HR Analytics: Data Insights in R

Data Visualization

Ki Hin Stephanie Cheng
January 7, 2025

EXECUTIVE SUMMARY

This report presents the visualization of employees on the dataset "HRDataset_v14.csv" with using R programming.

CONTENTS

DATA VISUALIZATION

Introduction

This part illustrates different principles within the organization, mainly with using `ggplot` and `lattice` libraries.

Finding / Visualization

<u>The relationship between employees' supervisors and their performance score</u>

There are 21 managers in the company, and the number of employees under each manger is as follows:

```
Alex Sweetwater              Amy Dunn  Board of Directors  Brandon R. LeBlanc
            9                      21                   2                   7
Brannon Miller        Brian Champaigne        David Stanley      Debra Houlihan
           22                       8                  21                   3
 Elijiah Gray            Eric Dougall          Janet King      Jennifer Zamora
           22                       4                  19                   7
   John Smith           Kelley Spirea        Ketsia Liebig       Kissy Sullivan
           14                      22                  21                  22
Lynn Daneault           Michael Albert        Peter Monroe          Simon Roup
           13                      22                  14                  17
Webster Butler
           21
```

Below is the plot of employee performance scores by manager:

Performance Score ID (4 = Exceeds , 3 = Fully Meets , 2 = Need Improvement ,1 = PIP)
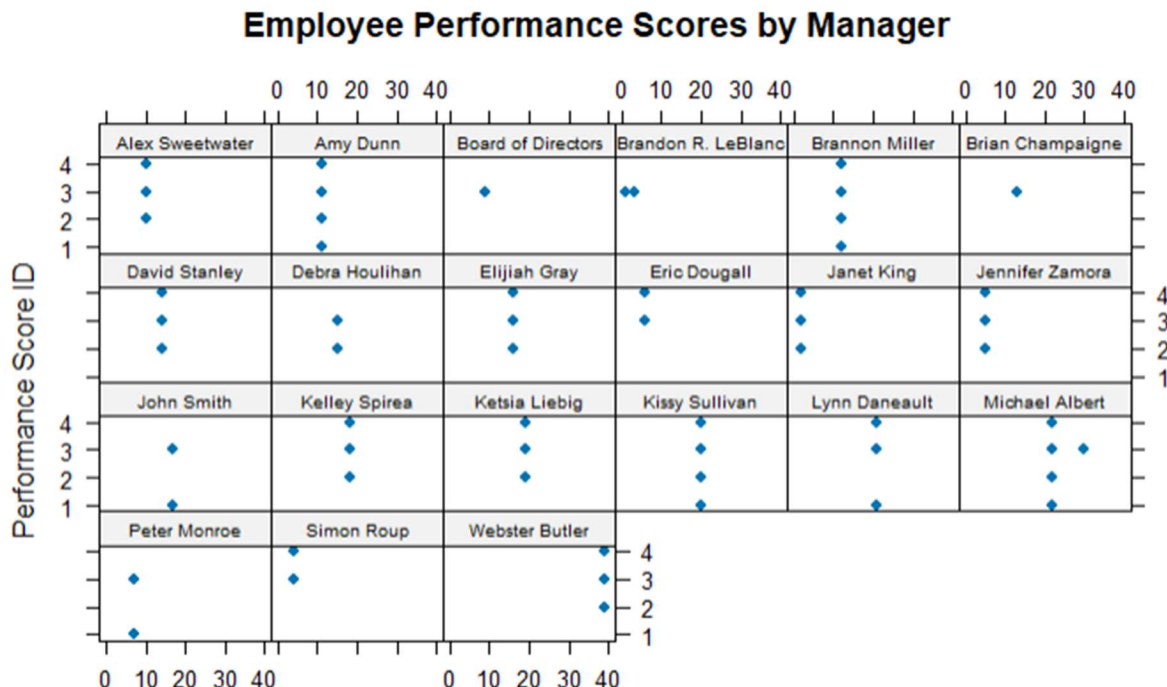
Employees working under **Manager Simon Roup and Eric Dougall** have the **highest average performance**, as their scores either exceed or fully meet expectations. Employees under **Managers Brandon R. LeBlanc, the Board of Directors, and Brian Champaigne** also perform well, with most of them **fully meeting expectations**.
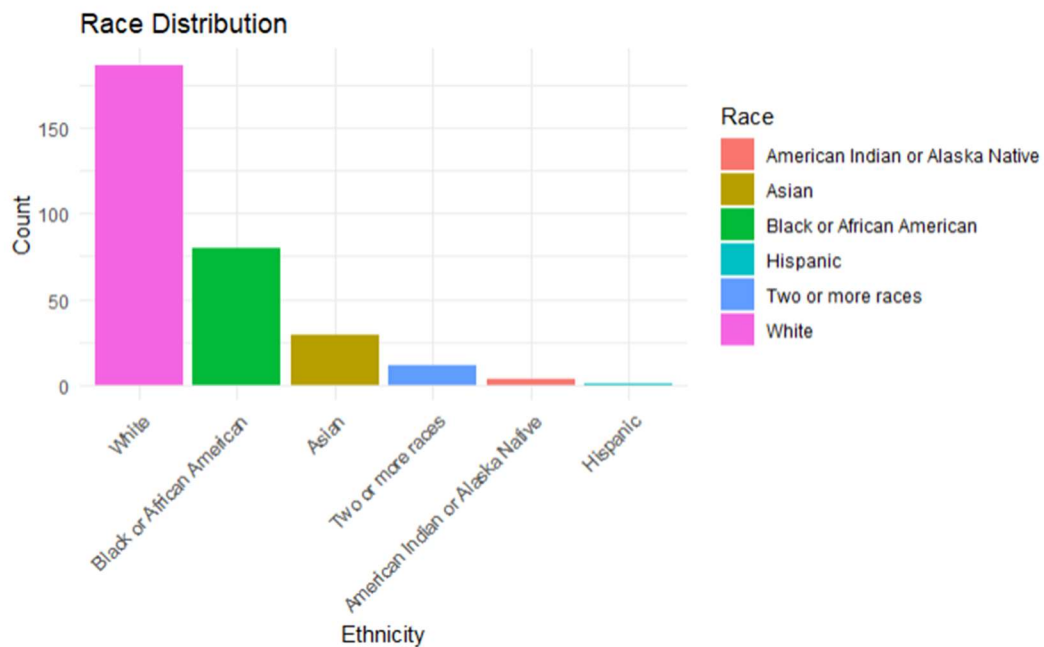
Diversity Profile

