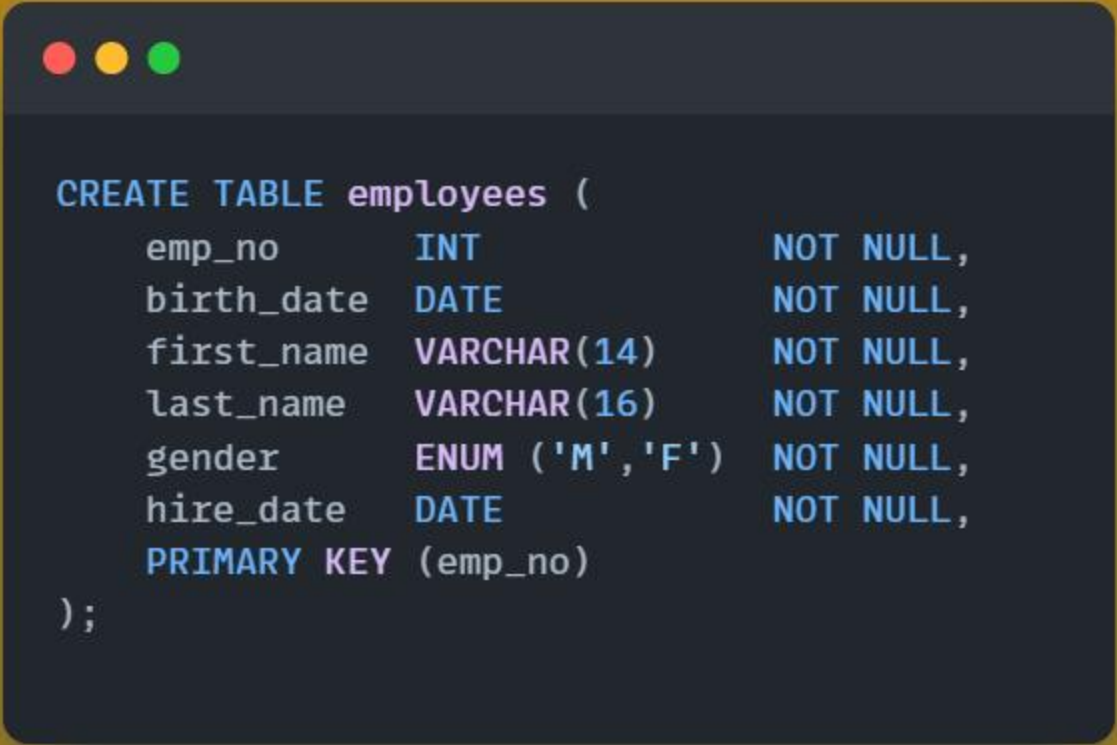
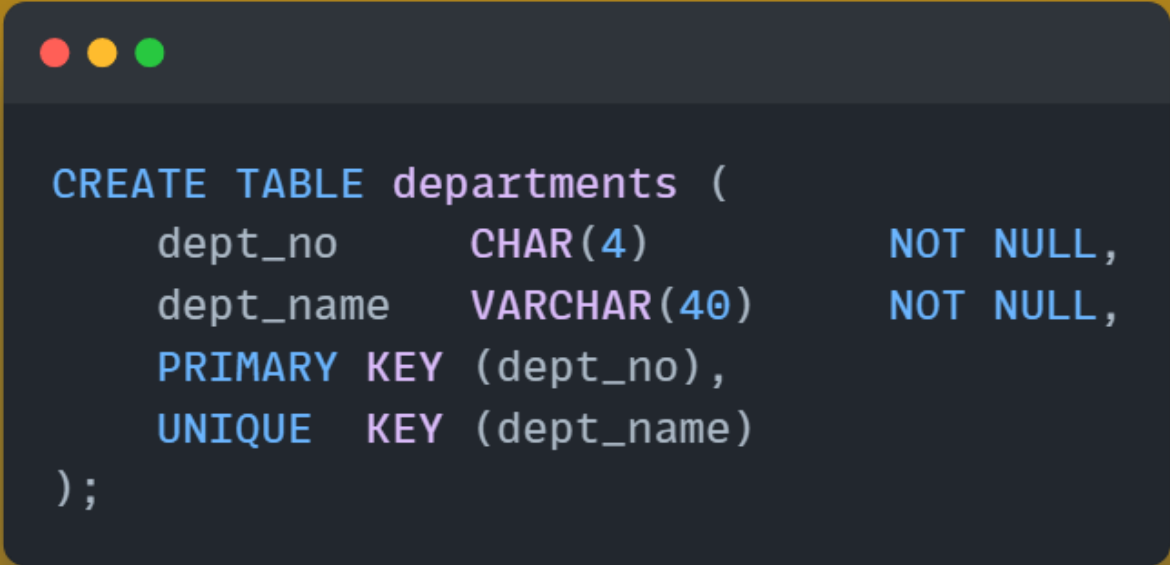


