



# **HUMAN RESOURCES DATASET REPORT**

Maria Clara Duque

A decorative graphic on the left side of the slide consisting of two blue squares. The top square is a lighter shade of blue and is positioned above the bottom square, which is a darker shade of blue. They are aligned to the left of the main text area.

# Human Resources Dataset

The dataset contains **various features** about a **fictitious company**, such as names, reasons for termination, department, position title, and performance score.

**Objective:** Apply analytics and data visualization to obtain insightful information.

**Dataset:**

<https://www.kaggle.com/rhuebner/human-resources-data-set>



# The Problem

The company HR manager needs to identify the relationships between the performance of its employees and who leads them.

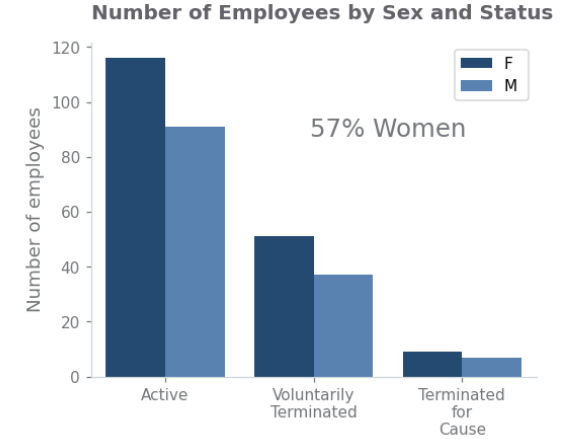
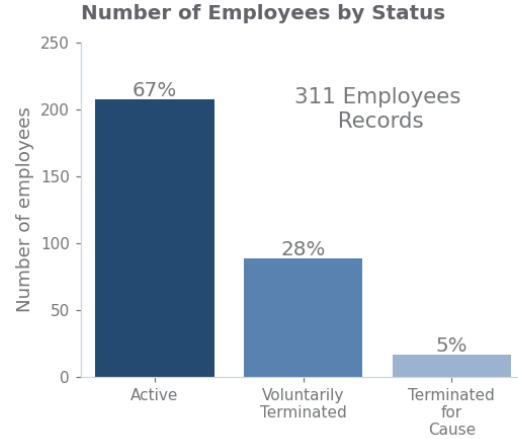
We want to answer the questions:

- What is the general profile of the organization?
- What are the best sources of recruitment?
- Suggestions and recommendations.

# General Information of Employees

- 67% of the 311 records are **ACTIVE** employees.
- Most employees are **WOMEN**.
- The Company is in **Massachusetts**.
- Most employees are **not married**.

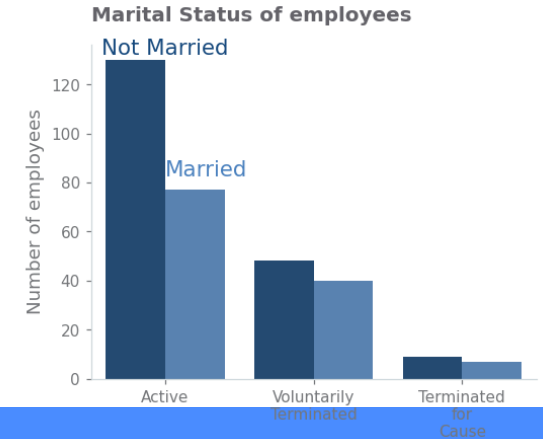
## General information of Employees



## US States Employees

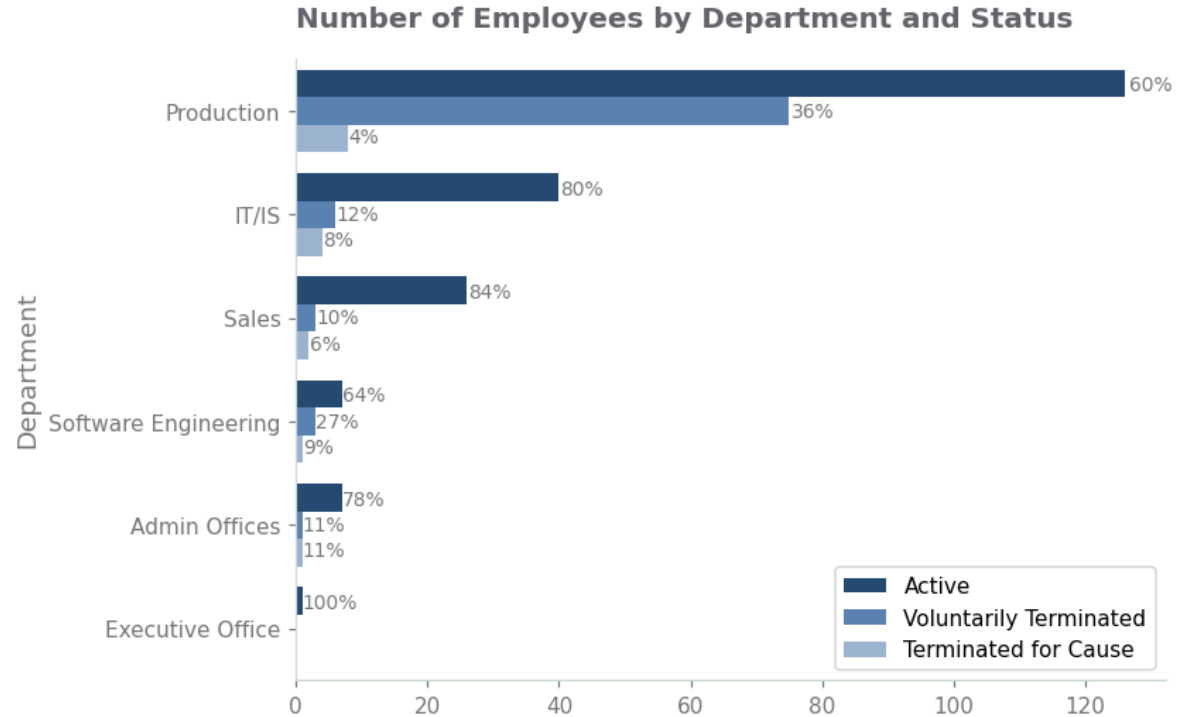
**89%** from **MASSACHUSETTS**

11% Other States.