Gender

Gender Distribution
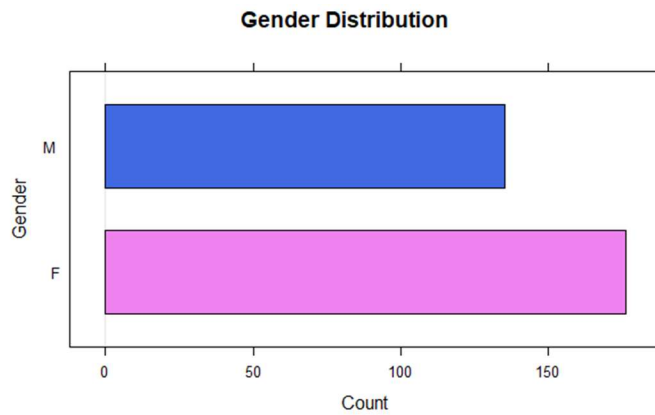
The number of **female employees** (176) **is higher** than that of male employees (135).

Race
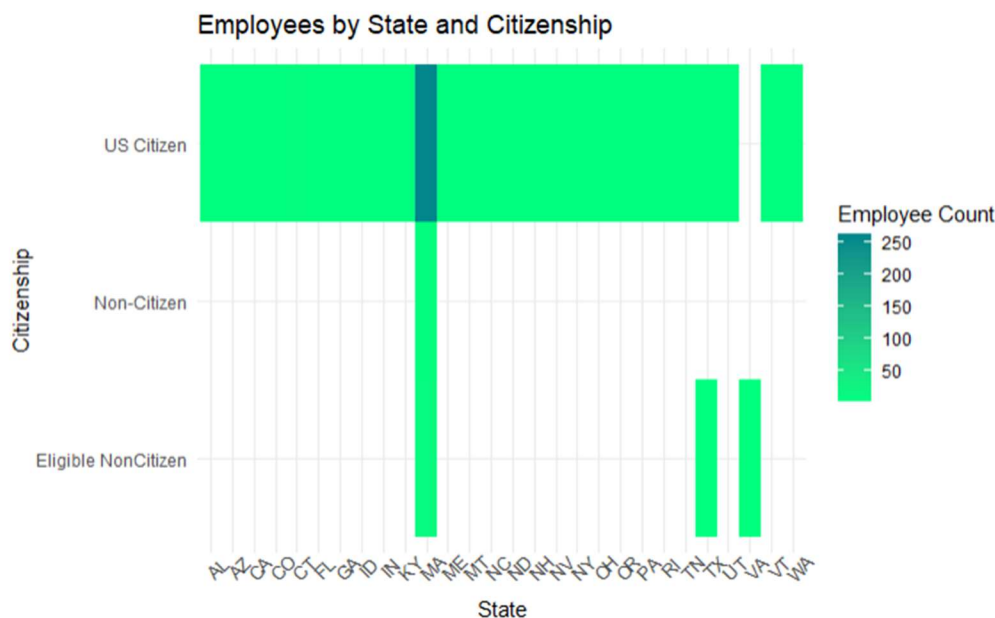
Race Distribution

White employees make up **60.13%** of the workforce, making them the largest racial group. The second-largest group is **Black or African American** employees, comprising **25.72%** of the company. In contrast, the **smallest** racial group is **Hispanic**.

Marital Status

Employee Distribution by Marital Status
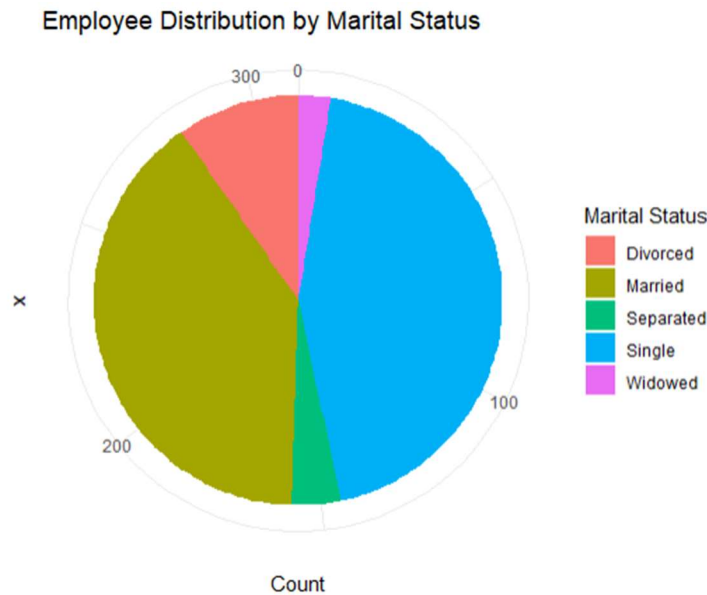
Most employees are **single**, while only **3% of the workforce is widowed**.

Citizen Information

Employees by State and Citizenship

The majority of employees are **U.S. citizens**, with most residing in **Massachusetts (MA)**. The distribution of employees across other states is **average**.

Very few employees are **non-citizens**, and all of them reside in **Massachusetts (MA)**. Eligible non-citizen employees are located in **Massachusetts (MA), Texas (TX), and Virginia (VA)**.