# SQL ANALYSIS using Employee Database



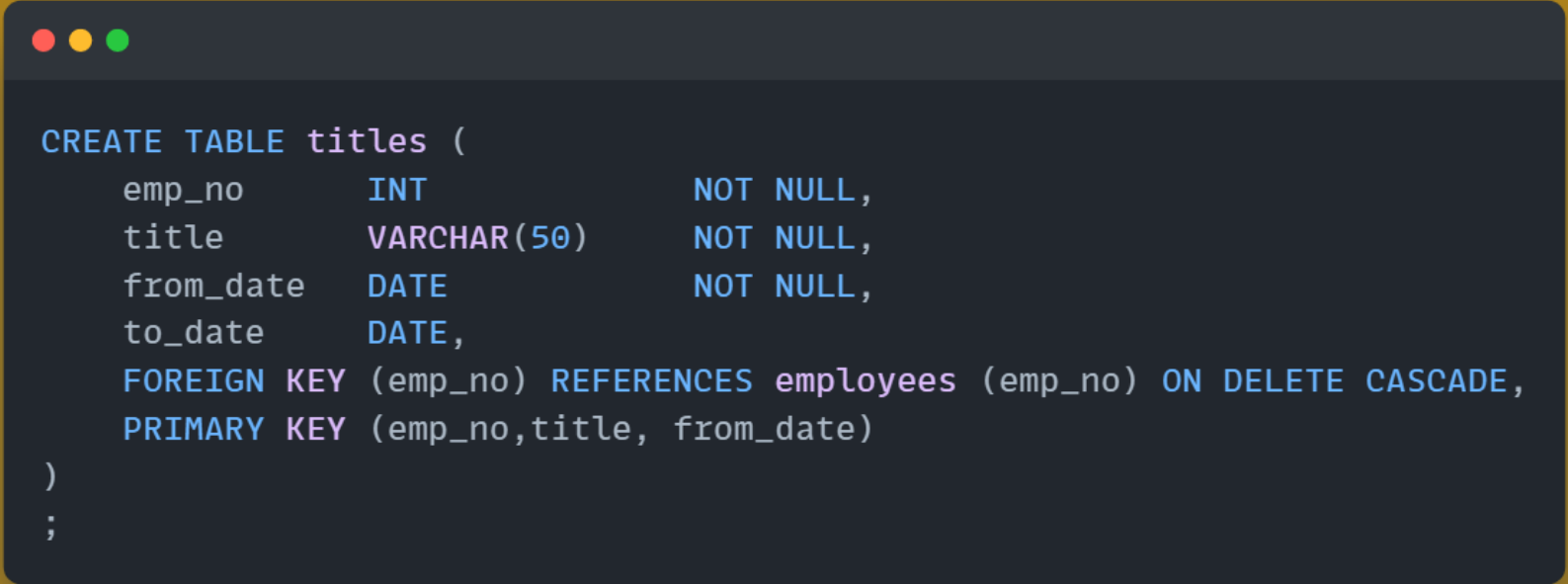
```
CREATE TABLE employees (  
    emp_no      INT          NOT NULL,  
    birth_date  DATE         NOT NULL,  
    first_name  VARCHAR(14)  NOT NULL,  
    last_name   VARCHAR(16)  NOT NULL,  
    gender      ENUM ('M','F') NOT NULL,  
    hire_date   DATE         NOT NULL,  
    PRIMARY KEY (emp_no)  
);
```



