HR CASE STUDY

Group Members:

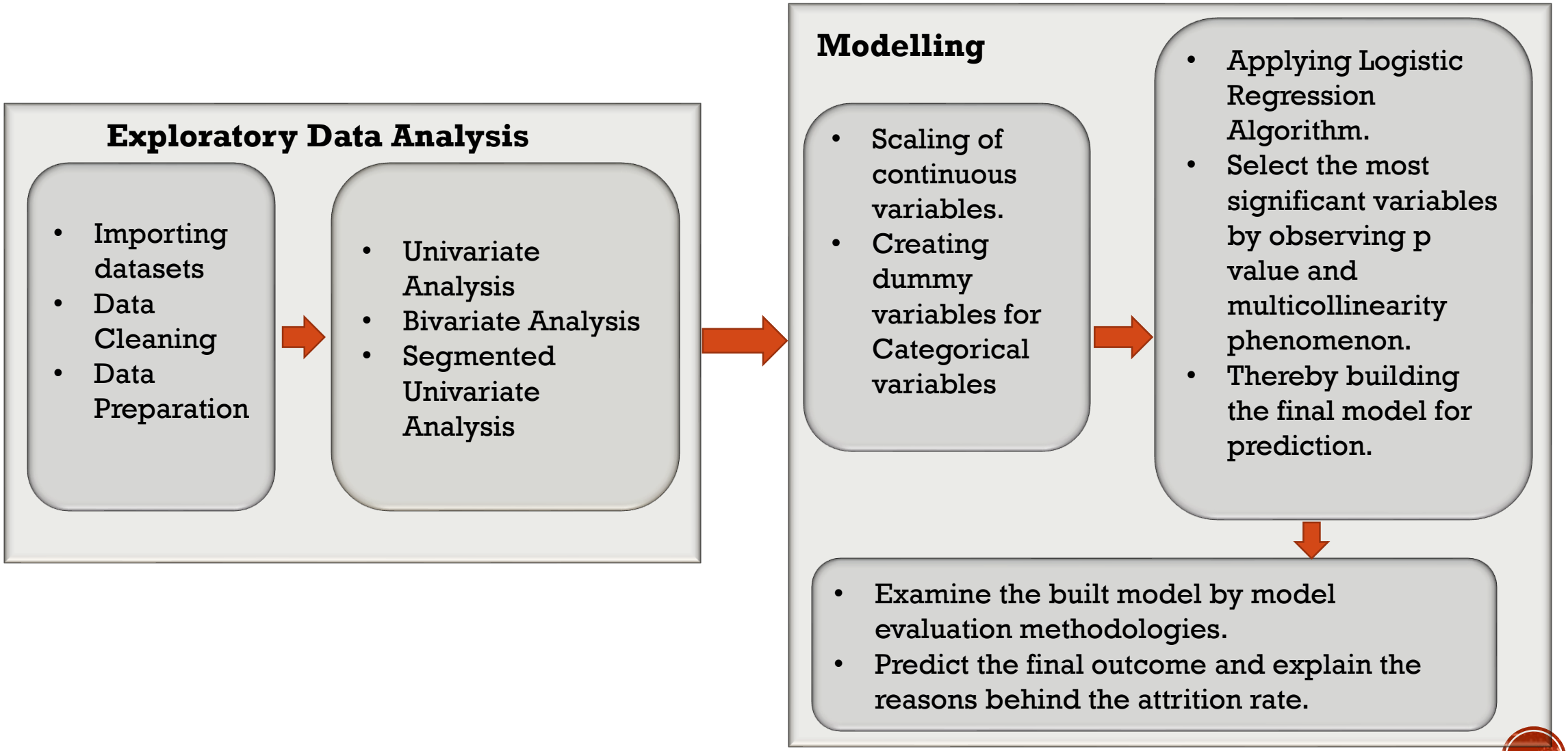
1. Varsha Venkapally
2. Ashish Kumar Korukonda
3. Shravani Kothur
4. Chayan Naskar

AGENDA

A large company named **XYZ**, employs, at any given point of time, around 4000 employees. However, every year, around 15% of its employees leave the company and need to be replaced with the talent pool available in the job market. The management believes that this level of **attrition** is bad for the company.

As an data analyst , we are supposed to find out what changes they should make to their workplace, in order to get most of their employees to stay. Also, to know the reasons behind the **attrition** and that needs to be addressed right away.

