INTERVIEW ASSESMENT

OVERVIEW

1. Data Loading and Inspection

Data Loading: The first step involved importing the dataset into a Python environment using the pandas library. This allowed us to load the data from its source, typically a CSV or Excel file, for further analysis.

Data Inspection: We performed an initial inspection of the dataset to gain insights into its structure and content. Various functions like head() , info(), describe() were employed to review the dataset, providing an overview of its format and data types.

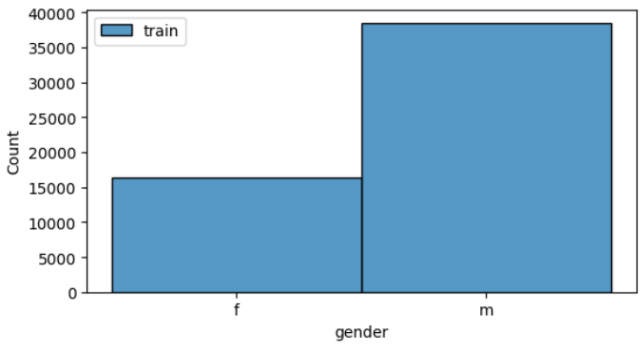
Checking for Missing Values: To ensure data completeness, we used functions like isnull() and info() to identify any missing values within the dataset. Detecting missing data is crucial for subsequent data cleaning steps.

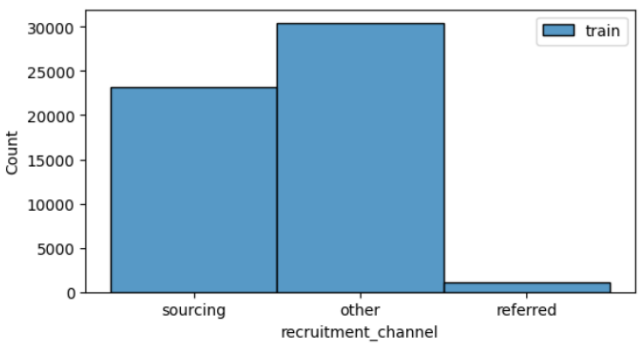
2. Data Cleaning

Handling Missing Values: The dataset contained missing or null values, which were addressed based on the nature of the missing data. Missing values were either filled with appropriate data, using strategies like mean or median imputation, 0, or rows/columns with excessive missing data were removed.

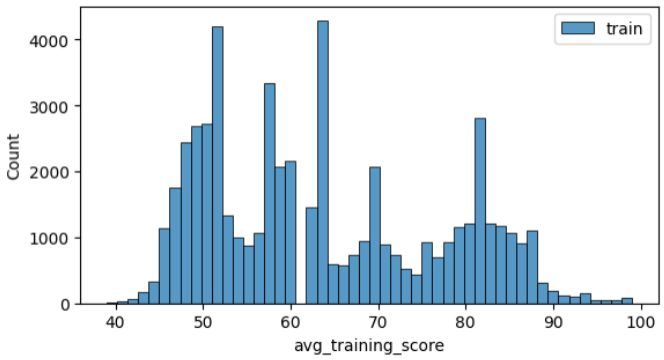
3. Data Visualization

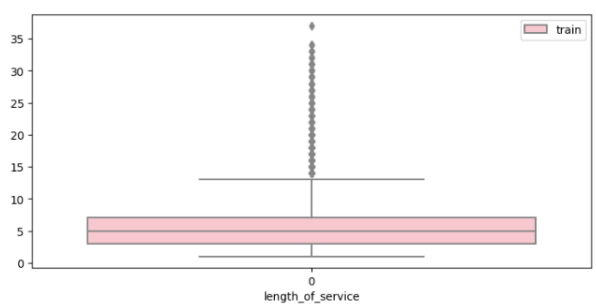
Visualizing Distributions: To gain a deeper understanding of employee information distributions, we employed data visualization techniques. Histograms and box plots were created to visualize the distributions, aiding in the identification of patterns, outliers, and trends within the data.

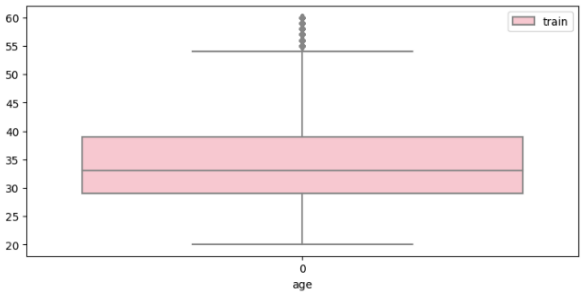




INTERVIEW ASSESMENT







4. Data Pipeline

Creating a Data Pipeline: A data pipeline was established using the pandas library to automate the data cleaning process. This pipeline streamlines and standardizes data cleaning procedures for use with new datasets.

