

BOOSTUP DIGITAL DATA **ANALYTICS TASK**



PROBLEM STATEMENT

We are provided with a dataset containing information about a company's employees and we need to need to come up with insights on the following questions:

- What is the average salary of employees by department?
- Which department has the highest number of employees?
- What is the distribution of gender in the company?
- Is there a correlation between years of experience and salary?
- Which department has the highest average salary?

TASK 1

After analysing the data we found that following are average salary by department:

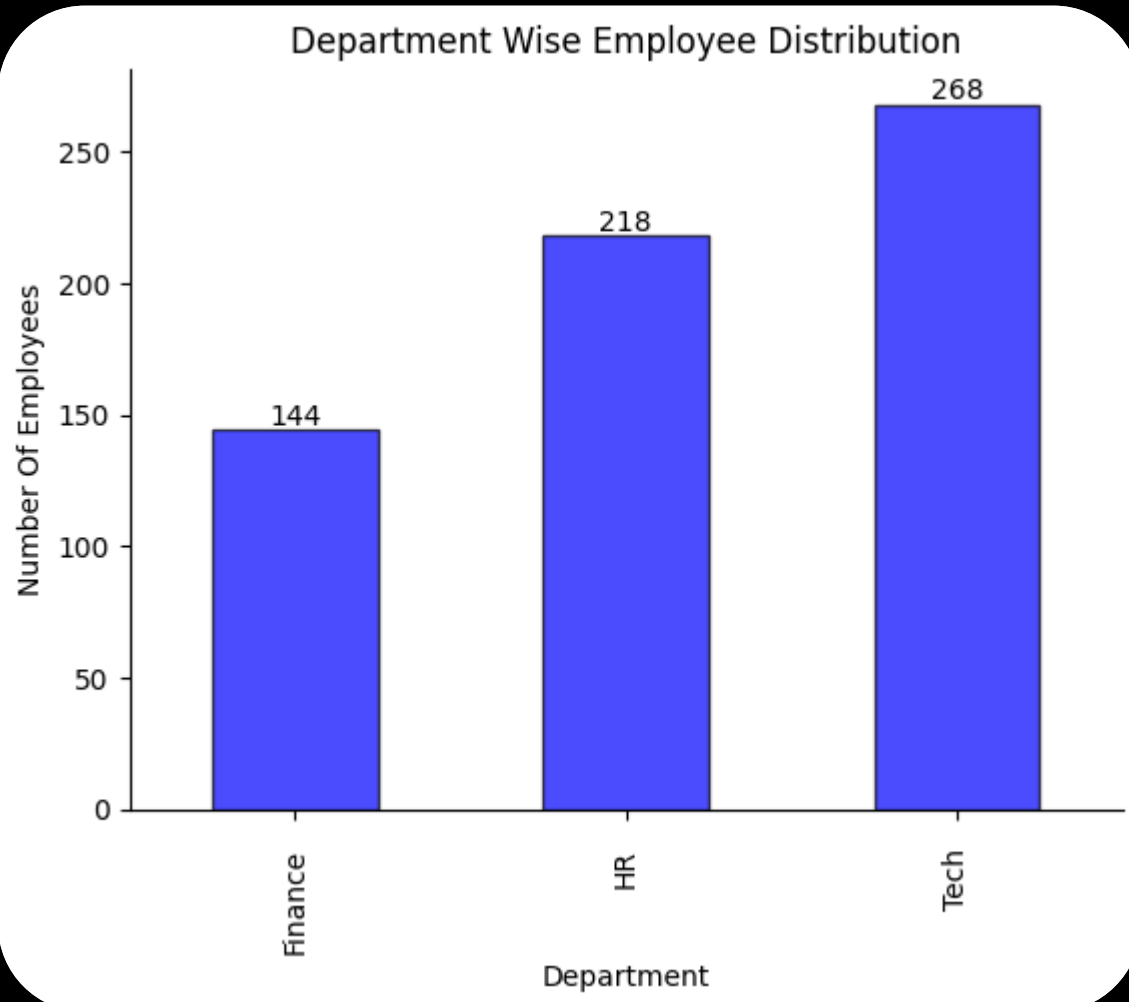
- Finance: Rs. 21,65,561
- HR: Rs. 19,58,208
- Tech: Rs. 20,77,257

```
avg_sal=df_fin.groupby(df_fin['Dept_name'])  
avg_sal['Salary'].mean()
```

