Hiring Process Analytics

Trainity Project No. 4
for DA Training

Prepared by:
Nilesh Kulkarni

Agenda

- Project Description
- Project Approach
- Tech Stack used
- Dataset Overview and Data Cleaning
- Insights
- Results

Prepared by:
Nilesh Kulkarni

Project Description

The project is based on provided data about the hiring process. The aim of project is to analyze the data using know-how of statistics as well as MS Excel and provide insights about the hiring process. The hiring process is a crucial function of any company, and understanding trends such as the number of rejections, interviews, job types, and vacancies can provide valuable insights for the hiring department. So the meaningful insights would help to understand the hiring process and help improve the same.

Following are the tasks/insights required post data cleaning and analysis:

A. **Hiring Analysis** - Determine the gender distribution of hires

B. **Salary Analysis** – Determine average salary offered by company

C. **Salary Distribution** - Create class intervals for the salaries in the company

D. **Departmental Analysis** – Chart to show the proportion of people working in different departments

E. **Position Tier Analysis** – Chart to show different position tiers within the company

Prepared by:
Nilesh Kulkarni

Project Approach

High level steps for the Project approach are as outlined below:

- **Data cleaning/preparation** – Work on identifying missing data, duplicate data, data errors and outliers. Additionally strategy to handle these data related issues.
- **Data Analysis**: Understand the given data set and it's attributes.
- **Insights Analysis**: Analyse each insights requirement in detail and prepare MS Excel formula or functions to extract insights. Select optimal and efficient approach.
- **Extract insights**: Use MS Excel as a tool to extract new insights as required including visuals/charts.
- **Review**: Review and cross check output to verify it matches with the requirements/insights required
- **Document**: Document the insights and results to be shared across business teams

Prepared by:
Nilesh Kulkarni

Tech Stack Used

- **Data Analytics tool: Microsoft Excel (Office 365)** Hiring data has been provided in Excel format and excel is further used for data cleaning, analysis and creating visuals/charts to demonstrate insights. Excel is user friendly and functionally rich tool to analyze, visualize and report the data insights.
- **Operating System: Microsoft Windows 11** Version 22H2
- **Documentation: Microsoft office 365 (Power Point) & Acrobat PDF**

Prepared by:
Nilesh Kulkarni

Dataset Overview

The dataset contains records of candidates who were interviewed previously with information about hiring status, hiring department, salary, Gender, Date/Time etc.
Please refer to XLS dataset at Dataset-Link-Google-Drive

The Dataset details are:
- **Number of Rows: 7,168**
- Number of Columns: 6
- Column Details:

1. **Application_id**: ID of the applicant
2. **Interview Taken on**: Date and time of the interview
3. **Status**: Hired or rejected
4. E**vent_name**: Gender of the applicant
5. **Department**: Name of the department for which interview was conducted
5. **Post Name**: Name of the post offered
6. **Offered Salary**: Salary offered for the job

