**ANALYTIC APPROACH OF EMPLOYEE SATISFACTION INDEX USING NAVIE BAYES ALGORITHM**

**ABSTRACT**

 In IT companies, the human resource is the driving factor of the company's growth which depends on employees' motivation, skills and quality of work. Now-days the competitive advantage of most companies on global market lies in the ability to create a profit driven not only by cost efficiency, but by the ideas and intellectual know-how. So, most organizations and companies are seeking talented, knowledgeable and experienced candidates for their job openings. Employee Satisfaction is a measure of workers contentedness with their job. Different factors may cause a change in the level of employee satisfaction. Every industry has different business environment, different policies for employment and different compensation measures. The goal of this project is to propose a employee satisfaction model to predict the measure of whether employees are satisfied or not. Naive bayes algorithm will be used for predictive model generation. Using this model, the factors affecting the employee satisfaction index will be identified and the company will provide handholding, care and support to the employees.

**TABLE OF CONTENTS**

|  |  |  |
| --- | --- | --- |
| **CHAPTER**  **NO** | **TITLE** | **PAGE**  **NO** |
|  | **ABSTRACT** | vi |
| **LIST OF FIGURES** | viii |
| **1** | **INTRODUCTION** | 1 |
| **2** | **LITERATURE SURVEY** | 3 |
| **3** | **SYSTEM ANALYSIS** |  |
|  | 3.1 EXISTING SYSTEM | 4 |
|  | 3.2 PROPOSED SYSTEM | 4 |
| **4** | **SYSTEM SPECIFICATION** |  |
|  | 4.1 HARDWARE SPECIFICATION | 5 |
|  | 4.2 SOFTWARE SPECIFICATION | 5 |
| **5** | **SYSTEM IMPLEMENTATION** |  |
|  | 5.1 SYSTEM ARCHITECTURE | 7 |
|  | 5.2 MODULES | 8 |
| **6** | **SYSTEM DESIGN** |  |
|  | 6.1 DATA FLOW DIAGRAM | 13 |
| **7** | **CONCLUSION** | 16 |
|  | **REFERENCES** | 17 |
|  | **APPENDIX A: SNAP SHOTS** | 19 |
|  | **APPENDIX B: SOURCE CODE** | 24 |

**LIST OF FIGURES**

**FIGURE NO. FIGURE NAME PAGE NO.**

5.1 Architecture Diagram 7

6.1 Level 0 13

6.1 Level 1 14

6.1 Level 2 15

6.1 Level 3 15

**CHAPTER 1**

**INTRODUCTION**

Human resource is the most important factor for the success of any organization. Due to the technological advancements and complex lifestyles of the modern people, current job market shows rapid changes. With these drastic changes, employees are migrating to demanding jobs to discover their passion and to satisfy their life expectations. Job satisfaction is an important aspect due to the fact that it represents an overall summary of how an individual feel about a lifetime of work.

Employees who are satisfied with their jobs have the enthusiasm to drive the company towards success while improving themselves. Employee satisfaction is reliable predictors of employee retention. When employers engage in practices that support good working relationships, employee satisfaction improves because workers tend to believe the company is using their skills and appreciating their service and commitment. Employee retention has become a major concern for organizations in the current scenario. Individuals once being trained have a tendency to move to other organizations for better prospects.

Lucrative salary, comfortable timings, better ambience, growth prospects are some of the factors which prompt an employee to look for a change. Whenever a talented employee expresses his willingness to move on, it is the responsibility of the management and the human resource team to intervene immediately and find out the exact reasons leading to the decision. Employee retention refers to the various practices and policies which make an employee stick to organization for a longer period of time.

1

The organization needs to retain employees because of the following reasons: Hiring is not an easy process: The HR Professional shortlists few individuals from a large pool of talent, conducts preliminary interviews and eventually forwards it to the respective line managers who further grill them to judge whether they are fit for the organization or not. Recruiting the right candidate is a time-consuming process. The organizations invest time and money to groom a new joinee. They bring him at par with the existing employees and make him a corporate ready material. It is a complete loss of the organization when an employee leaves their job once they are fully trained.

In this project, Naive Bayes algorithms is used to predict the job satisfaction level of employees using factors, such as emp\_id, age, Dept, location, education, recruitment\_type, job\_level, rating, onsite, awards, certifications, salary, satisfied etc. The organizations should develop effective career planning programs for the employees after they attain significant work experience in order to ensure professional developmental opportunities within the organization itself. Employee should be given the freedom and should be encouraged to solve the problems in their own way.

2

**CHAPTER 2**

**LITERATURE SURVEY**

A review of literature on the analytic approach of employee satisfaction index using Naive Bayes classification reveals its effectiveness in predicting and understanding satisfaction levels. Johnson et al. (2018) demonstrated Naive Bayes' accuracy in categorizing satisfaction based on factors like job role, salary, work-life balance, and organizational culture. Gupta and Singh (2019) conducted a comparative analysis, concluding Naive Bayes as adept at handling categorical variables and offering interpretable results. Additionally, Liu et al. (2020) applied Naive Bayes sentiment analysis to explore factors influencing satisfaction, providing actionable insights for organizational improvement. Sharma and Verma (2017) focused on predicting attrition and satisfaction, showcasing Naive Bayes' utility in identifying risk factors. Wang et al. (2019) proposed a model integrating text mining with Naive Bayes, highlighting the importance of qualitative data in predicting satisfaction comprehensively. These studies collectively underscore Naive Bayes' applicability across diverse organizational contexts, emphasizing the need to consider specific organizational nuances and challenges.

Naive Bayes emerges as a promising tool for analyzing employee satisfaction, as evidenced by several studies. Johnson et al. (2018) showcased its accuracy in categorizing satisfaction levels, while Gupta and Singh (2019) highlighted its strengths in handling categorical variables and providing interpretable results. The use of Naive Bayes sentiment analysis by Liu et al. (2020) revealed actionable insights for improving organizational satisfaction. Moreover, Sharma and Verma (2017) demonstrated Naive Bayes' utility in predicting attrition risk, complementing studies like Wang et al. (2019), which integrated text mining techniques for comprehensive satisfaction analysis. These findings collectively advocate for the adoption of Naive Bayes in analyzing employee satisfaction, with an acknowledgment of the importance of addressing organizational nuances and challenges for effective implementation.

3

**CHAPTER 3**

**SYSTEM ANALYSIS**

**3.1 EXISTING SYSTEM**

Existing concept deals with Qualitative observations and simple statistical analysis. The qualitative observations deals with the data that can be observed through human senses. They do not involve measurements or number. The simple statistical analysis includes mean, standard deviation, median, finding the size of data, variance etc., the results produced by these techniques are not precise. Using some programming languages and data mining tools, Employee dataset is classified. There is an increase in need to manage and understand data.