✓ 0.0s

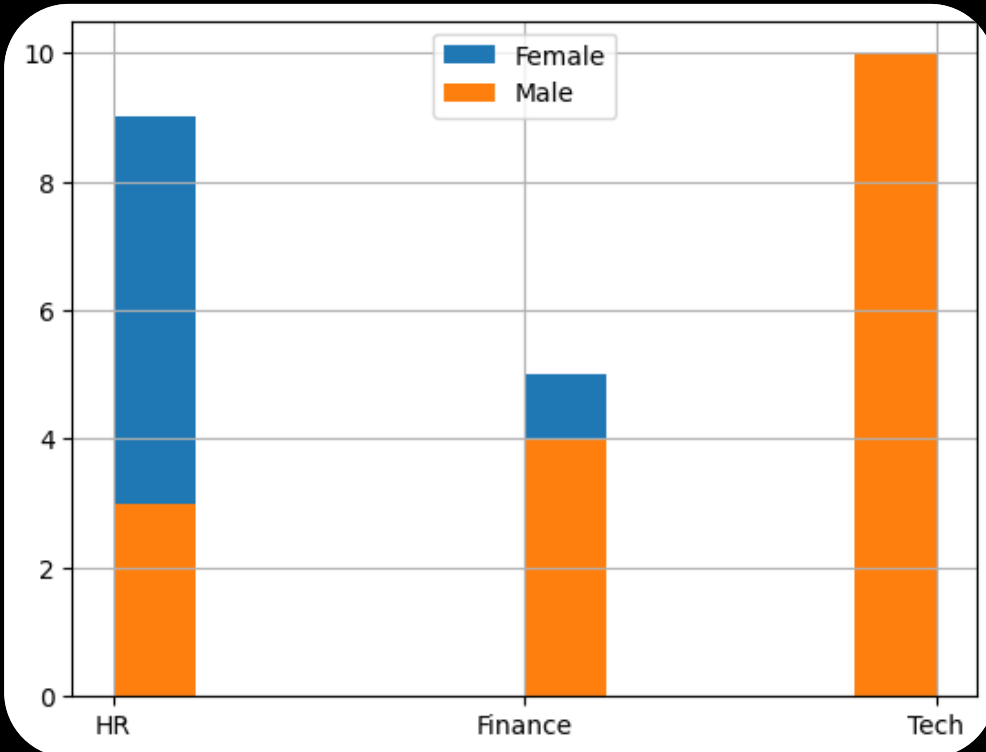
```
Dept_name  
Finance    2.165561e+06  
HR         1.958208e+06  
Tech       2.077257e+06  
Name: Salary, dtype: float64
```

TASK 2



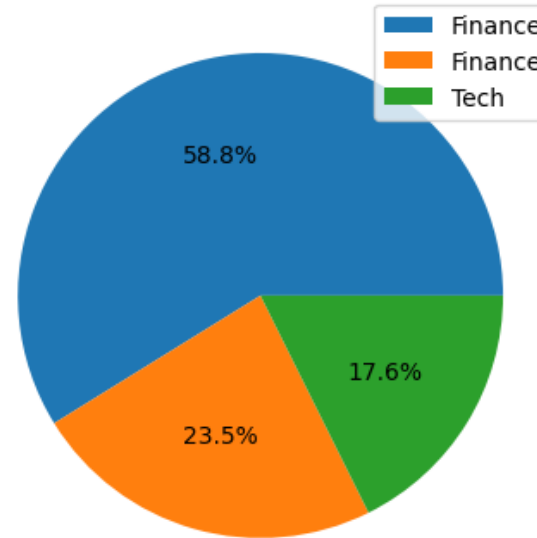
From the graph beside we can see that the highest number of employees are working in the Tech department then HR and then Finance. Even though Finance department has the least number of employees present but the organization spends/rewards them the highest, as we saw in our previous analysis

