi.org/10.54784/1990-6587.1205
- Khan, A., Alamdar, N., Nawaz, M., Aleem, M., & Hamed, W. (2014). Impact of job satisfaction on employee performance: An empirical study of autonomous Medical Institutions of Pakistan. Journal of International Studies, 7, 11. https://doi.org/10.14254/2071-8330.2014/7-1/11
- Karasek, R. A. (1979). Job Demands, Job Decision Latitude, and Mental Strain: Implications for Job Redesign. Administrative Science Quarterly, 24(2), 285-308.
- Koopmans, L., Bernaards, C. M., Hildebrandt, V. H., De Vet, H. C. W., & van der Beek, A. J. (2013). Measuring individual work performance: Identifying and selecting indicators. Work: A Journal of Prevention, Assessment, and Rehabilitation, 48, 241-252. https://doi.org/10.3233/WOR-131659.
- Lather, A. S., Malhotra, R., Saloni, P., Singh, P., & Mittal, S. (2019, November). Prediction of employee performance using machine learning techniques. In Proceedings of the 1st International Conference on Advanced Information Science and System (pp. 1-6).
- Liang, P. J., Rajan, M. V., & Ray, K. (2008). Optimal team size and monitoring in organizations. The Accounting Review, 83(3), 789-822.
- Livingstone, L. P., Roberts, J. A., & Chonko, L. B. (1995). Perceptions of internal and external equity as predictors of outside salespeoples' job satisfaction. Journal of personal selling & sales management, 15(2), 33-46.
- McDaniel, M. A., Schmidt, F. L., & Hunter, J. E. (1988). Job experience correlates of job performance. Journal of applied psychology, 73(2), 327.
- McEnrue, M. P. (1988). Length of experience and the performance of managers in the establishment phase of their careers. Academy of Management Journal, 31(1), 175-185.
- Medoff, J. L., & Abraham, K. G. (1980). Experience, performance, and earnings. The Quarterly Journal of Economics, 95(4), 703-736.
- Motyka, B. (2018). Employee engagement and performance: a systematic literature review. International Journal of Management and Economics, 54(3), 227-244.
- Nguyen, T. L., & Giang, P. X. (2019). Improving employee performance in industrial parks: an empirical case of garment enterprises in Binh Duong province, Vietnam. European Journal of Investigation in Health, Psychology and Education, 10(1), 44-58.
- Ng, T. W., & Feldman, D. C. (2013). Does longer job tenure help or hinder job performance?. Journal of Vocational Behavior, 83(3), 305-314.

- Page, N. C., & Nilsson, V. O. (2017). Active commuting: workplace health promotion for improved employee well-being and organizational behavior. Frontiers in psychology, 7, 1994.
- Prendergast, C., & Topel, R. (1993). Discretion and bias in performance evaluation. European Economic Review, 37(2-3), 355-365.
- Rasool, T., Warraich, N. F., & Sajid, M. (2022). Examining the impact of technology overload at the workplace: A systematic review. SAGE Open, 12(3), 21582440221114320.
- Ringle, C. M., Sarstedt, M., Mitchell, R., & Gudergan, S. P. (2020). Partial least squares structural equation modeling in HRM research. The International Journal of Human Resource Management, 31(12), 1617-1643.
- Riyanto, S., Endri, E., & Herlisha, N. (2021). Effect of work motivation and job satisfaction on employee performance: Mediating role of employee engagement. Problems and Perspectives in Management, 19(3), 162.
- Santos, J. P., Caetano, A., & Tavares, S. M. (2015). Is training leaders in functional leadership a useful tool for improving the performance of leadership functions and team effectiveness?. The Leadership Quarterly, 26(3), 470-484.
- Schmidt, S. (2007). The Relationship between Satisfaction with Workplace Training and Overall Job Satisfaction. Human Resource Development Quarterly, 18, 481-498. https://doi.org/10.1002/hrdq.1216.
- Schmidt, F. L., Hunter, J. E., & Outerbridge, A. N. (1986). Impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. Journal of applied psychology, 71(3), 432.
- Singh, H., Jamil, R., Baroto, M. B., Hussin, M., Boyi, Y., & Singh, S. (2017). Impact of financial and non-financial incentives on employee performance. Advanced Science Letters, 23(1), 146-150.
- Spies, M. (2006). Distance between home and workplace as a factor for job satisfaction in the North-West Russian oil industry. Fennia-International Journal of Geography, 184(2), 133-149.
- Sugiarti, E. (2022). The Influence of Training, Work Environment and Career Development on Work Motivation That Has an Impact on Employee Performance at PT. Suryamas Elsindo Primatama In West Jakarta. International Journal of Artificial Intelligence Research, 6(1).
- Suryadi, Y., FoEh, J. E., & Manafe, H. (2022). Employee Productivity Determination: In Work Life Balance (WLB), Work From Home (WFH), Information Technology (IT) and Work Flexibility. Indonesian Interdisciplinary Journal of Sharia Economics (IIJSE), 5(2), 730-750.
- Tamimi, M., & Sopiah, S. (2022). The Influence of Leadership Style on Employee Performance: A Systematic Literature Review. International Journal of Economics, Business and Management Research, 1(2), 128-138. https://doi.org/10.54099/ijebm.v1i2.360.
- Tett, R. P., & Meyer, J. P. (1993). Job satisfaction, organizational commitment, turnover intention, and turnover: Path analyses based on meta-analytic findings. Personnel Psychology, 46(2), 259-293.
- Tubre, T. C., & Collins, J. M. (2000). Jackson and Schuler (1985) revisited: A meta-analysis of the relationships between role ambiguity, role conflict, and job performance. Journal of management, 26(1), 155-169.
- Van Knippenberg, D., Nishii, L. H., & Dwertmann, D. J. (2020). Synergy from diversity: Managing team diversity to enhance performance. Behavioral Science & Policy, 6(1), 75-92.
- Wright, T. A., Cropanzano, R., & Bonett, D. G. (2007). The moderating role of employee positive well being on the relation between job satisfaction and job performance. Journal of occupational health psychology, 12(2), 93.
- Yuan, J. (2022, May). Big Data Analysis in Human Resources Management: Performance Prediction Based on Employee Network. In 2022 5th International Conference on Artificial Intelligence and Big Data (ICAIBD) (pp. 389-395). IEEE.
- Zhang, Y., Xu, S., Zhang, L., & Yang, M. (2021). Big data and human resource management research: An integrative review and new directions for future research. Journal of Business Research, 133, 34-50.
- Zohar D. (2010). Thirty years of safety climate research: Reflections and future directions. Accident Analysis & Prevention, 42, 1517–1522.