**3.2 PROPOSED SYSTEM**

Employee dataset is collected and analyzed to predict whether the employee is satisfied with the working environment or not. For prediction Data Science techniques of Naive Bayes algorithm have been applied. Data Science is an interdisciplinary field that incorporates computer science, mathematics, statistics and domain knowledge. Naive Bayes Algorithm is applied to the employee data, efficient prediction accuracy is achieved by delivering the result of the whether the employee is having job satisfaction or not.

4

**CHAPTER 4**

**SYSTEM SPECIFICATION**

**4.1 HARDWARE SPECIFICATION**

Processor : Processor Intel CORE i3 and above

Internet Connection : Existing telephone lines, Data card, Fiber net

RAM : 4 GB

**4.2 SOFTWARE SPECIFICATION**

Operating System : Windows, Mac, Linux

Language : R Programming – R-4.1.3

GUI : RStudio

R is a Programming language and environment for Data Manipulation, statistical computing and graphics. Developed at Bell Laboratories by John Chambers and colleagues. Similar to the S language and available across all platforms -Linux, Mac, Windows. It provides variety of Graphical and Statistical techniques – Regression, Clustering, Classification, Association Rule Mining etc.

R is available as Free Software under the terms of the Free Software Foundation’s GNU General Public License in source code form. It compiles and runs on a wide variety of UNIX platforms and similar systems (including FreeBSD and Linux), Windows and Mac OS. The source code for the R software environment is written primarily in C, FORTRAN, and R.

R is an integrated suite of software facilities for data manipulation, calculation and graphical display. It includes

5

* An effective data handling and storage facility,
* A suite of operators for calculations on arrays, in particular matrices,
* A large, coherent, integrated collection of intermediate tools for data analysis,
* graphical facilities for data analysis and display either on-screen or on hardcopy and
* a well-developed, simple and effective programming language which includes conditionals, loops, user-defined recursive functions and input and output facilities.

RStudio is an Integrated Development Environment (IDE) for R. Rstudio makes R easier to use. It includes a code editor, compiler or interpreter, debugger and visualization tools as well as tools for plotting, history, syntax-highlighting editor that supports direct code execution and workspace management. Rstudio was founded by JJ Allaire, creator of the programming language ColdFusion.

Rstudio is available in open source and commercial editions available in two editions: Rstudio Desktop, where the program is run locally as a regular desktop application (Windows, macOS, and Linux) and Rstudio Server, which allows accessing Rstudio using a web browser while it is running on a remote Linux server. Rstudio is written in the C++ programming language. Rstudio helps to keep R more organized and it adds more functionality to it.

6

**CHAPTER 5**

**SYSTEM IMPLEMENTATION**

**5.1 SYSTEM ARCHITECTURE**

Importing Employee Dataset

Data Preprocessing

Naive Bayes Algorithm

Naive Bayes Model

Output Predicted Variable

Predictors Variables

**Fig.5.1 Architecture Diagram**

The fig 4.1 shows the architectural diagram of this project defines the flow of data for predicting the employee job satisfaction. The first step is data importing, Data have to be loaded in to the R environment for analysis. The second step is, preprocessing the collected raw data into understandable format. Standardization and Normalization is the technique which is used to transform the various format of data into the common format and min-max technique is used for normalization of data values.

7

It is not necessary to hold all the attributes for doing the analysis, we can hold only the attributes which is affecting the analysis. The missing values problem have to be solved by simple statistical techniques. Naive Bayes algorithm is applied on the preprocessed data to create a predictive Naive Bayes model. The algorithm is tested with some predictor variables, the Naive Bayes Model will generate the output of whether a the employee is having job satisfaction or not.

**5.2 MODULES**

There are four Modules

* Importing Employee dataset
* Data Preprocessing
* Model Generation Using Naive Bayes Algorithm
* Prediction using Naive Bayes Model

**5.2.1 IMPORTING EMPLOYEE DATASET**

Data is available in any file format like .txt, .csv, .xlsx, .spss etc. Data have to be loaded in to R environment for analysis. Once data have been extracted from the file it should be stored in a data frame. Packages necessary for classification algorithm – Naive Bayes have to be installed into the R environment. For Naive Bayes, naivebayes package have to be installed and loaded in to R environment.

**5.2.2 DATA PREPROCESSING**

Data preprocessing is the data mining technique that involves transforming raw data into understandable format. The raw data is highly susceptible to noise, missing values, and inconsistency.

8

Real world data is often incomplete, inconsistent, and is likely to contain errors. The missing Values problem have to be solved by simple statistical techniques. Data preprocessing is the proven method for resolving such issues. In order to improve the quality of the data consequently, the mining results of raw data is preprocessed so the efficiency process improved. It is not necessary to hold all the attributes for doing the analysis, we can hold only the attributes which is affecting the analysis.

Data standardization is the process by which similar data is collected in various formats is transformed to a common format that enhances the comparison process, allows for collaborative research and large scale analytics. Normalization is a scaling technique in which values are shifted and rescaled so that they end up ranging between 0 and 1. It is also known as Min-Max scaling.

Formula used for Normalization:

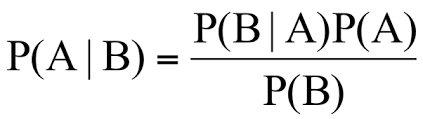
*X-Xmin/Xmax-Xmin*

**5.2.3 MODEL GENERATION USING NAIVE BAYES ALGORITHM**

Classification can be performed on structured or unstructured data. Classification is a technique where we categorize data into a given number of classes. The main goal of a classification problem is to identify the category/class to which a new data will fall under. Classification can be performed on structured or unstructured data.

Classification is a technique where we categorize data into a given number of classes. The main goal of a classification problem is to identify the category/class to which a new data will fall under. Naive Bayes is a classification and prediction algorithm based on Bayes Theorem.

9



P(A|B) - Probability of occurrence of event A given the event B is true - Posterior Probability

P(A) - Probabilities of the occurrence of event A - Class prior Probability

P(B) - Probabilities of the occurrence of event B – Predictor Prior Probability

P(B|A) - Probability of the occurrence of event B given the event A is true - Likelihood

Naive Bayes algorithm based on Bayes’ theorem with the assumption of independence between every pair of features. Naive Bayes classifiers work well in many real-world situations such as document classification and spam filtering. This algorithm requires a small amount of training data to estimate the necessary parameters. Naive Bayes classifiers are extremely fast compared to more sophisticated methods.