Implementing Functions: We developed functions that encapsulated the data cleaning steps, and ease of application to new datasets. User can input a .csv file and these functions encompassed tasks such as handling missing values, duplicate removal, and other preprocessing steps and returns and clean dataframe.

INTERVIEW ASSESMENT

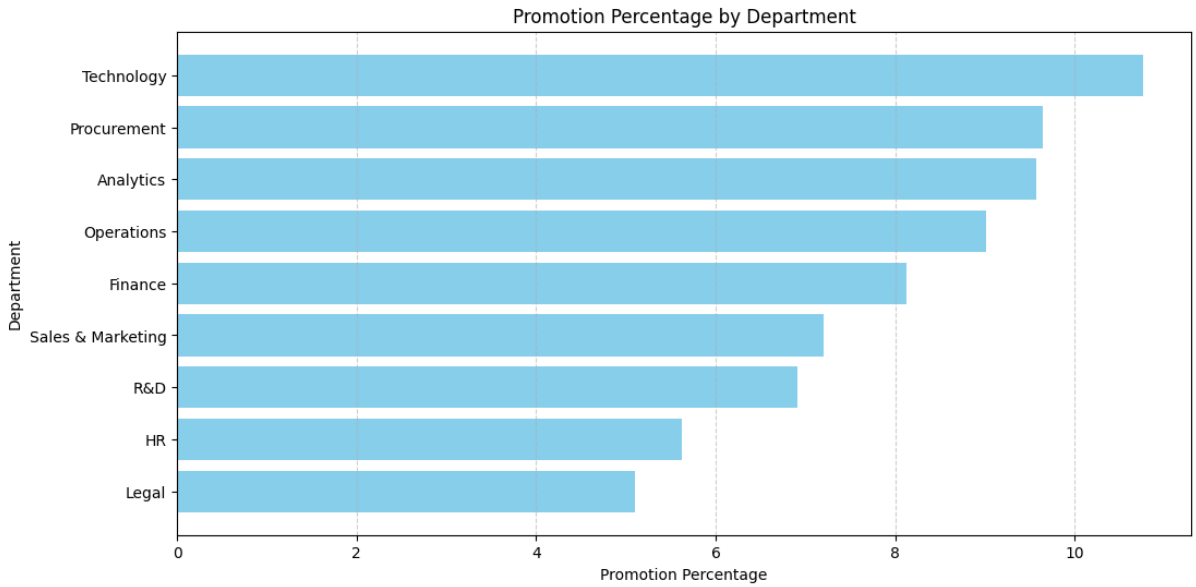
Answer -1

In my analysis of promotions within the organization, it becomes evident that the Technology department boasts the highest promotion ratio among all departments. Within the Technology department, approximately 10.7% of employees have received promotions, showcasing a robust commitment to recognizing and advancing talent in this department.

Conversely, the Legal department reports a comparatively lower promotion ratio. Here, around 5.1% of employees have secured promotions, suggesting a different dynamic within this division in terms of career advancement.

It's important to note that while there is a noticeable difference in promotion percentages between these two departments, the disparity is not significantly significant. Such variations may be attributed to the unique nature of roles and responsibilities within each department, as well as specific promotion criteria and opportunities associated with certain fields within the company.

These findings unveil valuable insights into the promotion landscape within an organization, indicating areas where promotions are relatively more common and those where they may be less frequent. They can serve as a foundation for further investigation and strategic decision-making, helping to foster equitable career development and promotion practices across all departments.