Recruitment Source Trend

Given : RecruitmentSource
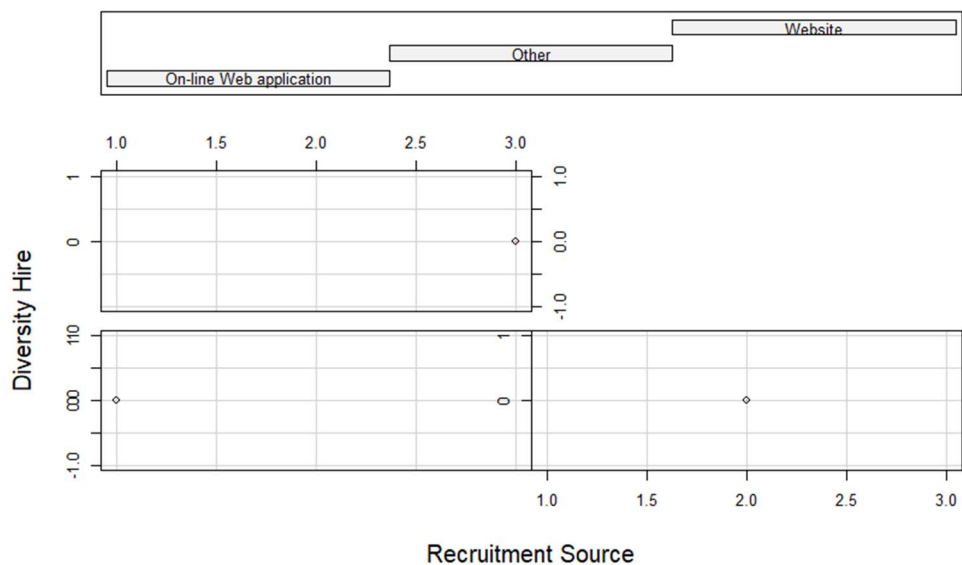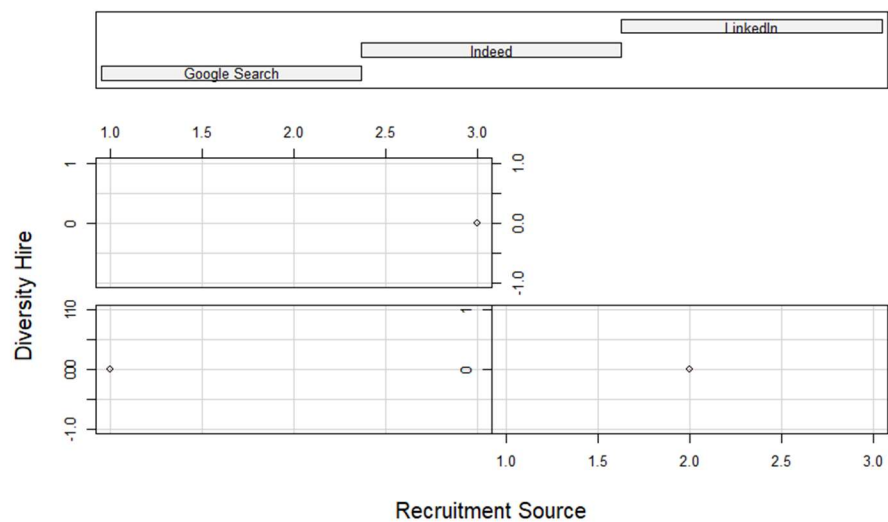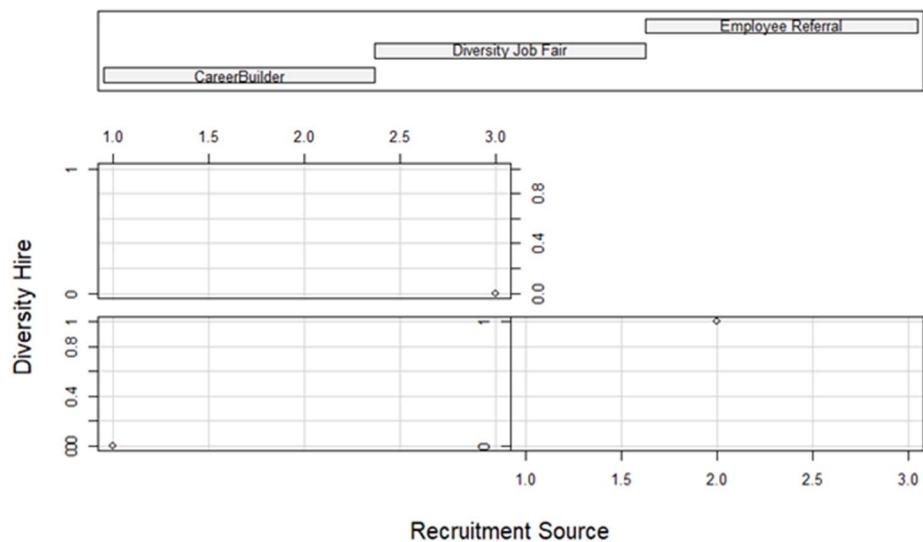
Diversity hires(Diversity hires = 1) mainly come from **Diversity Job Fairs**, while **non-diversity hires (Diversity hires = 0)** are sourced from various channels, including **Employee Referrals, Career Builder, others, etc**.