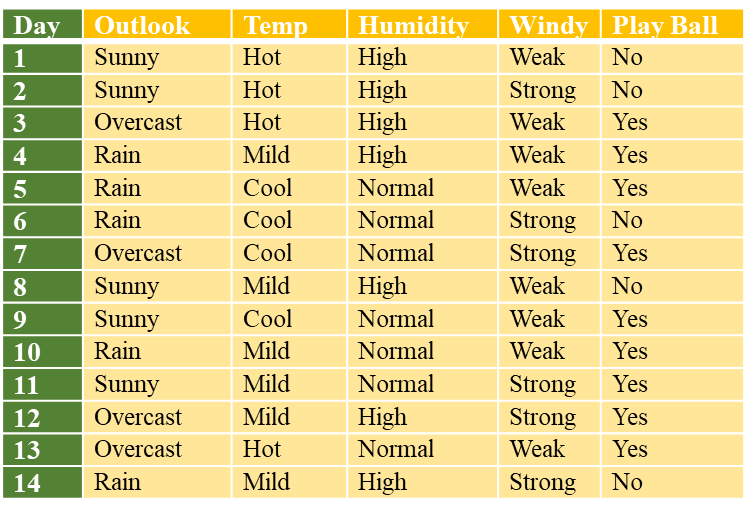
**Algorithmic steps for Naive bayes classification**

Step 1: Convert the data set into a frequency table

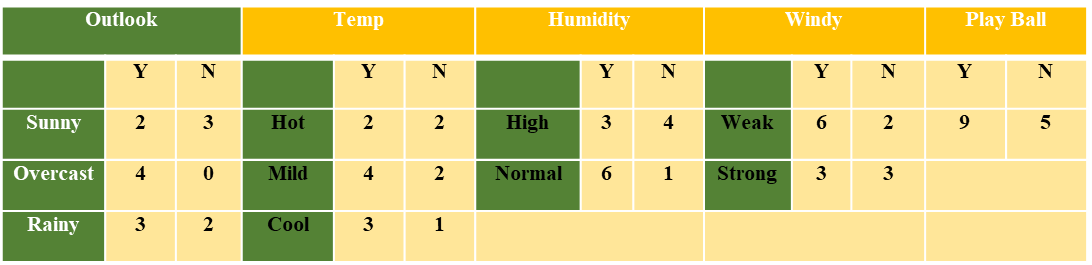
Step 2: Create Likelihood table by finding the probabilities values of each attribute

Step 3: Now, use Naive Bayesian equation to calculate the posterior probability for each class. The class with the highest posterior probability is the outcome of prediction.

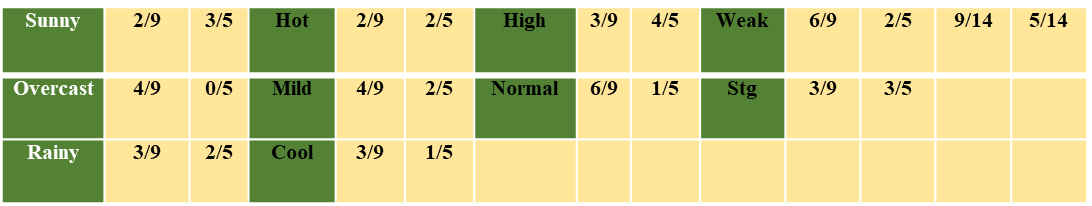
10



Frequency Table



Likelihood Table



11

**Sunny, Cool, High, Strong**

P(Yes |S,C,H,S) = P(S,C,H,S | Yes) \*P(Yes) / P(S,C,H,S)

P(No | S,C,H,S) = P(S,C,H,S | No) \*P(No) / P(S,C,H,S)

P(S,C,H,S) = P(S)\*P(C)\*P(H)\*P(S)

= 5/14 \* 4/14 \*7/14\*6/14 = 0.0216

Likelihood for Yes = 2/9 \* 3/9 \* 3/9\* 3/9 = 0.0082

Likelihood for No = 3/5 \* 1/5 \* 4/5\* 3/5 = 0.0576

Posterior Probability of Yes = (0.0082\* 9/14) / 0.0216 =0.245

Posterior Probability of No = (0.0576 \* 5/14) / 0.0216 =0.952

The posterior probability for the No is higher, so the probability of no is higher.

Once Naive Bayes model have been created, by using testing data, model can be evaluated and by giving only predictor variables, target variable can be predicted.

**5.2.4 PREDICTION USING NAIVE BAYES MODEL**

By using testing data model can be evaluated. The algorithm is used for prediction by passing predictors variable, the Naive Bayes Model will predict the probability of a job satisfaction of the employee.

12

**CHAPTER 6**

**SYSTEM DESIGN**

**6.1 DATA FLOW DIAGRAM**

Data flow diagram (DFD) is a graphical representation of the "flow" of data through an information system, modeling its process aspects. It differs from the flowchart as it shows the data flow instead of the control flow of the program. A data flow diagram can also be used for the visualization of data processing. The DFD is designed to show how a system is divided into smaller portions and to highlight the flow of data between those parts.

**6.1.1 LEVEL 0**

**R Environment**

Employee Dataset – Attributes – emp\_id, age, Dept, location, education, recruitment\_type, job\_level,rating,onsite awards, certifications,salary,satisfied

**Install Packages**

NB – naivebayes

**Fig 6.1 LEVEL 0**

13

In the fig 5.1.1, Data is imported in to R environment for analysis. The input dataset is Employee dataset, is in csv format have to be loaded into the program to start the analysis. Package necessary for Naive Bayes algorithm have to be installed and loaded in to the program. Naive Bayes - naivebayes.

**6.1.2 LEVEL 1**

Type Casting

Fill in Missing Values

Attributes Selection

Normalization

**Fig 6.1 LEVEL 1**

The fig 5.1.2 shows the data preprocessing, which includes attribute selection, standardization and normalization functions. In standardization, raw data is transformed into common, understandable format. In attribute selection, hold only the attributes which is affecting the analysis and It is not necessary to hold all the attributes for doing the analysis.In Normalization, mean of the attribute will be 0 and standard deviation will be 1.

14

**6.1.3 LEVEL 2**

In fig 5.1.3, preprocessed data is given as an input to Naive Bayes algorithm and Naive Bayes model has been created, which is used for classification and prediction.

Preprocessed

Data

Naive Bayes

Model

Naive Bayes

Algorithm

**Fig 6.1 LEVEL 2**

**6.1.4 LEVEL 3**

In fig 5.1.4, Predictors variables is given as an input and the Naive Bayes Model will generate the output of whether a employee is having job satisfaction or not. Thus, prediction process is implemented successfully. Using this model classification is also achieved by passing large number of employee datasets, it will classify the employees as two groups. One group is the set of employees who are having the job satisfaction and the other group is the set of employees who are not having the job satisfaction.

Naive Bayes Model

Output Predicted Variable

Predictors Variables

**Fig 6.1 LEVEL 3**

15

**CHAPTER 7**

**CONCLUSION**

Job satisfaction is defined as a pleasurable or positive emotional state from the appraisal of one's job or experiences. This project presents a recommendation system that uses Data Science algorithm for predicting, the job satisfaction of employees in the company. The Naive Bayes Data Science algorithm is used for predicting whether the employee is having job satisfaction or not. Thus, this project concludes, Naive Bayes Algorithm performs better and faster when compared to other statistical techniques. This project can be further extended using some more data science algorithm and deep learning algorithms to improve the performance of prediction system. Foreseeing the employees' job satisfaction makes it easy for a company to take swift actions to improve the job satisfaction of its employees.