Prepared by:
Nilesh Kulkarni

Data Cleaning

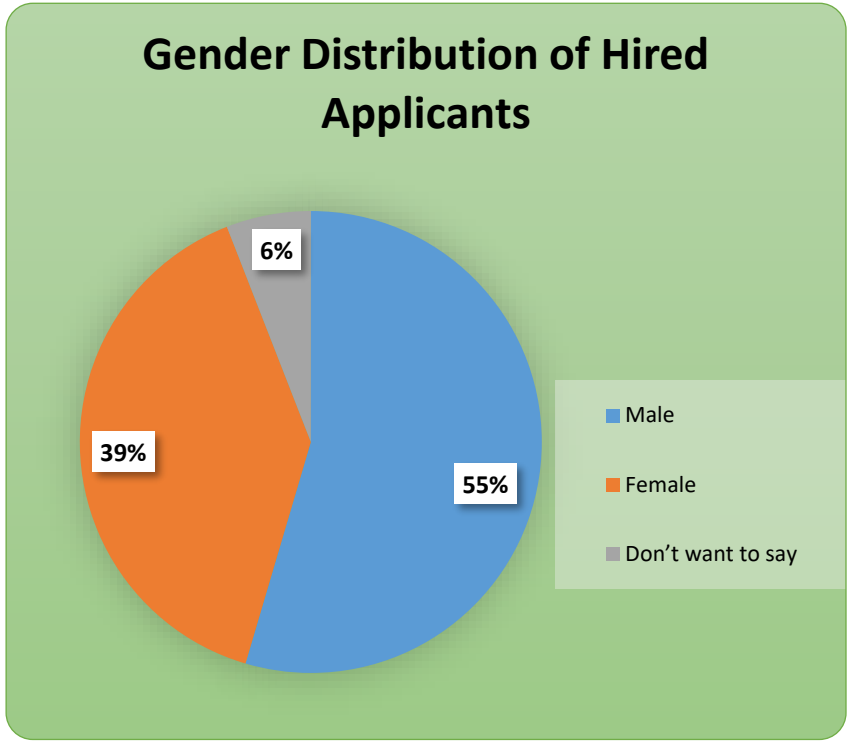
- Missing Values
  - Column event_name has 15 rows with “-“ as its values. It is now replaced with “Don’t want to say” as they both implies the same thing in context of project i.e. gender of the candidate is not known.
  - One row with Salary as blank with post i7 in Sales Department so we set it to median salary of i7 positions in Sales Department = 45400.
  - One row with 'Post Name' as '-' in Sales Department so changed it’s post to C9 based on most matching data with salary in range close to 85914 and department as 'Sales Department’
- Outliers
  - 3 records have ‘Offered Salary’ as 200000, 300000 & 400000. These Seems outliers, but since it might not be appropriate to set those to average/median, this has been left those as it is.
- Duplicate Values & Data Error
  - There are 54 duplicate values based on Application ID, since the complete data(row) was not duplicated, no action/correction was made in data.
  - There was one row with post name as n10, It has been rectified as n-10 based on other 232 rows for c-10.

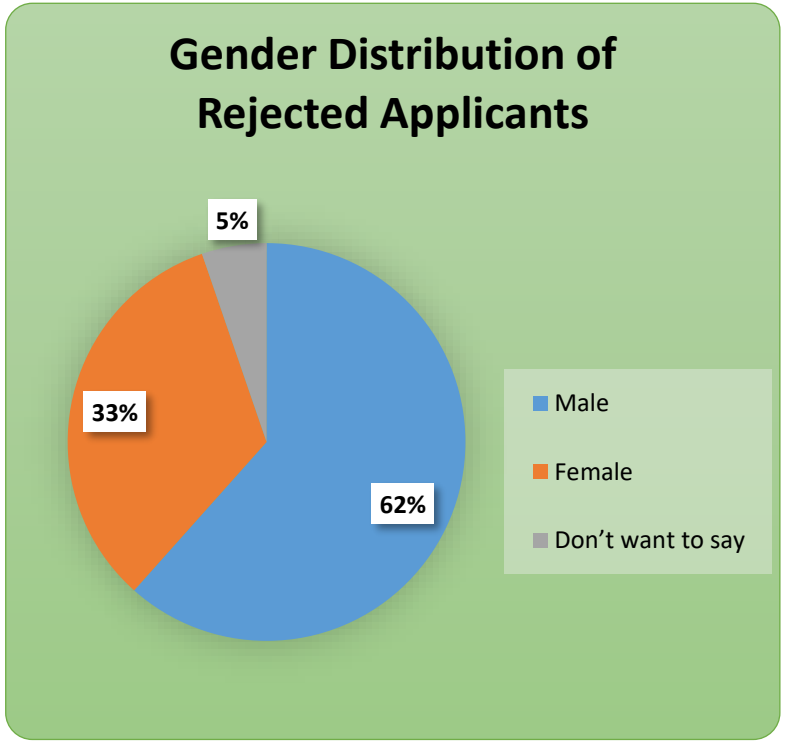
Prepared by:
Nilesh Kulkarni

Task: Determine the gender distribution of hires. How many males and females have been hired by the company

Data and Pie charts as shown for Gender Distribution of hired as well as rejected applicants

Gender	Count of Hired applicants	% of Total	Count of Rejected applicants	% of Total
Male	2563	**54.57%**	1522	62%
Female	1856	**39.51%**	819	33%
Don't want to say	278	**5.92%**	130	5%