ISSN: 2792-3770

8. Appendix: additional tables and figures

Variable	MOTIVATIONS FOR INCLUSION IN THE STUDY
tenure	*Tenure in the company has been previously investigated in relation to employee performance (Ng & Feldman, 2013; McDaniel et al., 1998; Schmidt et al., 1986). According to Schmidt et al. (1986), seniority in the company may be positively related to performance as employees with more time in the company may have a better understanding of processes and greater loyalty to the organization. However, Medoff and Abraham (1980) suggest that the relationship between seniority and performance may be more complex as employees with very long seniority may lose motivation and productivity.*
professional category	*Some studies suggest that professional category is related to job performance (McDaniel et al., 1998; Tubre & Collins, 2000). Specifically, it has been found that employees with a higher professional category tend to have better job performance than those with a lower category.*
training	*Previous studies suggest that employee training is positively related to job performance in the highly specialized chemical industry (Al Karim, 2019; Kariuki & Murimi, 2015). Schmidt (2007) found that training is related with job performance and job satisfaction, as well as their motivation and commitment to the organization. Al Karim (2019) argue that training is essential to improving employee performance, especially in complex jobs that require technical skills and decision-making abilities.*
team size	*While there are no studies indicating a direct relationship between the number of employees reporting to a manager and employee performance, research suggests that team size can impact job satisfaction and perception of the work environment. Griffin et al.'s 2001 study found that managers with larger teams reported lower levels of job satisfaction, while Liang et al.'s 2008 study suggested that team size can impact perceptions of the work environment, which in turn may affect team performance.*
career mobility	*The length of time in a position may be related to the acquisition of specific skills and knowledge needed to perform a job successfully. Furthermore, McEnrue (1988) study found that the average length of a job position can impact employee performance as it may influence their level of experience and knowledge in their area of work. Therefore, the amount of time employees take to switch positions within the same company could be related to their job performance.*
distance from home to work	*Limited and contradictory evidence exists regarding the relationship between home-to-work distance and employee performance. Spies (2006) found that greater distance negatively affected job satisfaction. Similarly, Amponsah-Tawiah, et al. (2016) observed a relationship between home-to-work distance, job satisfaction, and employee retention. However, these studies primarily focus on the impact on job satisfaction and retention rather than direct employee performance.*
salary	*There are studies that have found a relationship between salary and employee performance. For example, a study by Singh et al. (2017) found that financial and non-financial incentives have a strong influence on the respondents' performance. Additionally, a study by Livingstone et al. in 1995 found that the perception of salary equity was related to job satisfaction and commitment.*
teleworking	*Telecommuting has been studied in relation to job performance in several research works (Allen et al., 2015; Suryadi et al., 2022). According to Suryadi et al. (2022), telecommuting can enhance employees' job performance as it enables them to work in a more comfortable and distraction-free environment, which can increase their motivation and productivity. Additionally, telecommuting can allow for greater flexibility in work schedules and reduce stress associated with daily commuting.*
Manager Program & Middle Manager Program	*For example, Santos et al. (2015) notes that leadership training can improve managers' ability to motivate and direct their employees, which in turn can lead to greater productivity and performance. In addition, Alexander et al. (1992) found that managers who have better communication and teamwork skills had more productive teams.*

Table 4.1 Independent variables. Source: own elaboration

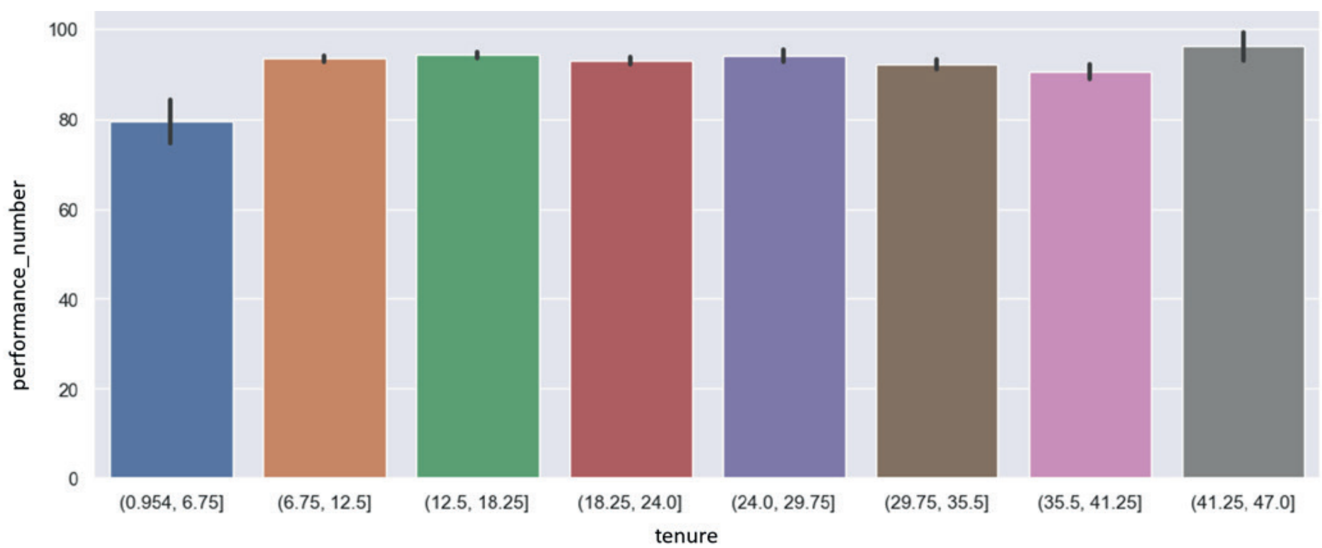

Figure 5.1 Relationship between Performance and Tenure (Group A)