INTERVIEW ASSESMENT

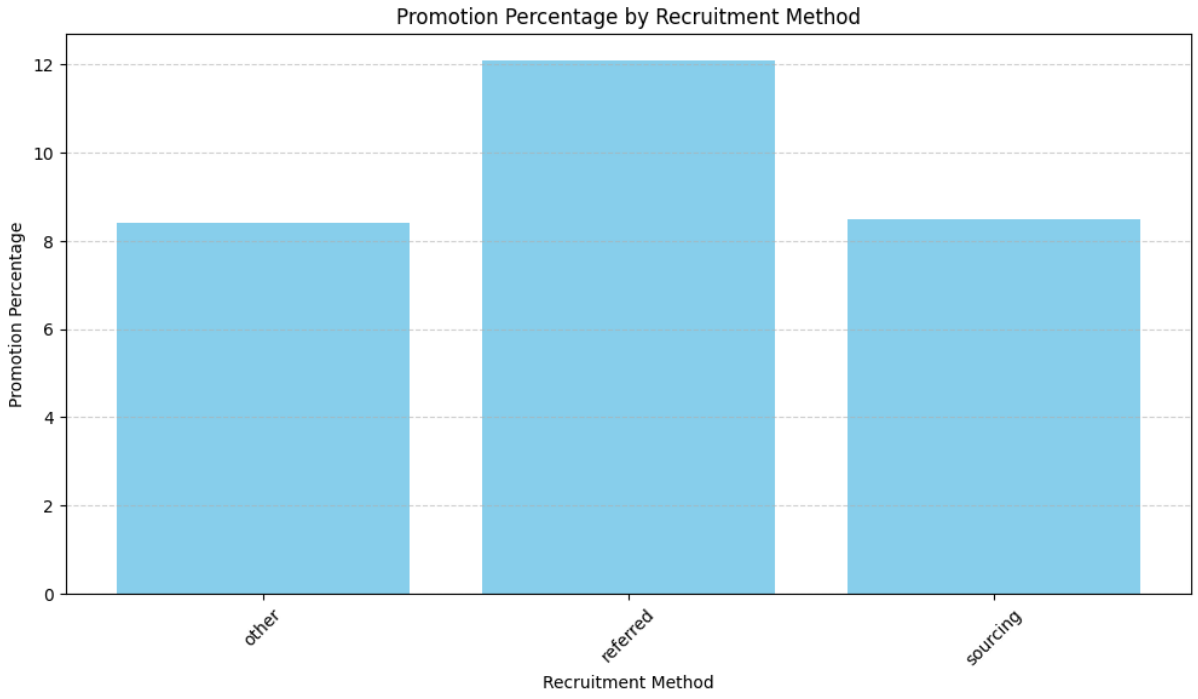
Answer – 2

In my analysis of promotion percentages based on recruitment methods, I observe distinct differences that shed light on the effectiveness of each approach. The data shows that employees recruited through referrals, have achieved a notably higher percentage of promotions compared to other recruitment methods.

Specifically, the promotion percentage for employees recruited through referrals stands at 12.08%. This means that approximately 12.08% of employees who joined the organization through referrals have been promoted during their tenure. This is a substantial and statistically significant increase compared to the promotion percentages associated with other methods. For instance, the "other" recruitment method reports a promotion percentage of 8.40%, while the "sourcing" method shows a promotion percentage of 8.50%.

The significance of these differences cannot be understated. The higher promotion percentage for referrals indicates that this recruitment method has been particularly effective in identifying and nurturing talent within the organization. It suggests that employees who join through referrals not only integrate well into the company culture but also excel in their roles, leading to a higher likelihood of career advancement.

This finding underscores the importance of employee referrals as a valuable recruitment channel for the organization. It highlights the potential benefits of encouraging and incentivizing employees to refer suitable candidates, as it not only contributes to workforce diversity but also enhances the chances of identifying high-potential individuals who can make significant contributions to the organization's success.