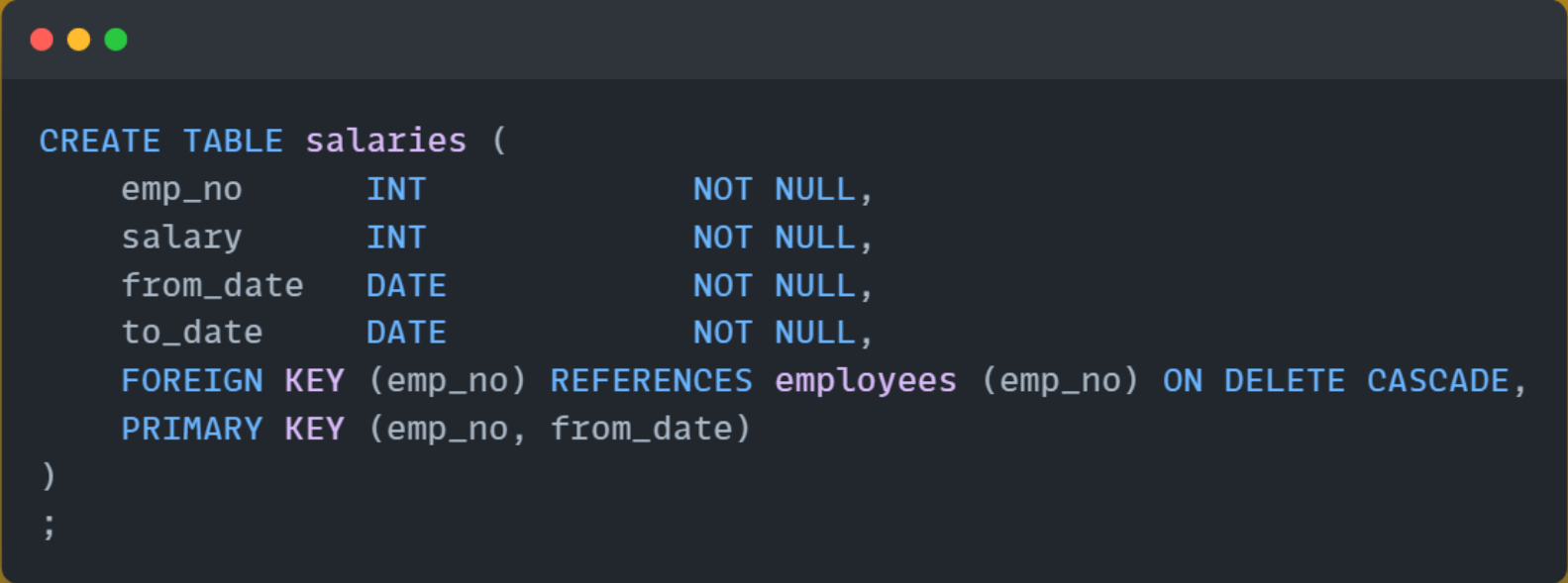
```
CREATE TABLE departments (  
    dept_no      CHAR(4)          NOT NULL,  
    dept_name    VARCHAR(40)     NOT NULL,  
    PRIMARY KEY (dept_no),  
    UNIQUE KEY (dept_name)  
);
```

```
CREATE TABLE dept_manager (  
    emp_no      INT          NOT NULL,  
    dept_no     CHAR(4)      NOT NULL,  
    from_date   DATE         NOT NULL,  
    to_date     DATE         NOT NULL,  
    FOREIGN KEY (emp_no) REFERENCES employees (emp_no) ON DELETE CASCADE,  
    FOREIGN KEY (dept_no) REFERENCES departments (dept_no) ON DELETE CASCADE,  
    PRIMARY KEY (emp_no,dept_no)  
);
```

```
CREATE TABLE dept_emp (  
    emp_no      INT          NOT NULL,  
    dept_no     CHAR(4)      NOT NULL,  
    from_date   DATE         NOT NULL,  
    to_date     DATE         NOT NULL,  
    FOREIGN KEY (emp_no) REFERENCES employees (emp_no) ON DELETE CASCADE,  
    FOREIGN KEY (dept_no) REFERENCES departments (dept_no) ON DELETE CASCADE,  
    PRIMARY KEY (emp_no,dept_no)  
);
```



```
CREATE TABLE titles (  
    emp_no      INT          NOT NULL,  
    title       VARCHAR(50)  NOT NULL,  
    from_date   DATE         NOT NULL,  
    to_date     DATE,  
    FOREIGN KEY (emp_no) REFERENCES employees (emp_no) ON DELETE CASCADE,  
    PRIMARY KEY (emp_no,title, from_date)  
);
```