This trend suggests that the organization should **invest more in Diversity Job Fairs** if it consider expanding diversity-focused recruitment. Partnering with diversity-focused organizations, expanding outreach to underrepresented groups, and ensuring inclusive job postings and interview process could further support diversity hiring outcomes.

Termination
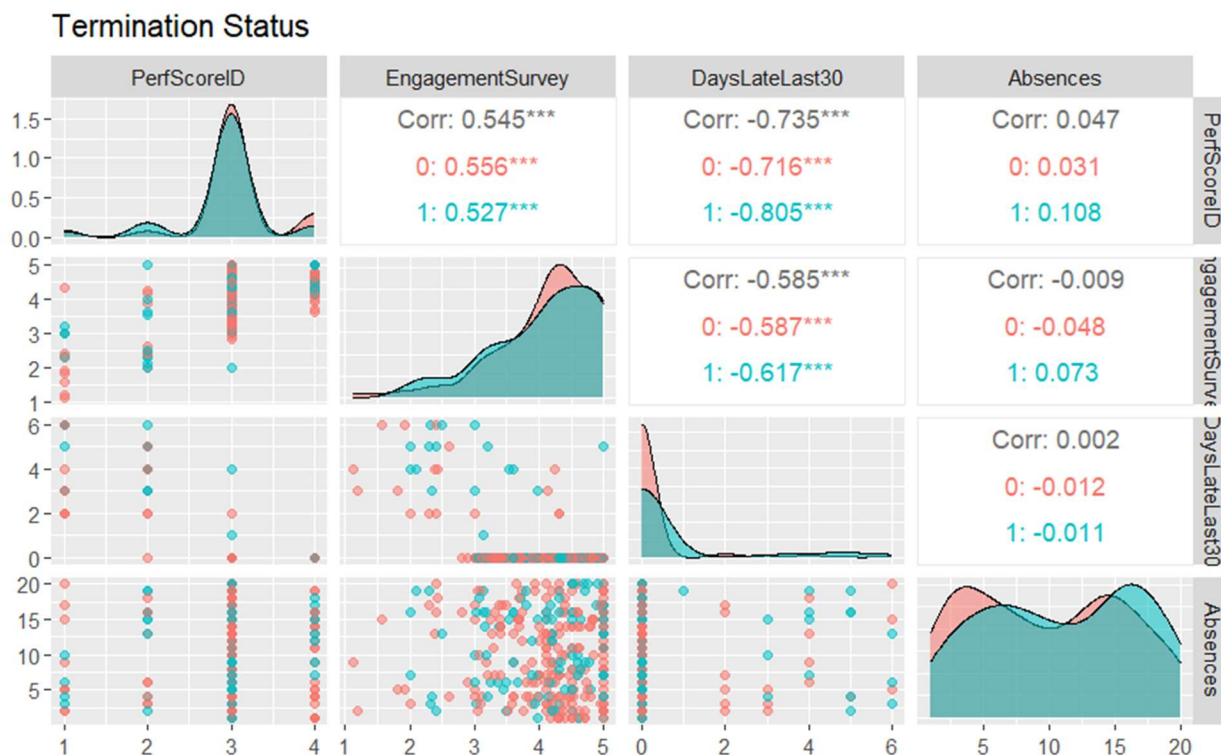
This plot shows a positive correlation (**0.545**) between **PerfScoreID** and **Engagement Survey**, indicating that employees with high performance and engagement scores are less likely to be terminated. A strong negative correlation (**-0.805**) between **DaysLateLast30** and retention suggests that frequent lateness significantly increases the likelihood of termination. While the correlation between **Absences** and retention is weak (**0.002**), employees with fewer absences (around 10 or less) tend to have better retention rates.