INTERVIEW ASSESMENT

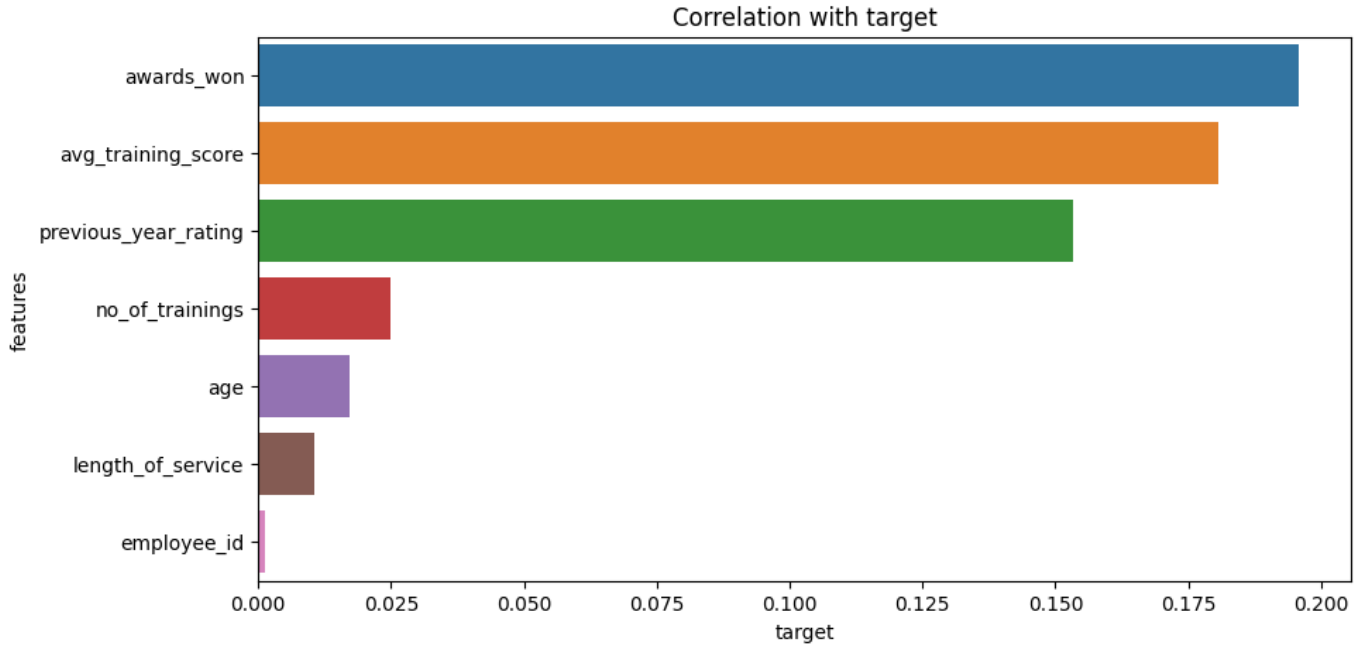
Answer – 3

Our analysis reveals compelling insights into the relationship between the previous year's performance rating and the likelihood of employees getting promoted.

A statistical examination using the Point-Biserial Correlation Coefficient yielded a positive correlation coefficient of 0.15. This coefficient signifies a positive linear relationship between the two variables, indicating that as the previous year's performance rating increases, the likelihood of an employee getting promoted also tends to increase.

Furthermore, the associated p-value, which is reported as 0.0000, is well below the common significance threshold of 0.05. This suggests that the correlation observed is not due to random chance but is indeed statistically significant.

Therefore, there exists a statistically significant positive correlation between an employee's previous year's performance rating and the probability of them receiving a promotion. This finding underscores the importance of performance assessments in the promotion process, with higher ratings generally associated with an elevated likelihood of career advancement within the organization.

