**Group 16 - Mental Health Illness In Tech**

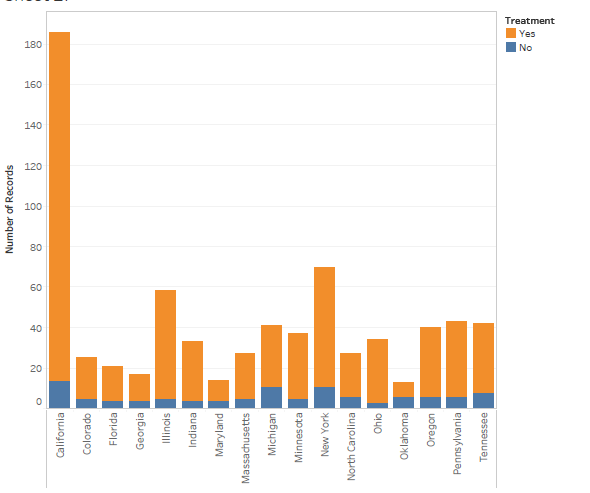
**Exploratory Analytics and Classification**

**D.Chakravardhan Reddy**  **IMT2014015**

**Target users:** Doctors, state administrators

**Report - 1**

* Among the people who have mental illness, how many people have sought treatment?
* The following graph indicates the number of people who have sought treatment among the people with mental illness.

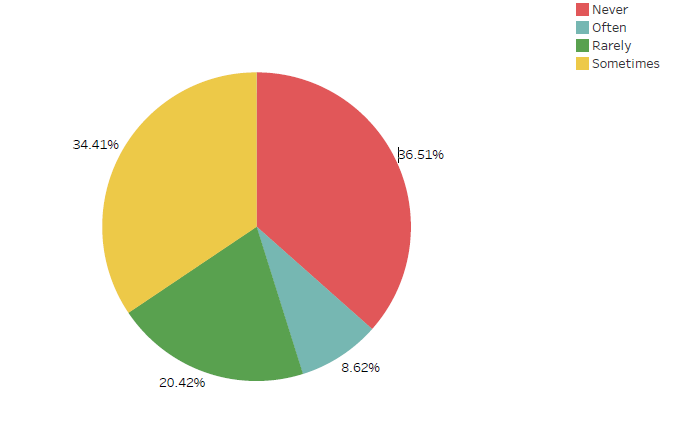


* We can see that there are few people who have mental illness and who have not sought for treatment. But the number is few compared to the people who have sought treatment. So, we can say that there is awareness about the mental illness.

**Target user:** Higher officials in the company's (HR, CXO’s)

**Report -** 1

* If you have mental illness, how does it interfere your work?
* The following graph indicates how mental illness vary across work interference.



* We can see that nearly 50% of the people have sometimes or often encounter interference with their work due to a mental health condition, whereas the other 50% did not answer, or indicated they never or rarely encounter such interference.

**Classification:**

**Report - 1**

**Target user:** All the employees and higher officials.

**Problem formulation:** How likely it is for an employee with certain features to have mental illness?

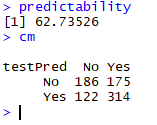
**Data preparation:**

* Removed NA values in our target attribute
* Selected some important attributes gained from data preparation stage
* Split training and test data in the ratio 7:3

**R script:**

**Fig 1.**

**Evaluation:**



**Accuracy : 62%**



F1 score =

* The higher officials will get to know whether an employee with certain features is affected by mental illness or not.
* The employees will get to know whether they have mental illness or not.

**Report-2**

**Target user:** All the employees and higher officials

**Problem formulation:** How likely it is for an employee with certain features to have mental illness?

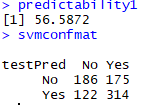
**Data preparation:**

* Removed NA values in our target attribute
* Selected some important attributes gained from data preparation stage
* Split training and test data in the ratio 7:3

**R script: Fig 1.**

**Evaluation:**

**Accuracy: 56%**





**F1 score =**

* The higher officials will get to know whether an employee with certain features is affected by mental illness or not.
* The employees will get to know whether they have mental illness or not.