Equality of Salary
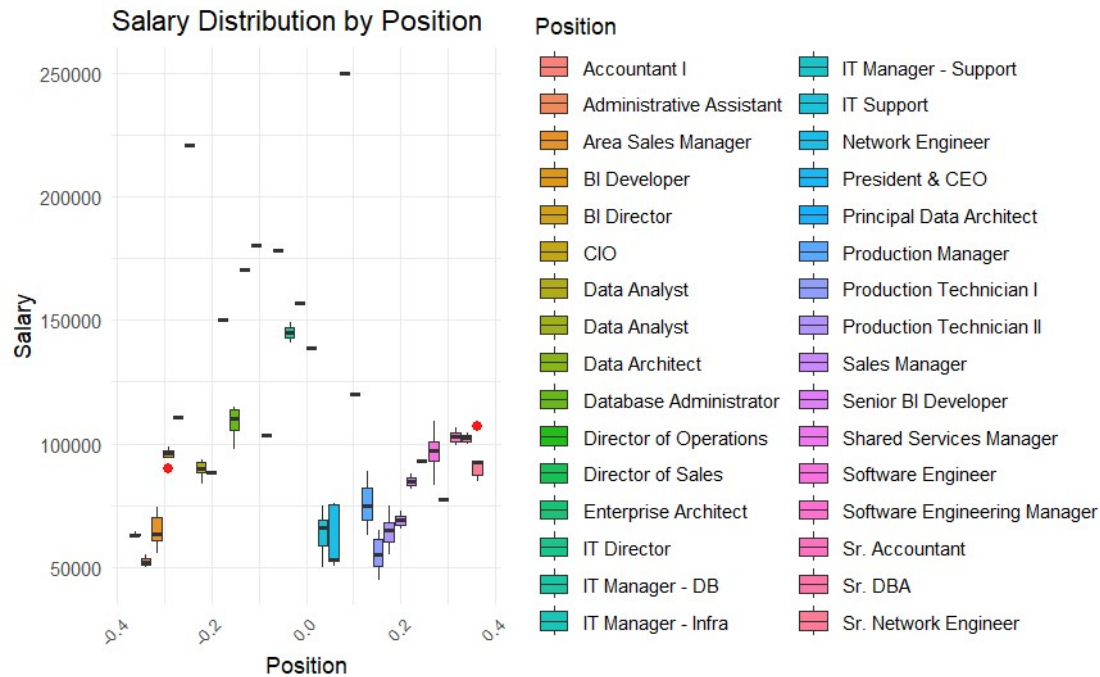
The **President & CEO** has the highest salary of **$250,000**, followed by the **CIO** and then the **Director of Sales**. For positions with higher salaries, the number of employees in those roles does not exceed five. On the other hand, **Production Technician I** earns the lowest salary and has the highest number of employees.
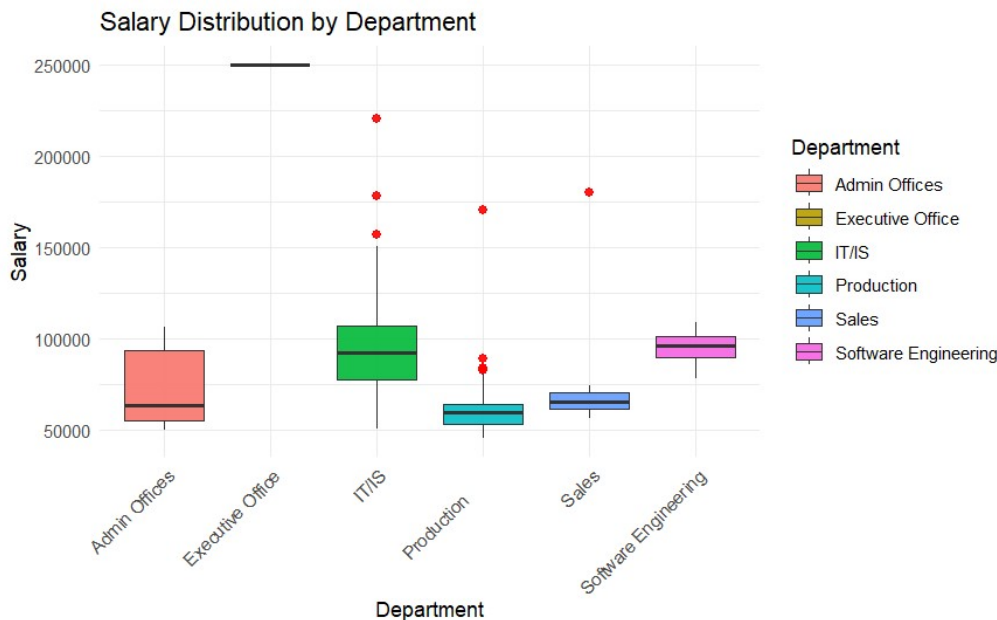
The **Executive Office** has the highest salary range. Most **IT/IS** salaries range between **$76,000 and $105,000** , making it the second or third highest salary range. However, a few individuals in this department earn significantly more, which explains the presence of **outliers** in the **IT/IS** section. The **Production** department has the **lowest salary range**.