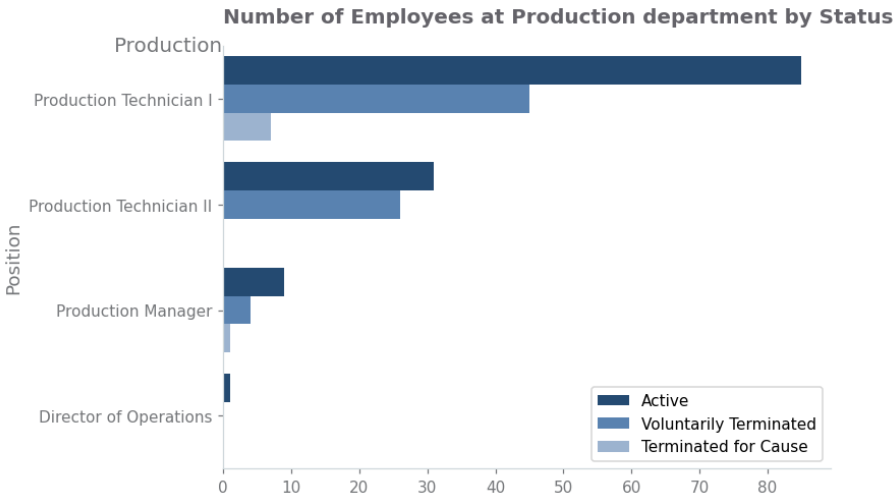
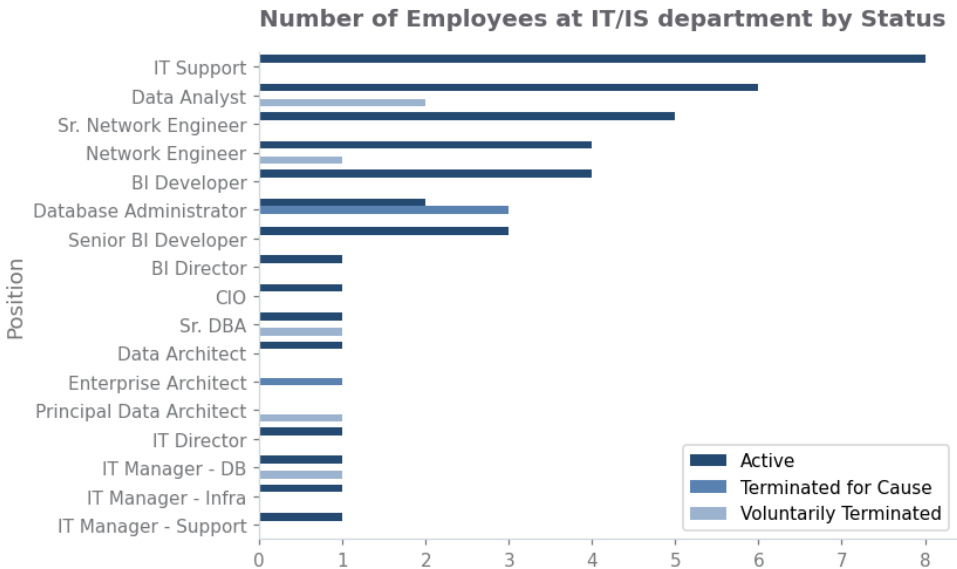
## General Information of Employees

- **Production** is the department with the highest number of employees.
- Department with highest absence level:
  - **Production (40%)**
  - Software Engineering (36%)
  - Admin Office (22%)
  - IT/IS (20%)
  - Sales (16%)



# Production and IT/IS Department

Production: 4 different positions.

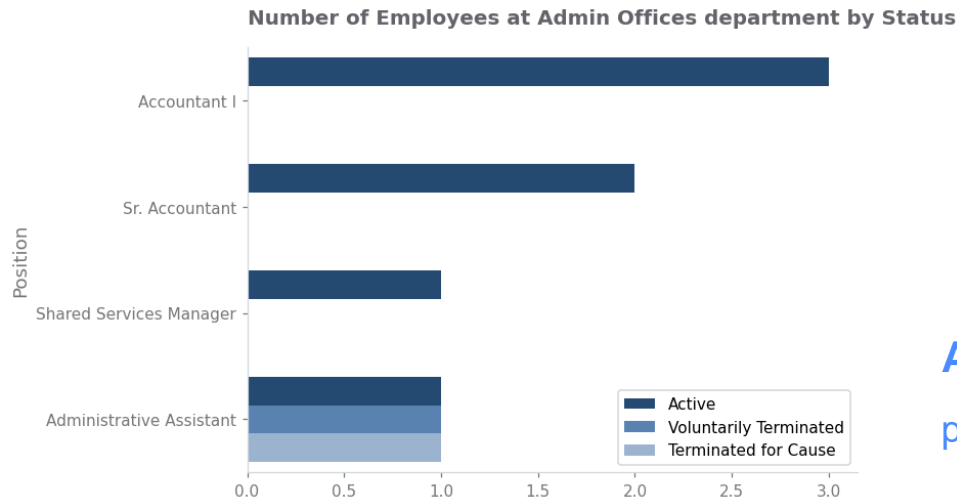
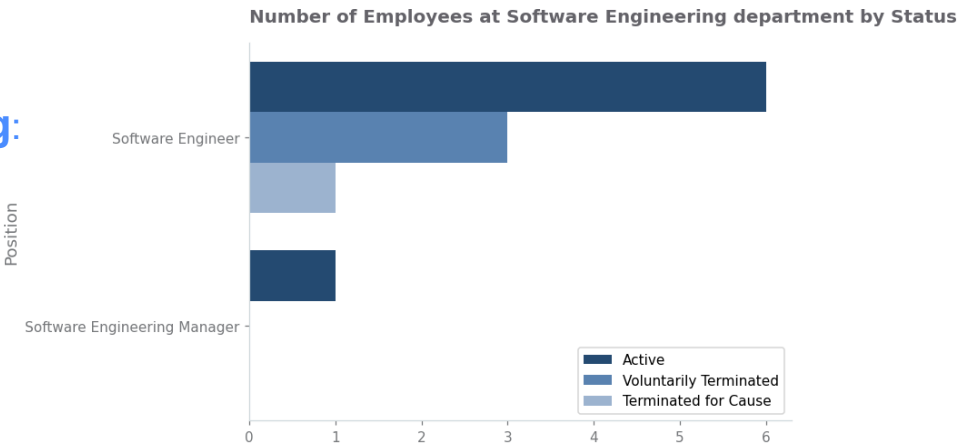


IT/IS: 17 different positions.

Note that 3 different **Database Administrator** were terminated for cause.