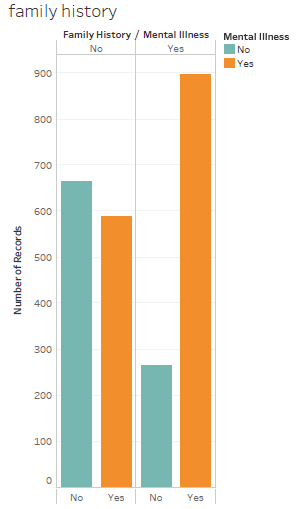
From the above two classifiers, we can see that decision tree has the highest accuracy.

**Abhijnu Yadlapalli IMT2014065**

**Exploratory Analysis:**

**Target User :** All the employees

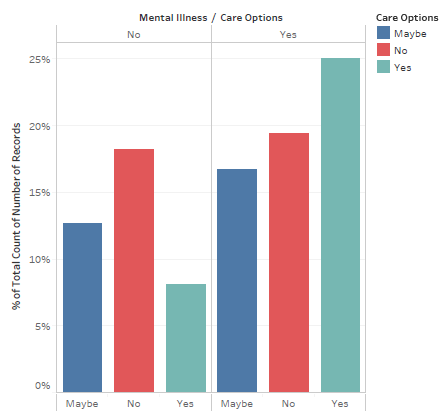
* How does the family history affect the mental illness of the people?
* The following graph shows how mental illness varies with the family history.



* We can observer from the above graph that the chance of mental illness is more when there is a family history of mental illness. And also people without family history have less chance of getting mental illness. So this helps all the employees to be careful and get checked up for mental illness if there is a family history of mental illness.

**Target User:** Higher officials in the company's (HR, CXO’s)

* How many people know that care options have been provided by their company for mental illness?
* The following graph shows the comparison of people with and without mental illness whether they know the care options provided by their company or not.



* From the above graph we can observe that a large proportion of people with mental illness don’t know or they are not sure that their company is providing care options. And even in the case of the people without mental illness the condition is same. So here the higher officials of the company should try to increase the awareness of the care options provided by their company to help the employees.

**Classification:**

***Recursive Partitioning:***

**Target user:** All the employees and higher officials.

**Problem formulation:**

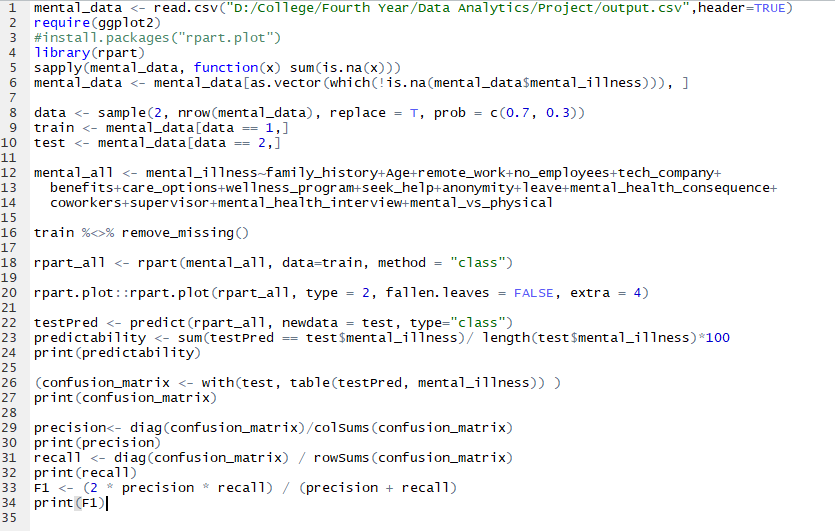
How likely it is for an employee with certain features to have mental illness?