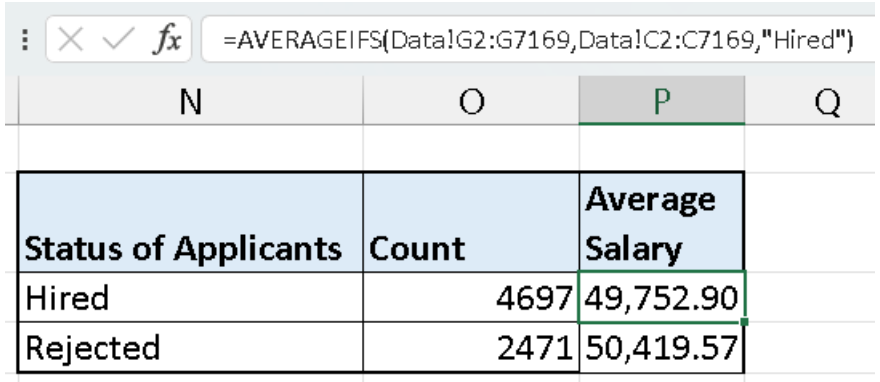
Insights: The Company hired more male (55%) compared to females(39%). Focus on gender diversity might be evaluated.

Task A - Hiring Analysis

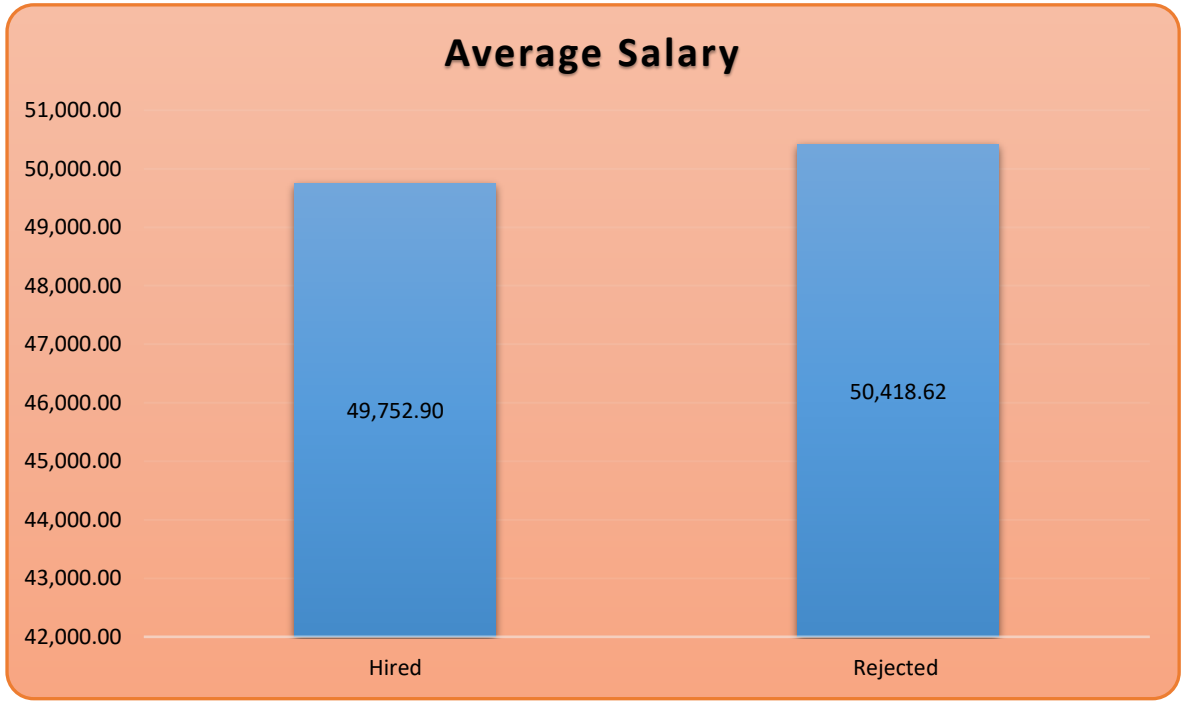
Prepared by:
Nilesh Kulkarni

Task: What is the average salary offered by this company? Use Excel functions to calculate this.

Average salary for Hired as well as Rejected applicants is shown below along with the Excel function in formula bar:



Status of Applicants	Count	Average Salary
Hired	4697	49,752.90
Rejected	2471	50,419.57



Insights: Average salary of hired as well as rejected applicants is in the range of 49,500 to 50,500. It can be evaluated if this matches with industry standards via benchmarking.