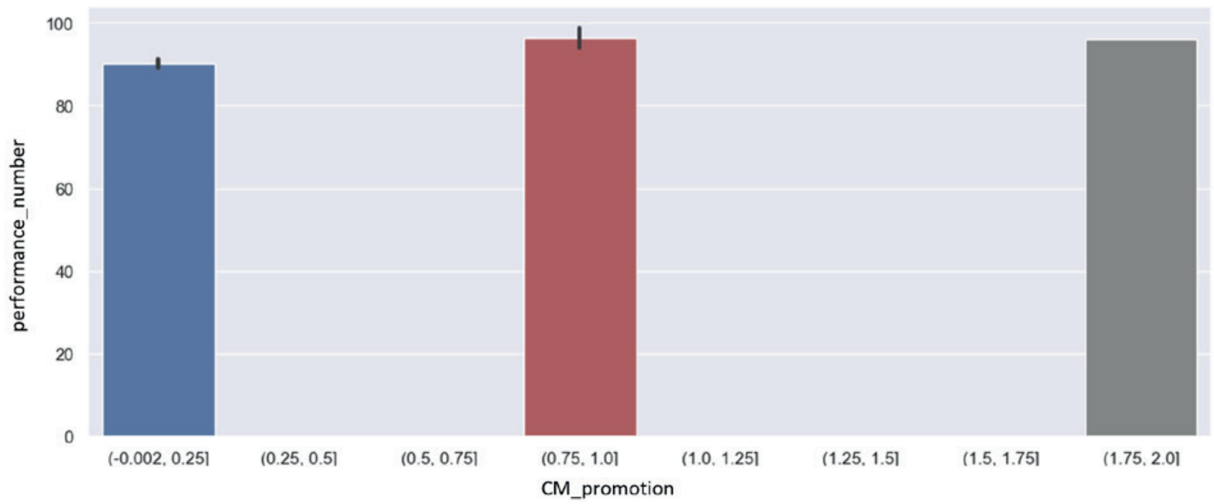

Figure 5.2 Relationship between Performance and Promotion (Group A)

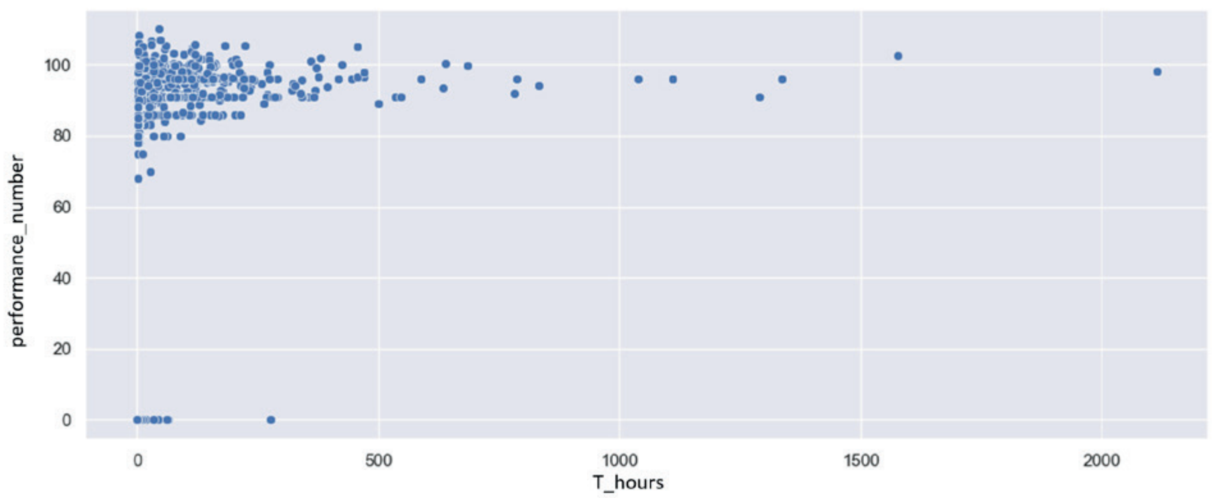

Figure 5.3 Relationship between Performance and Training hours (Group A)

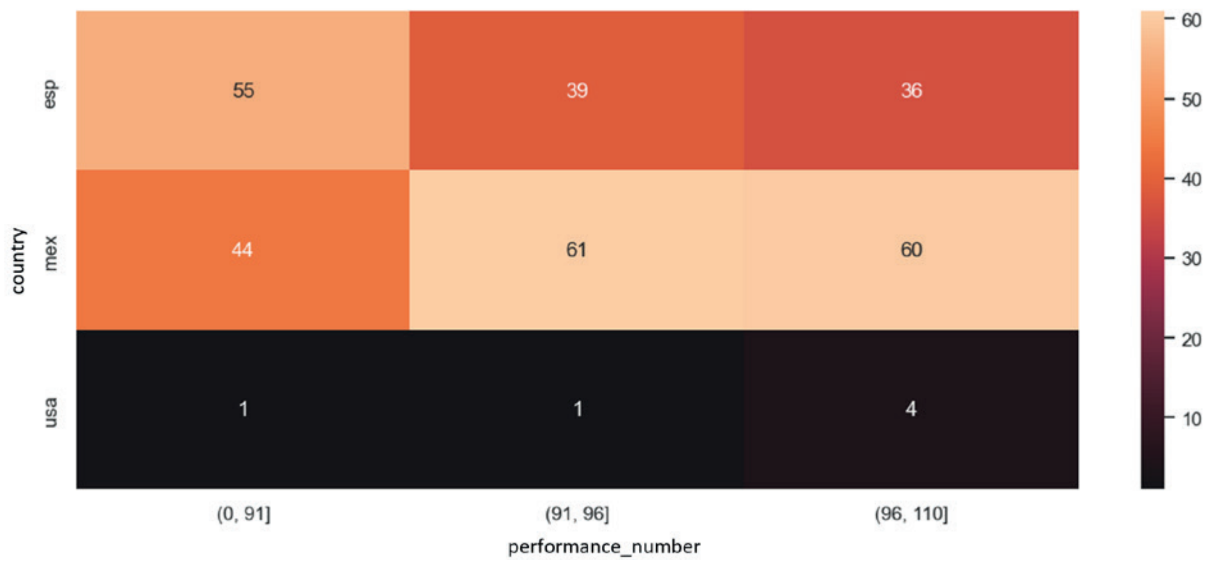

Figure 5.4 Relationship between Performance and Country (Group A)