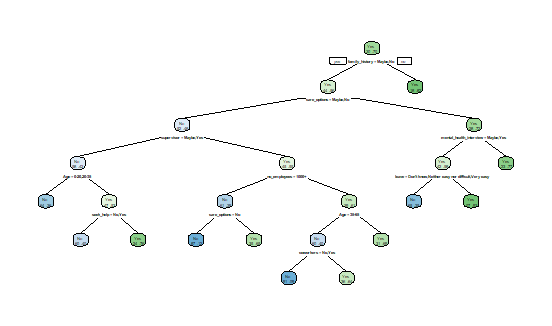
**Data Preparation Steps:**

In our target variable there are 8 NA values. First we need to remove those NA rows. Then the data is divided into training(70%) and testing data(30%).

We selected some important attributes gained from the preparation stage for our model. Then we need to remove all the missing values from the trained data.

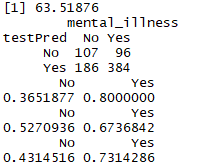
**R Code:**

****

****

**Evaluation:**

The predictability of the model is 63.518. Given below is the confusion matrix. The F1 score is given in the last line of the image.

****

The higher officials will get to know whether an employee with certain features is affected by mental illness or not.

The employees will get to know whether they have mental illness or not.

***Random Forest:***

**Target user:** All the employees and higher officials.

**Problem formulation:**

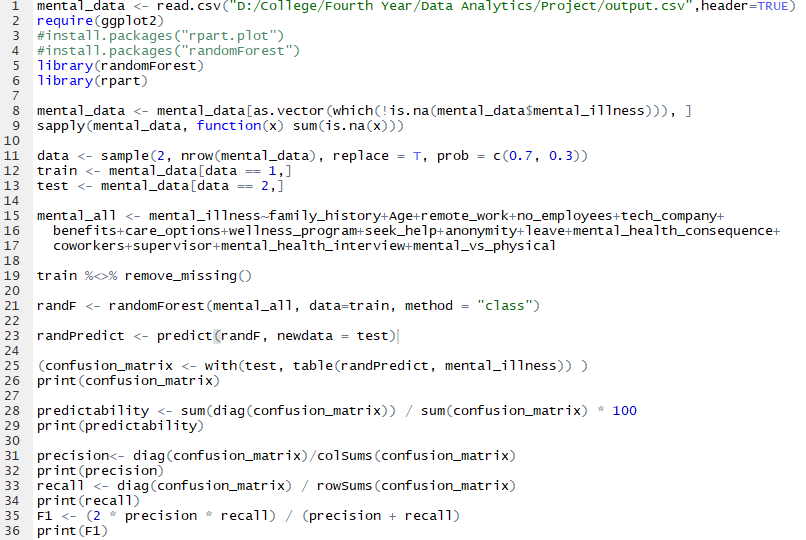
How likely it is for an employee with certain features to have mental illness?

**Data Preparation Steps:**

In our target variable there are 8 NA values. First we need to remove those NA rows. Then the data is divided into training(70%) and testing data(30%).

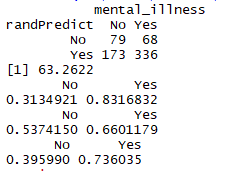
We selected some important attributes gained from the preparation stage for our model. Then we need to remove all the missing values from the trained data.

**R Code:**



**Evaluation:**

The predictability of the model is 63.262. Given below is the confusion matrix. The F1 score is given in the last line of the image.



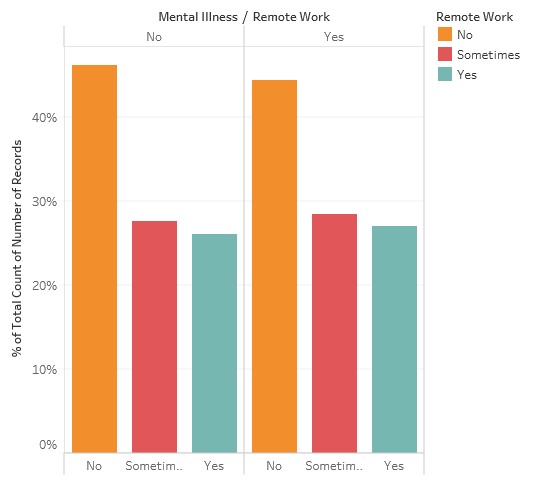
The higher officials will get to know whether an employee with certain features is affected by mental illness or not.