TASK 3

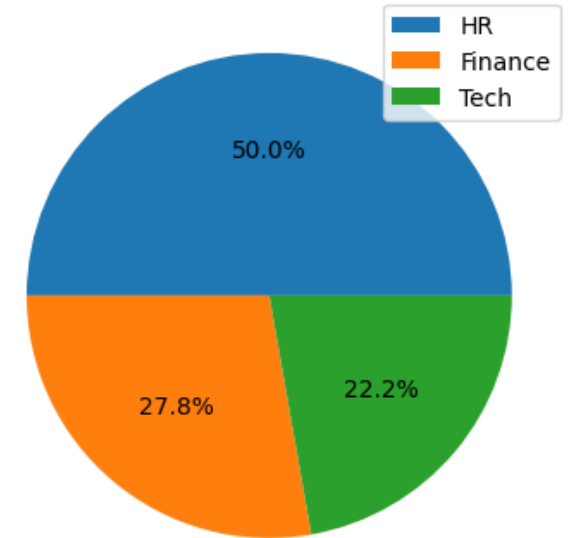


Gender Distribution by Department

Male Distribution by Department

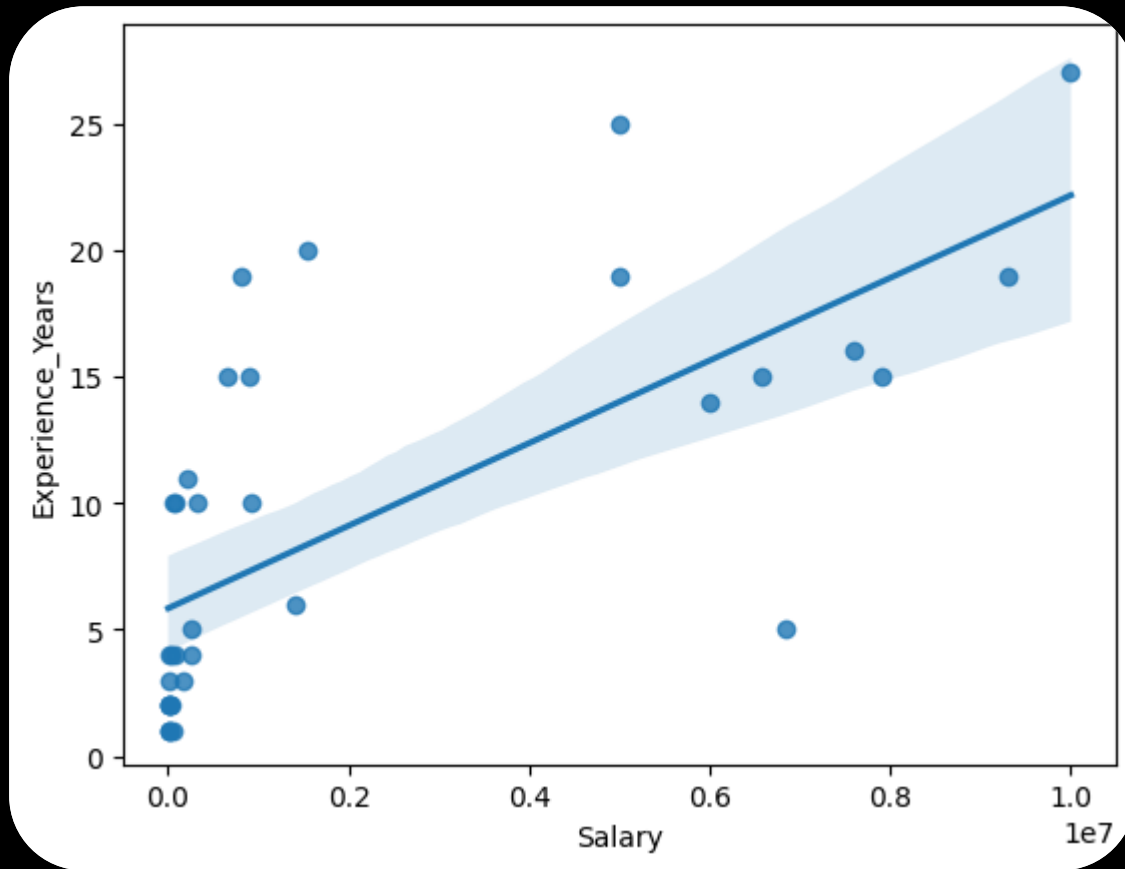


Female Distribution by Department



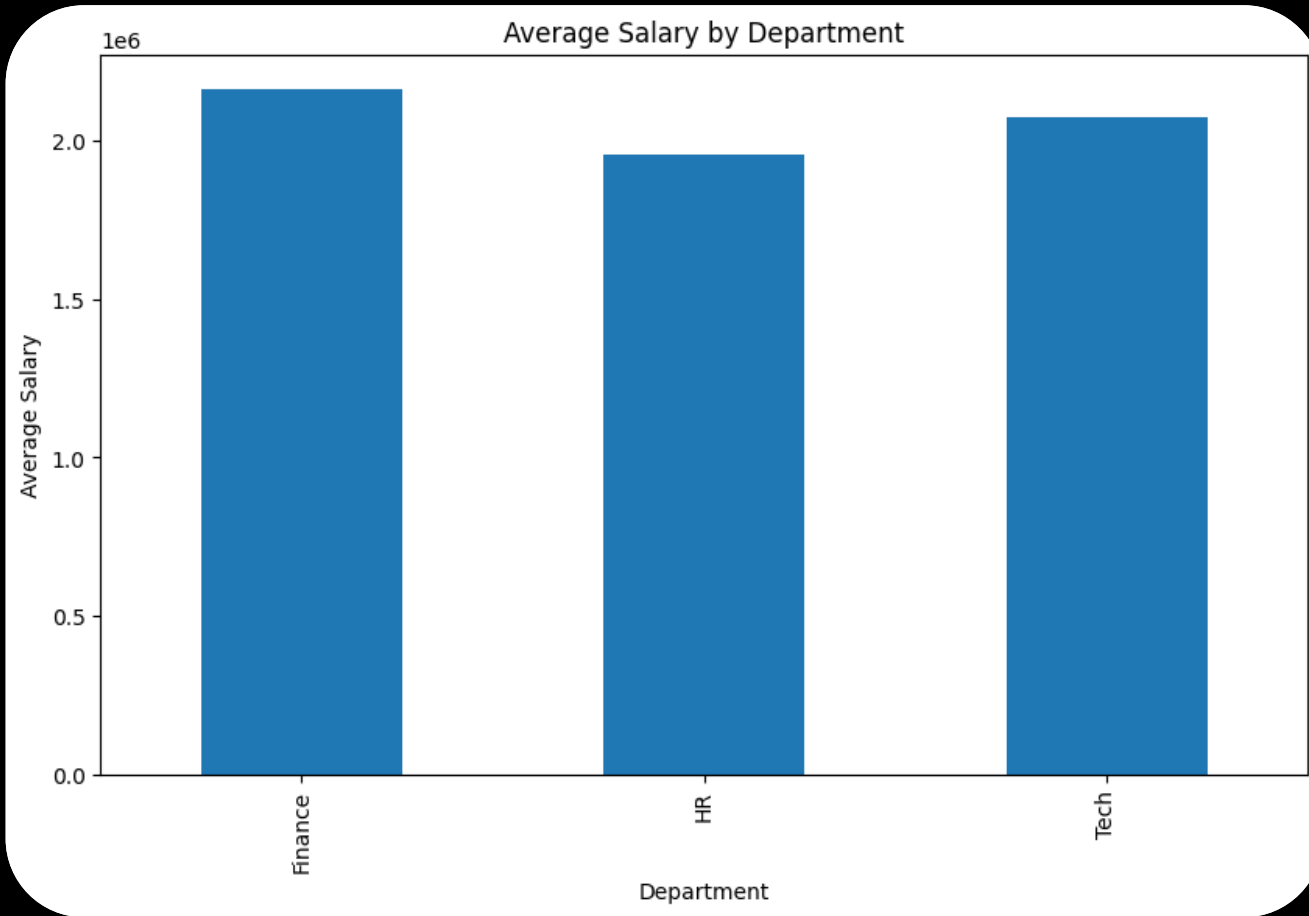
The charts provides us with the gender distribution within the organization/departments and we can observe that the Tech Department is dominated by male employees whereas HR Department is dominated by female employees and Finance has a fare share of both