# Production and IT/IS Department

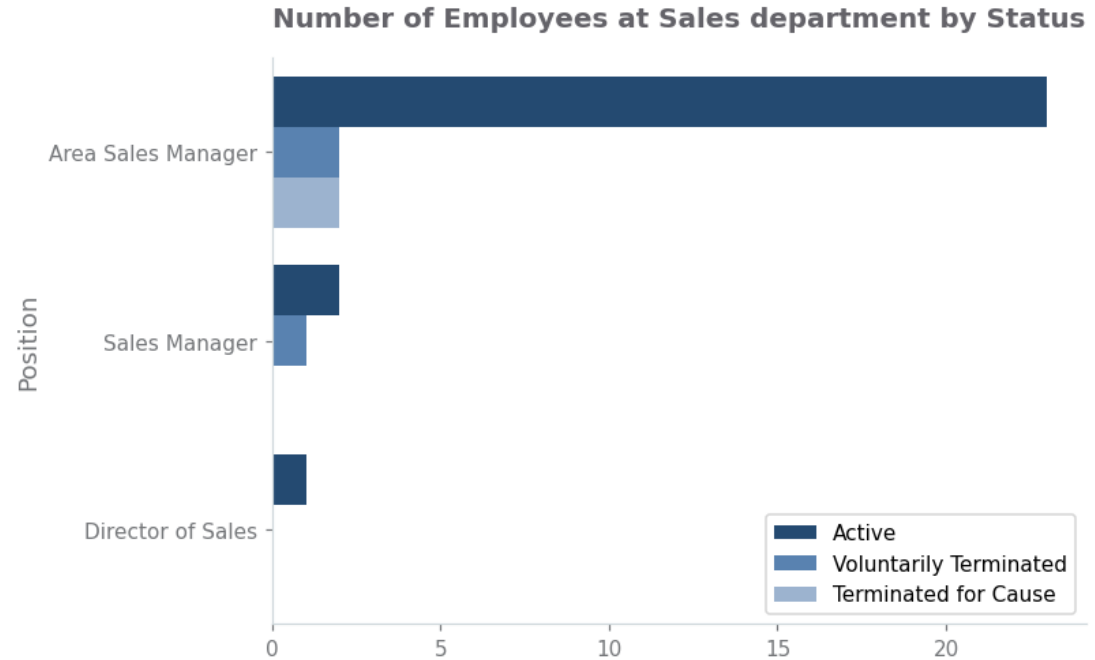
Software Engineering:  
2 different positions.



Admin Office: 4 different  
positions.

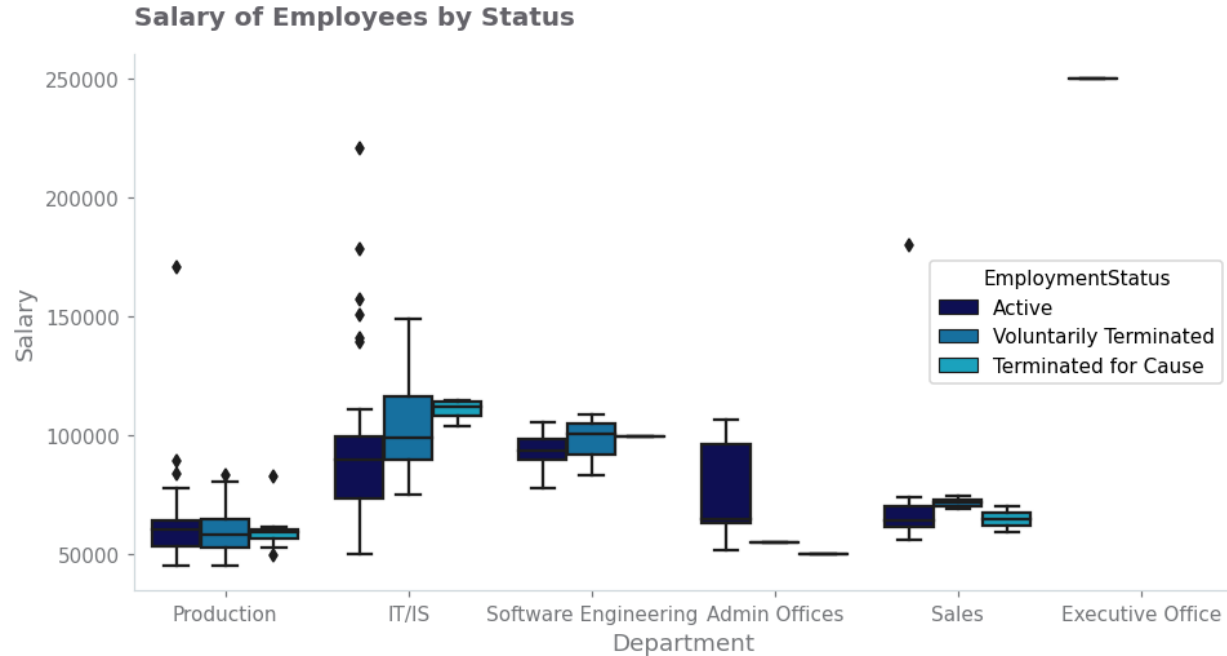
# Sales

Sales: 3 different positions.





# Salary by Department

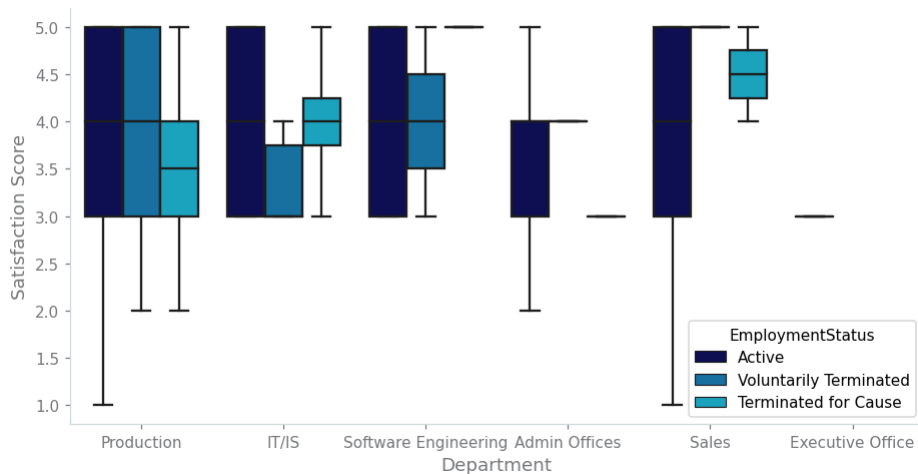


- IT/IS with **greater** variation in salaries due to different positions.
- Executive Office with the highest average salary, followed by IT/IS and Software Engineering.