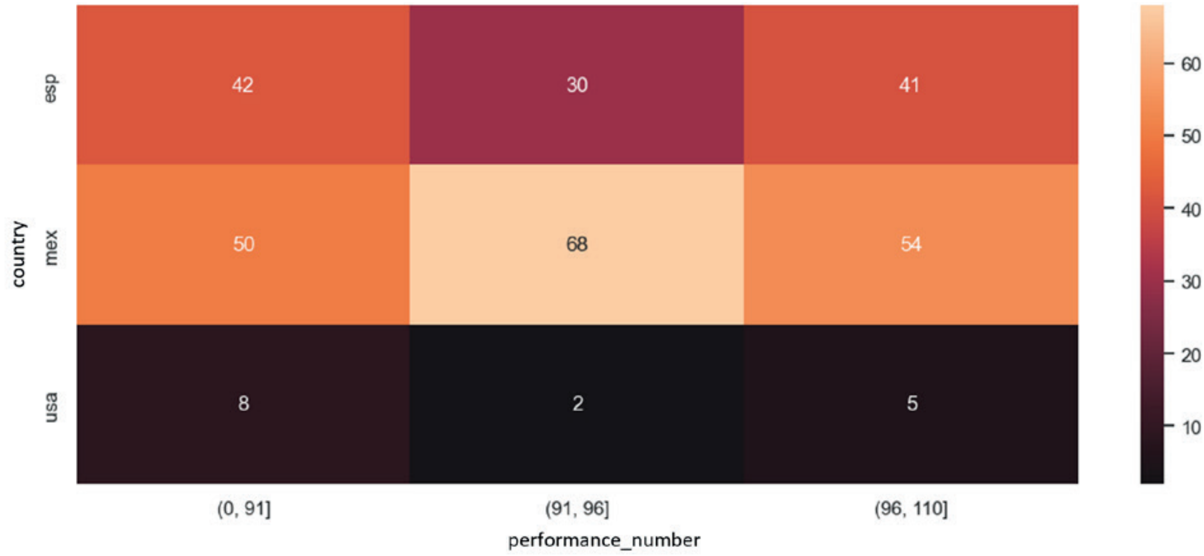

Figure 5.5 Relationship between Performance and Country (Group A1

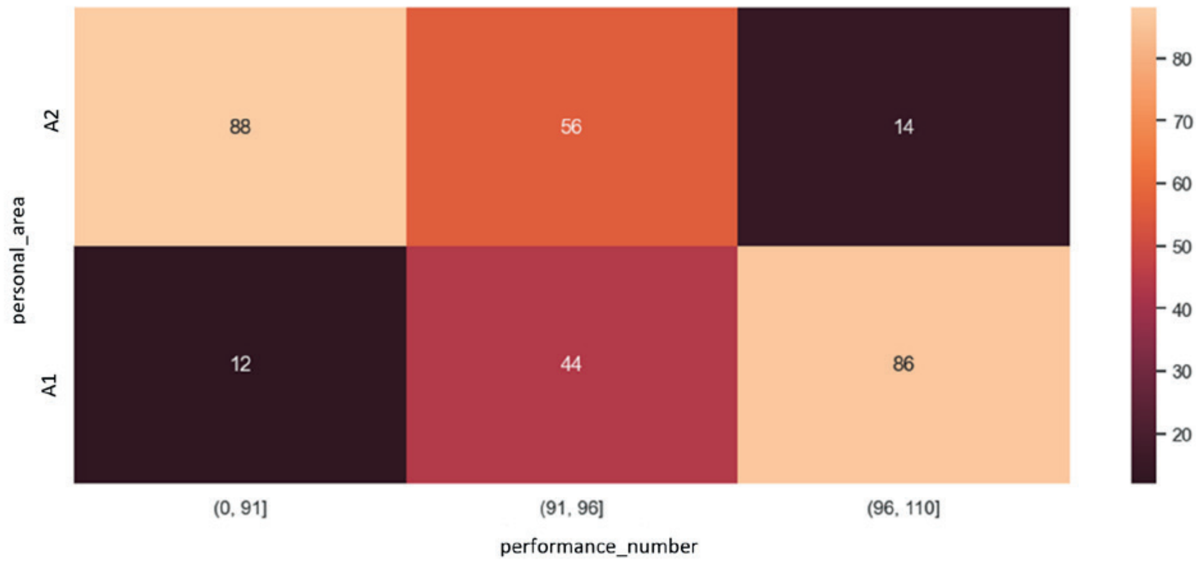

Figure 5.6 Relationship between Performance and Personal Area (Group A)

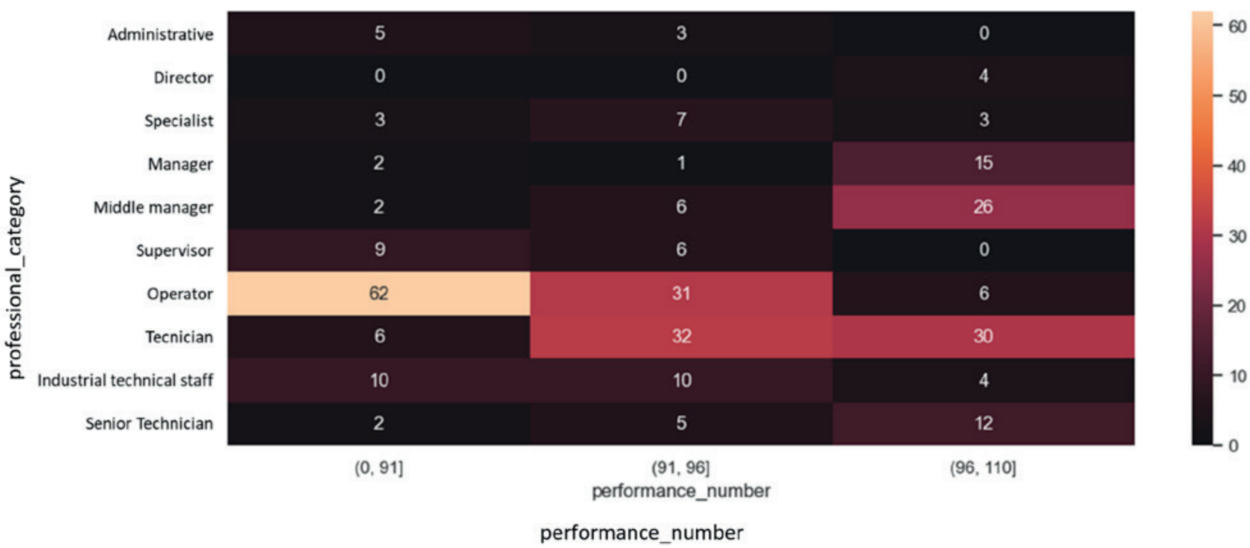

Figure 5.7 Relationship between Performance and Professional category (Group A)