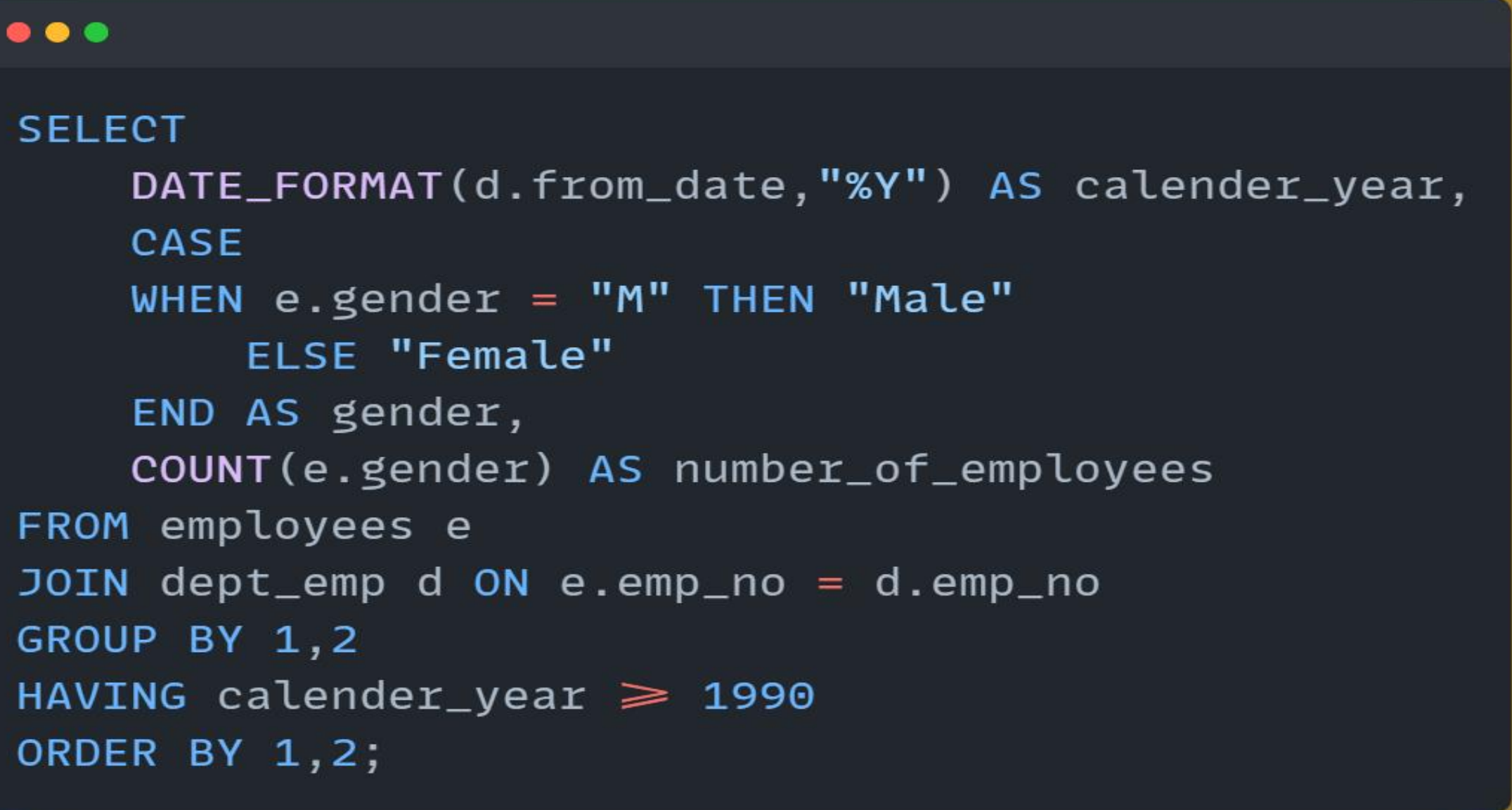
```
CREATE TABLE salaries (  
    emp_no      INT          NOT NULL,  
    salary      INT          NOT NULL,  
    from_date   DATE         NOT NULL,  
    to_date     DATE         NOT NULL,  
    FOREIGN KEY (emp_no) REFERENCES employees (emp_no) ON DELETE CASCADE,  
    PRIMARY KEY (emp_no, from_date)  
);
```