16

**REFERENCES**

1. Nimasha Arambepola; Lankeshwara Munasinghe,“ What makes job satisfaction in the information technology industry? ", International Research Conference on Smart Computing and Systems Engineering (SCSE), 2021.
2. Jia Yuan, “Research on Employee Turnover Prediction Based on Machine Learning Algorithms ", IEEE International Conference on Intelligent Engineering and Managemen International Conference on Artificial Intelligence and Big Data (ICAIBD), 2021.
3. Jinny Rhee; Sheri Sheppard, " Impact of Lifestyle and Working Process Organization on the Job Satisfaction Level of Software Engineers", IEEE/ACM 41st International Conference on Software Engineering: Companion Proceedings (ICSE-Companion).
4. Jieshi Chen, Aldabbagh Ghadah, “Confucian values and job performance of IT employees in Guangdong-Hong Kong-Macao Greater Bay Area ", IEEE International Conference on Education, Knowledge and Information Management (ICEKIM), 2021.
5. K. Ramamurthy et al., "Identifying employees for re-skilling using an analytics-based approach", IEEE ICDM, pp. 345-354, 2015.
6. P. Domingos and M. Pazzani, "Beyond Independence: Conditions for the Optimality of the Simple Bayesian Classifier," in Proceedings of the 13th International Conference on Machine Learning (ICML-1996), 1996.

17

1. C. D. Manning, P. Raghavan, and H. Schütze, Introduction to Information Retrieval. Cambridge University Press, 2008. [Chapter 13: Naive Bayes and Text Classification]
2. D. Jurafsky and J. H. Martin, Speech and Language Processing, 3rd ed. Pearson, 2020. [Chapter 4: Naive Bayes and Sentiment Classification]
3. T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, 2009. [Chapter 6: Naive Bayes Methods]
4. Y. Freund and R. E. Schapire, "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting," Journal of Computer and System Science.

