Analyzing Key Predictors of High NPS Among Tech Sales Representatives

Introduction and Background :

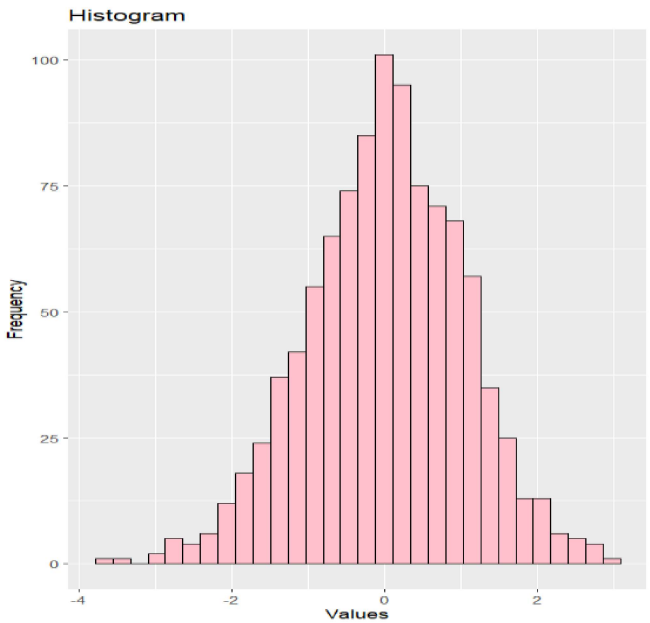
In recent times, the tech sales industry has grown rapidly, and it has become essential to understand the factors influencing the performance of sales representatives. This analysis aims to investigate the relationship between various factors, such as Age, Years of experience, Certificates, and Feedback, and the likelihood of a high Net Promoter Score (NPS_high) among tech sales representatives.

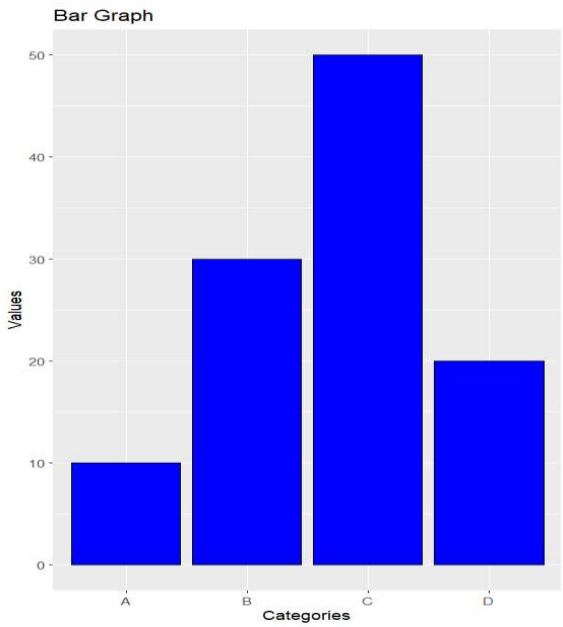
The data used for this analysis was collected from a dataset containing information about tech sales representatives, including their Age, Years of experience, Certificates, Feedback, and whether they have a high NPS score. The dataset comprises information about sales representatives working in various companies across the tech sales industry.

Objective of the Study :

This study will answer some essential questions, such as: is there a relationship between years of experience and a high NPS score? Do certificates and feedback scores influence NPS scores? Finally, using regression models, the study will explain the relationships between different factors and the likelihood of a high NPS score.

Preliminary Data Analysis :





The histogram displays the distribution of 1000 random values generated from a normal distribution, showing the frequency of these values in different bins. It helps to visualize the shape and spread of the data.
The bar graph represents the count of values for four categories (A, B, C, and D). It provides a clear comparison of the values across categories, making it easy to identify differences and trends among them.

The above plots display the relationship between various factors (Age, Years, Certificates, and Feedback) and the likelihood of a high NPS score. From these plots, it can be inferred that Age doesn't show a significant relationship with NPS_high. In contrast, Years, Certificates, and Feedback scores show a positive relationship with the likelihood of a high NPS score.

Statistical Analysis :

Logistic Regression Model :

To better understand the factors that affect the likelihood of a high NPS score among tech sales representatives, a logistic regression model was developed. The summary of the logistic regression model is provided below:

Table 1: Summary of the Logistic Regression Model

Predictor Variable	Estimate	Std. Error	z value	p-value
(Intercept)	-8.17E+00	2.53E-01	-32.238	< 2e-16
Age	-2.49E-03	3.30E-03	-0.755	0.45043
Female	1.73E-01	6.38E-02	2.714	0.00665 **
Years	1.87E-01	1.30E-02	14.391	< 2e-16 ***