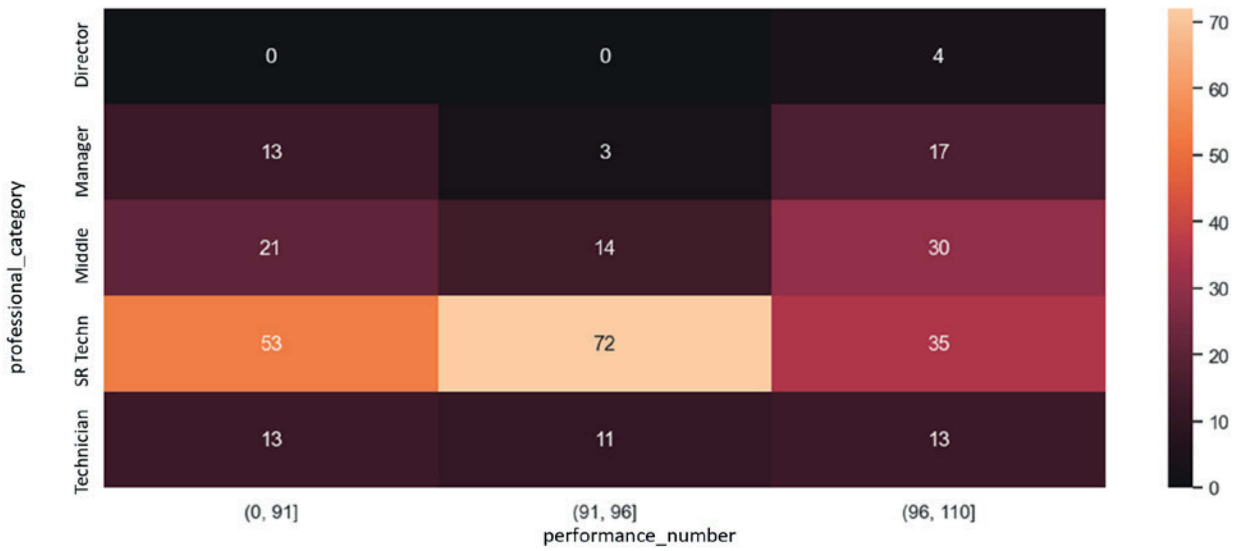

Figure 5.8 Relationship between Performance and Professional category (Group A1)

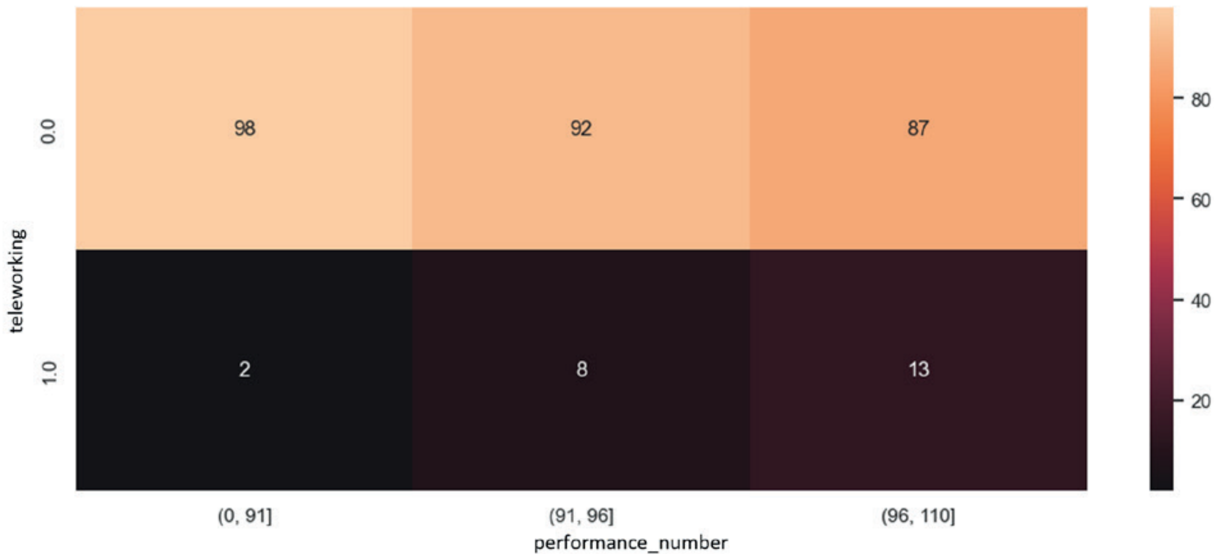

Figure 5.9 Relationship between Performance and Teleworking (Group A)