18

**APPENDIX A**

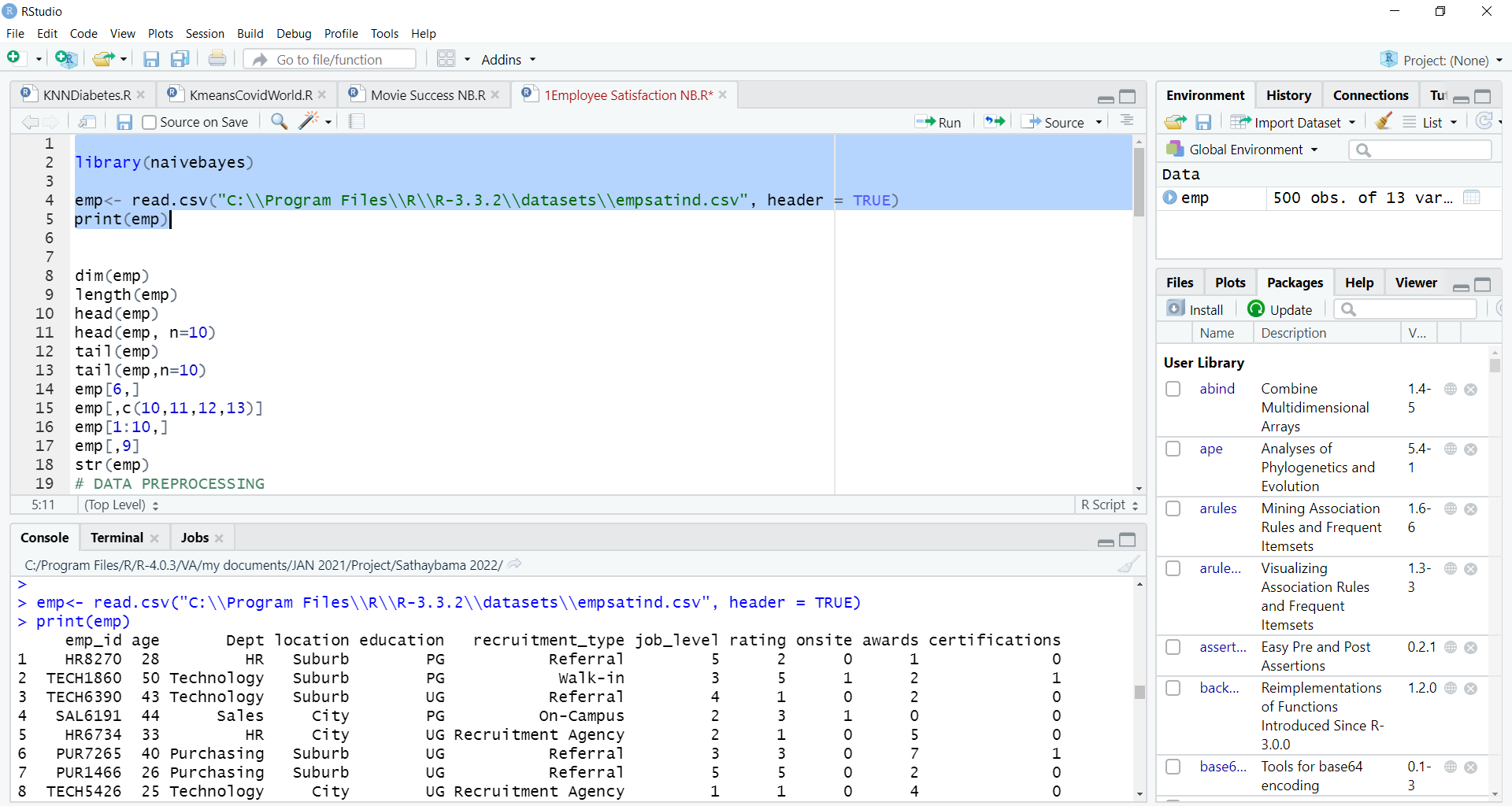
**SNAP SHOTS**

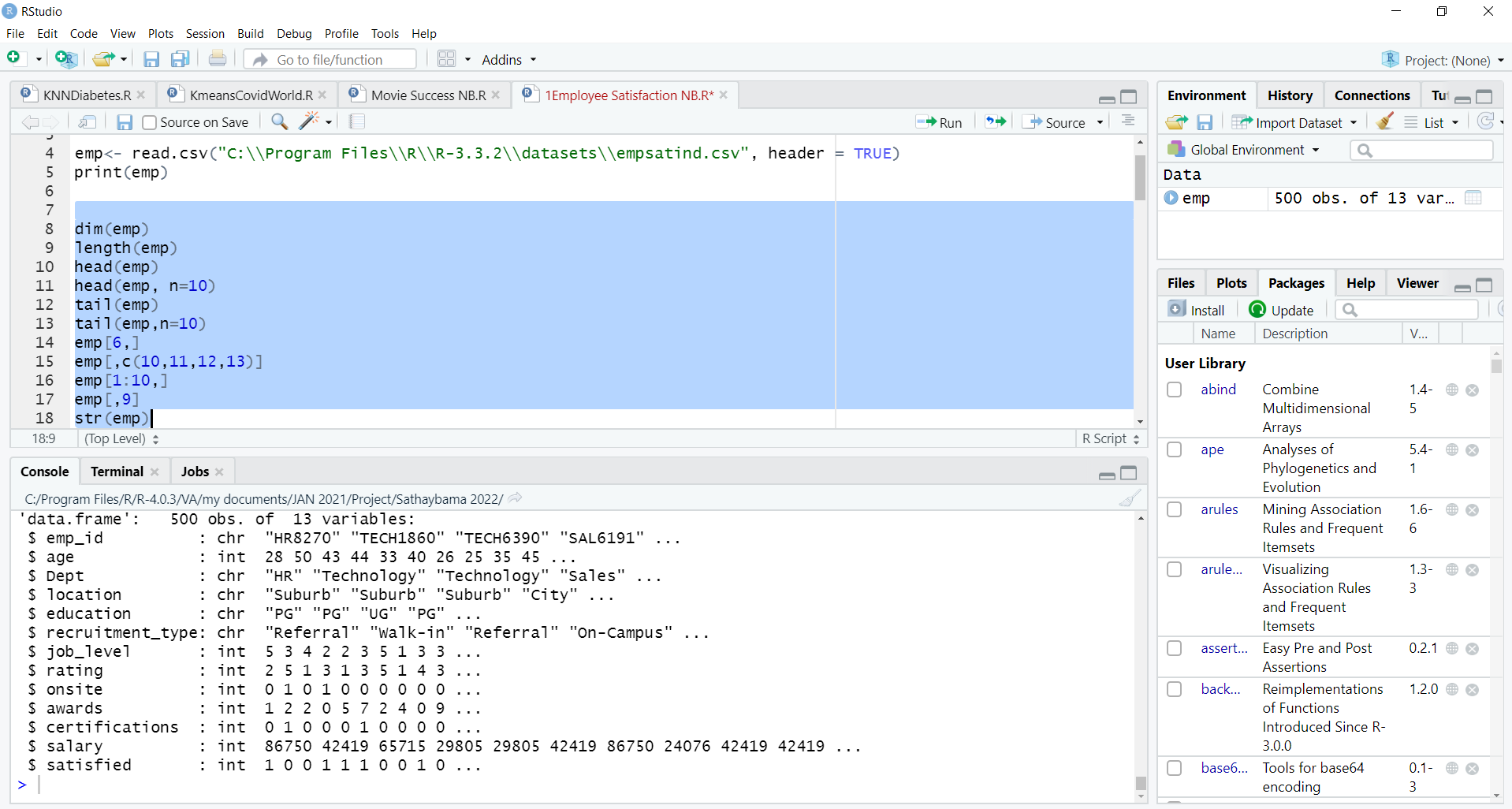
**SAMPLE DATASETS**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| emp\_id | age | Dept | location | education | recruitment\_type | job\_level | rating |
| HR8270 | 28 | HR | Suburb | PG | Referral | 5 | 2 |
| TECH1860 | 50 | Technology | Suburb | PG | Walk-in | 3 | 5 |
| TECH6390 | 43 | Technology | Suburb | UG | Referral | 4 | 1 |
| SAL6191 | 44 | Sales | City | PG | On-Campus | 2 | 3 |
| HR6734 | 33 | HR | City | UG | Recruitment Agency | 2 | 1 |
| PUR7265 | 40 | Purchasing | Suburb | UG | Referral | 3 | 3 |
| PUR1466 | 26 | Purchasing | Suburb | UG | Referral | 5 | 5 |
| TECH5426 | 25 | Technology | City | UG | Recruitment Agency | 1 | 1 |
| HR6578 | 35 | HR | City | PG | Referral | 3 | 4 |
| TECH9322 | 45 | Technology | City | PG | Referral | 3 | 3 |
| MKT2685 | 31 | Marketing | City | UG | Walk-in | 4 | 4 |
| TECH1769 | 43 | Technology | Suburb | PG | Recruitment Agency | 2 | 1 |
| TECH7949 | 28 | Technology | City | UG | On-Campus | 3 | 4 |
| PUR3433 | 48 | Purchasing | Suburb | PG | Referral | 3 | 4 |
| PUR6311 | 52 | Purchasing | City | PG | Recruitment Agency | 5 | 1 |
| PUR6051 | 50 | Purchasing | City | UG | Walk-in | 5 | 5 |
| MKT7420 | 34 | Marketing | City | PG | On-Campus | 1 | 4 |
| PUR2184 | 24 | Purchasing | Suburb | UG | Recruitment Agency | 4 | 4 |
| HR5555 | 54 | HR | Suburb | PG | On-Campus | 1 | 5 |
| SAL4385 | 25 | Sales | City | UG | On-Campus | 5 | 4 |
| HR7396 | 25 | HR | Suburb | UG | On-Campus | 2 | 4 |
| HR9666 | 50 | HR | Suburb | PG | Referral | 5 | 4 |
| PUR3558 | 34 | Purchasing | City | PG | Referral | 4 | 2 |
| SAL8849 | 37 | Sales | City | UG | Walk-in | 5 | 5 |
| SAL3047 | 31 | Sales | Suburb | PG | Walk-in | 4 | 4 |