# Employee Satisfaction Survey

## Internal Survey

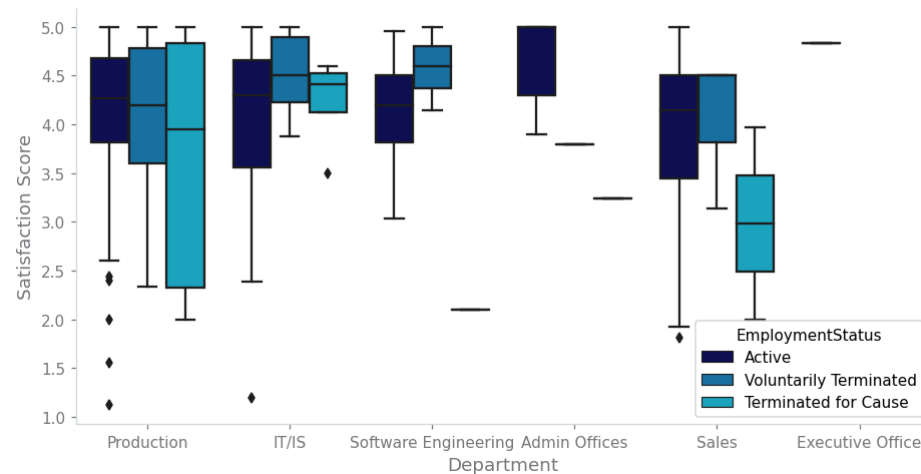
Employee Satisfaction Survey Result by Status