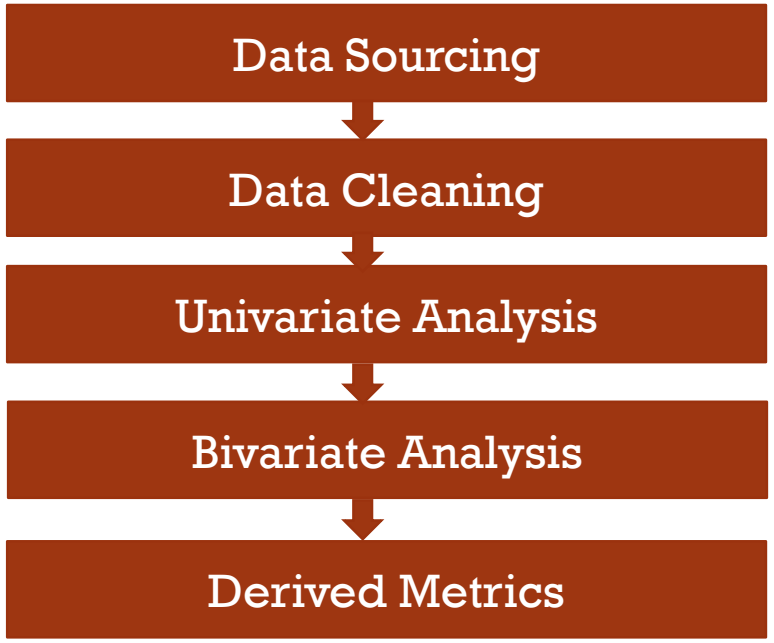
PROBLEM SOLVING METHODOLOGY



EXPLORATORY DATA ANALYSIS (EDA)

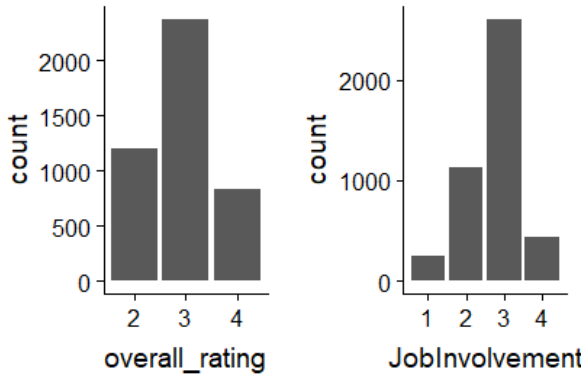
We would be looking into the dataset provided and use EDA approach to get insights from the data and accordingly find out the reasons behind the attrition rate and the significant variables behind it.

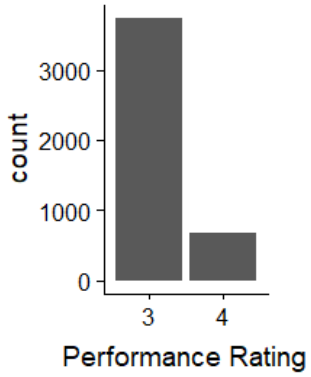
Below are the steps we can follow to get insights from the data:

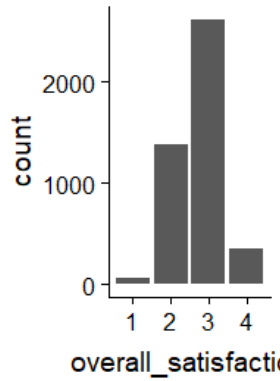


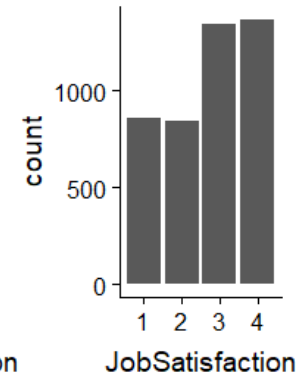
DATA CLEANING AND DATA PREPARATION

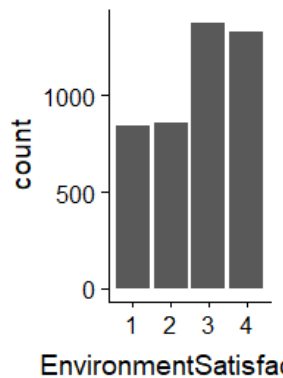
- As data is the key : **the better the data, the more insights you can get out of it.**
- After we import the datasets below are the steps followed to have clean data:
  - Handling missing values
  - Removing duplicated data
  - Standardising the values
  - Filtering the unwanted data and fixing the invalid values
  - Treating outliers.
- After the data cleaning is done, we have prepared the data for analysis and below are some of the steps:
  - Deriving new variables
  - Merge all the datasets to a single data frame
  - Check for any anomalies in the merged data.