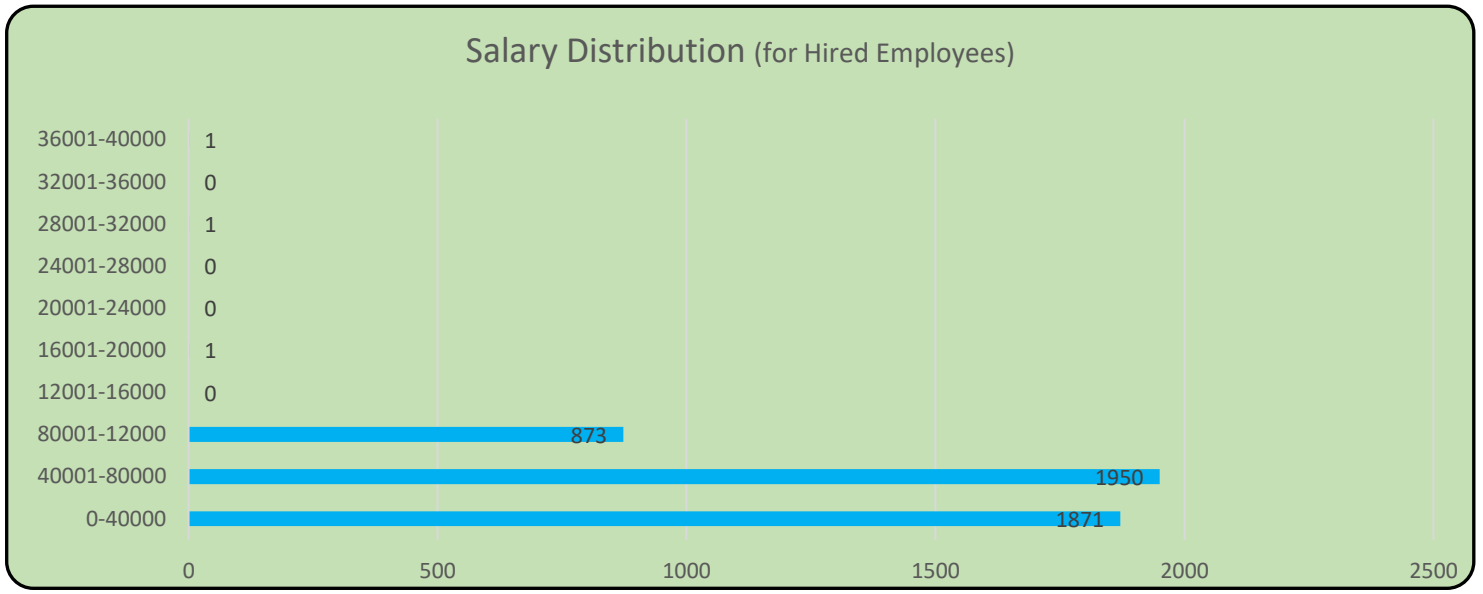
Task B – Salary Analysis

Prepared by:
Nilesh Kulkarni

Task: Create class intervals for the salaries in the company. This will help understand the salary distribution.

Salary distribution data for class intervals as well as chart is shown

Class Intervals	Frequency	% of Total
0-40000	1871	39.83%
40001-80000	1950	41.52%
80001-12000	873	18.59%
12001-16000	0	0.00%
16001-20000	1	0.02%
20001-24000	0	0.00%
24001-28000	0	0.00%
28001-32000	1	0.02%
32001-36000	0	0.00%
36001-40000	1	0.02%
Total	4697	100%



Insights: Salary of more than 99% of employees falls within 0-1,20,000 and 80% employees falls within 0-80,000. The salary distribution should be analyzed and checked if this is as per industry standards. Also 3 records of more than 1,20,000 salary should be evaluated if these are outliers.