Mean	3,89
std	0,90

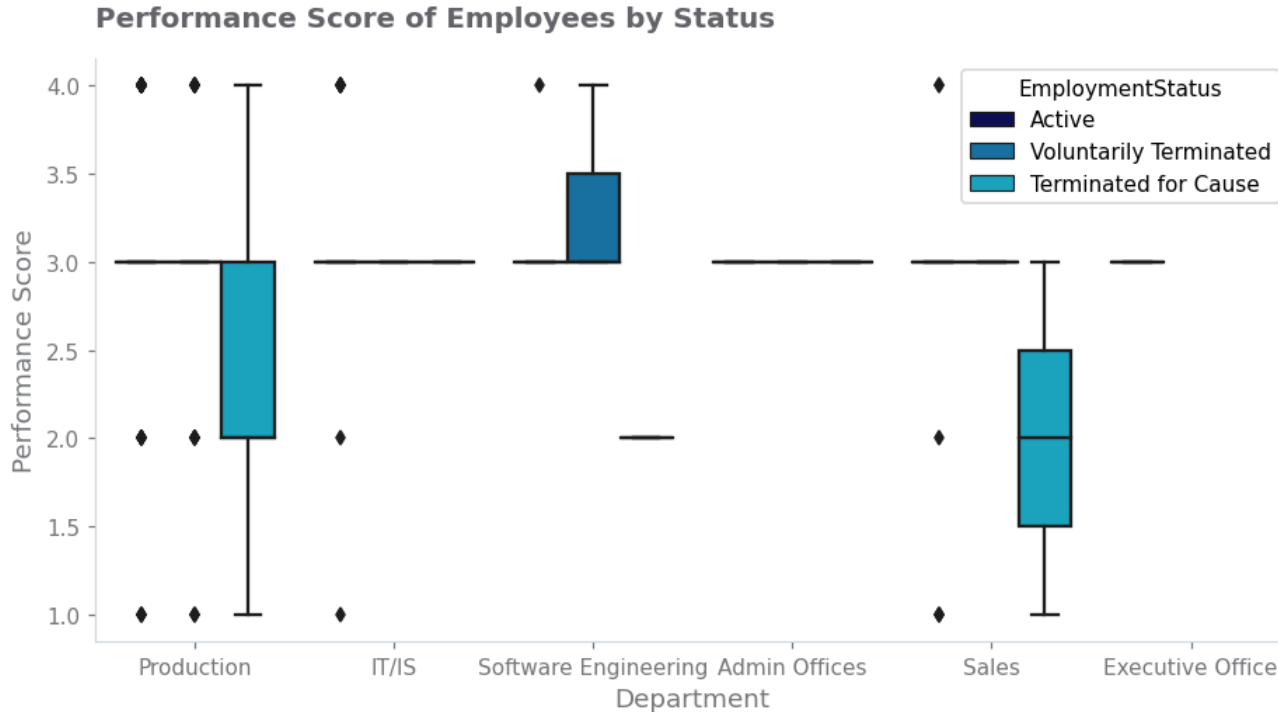
## External Survey

Employee Satisfaction from External Partner by Status



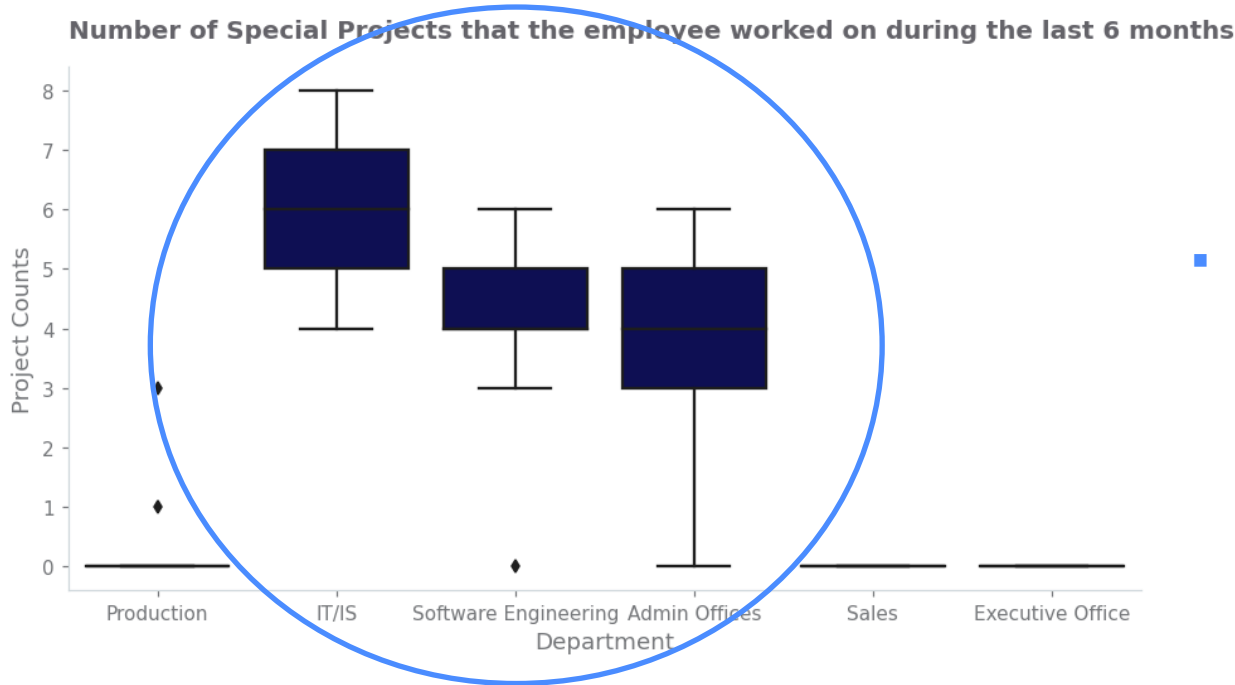
Mean	4,11
std	0,79

# Performance Score of Employees



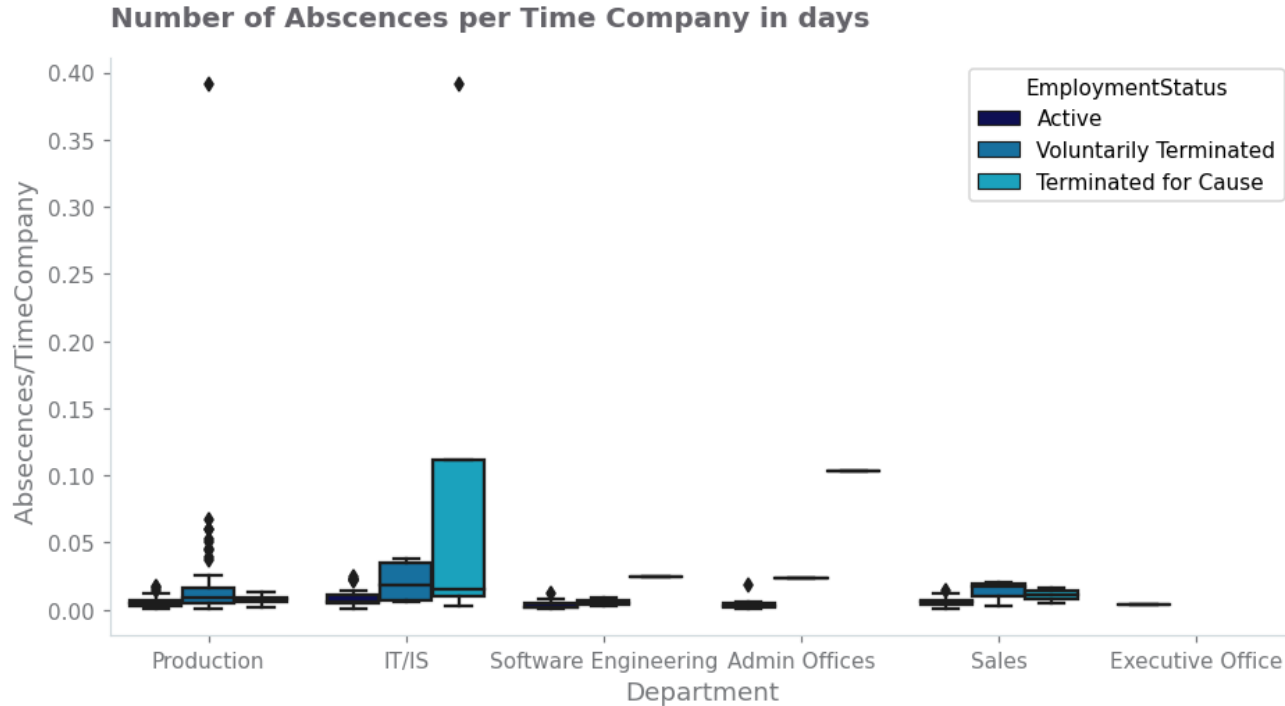
- **PRODUCTION AND SALES:**  
As expected, employees with lowest grade were terminated for cause.
- **SOFTWARE ENGINEERING:**  
Interestingly, voluntarily terminated employees had a better grade.

# Special Projects during the last 6 months



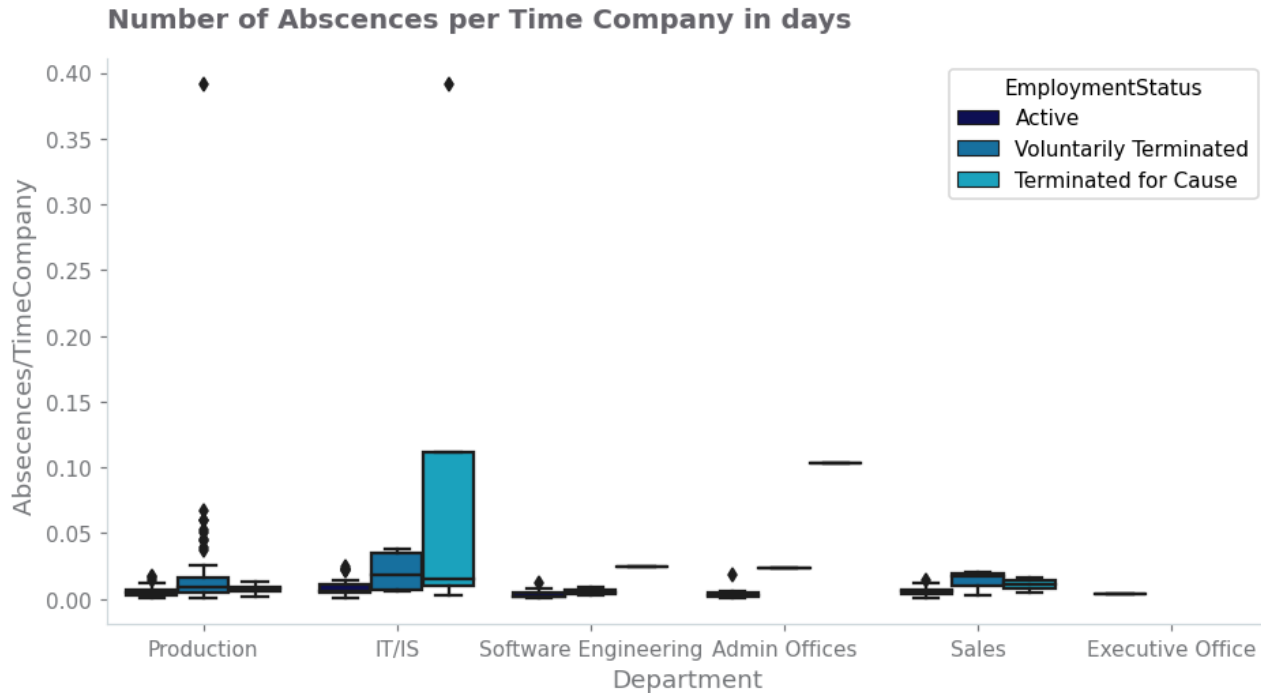
- Look at **IT/IS, SOFTWARE ENGINEERING AND ADMIN OFFICES.**

# Special Projects during the last 6 months



- Look at IT/IS, SOFTWARE ENGINEERING AND ADMIN OFFICES.