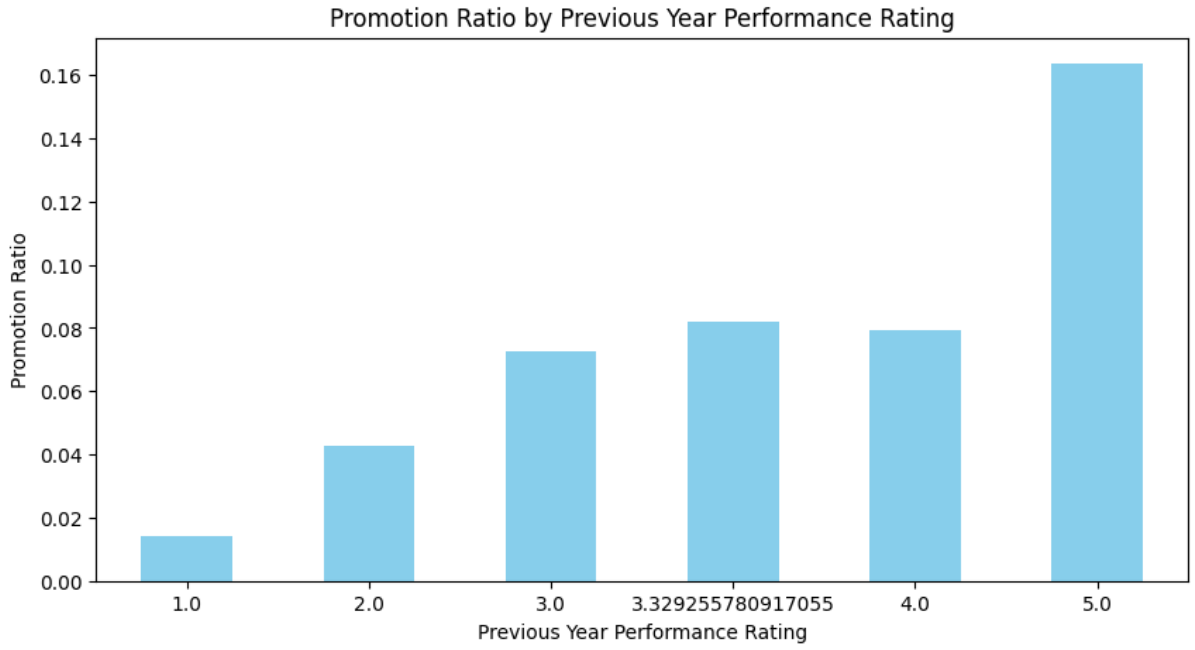


INTERVIEW ASSESMENT

Answer – 4

The analysis of the data indicates a clear trend that the ratio of promoted employees increases significantly with higher previous year's performance ratings. In this examination, I calculated the promotion ratio for each distinct performance rating, how an employee's past performance is linked to their likelihood of receiving a promotion.

The graph demonstrates that employees with higher previous year's performance ratings have notably higher promotion ratios. This means that the employee who received better performance ratings in the previous year are more likely to be promoted in the current year. This relationship suggests a positive correlation between previous year's performance and promotion. Such insights are invaluable for human resource management and organizational decision-making, as they provide empirical evidence that performance evaluation plays a substantial role in the promotion process.



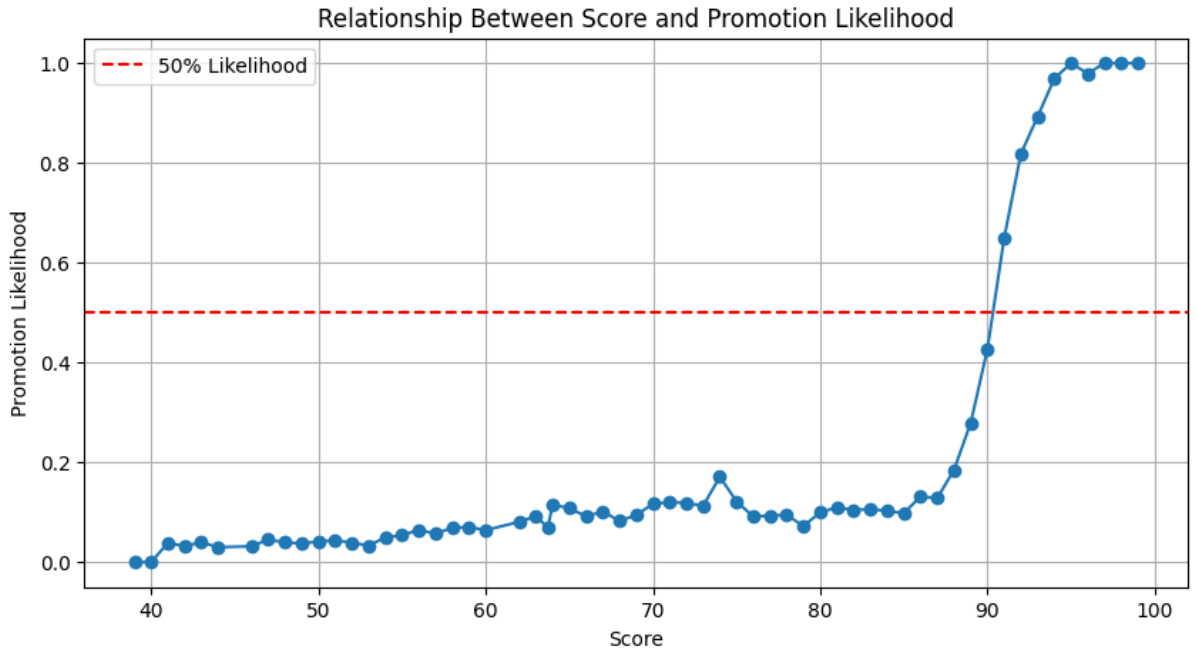
INTERVIEW ASSESMENT

Answer – 5

Analysis of the data reveals an interesting trend regarding the relationship between average training scores and the likelihood of getting promoted. It is evident that as the average training score increases, so does the probability of receiving a promotion. This suggests a positive correlation between training performance and promotion likelihood, indicating that employees who excel in their training are more likely to be promoted.