The employees will get to know whether they have mental illness or not.

**Sandeep Tadepalli IMT2014057**

**Target users:**all employees who work in tech or non-tech companies

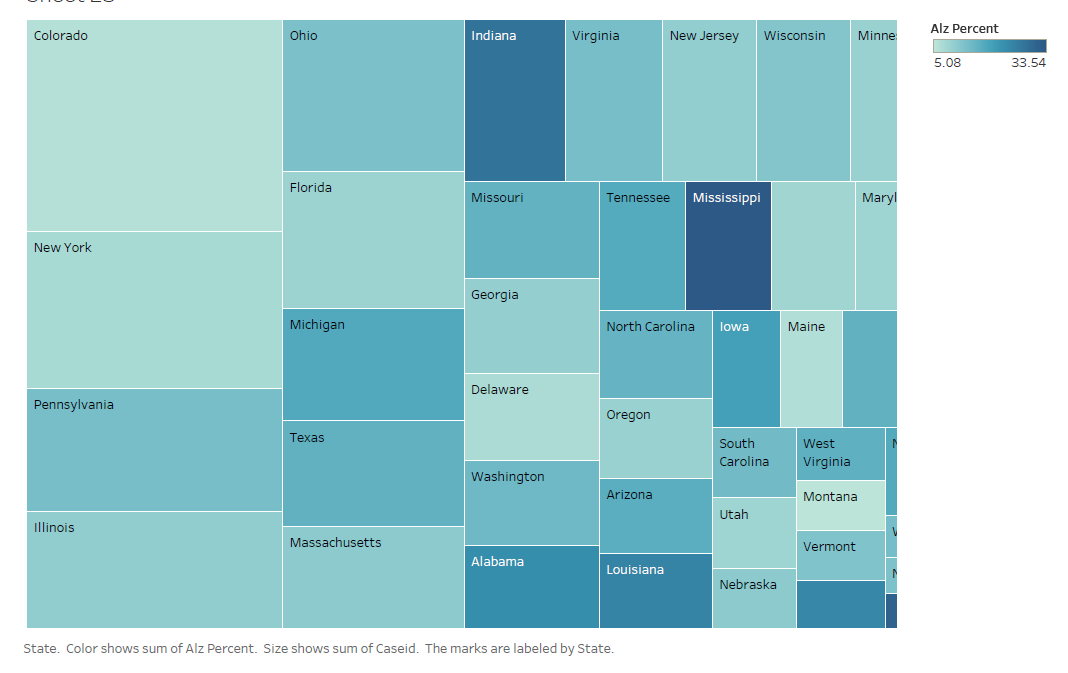
* Among the people who have mental illness, how many people work remotely?
* The following graph indicates remote\_work vs mental\_illness

****

* We can see that most of the people who have mental illness doesn’t or sometimes work remotely than the people who don’t have mental illness

**Target User:** State administration

* How much percentage of health care centers in a state provide treatment for Alzheimer's?
* The following treemap shows the ease of availability of treatment for Alzheimer’s.



* From the above treemap we can see that in most of the states there are not many treatment centers. So the state administration need to increase the treatment facilities in their states.

**Classification:**

Target User: Higher Officials.

Problem Formulation: How likely it is for an employee to discuss his/her mental condition with his supervisors

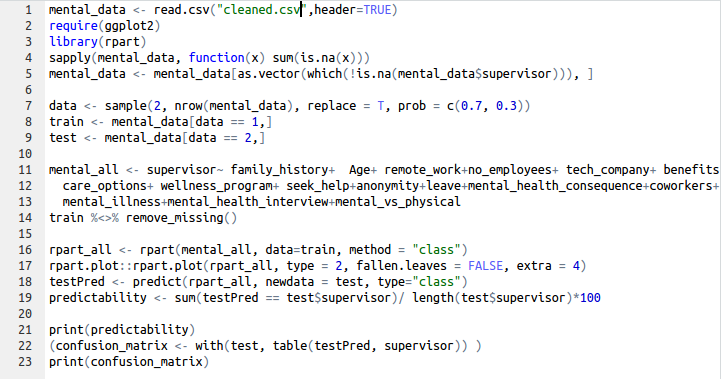
Data Preparation:

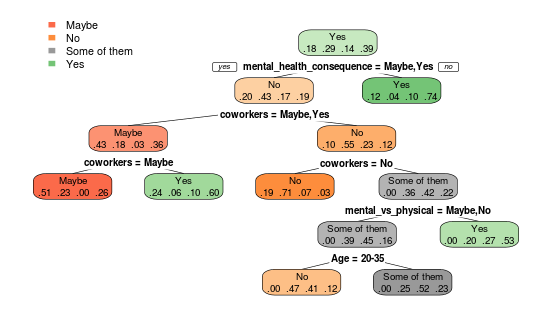
In our target variable there are some NA values. First we need to remove those NA rows. Data is divided into Testing (70%) and training (30%)

For the feature set we chose some attributes which were chosen as important attributes at the data preparation stage via covariance matrix and some other important attributes then the missing values are removed in all those attributes.

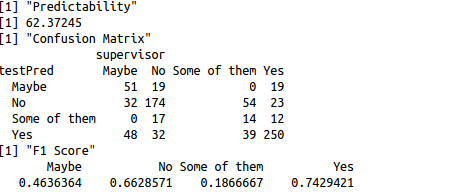
Classification Models:

**RPart:**

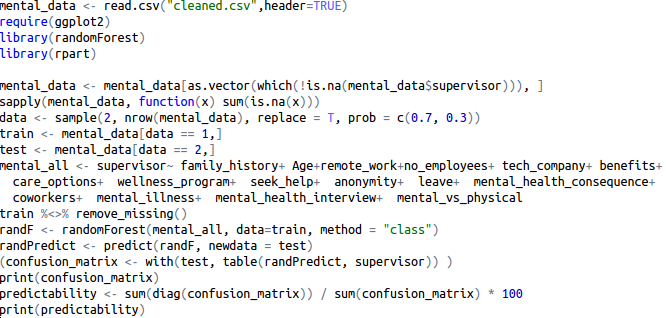




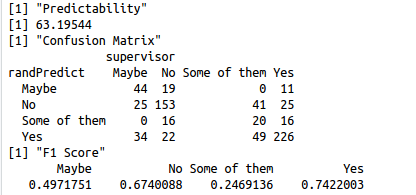
**Evaluation:**



**Random Forest:**

****

**Evaluation:**

****

So both random forest and Decision tree almost game same precision so there is no clear winner in this.

Based on this the higher officials can guess how much likely he is going to discuss his mental illness with them So they can accordingly take steps for the well being of the employee.