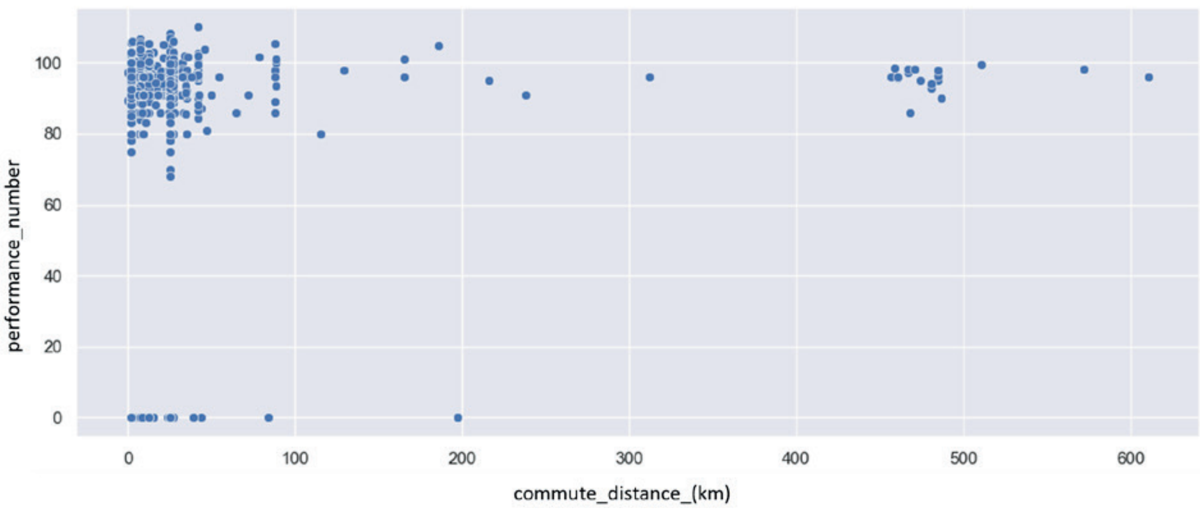

Figure 5.10 Relationship between Performance and Commute distance (Group A)

ISSN: 2792-3770

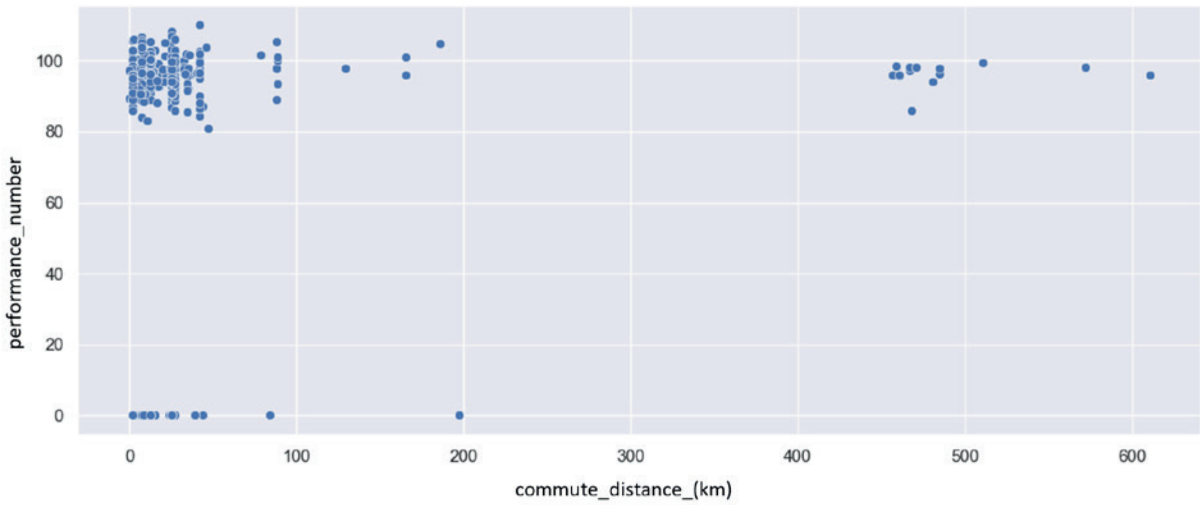

Figure 5.11 Relationship between Performance and Commute distance (Group A1

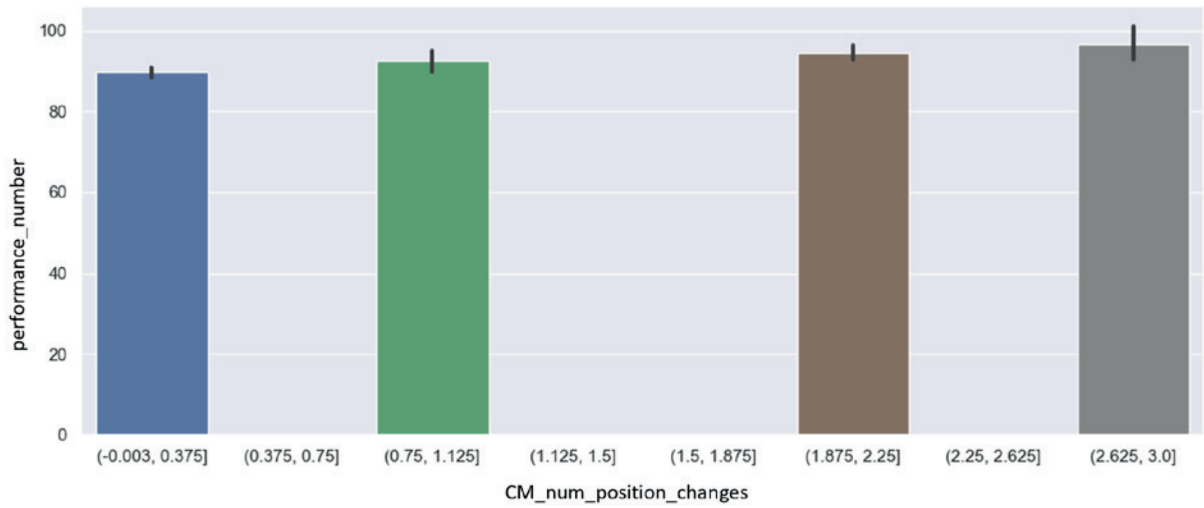

Figure 5.12 Relationship between Performance and CM_num_position_changes (Group A)