However, it's noteworthy that this correlation is not linear across the entire range of training scores. Instead, we observe that the likelihood of getting promoted reaches its maximum within a specific range of training scores, typically between 90 and 100. Within this range, employees who consistently achieve high training scores have the highest probability of being promoted.

These findings underscore the importance of training performance as a predictor of promotion within the organization. Managers and decision-makers may find it beneficial to pay particular attention to employees who fall within the peak range of training scores, as they are the most likely candidates for advancement.



INTERVIEW ASSESMENT

Answer – 6

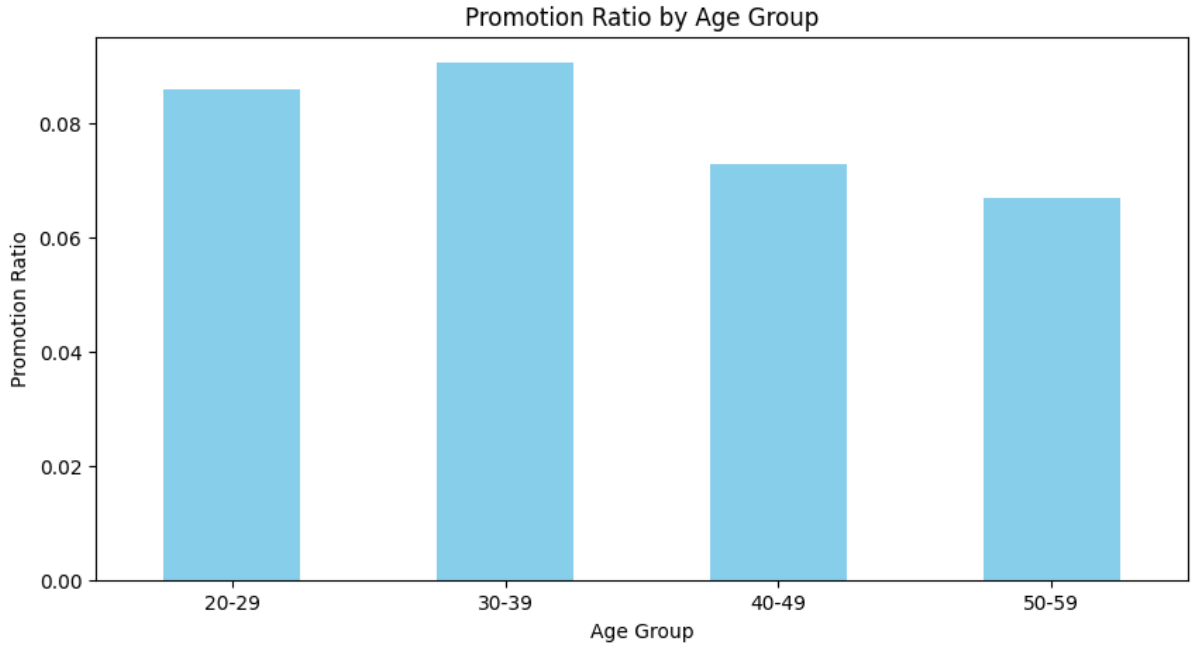
In this analysis, upon conducting a statistical examination, I found that the Point-Biserial Correlation Coefficient between age and promotion ratio is approximately -0.02. This coefficient represents the strength and direction of the relationship between age and promotion ratio. In this case, the negative sign (-0.02) indicates a negative correlation, suggesting that as age increases, the promotion ratio tends to decrease slightly.

The accompanying p-value, which is reported as 0.0001, is well below the conventional significance threshold of 0.05. This indicates that the observed correlation is statistically significant and not a result of random chance.

In practical terms, the negative correlation suggests that there is a slight tendency for older employees to have a marginally lower promotion ratio compared to younger employees. However, it's important to note that the correlation coefficient is very close to zero, indicating a weak correlation. This implies that age alone has a minimal impact on promotion ratios, within the organization.

Additionally, when we visualize the data, the graph illustrates that the promotion ratio remains relatively consistent across different age groups. The absence of significant variations in the promotion ratio among age groups further supports the notion that age is not a predominant factor influencing promotion outcomes.

This analysis reveals a statistically significant, albeit weak, negative correlation between age and promotion ratio. While there is a slight trend of lower promotion ratios among older employees, it is not a dominant or substantial factor in promotion decisions. Other variables and criteria likely contribute more significantly to the promotion process within the organization.