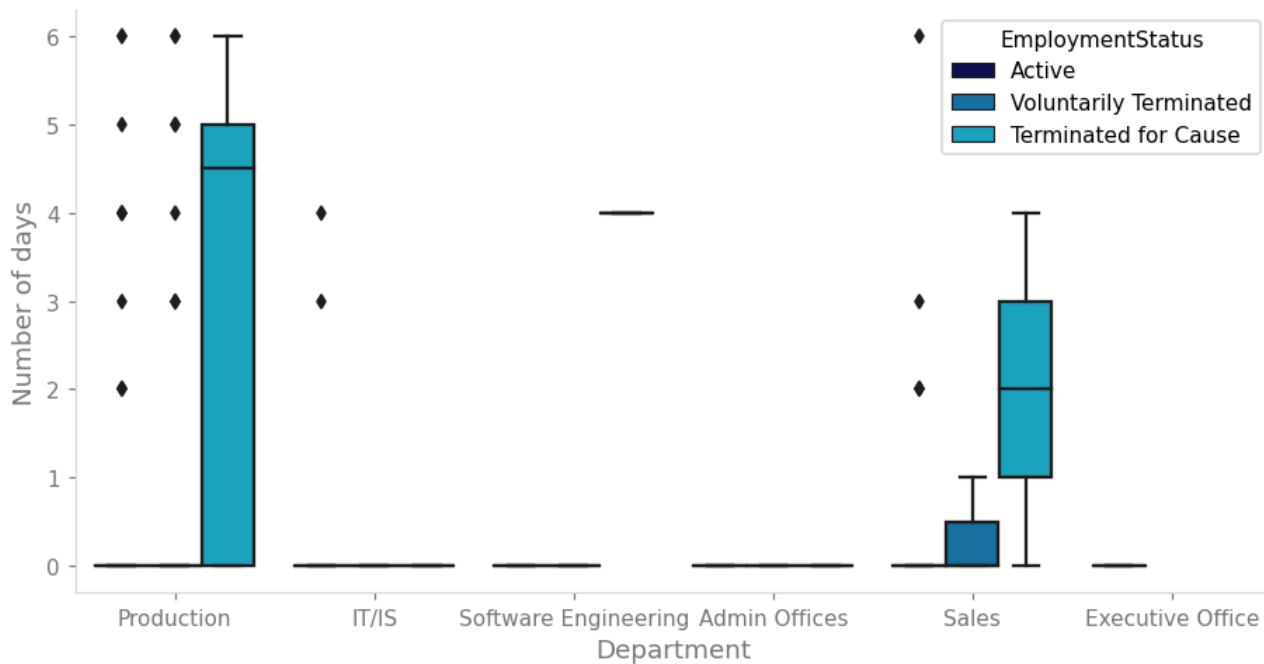
# Absences per Time Company (Days)



- Look at IT/IS - Terminated for Cause.

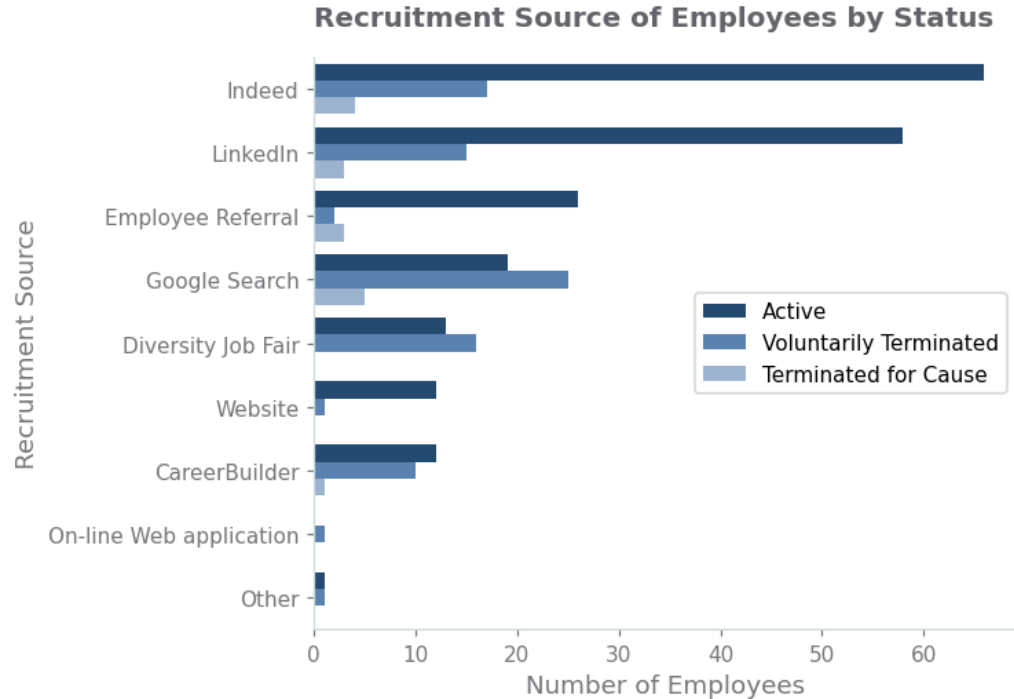
## Number of times that the employee was late

**Number of times that the employee was late to work during the last 30 days**



Terminated for cause  
Production, Software  
Engineering and Sales

# Recruitment Source of Employees



Will the employee continue at the Company?

**Indeed, LinkedIn and Website:**

- **Best** Recruitment Source. The most active employees.

**Google Search and Diversity**

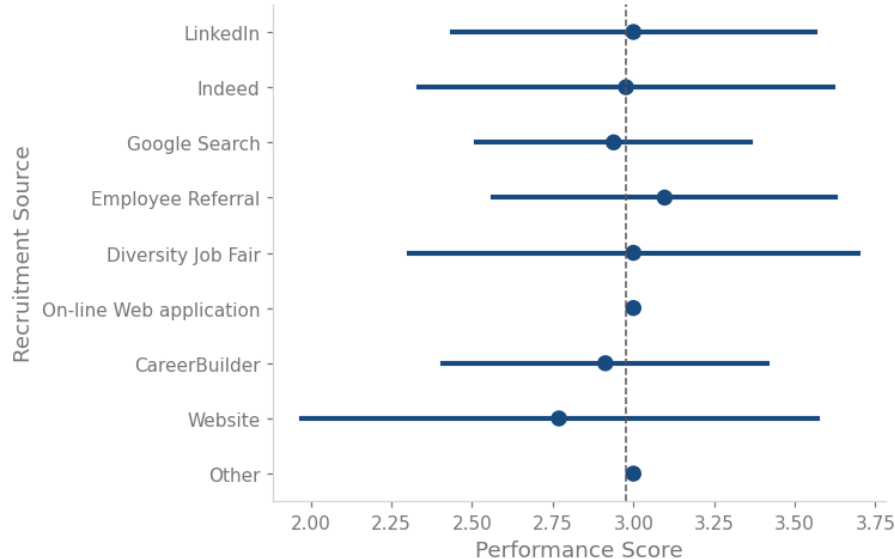
**Job Fair:**

- **Worst** Recruitment Source. Most of Voluntarily Terminated Employees.