Salary Distribution by Gender
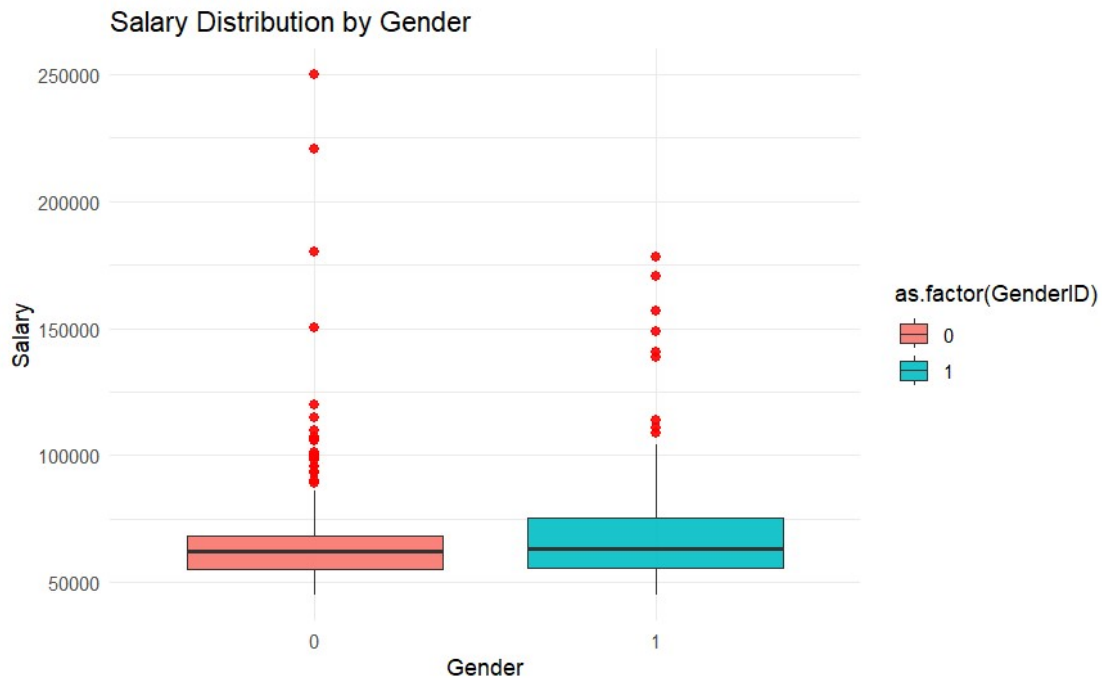
Gender (0 = Male, 1 = Female)

Most **male employees** earn more than **female employees** in individual comparisons. However, on average, **female employees** earn more than **male employees** overall.