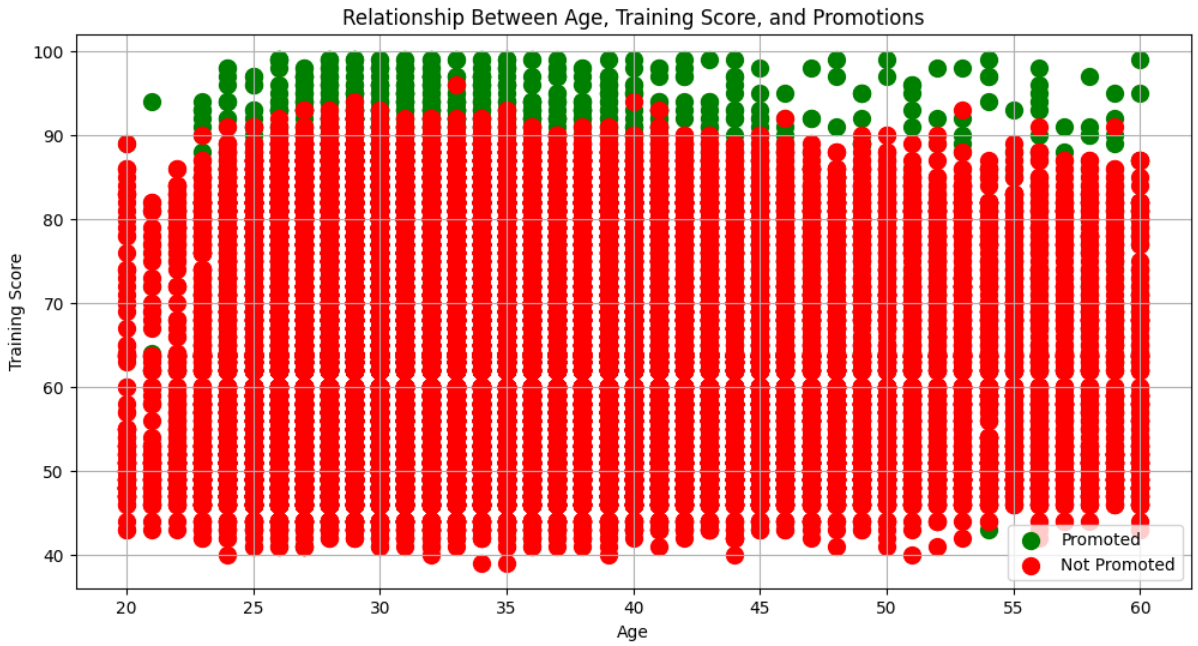
INTERVIEW ASSESMENT

Answer – 7

This analysis leads us to a conclusion that promotions within our organization do not appear to be significantly influenced by age. The statistical examination and visualization of promotion rates across different age groups have shown that there is no pronounced variation in promotion likelihood based solely on age. This suggests that age, as a standalone factor, may not be a significant determinant in our promotion decisions.

However, the analysis uncovers a contrasting narrative when it comes to average training scores. The visualization provides compelling evidence of the impact of training scores on promotions. The scatter plot, where green dots represent promoted candidates, clearly illustrates that most promoted employees are concentrated in the upper portion of the graph along the y-axis throughout the x-axis range of training scores. This visually striking pattern suggests that employees who consistently achieve higher training scores are more likely to receive promotions. Thus, it is evident that training performance is a pivotal factor influencing promotion decisions within our organization.

In summary, our analysis highlights that while age does not significantly affect promotion prospects, average training scores play a substantial role. The visual representation reinforces this insight, underlining the importance of consistently high training performance as a key driver of promotions. These findings can guide our organization in refining promotion criteria and focusing efforts on nurturing and recognizing employees with exceptional training accomplishments.



INTERVIEW ASSESMENT

Answer – 8

Our analysis of the employee promotion dataset has uncovered a notable issue: a substantial number of missing values in the previous year's rating column. As indicated in the data visualization, a considerable portion of this column consists of null values, which can be attributed to various factors, including potential manual errors during data entry or incomplete data collection processes.

To address this challenge and ensure the dataset's reliability and completeness, several data management recommendations can be made. First and foremost, it is essential to conduct a thorough data cleaning process. This involves identifying and rectifying any discrepancies or inaccuracies in the dataset. Data entry procedures should be meticulously reviewed, and measures should be implemented to minimize the chances of manual errors during data input.