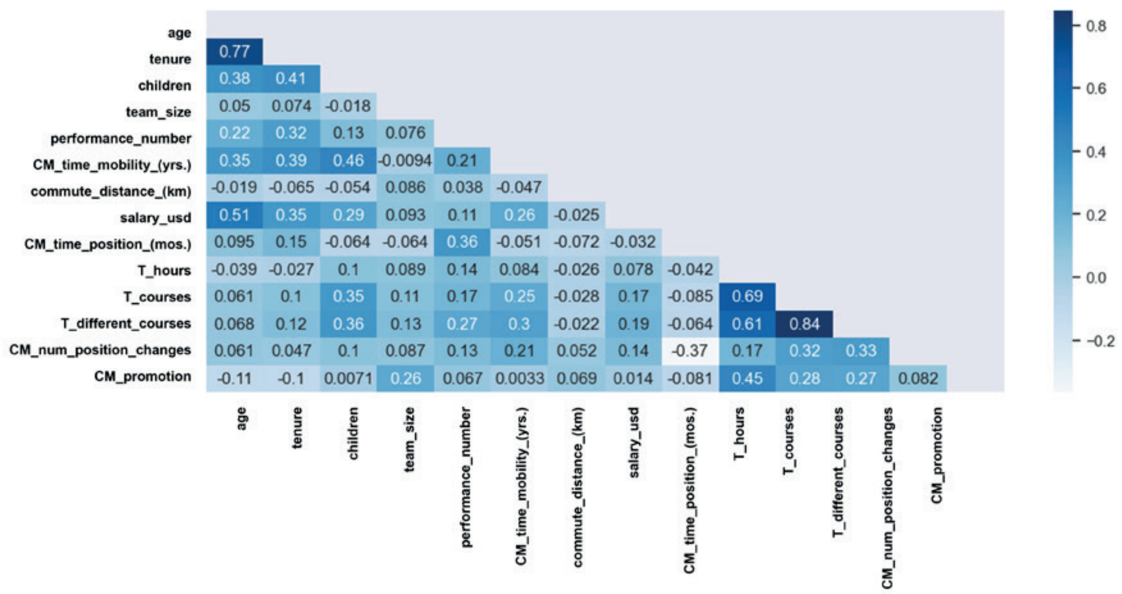

Figure 5.14 Correlation Matrix A1 Numerical Variables

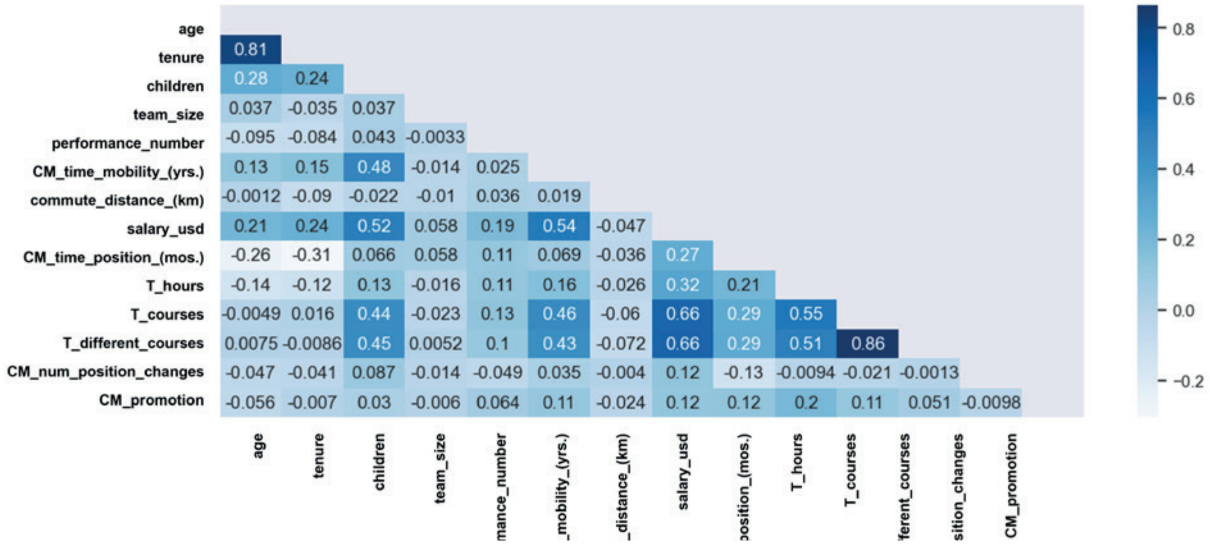

Figure 5.15 Correlation Matrix A2 Numerical Variables

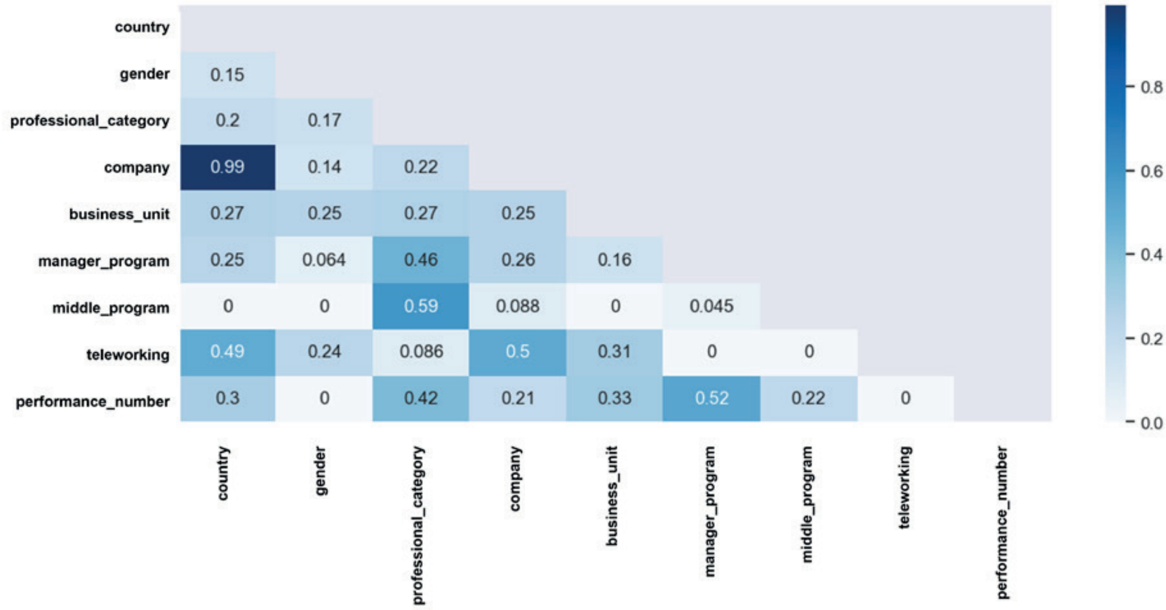

Figure 5.17 Correlation Matrix A1 Categorical Variables

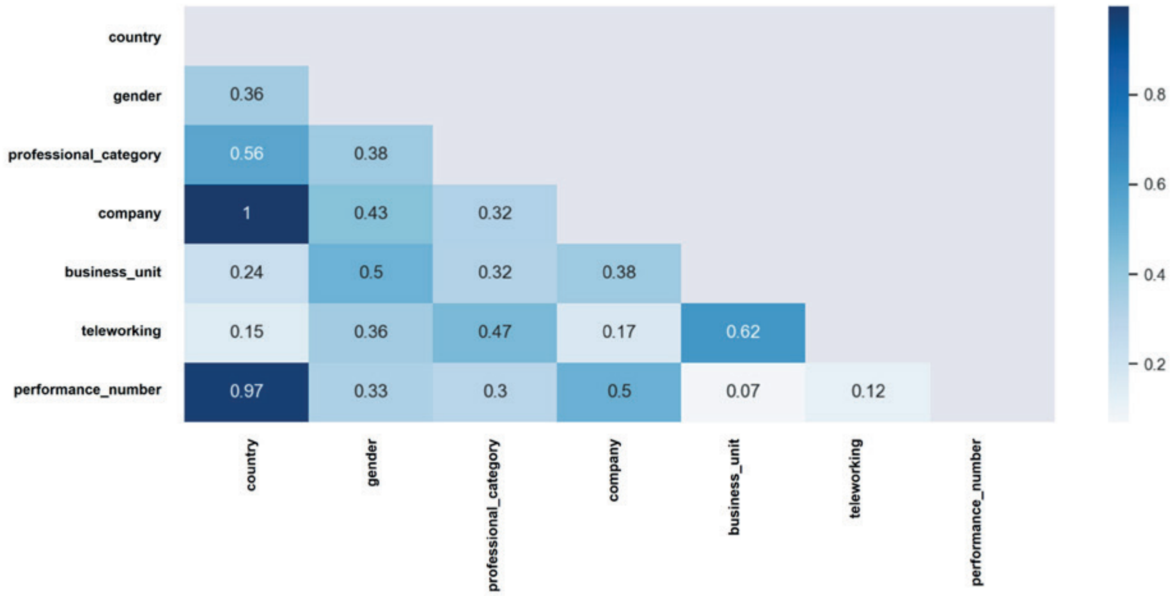

Figure 5.18 Correlation Matrix A2 Categorical Variables

.02

iic
instituto de ingeniería
del conocimiento

INSTITUTO DE INGENIERÍA DEL CONOCIMIENTO (IIC)

PIONEROS EN INTELIGENCIA ARTIFICIAL DESDE 1989

El Instituto de Ingeniería del Conocimiento (IIC) es un centro privado de I+D+i fundado en 1989 y pionero en Inteligencia Artificial. Llevamos más de 30 años trabajando en análisis Big Data y diferentes técnicas de **Machine Learning** y **Procesamiento del Lenguaje Natural (PLN)**, entre otras.

Nuestra apuesta de valor es el desarrollo de algoritmos y soluciones tecnológicas adaptadas a diferentes sectores y basadas siempre en la investigación aplicada. Como centro sin ánimo de lucro, comercializamos los productos y servicios fruto de nuestro trabajo para reinvertir los beneficios en más investigación e innovación.

En el IIC estamos especializados en tecnologías de analítica avanzada y en el desarrollo de **modelos de Inteligencia Artificial** para las áreas de Banca, RR. HH., Energía, Salud, Seguros e Inteligencia de Cliente.

Además, destacamos por ser expertos en la preparación y tratamiento de todo tipo de datos para su análisis y por la metodología que utilizamos para acompañar a nuestros clientes durante todo el proceso.

ENTORNO BANCARIO | ENTORNO RR.HH. | ENTORNO SEGUROS | ENTORNO ENERGÍA | ENTORNO SALUD | ENTORNO INTELIGENCIA DE CLIENTE

TECNOLOGÍA Y ANÁLISIS DE DATOS AL SERVICIO DE RR. HH.