# Recruitment Source of Employees

Recruitment Source of Employees by Performance Score



Which of them have the best performance score?

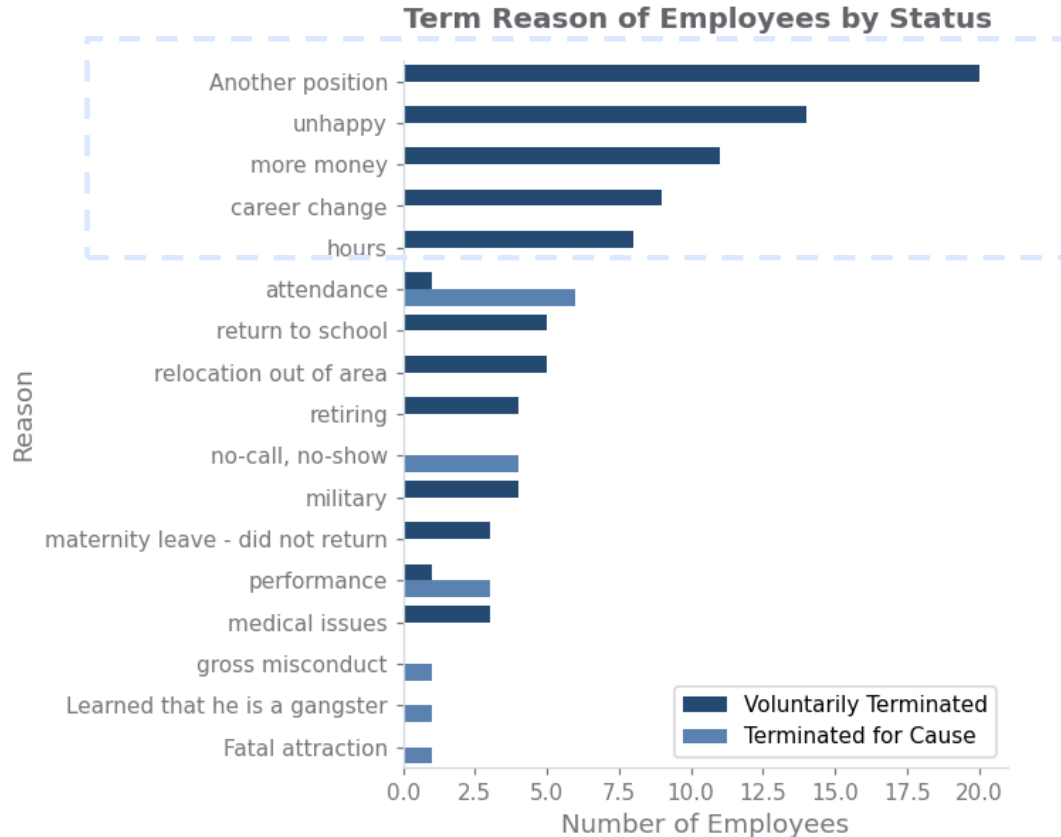
## Employee Referral:

- **Best** Recruitment Source. Highest mean performance score.

## Website:

- **Worst** Recruitment Source. Lowest mean performance score and highest standard deviation.

# Term Reason of Employees



## Voluntarily Terminated:

- **Another position**, **unhappy** and **more money** are the most reasons.

## Terminated for Cause:

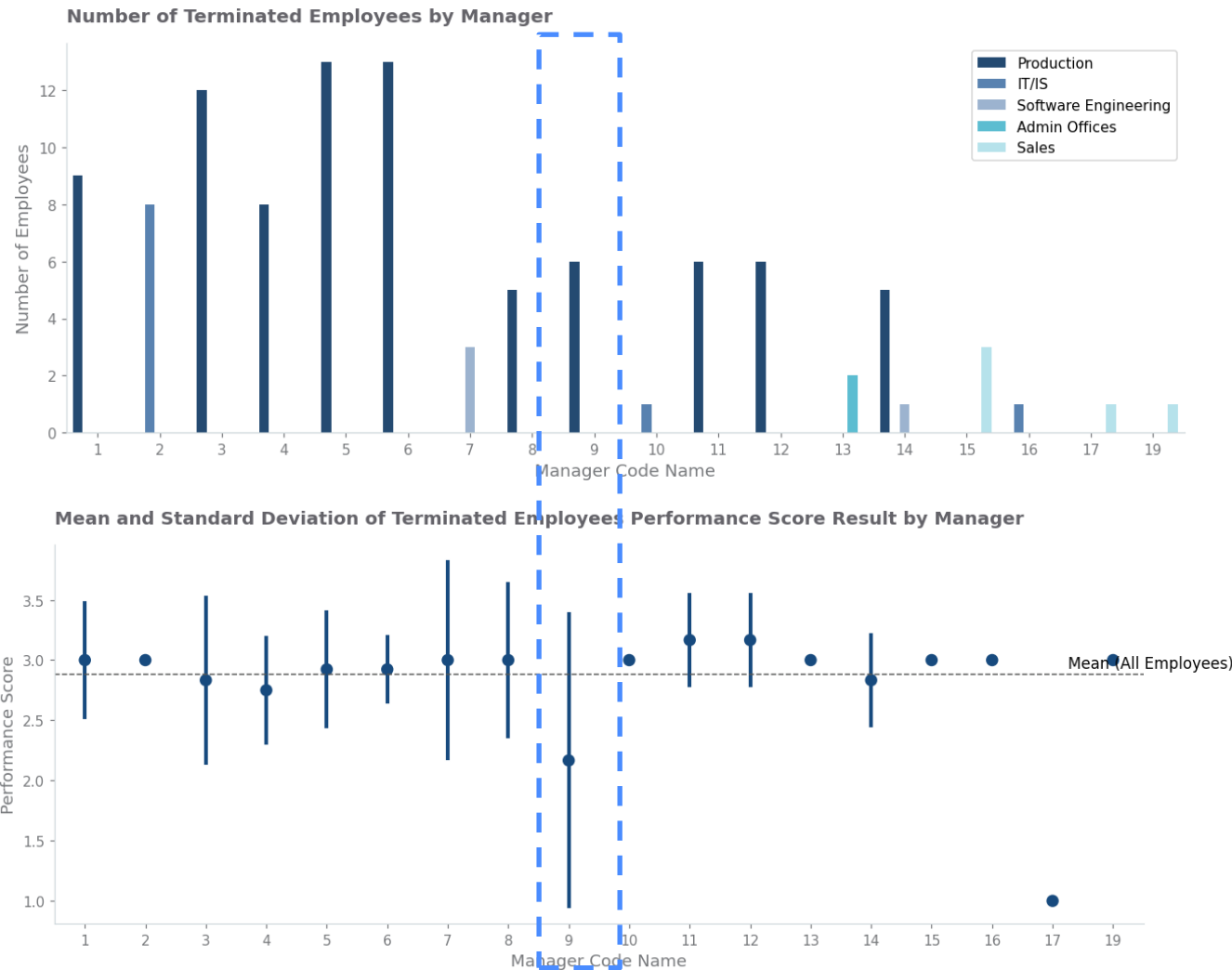
- **Attendance** and **performance** are the most reasons.