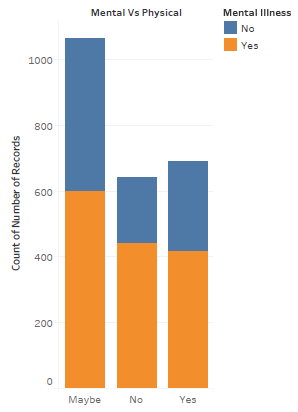
**V. Keerthi Chandra**

**IMT2014064**

**Exploratory Analysis**

**Target Users**: Employer as well as the employee of a company

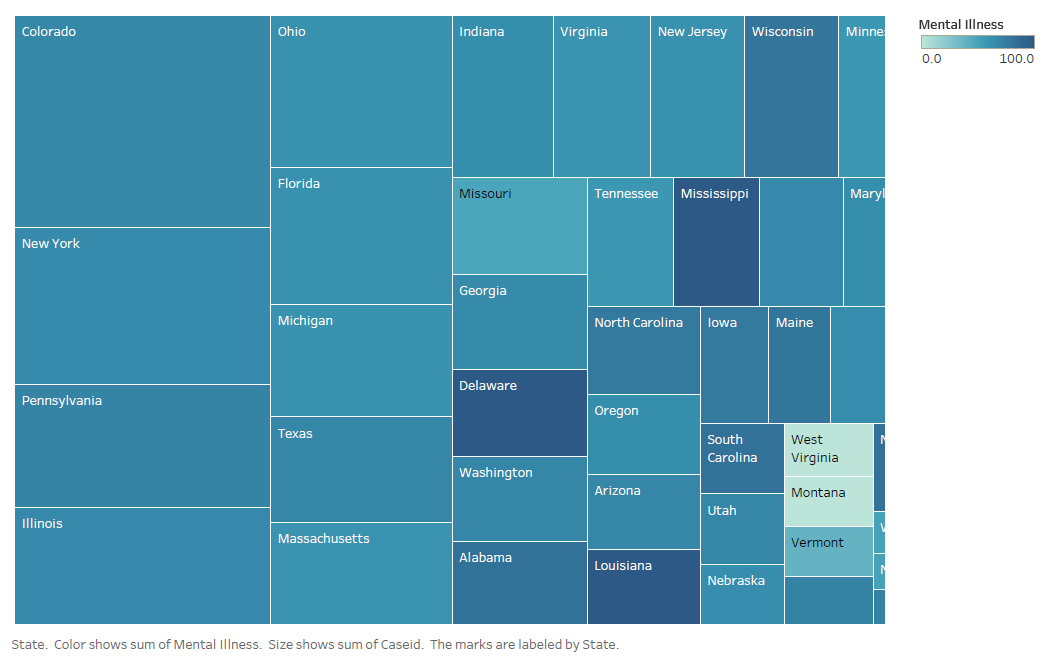
* The plot below shows you the number of people who are actually aware that their company takes both mental health as seriously as physical health.



* From the above graph, we can observe that a large proportion of people with mental illness don’t know or they are not sure that their company gives equal importance of mental and physical health. And even in the case of the people without mental illness the condition is same. So here the higher officials of the company should try to increase the awareness of the importance to mental health provided by their company to help the employees.

**Target User:** State Administration

* How much awareness is provided by the states about mental illness to their people?
* The following treemap shows the awareness percentage in each state.



* From the above treemap, we can observe that many states have good awareness programmes or something similar. But some of the states lack in their awareness programmes. So for these states, the state administration need to improve the awareness programmes.

**Classification**

**Target User:** Employer as well as employee of the company

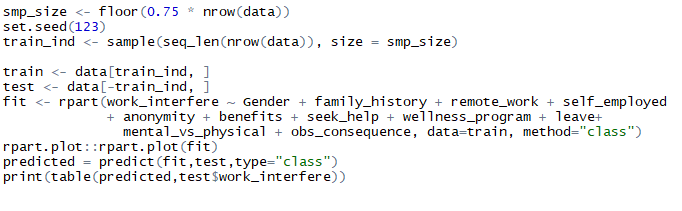
**Problem Formulation**: Is the work interference a major factor for the mental health of an employee?

**Data Preparation**:

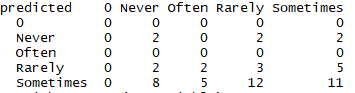
* The dataset is divided into two parts, training and testing data with a predefined ratio but at the same time preserving the ratios of different labels in the data.
* All the Na values are removed from the target variable

**Model:** Decision Trees

**Rcode**:

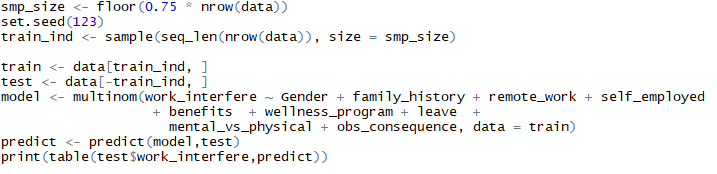


**Evaluation Metric**: Confusion Matrix

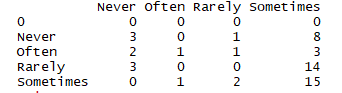


**Model:** Logistic Regression

**Rcode**:



**Evaluation** : Confusion Matrix



**Use Case**: So now we know the effect of the work interference in contributing to the mental health of an employee. So higher officials can help their employees by providing flexible work hours or some sort of medical incentives for their employees.