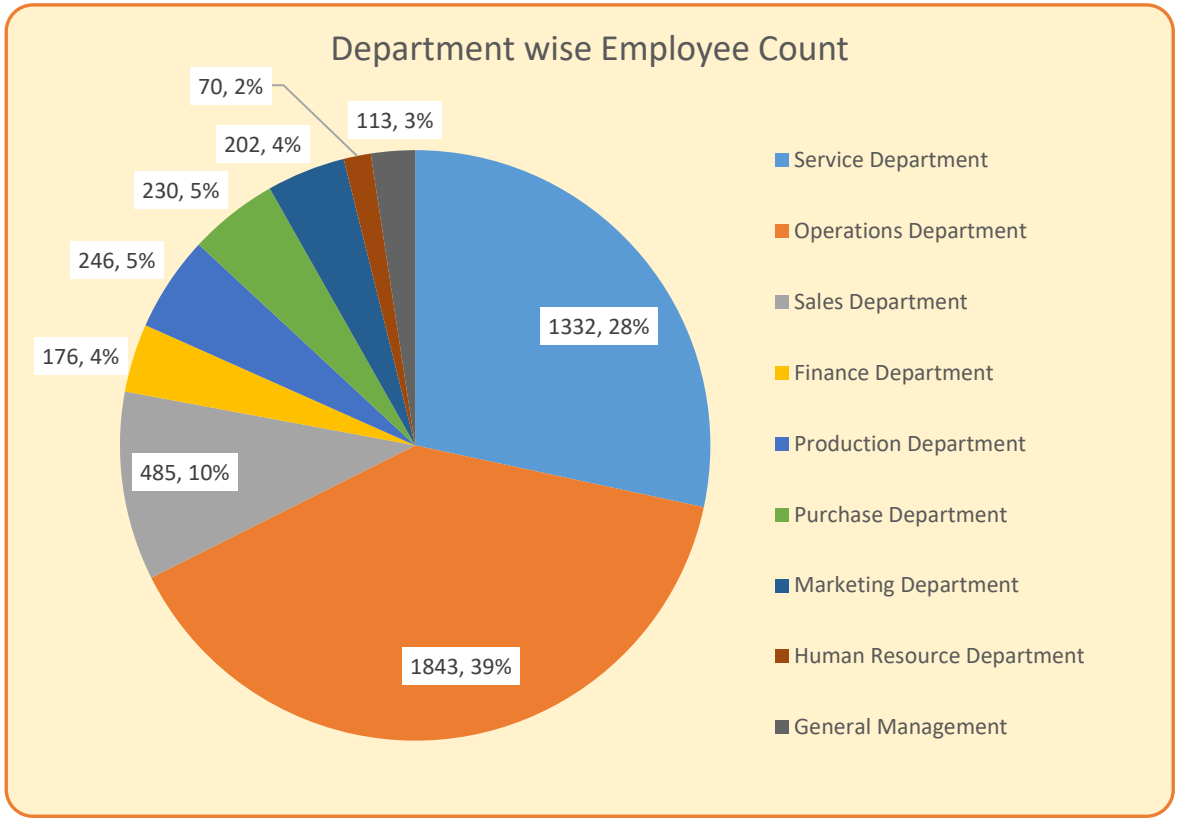
Task C – Salary Distribution

Prepared by:
Nilesh Kulkarni

Task: Use a pie chart, bar graph, or any other suitable visualization to show the proportion of people working in different departments.

Department wise count of employees as given below and pie chart as shown:

Department Name	Count of Employees	% of Total
Service Department	1332	28.36%
Operations Department	1843	39.24%
Sales Department	485	10.33%
Finance Department	176	3.75%
Production Department	246	5.24%
Purchase Department	230	4.90%
Marketing Department	202	4.30%
Human Resource Department	70	1.49%
General Management	113	2.41%



Insights: More than 67% of employees work in the "Operations" and "Service" departments. Evaluate distribution to ensure that departments have the required resources and identify if any hiring issues for departments with low % of total count.