TASK 4

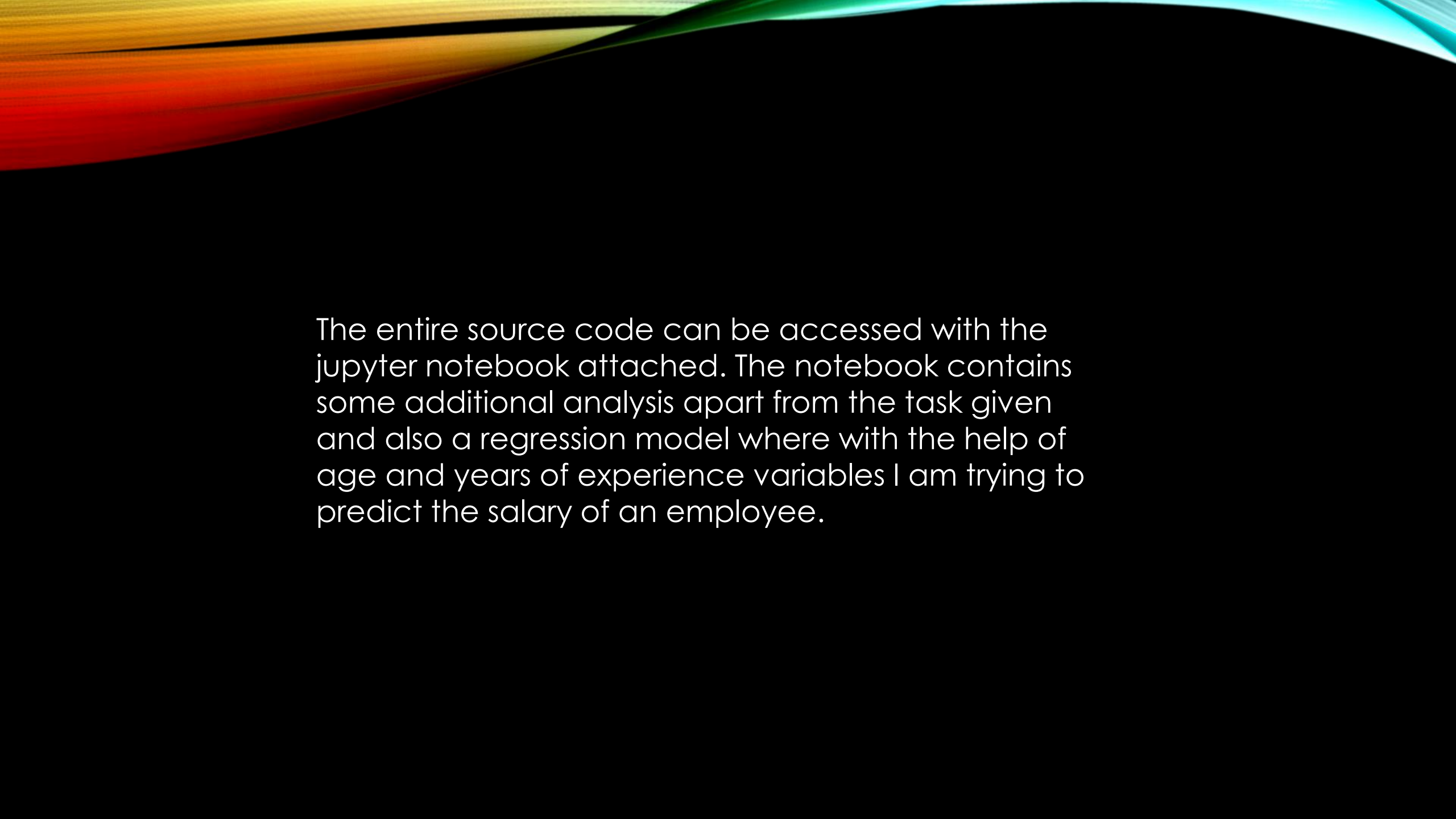


The graph beside demonstrates the correlation between years of experience and salary. The points are somewhat seen to be increasing i.e. there seems to be some correlation between the two variables. Exactly the correlation comes out to be 68.56%, which is decent but it tells us that the salary does not always depends on years of experience.

TASK 5



We can see that the highest average salary is rewarded to the Finance department which comes out to be Rs. 21,65,561



The entire source code can be accessed with the jupyter notebook attached. The notebook contains some additional analysis apart from the task given and also a regression model where with the help of age and years of experience variables I am trying to predict the salary of an employee.



THANK YOU

Ayush Kumar Panda

ayushpanda2301@gmail.com

9871493424