Furthermore, for cases where missing values in the previous year's rating are inevitable due to incomplete data, it is advisable to explore data imputation techniques. Imputation methods, such as mean imputation or regression imputation, can be employed to estimate missing values based on available data, thereby enhancing the dataset's completeness and analytical potential.

The presence of missing values is an important problem for robust data management practices. By assuring good quality data, we can enhance the dataset's integrity, ultimately ensuring that our analyses and decision-making processes are based on the most reliable and complete information available.

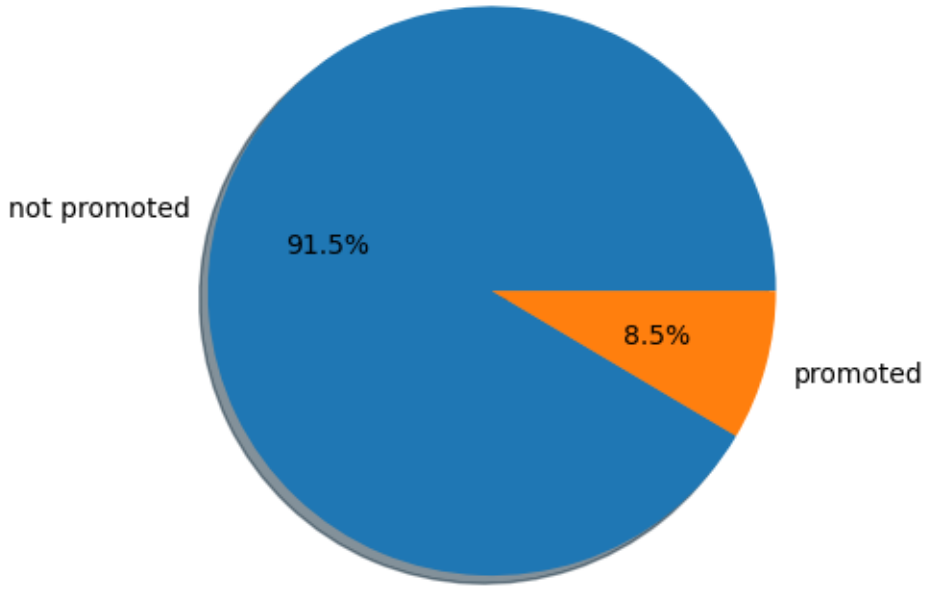
```
0s
df.isnull().sum()

employee_id              0
department               0
region                   0
education             2409
gender                   0
recruitment_channel      0
no_of_trainings          0
age                      0
previous_year_rating    4124
length_of_service        0
awards_won               0
avg_training_score    2560
is_promoted              0
dtype: int64
```

INTERVIEW ASSESMENT

Answer – 9

Despite the inherent imbalance in our dataset, where only a relatively small proportion, approximately 8.5%, of employees have been promoted, valuable insights can still be derived to identify the factors that are more likely to contribute to an employee's promotion within the organization.



One of the most revealing observations from our analysis is the identification of key features that exhibit a correlation with employee promotions. Among these features, three standout as particularly influential:

Awards Won: Our data analysis highlights that the number of awards an employee has won is positively correlated with the likelihood of receiving a promotion. This suggests that employees who have been recognized and awarded for their contributions tend to have a higher chance of advancement within the organization.

Average Training Score: Another prominent factor is the average training score of employees. Those with higher training scores are more likely to be promoted. This correlation underscores the significance of continuous learning and skill development in an employee's career progression.

Previous Year's Rating: A strong correlation is also observed between an employee's previous year's performance rating and their chances of promotion. Higher ratings in the

INTERVIEW ASSESMENT

previous year are associated with a greater likelihood of promotion, emphasizing the importance of consistent performance and appraisal.

These insights, even within an imbalanced dataset, suggest that a combination of factors, including recognition through awards, investment in training and skill enhancement, and consistent high performance, can increase an employee's likelihood of being promoted. While promotions may be less common due to the dataset's imbalance, focusing on these correlated features can guide organizations in identifying and nurturing talent with a higher potential for advancement.

Although promotions remain a relatively rare occurrence within our dataset, the correlation analysis highlights key features that organizations can prioritize when assessing employee’s readiness for career progression. By recognizing and fostering these qualities, organizations can work towards a more equitable and strategic approach to promotions, benefiting both employees and the organization.