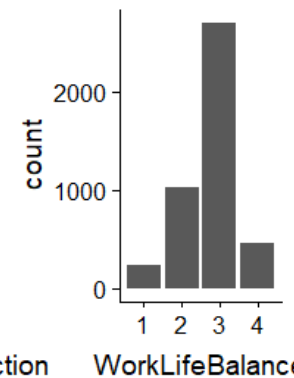


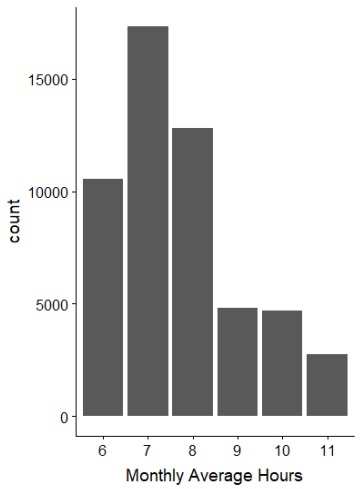


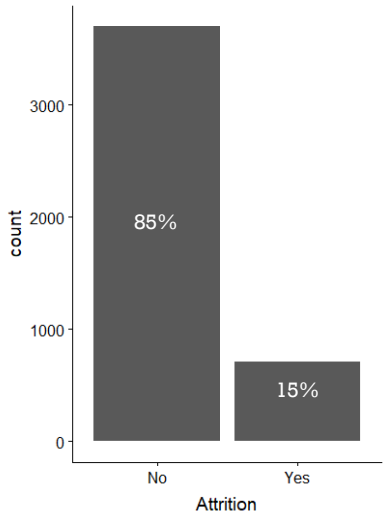












INSIGHTS USING DATA ANALYSIS METHODOLOGIES

- Average monthly hours is around 7 hours.
- Based on the employee survey we can see that overall satisfaction on an average of an employee is "High". We can thus say that on an average the employees are satisfied with work aspects such as job satisfaction, environment satisfaction and work-life balance.
- Based on the manager survey we can see that overall rating on an average is given "Excellent" to an employee. We can see that on a average every employee performs well at work.
- Attrition is 15% this year in the company.

MODELLING-I

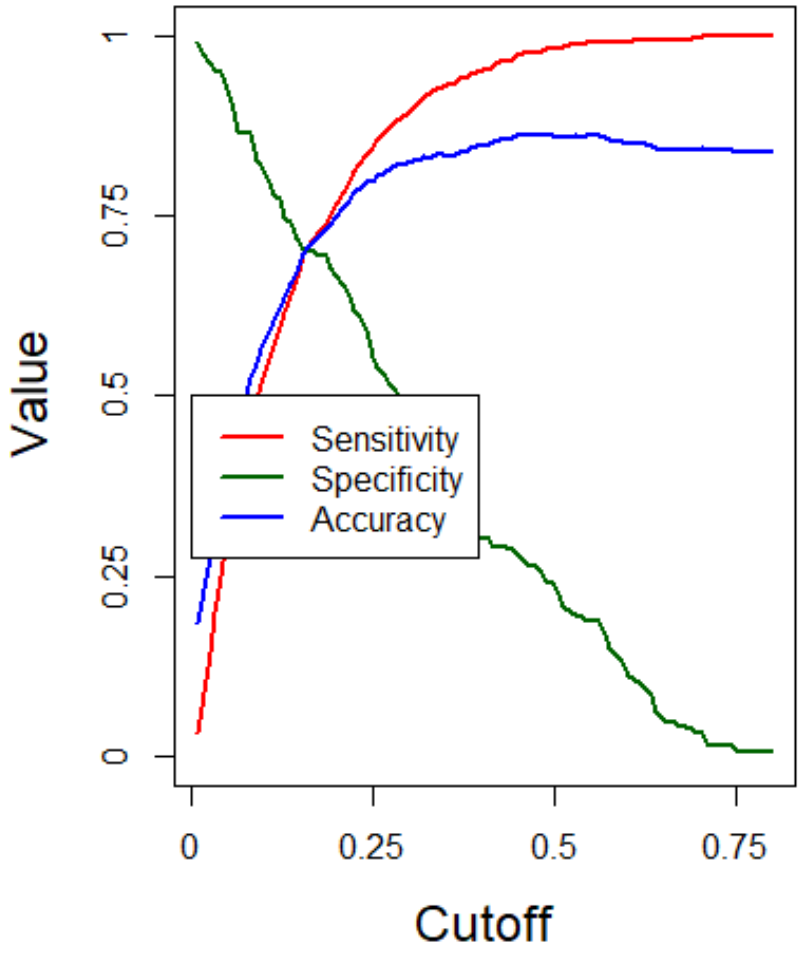
- To predict the attrition rate of the employees we have used Logistic Regression Algorithm.
- In order to perform Logistic Regression algorithm we have to make the final data obtained suitable for modelling.
- Below are the steps we perform to ensure the data is ready for modelling:
  - Perform scaling on continuous numeric variables. Scaling is performed to standardise the units/scale of the numeric variables to a universal scale. Therefore increasing the predictive power of the model by removing the dominating aspect of one single variable over other variables because of its comparatively higher scale. Some of the variables are : PercentageSalaryHike, MonthlyIncome etc.
  - Creating dummy variables on categorical variables. As regression analysis treats only numeric variables or values, therefore its must for us to convert the categorical variables to numeric variables. This is done by converting the categorical variables into 0's and 1's.
    - Two factor categorical variables such as Yes and No's are directly assigned 0's and 1's.
    - Variables having more than 2 factors are treated by creating separate columns for each category and assigned 0's and 1's.