HR ANALYTICS

El análisis de los datos de RR. HH. permite obtener información de valor para una mejor gestión del talento. El IIC aplica **analítica descriptiva y predictiva** para optimizar procesos de selección, predecir el absentismo o la rotación e identificar a los profesionales con más potencial, entre otros proyectos.

EVALUACIÓN DE COMPETENCIAS

Dentro de la **plataforma online eValue,** desarrollamos pruebas objetivas y fiables para evaluar las competencias transversales, el nivel de inglés o las motivaciones de candidatos y empleados. Además de tomar mejores decisiones, se obtienen **datos de calidad** para analizar, por ejemplo, sus necesidades de formación.

ANÁLISIS DE REDES ORGANIZACIONALES

Los proyectos AROS permiten analizar las relaciones de trabajo y las interacciones entre los profesionales. Representadas visualmente en un grafo, se pueden identificar **redes informales, referentes ocultos o cuellos de botella** en la organización, para emprender acciones de mejora.

NUESTROS ASOCIADOS:

Santander | MAPFRE | IBM | UAM Universidad Autónoma de Madrid

Instituto de Ingeniería del Conocimiento

C/ Francisco Tomás y Valiente, 11 EPS, Edificio B, 5ª planta UAM Cantoblanco. 28049 Madrid

http://www.iic.uam.es/empleo-iic/

contacto.iic@iic.uam.es

(+34) 91 497 2323

www.linkedin.com/company/instituto-de-ingenier-a-del-conocimiento---iic

www.twitter.com/IIConocimiento

www.youtube.com/IIConocimiento

INNOVADATA

instituto
de ingeniería
del conocimiento

www.iic.uam.es

Instituto de Ingeniería del Conocimiento

C/ Francisco Tomás y Valiente, nº 11
Escuela Politécnica Superior (EPS),

Edificio B, 5ª planta
Universidad Autónoma de Madrid (UAM).

28049 Cantoblanco, Madrid

T. (+34) 91 497 23 23