19

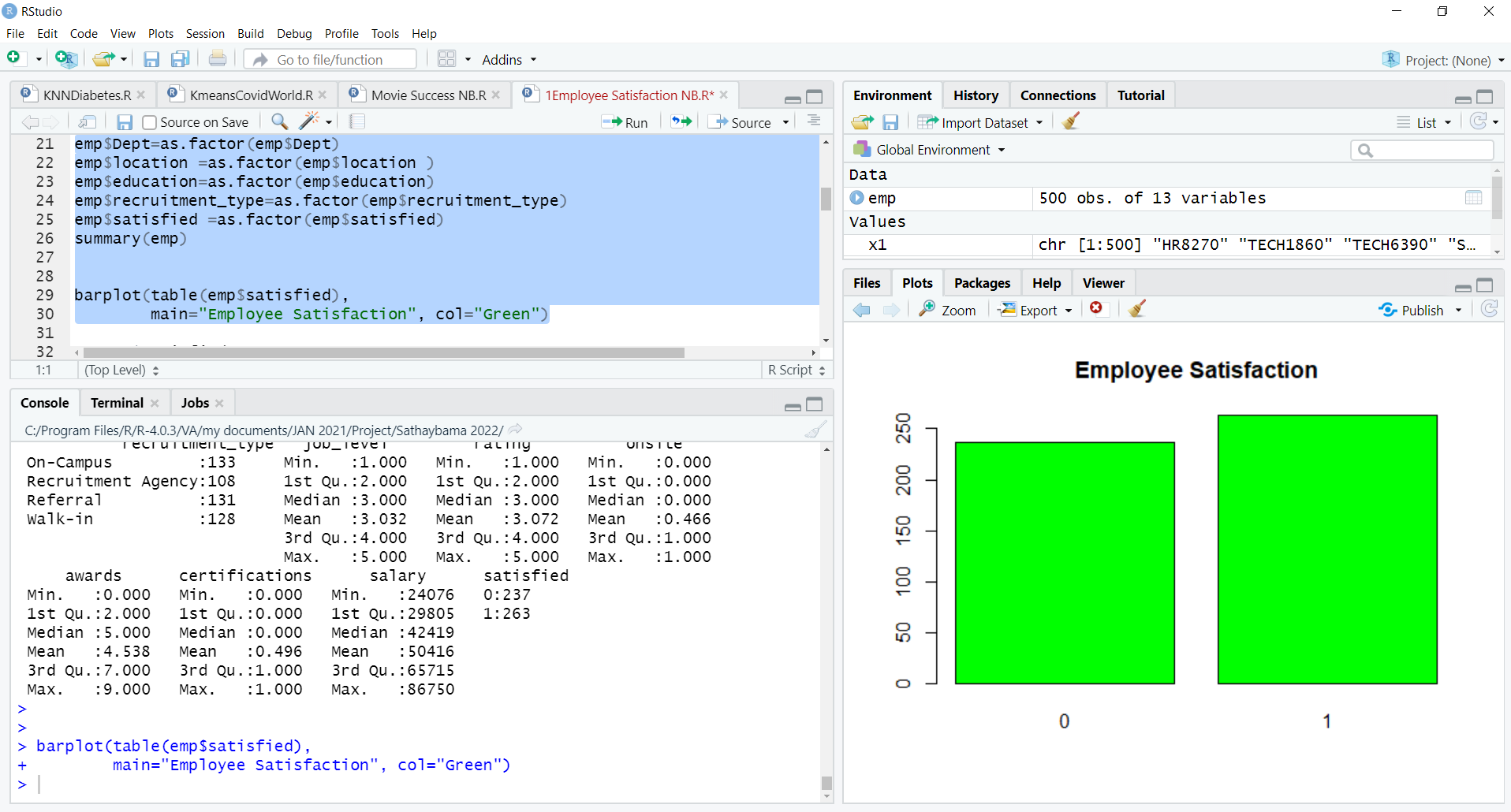
**DATA IMPORTING**

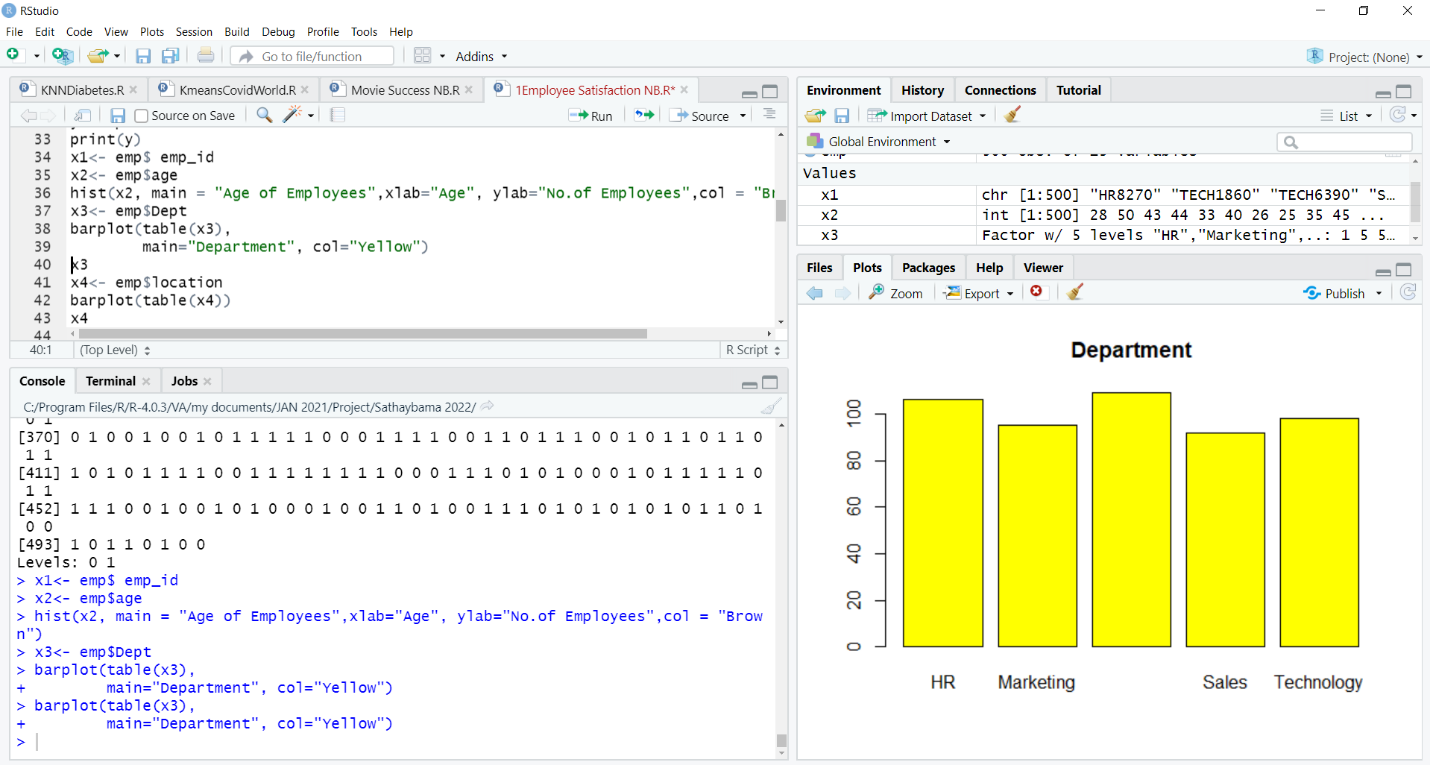
****

****

20

**DATA PREPROCESSING**

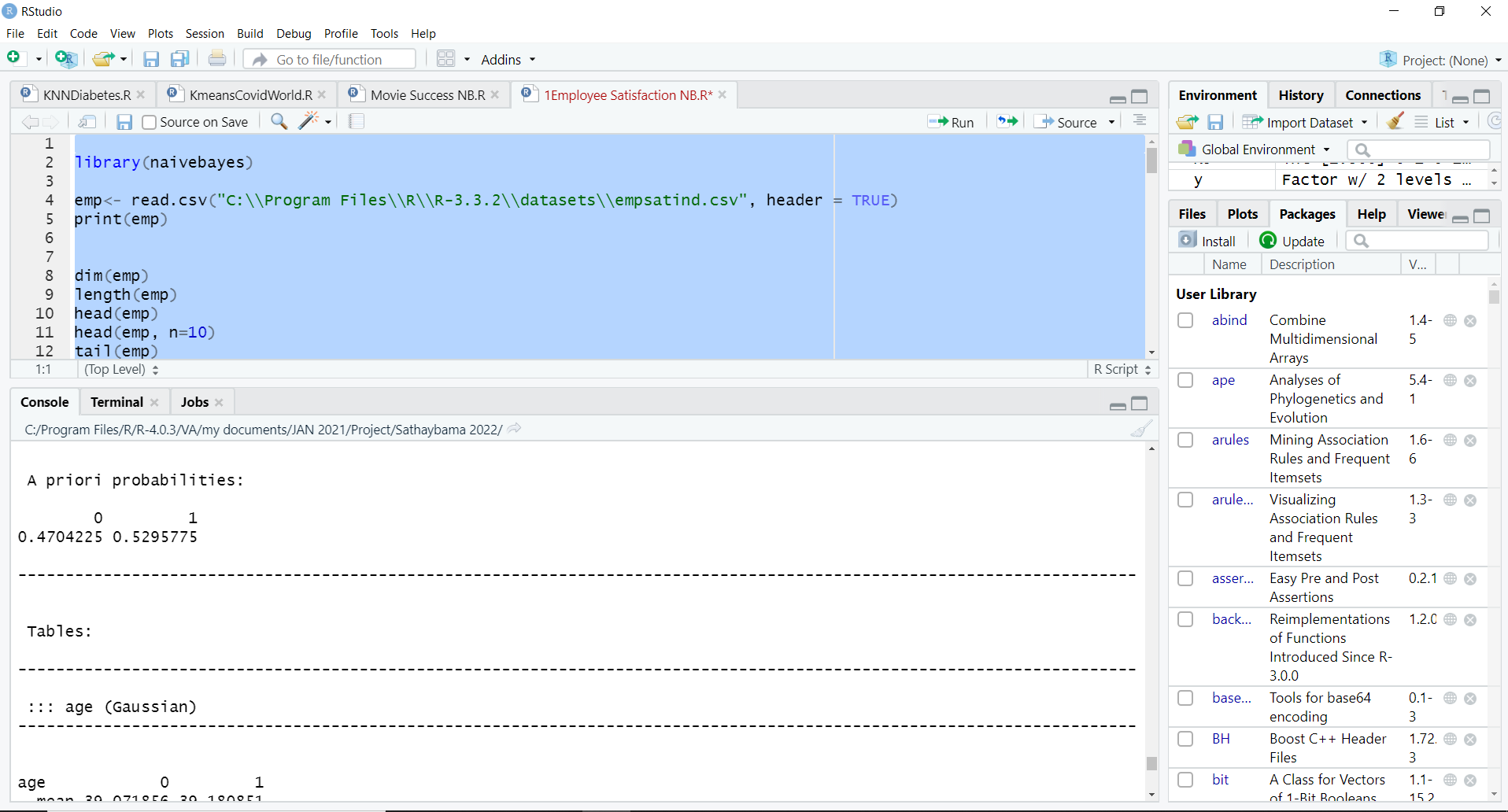
****

****

21

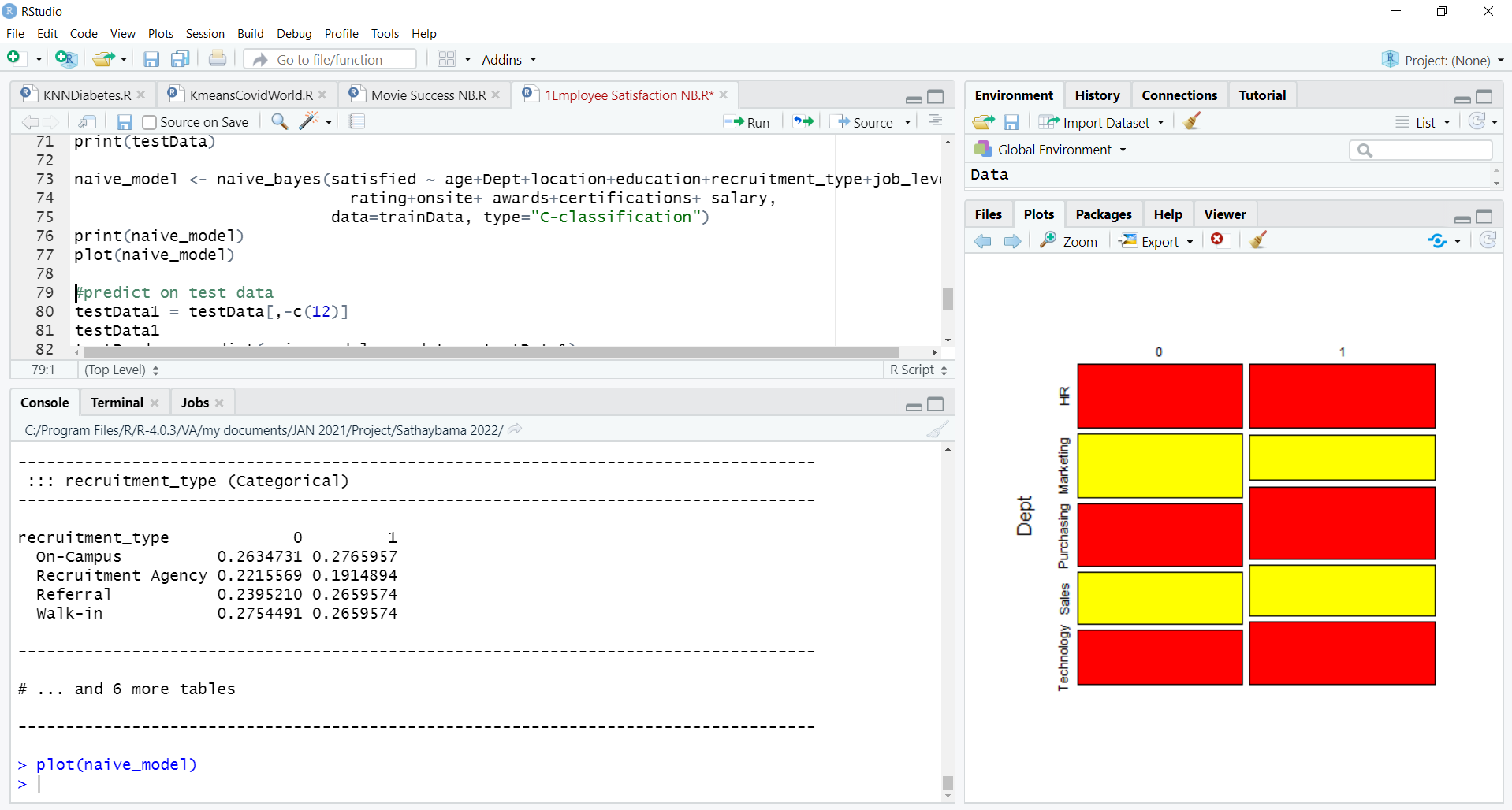
**NAIVE BAYES ALGORITHM**

**Naive Bayes Classification Model**

****

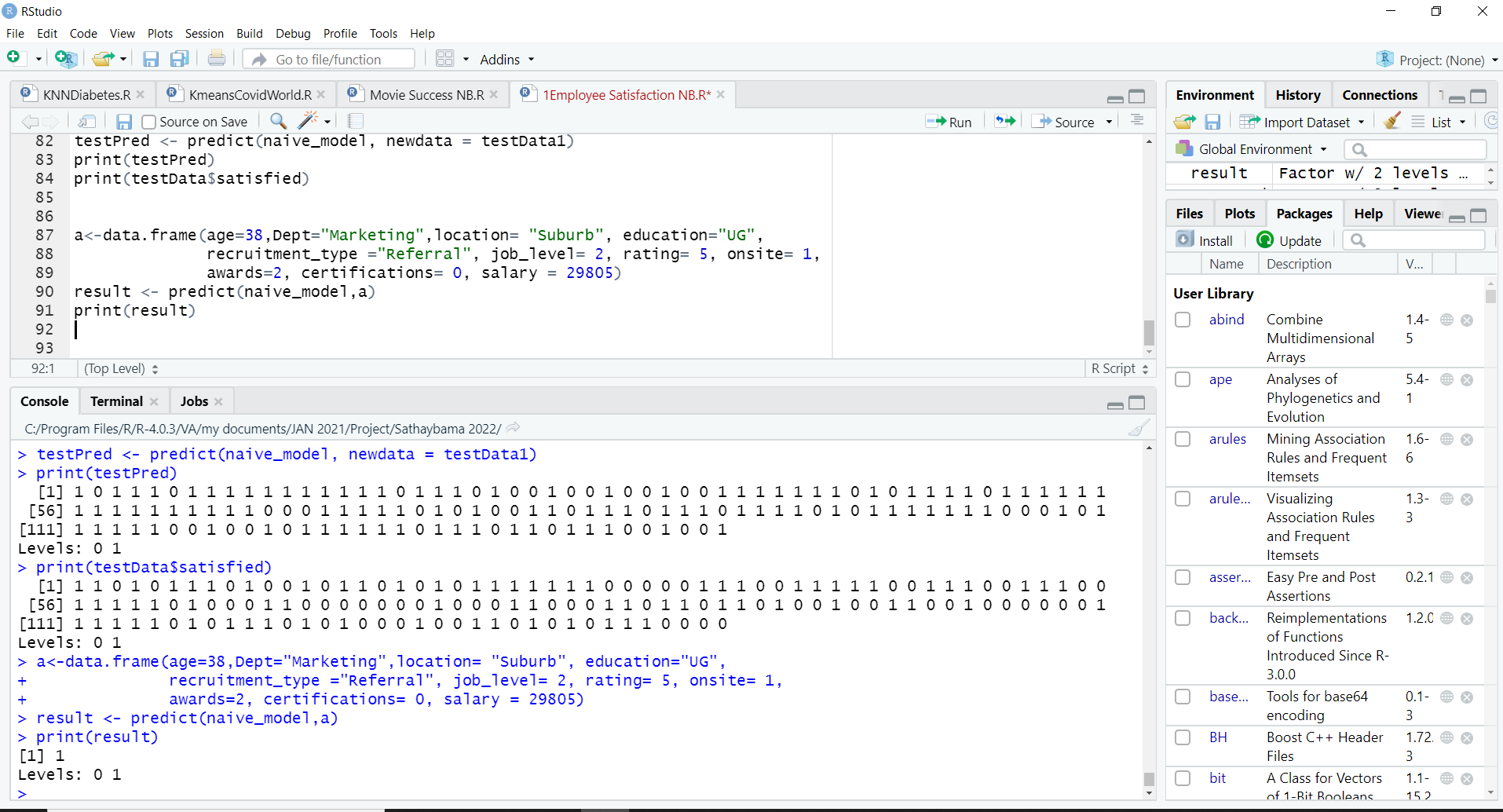
**NAIVE BAYES ALGORITHM**

**NB Model Inference**

****

22

**Naive Bayes Prediction Model**

****

23

**APPENDIX B**

**SAMPLE CODING**

library(naivebayes)

emp<- read.csv("C:\\Program Files\\R\\R-3.3.2\\datasets\\empsatind.csv", header = TRUE)

print(emp)

dim(emp)

length(emp)

head(emp)

head(emp, n=10)

tail(emp)

tail(emp,n=10)

emp[6,]

emp[,c(10,11,12,13)]

emp[1:10,]

emp[,9]

str(emp)

# DATA PREPROCESSING

summary(emp)

emp$Dept=as.factor(emp$Dept)