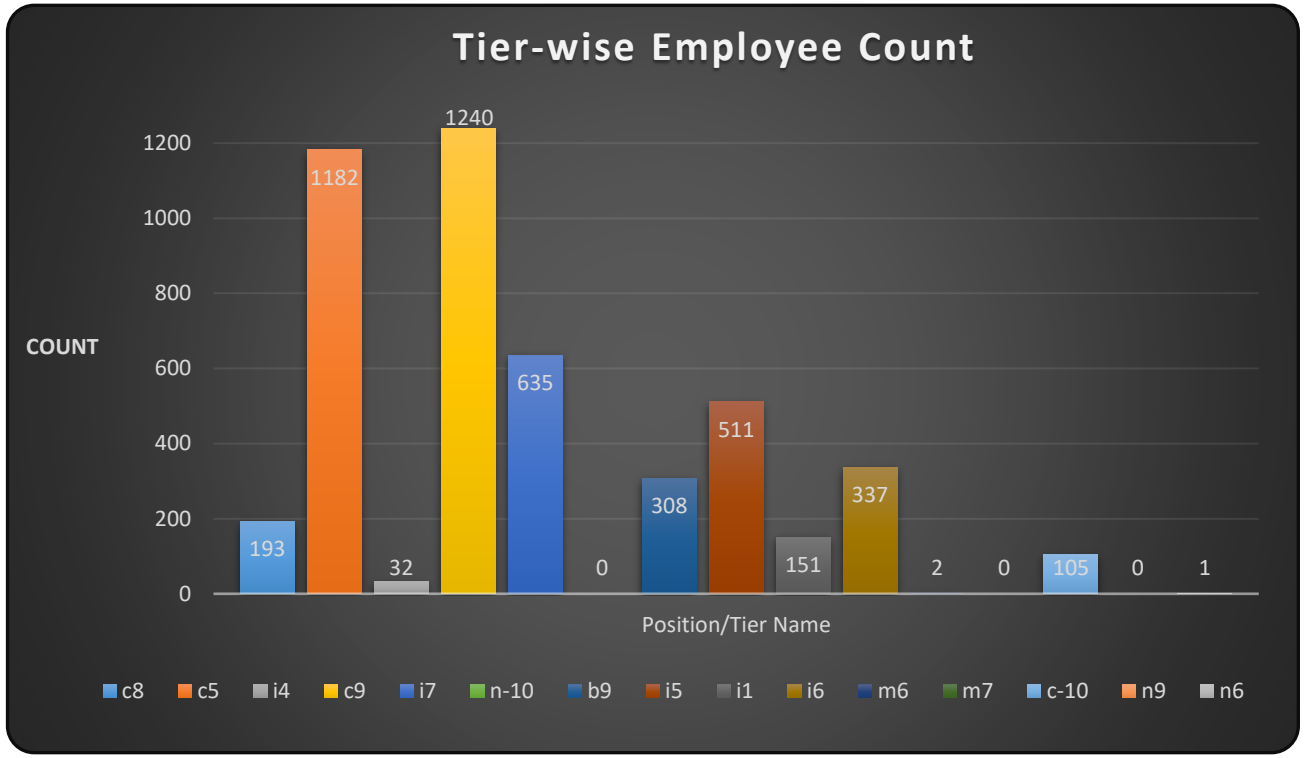
Task D – Departmental Analysis

Prepared by:
Nilesh Kulkarni

Task: Use a chart or graph to represent the different position tiers within the company. This will help understand the distribution of positions across different tiers.

Tier/Position wise count of hired employees is given and chart is shown:

Position/Tier Name	Count of hired Employees	% of Total
c8	193	4.11%
c5	1182	25.16%
i4	32	0.68%
c9	1240	26.40%
i7	635	13.52%
n-10	0	0.00%
b9	308	6.56%
i5	511	10.88%
i1	151	3.21%
i6	337	7.17%
m6	2	0.04%
m7	0	0.00%
c-10	105	2.24%
n9	0	0.00%
n6	1	0.02%



Insights: More than 50% of employees are in Tier “c9” and “c5”. Evaluate if there are enough career progression opportunities across tiers and this tier wise distribution matches with business goals and growth plans.

Task E – Position Tier Analysis

Prepared by:
Nilesh Kulkarni

Insights - Summary

- Data quality can be improved at source since there are few issues in hiring data like missing value/potential outliers and duplicates/errors.
- Need to focus on **Gender Diversity as % of females hired (39%)** is less than males (55%)
- **Average Salary** of hired & rejected candidate is in the range of **49,500 to 50,500**
- More than **99% of employees fall within salary range of up to 1,20,000.**
- Benchmarking can be done for Salary analysis & Salary class intervals to evaluate if salary offered as well as class intervals matches with industry standards.
- Evaluate departments with low% of total employees for any lack of resources. Additionally, evaluation needed with comparing department wise employee distribution is standard as per industry norms. **More than 67% employees work in only 2 departments** ('Service' & 'Operations').
- Based on tier wise distribution of hired employees, evaluation needed to ensure that employees have enough growth opportunities via career progression/promotion (since **more than 50% employees fall within only 2 tiers**).

Prepared by:
Nilesh Kulkarni

Results

- By completing this project, I have got exposure to business domain of hiring (Human Resources) which adds value to my profile.
- I learned how to clean/prepare given data before extracting insights. It also added skills on how to extract business insights from MS Excel data.
- I learnt MS Excel formulas & functions with know-how on how those can help to extract required business insights with practical hands on.
- I got exposure to document extracted insights in appropriate visuals including pie chart, bar chart etc.
- I also gained basic know how about concepts in statistics and their use in Data Analytics.
- Overall, it has been great learning experience on this project with Trainity.

Prepared by:
Nilesh Kulkarni

Thank you

Nilesh Kulkarni
inileshkulkarni@gmail.com