MODELLING-II

- After the dataset is ready for modelling we split the data into Test and Train datasets. We use Train dataset to build the model and Test dataset to examine the model by predicting the outcome on different dataset.
- We create the first model using the below code in R.

```
model_1 = glm(Attrition ~ ., data = train, family = "binomial")
```

- We use stepAIC function to remove the insignificant variables from the model.
- With the list of significant variables, we then perform more filtering my removing the variables with high multicollinearity and high p-values.
  - If VIF>2 then it refers that the variable is highly collinear. Thereby removing the variables to reduce the collinearity.
  - If the p-values>0.05 then we consider that variable as less significant as our null hypothesis is p<0.05.



MODEL EVALUATION & PREDICTION

Below variables which describe the discriminative power of a model and thereby evaluating a model's performance:

1. Accuracy – 71%
2. Specificity – 71%
3. Sensitivity –74%
4. KS Statistic – 41%

- As we can see in the graph the point where Accuracy, Specificity and Sensitivity meet – that is the optimal cut-off value to predict the probabilities of attrition.
- Model can be considered to be performing good if the KS Statistic value is equal or grater than 40%.

DRIVER VARIABLES TO INDICATE THE ATTRITION RATE:

No	Driver Variable	Reasons for Attrition
1.	Number of Companies worked	If an employee has less experience and has changed his company quite frequently then chances of changing the company after joining can be very high.
2.	Environment Satisfaction	If the employee is not satisfied with the environment such as workplace incentives, Supervisor support, performance feedback etc is likely to leave the company.
3.	Job Satisfaction	Factors during his work such as Purpose of work, A balanced lifestyle, self confidence during his work etc are related to job satisfaction which highly relate to attrition rate.
4.	Years with current Manager	It is very much known that people leave or stay because of their bosses and not because of the company. A good manager-employee relationship is crucial for the satisfaction and retention of the worker.
5.	Total Work Hours	Employees who have work exp ranging from 0-4 tend to change the company for various reasons such as greater hike and role changes. Whereas employees who are more then 4 years work exp tend to settle their career in a good and friendly work place and will have a lower probability to change.
6.	Average work hours	This attribute can help us if an employee is over stressed or who works very less, thereby predicting the chances of attrition of an employee.
7.	Years since last promotion	Based on his/her performance an employee can be eligible for a promotion for his current role. If proper implementation and guidance is given then the employee is more likely to be satisfied and less likely to leave the company.

RECOMMENDATIONS

- Frequency of radical changes of companies of an employee should be considered to be a red flag if it is very high – chances of attrition is relatively high.
- Even if benefits are not the primary reason for the employee to stick with the company, the benefits offered by the company cannot be marked worse than the one offered by your competitors. Thereby improving environment satisfaction.
- Fair compensation single-handedly will not guarantee the loyalty of the employee, but offering wages below market makes it more likely that the employee would look for a job at some other place. Thereby improving job satisfaction.
- To promote the loyalty of the employee implement a career ladder and ensure that employees are aware of what they need to do to earn promotion. To identify the strong point and weakness of the workforces conduct performance reviews regularly. Assist them in improving the areas that will lead them to seek advancement in their job. A proper professional development plan provides the employee a reason to stick around.
- Ensure that the managers in the organization are not driving away the technologists. Provide them with the necessary training that is required to develop people management skills and good supervisory.
- It is good to recognize outstanding achievements in public and on a regular basis but also take the time to comment on the small contributions that are made by the staff on a daily basis to meet the mission of the organization. Remember the staff makes the company look good. Give the employees best equipment and supplies. Upgrade the systems, machinery, and software so they could work efficiently. Taking care of the needs of the employee is an indication that the company values them.
- It is time for changes if the technologists spend much time in the filing of paperwork. The pressure of Paperwork adds to the stress and burnout the employees. Eliminate the pointless paperwork, convert the paperwork into electronic format. Hire non-tech administrative staff to take over as much of the burden related to the paperwork as per the regulatory or legal restrictions. Thereby reducing the time they spend in office and proportionally reduces the burden of an employee.