emp$location =as.factor(emp$location )

emp$education=as.factor(emp$education)

emp$recruitment\_type=as.factor(emp$recruitment\_type)

emp$satisfied =as.factor(emp$satisfied)

summary(emp)

barplot(table(emp$satisfied),

main=”Employee Satisfaction”, col=”Green”)

24

y<-emp$satisfied

print(y)

x1<- emp$ emp\_id

x2<- emp$age

hist(x2, main = “Age of Employees”,xlab=”Age”, ylab=”No.of Employees”,col = “Brown”)

x3<- emp$Dept

barplot(table(x3),

main=”Department”, col=”Yellow”)

x3

x4<- emp$location

barplot(table(x4))

x4

x5<- emp$education

barplot(table(x5))

x6<- emp$recruitment\_type

barplot(table(x6))

x7<- emp$job\_level

x8<- emp$rating

x9<- emp$onsite

x10<- emp$awards

x12<- emp$certifications

x13<- emp$salary

emp<- emp[,c(2,3,4,5,6,7,8,9,10,11,12,13)]

print(emp)

# Sampling of Datasets

set.seed(1)

ind <- sample(2, nrow(emp), replace=TRUE, prob=c(0.70, 0.30))

trainData <- emp[ind==1,]

testData <- emp[ind==2,]

25

print(trainData)

print(testData)

naïve\_model <- naïve\_bayes(satisfied ~ age+Dept+location+education+recruitment\_type+job\_level+

rating+onsite+ awards+certifications+ salary,

data=trainData, type=”C-classification”)

print(naïve\_model)

plot(naïve\_model)

#predict on test data

testData1 = testData[,-c(12)]

testData1

testPred <- predict(naïve\_model, newdata = testData1)

print(testPred)

print(testData$satisfied)

a<-data.frame(age=38,Dept="Marketing",location= "Suburb", education="UG", recruitment\_type ="Referral", job\_level= 2, rating= 5, onsite= 1, awards=2, certifications= 0, salary = 29805)

result <- predict(naive\_model,a)

print(result)

26