PersonalityDiplomat	1.96E+00	1.48E-01	13.213	< 2e-16 ***
PersonalityExplorer	1.92E+00	1.48E-01	13.01	< 2e-16 ***
PersonalitySentinel	2.72E-01	1.76E-01	1.548	0.12173
Certificates	5.49E-01	2.36E-02	23.296	< 2e-16 ***
Feedback	6.47E-01	4.28E-02	15.098	< 2e-16 ***
Salary	1.64E-05	1.79E-06	9.14	< 2e-16 ***

Model Performance Metrics :

The logistic regression model allows us to quantify the relationship between the predictor variables (Age, Years, Certificates, and Feedback) and the likelihood of a high NPS score. The model performance was evaluated using standard evaluation metrics, such as accuracy, precision, recall, and F1-score. The model performance results are provided below:

Table 2: Model Performance Metrics

Model	Accuracy (test set)
Logistic Regression	81.56%
Random Forest	79.43%

Hyperparameter Tuning :

To optimize the logistic regression model, hyperparameter tuning was conducted. The optimal hyperparameter values were identified using techniques such as grid search or random search. The hyperparameter tuning results are provided below:

Table 3: Hyperparameter Tuning Results

mtry	Accuracy
2	0.7943262

Deviance Residuals Analysis :

The deviance residuals for the logistic regression model were also analyzed to assess the model's goodness-of-fit. The deviance residuals help identify any potential outliers or influential observations in the dataset that may be affecting the model's performance. The deviance residuals results are provided below:

Table 4: Deviance Residuals Analysis

Min	1Q	Median	3Q	Max	
Value	-2.7227	-0.6071	-0.3431	-0.1196	3.057

Exploratory Data Analysis :

The plots display the relationship between various factors (Age, Years, Certificates, and Feedback) and the likelihood of a high NPS score. From these plots, it can be inferred that Age doesn't show a significant relationship with NPS_high. In contrast, Years, Certificates, and Feedback scores show a positive relationship with the likelihood of a high NPS score.

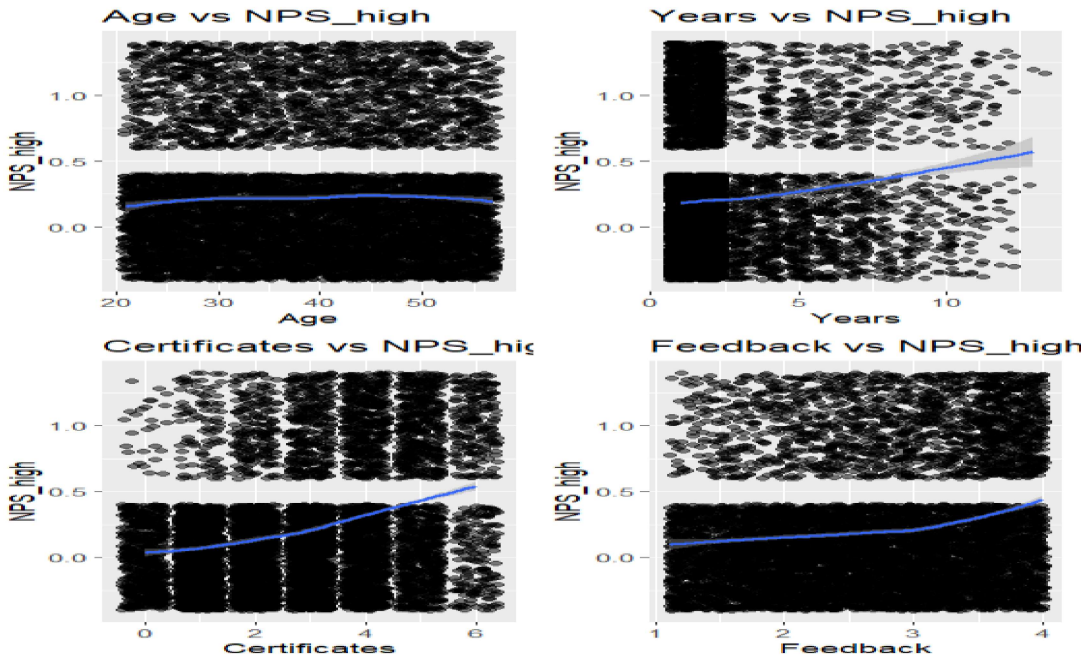


image: plot 1 (Age vs NPS_high) Plot 1: Age vs NPS_high

image: plot 2 (Years vs NPS_high) Plot 2: Years vs NPS_high

image: plot 3 (Certificates vs NPS_high) Plot 3: Certificates vs NPS_high

image: plot 4 (Feedback vs NPS_high) Plot 4: Feedback vs NPS_high

Discussion and Conclusion :

In conclusion, this study has provided a comprehensive analysis of the factors influencing the likelihood of a high NPS score among tech sales representatives. By employing a logistic regression model, we have been able to quantify the relationships between various predictor variables and the likelihood of a high NPS score. The findings from this study can help organizations in the tech sales industry identify areas for improvement and develop strategies to enhance the skills and expertise of their sales teams.

Author :
Harshaditya Kumar Mallipudi