# Terminated Employees by managers

## Manager 9:

- Name: Brannon Miller
- Sector: Production

6 demission at Production department and the lowest Mean Performance Score of the terminated employees.

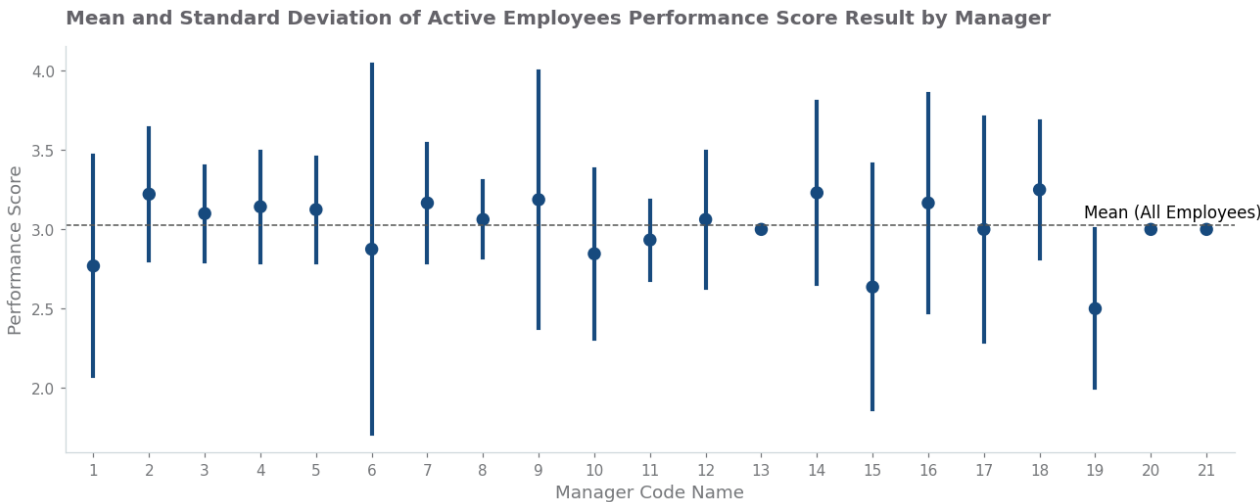
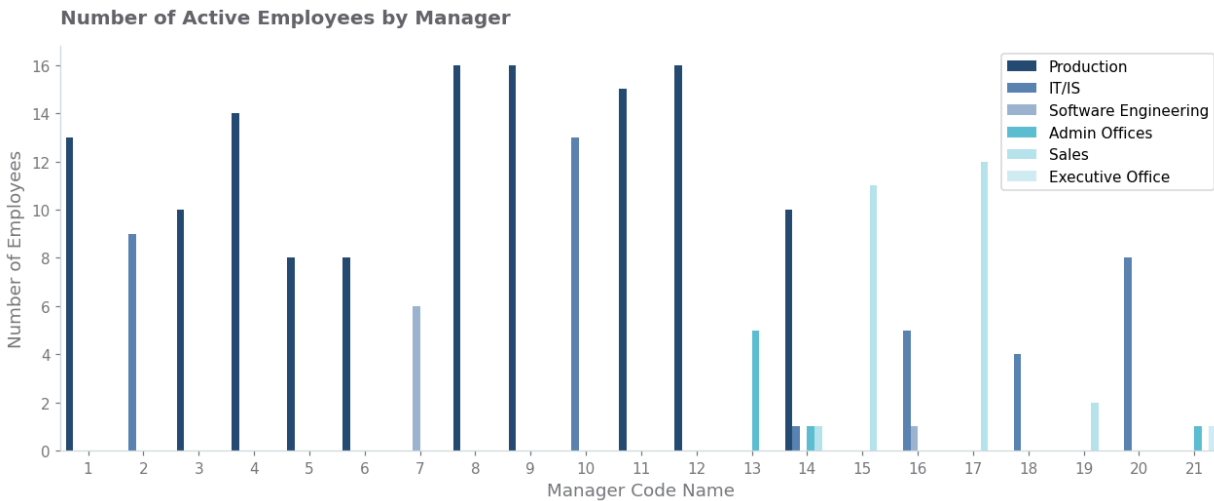


# Active Employees and Managers

Managers 1, 6, 10, 11, 15 and 19:

- Michael Albert, Amy Dunn, Peter Monroe, David Stanley, John Smith, Debra Houlihan.

Mean Performance Score lower that the mean score of all active employees.



# Conclusions

## Company Profile:

The analyzed company is great, with more than **200 active employees**. 89% of them are from **Massachusetts**. The company is divided into the departments: **Production, Software Engineering, Admin Office, IT/IS** and **Sales**. Most of the employees are from production and IT/IS.

## Recruitment Source of Employees :

- Employees remain in the company: Indeed, LinkedIn and Website
- Best mean performance score: Employee Referral

# Recommendations

## Employees and Managers:

- Track and monitoring managers in which their employees had a lower performance average than the general one.

## Employee Referral Program:

- *Employee Referral* is the recruitment source with the best mean performance metric. It is recommended to create programs to encourage that.
- *Indeed* and *LinkedIn* are the best recruitment sources when it is expected to the employee continues at the Company.

# Future Works

## **Classification model to predict if the employee will be terminated or not.**

1. Mapping the variables;
  2. Feature Study to select the most important variables;
  3. Divide in training and test set.
  4. Test different supervised classification models and optimize parameters by applying cross-validation. Logistic regression, Naïve Bayes, SVM are some models that can be tested.
  5. Evaluate the results with metrics, such as ROC, accuracy.
- *It will support the company at the decision of hire or not one new employee.*

**Thank you!**