**TASK: 1**





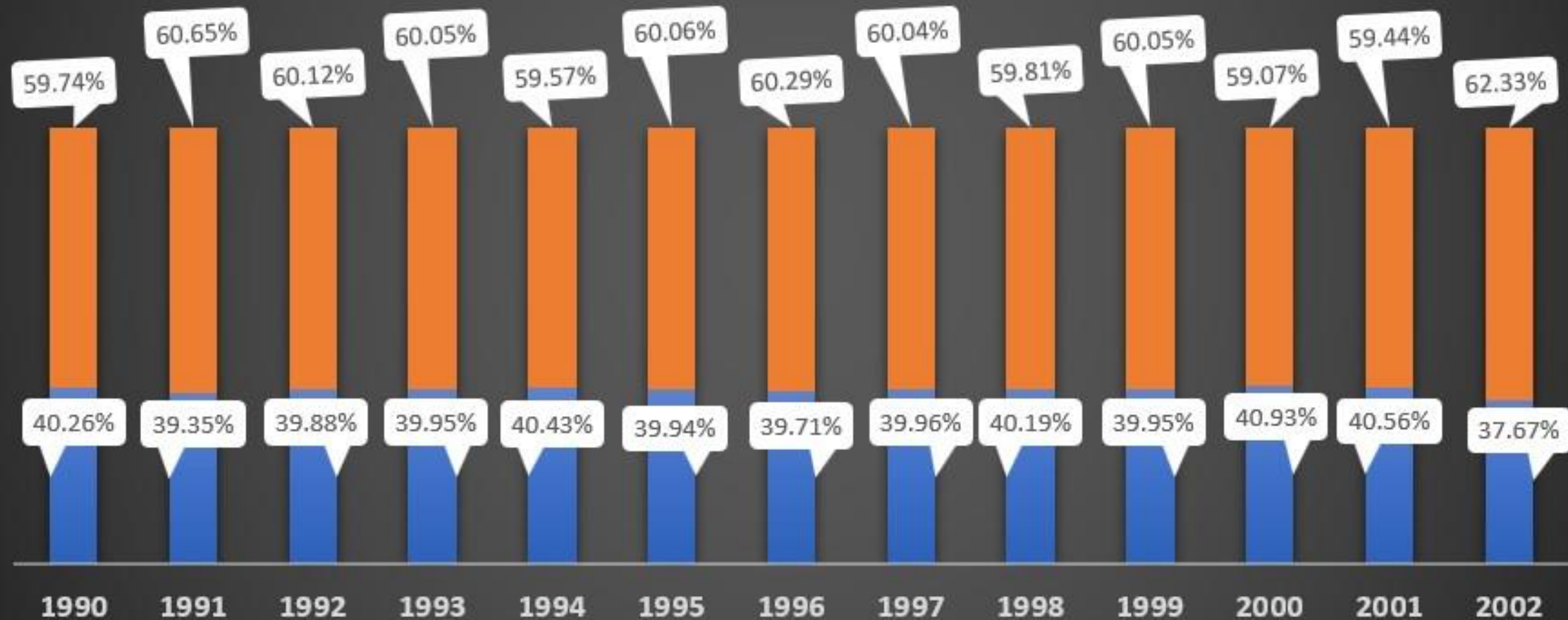
Create a visualization that provides a breakdown between the male and female employees working **in** the company each year, starting from 1990.



```
SELECT
    DATE_FORMAT(d.from_date,"%Y") AS calender_year,
    CASE
        WHEN e.gender = "M" THEN "Male"
        ELSE "Female"
    END AS gender,
    COUNT(e.gender) AS number_of_employees
FROM employees e
JOIN dept_emp d ON e.emp_no = d.emp_no
GROUP BY 1,2
HAVING calender_year ≥ 1990
ORDER BY 1,2;
```

## Gender Count Breakdown By Year

Female Male



1. The number of male employees has always remained  $>$  the number of female employees during the years 1990-2001.
2. During the years 1990-2001, there has been an increase in the number of male employees with  
  
**minimum value: 18.13 %**  
**maximum value: 21.30 %**
3. The year 2002 has seen the highest difference with a value of 24.65%.

# TASK: 2



Compare the number of  
male managers to the number of female managers  
from different departments for each year,  
starting from 1990

```
SELECT dept_emp.dept_name, dept_emp.gender,dept_emp.emp_no,dept_emp.from_date,dept_emp.to_date,e.calendar_year,
CASE
    WHEN YEAR(dept_emp.from_date) ≤ e.calendar_year AND YEAR(dept_emp.to_date) ≥ e.calendar_year THEN 1
    ELSE 0
END AS active_as_manager
FROM
    (SELECT YEAR(e.hire_date) AS calendar_year
    FROM employees e
    GROUP BY calendar_year
    ORDER BY 1) e
CROSS JOIN
    (SELECT d.dept_name, ee.gender,dm.emp_no,dm.from_date,dm.to_date
    FROM dept_emp de
    JOIN departments d ON de.dept_no = d.dept_no
    JOIN employees ee ON ee.emp_no = de.emp_no
    JOIN dept_manager dm ON dm.emp_no = ee.emp_no) dept_emp
ORDER BY dept_emp.emp_no, e.calendar_year;
```

## No. of Active Managers By Gender

Female Male

10  
9  
8  
7  
6  
5  
4  
3  
2  
1  
0

	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Female	4	4	4	6	6	5	7	9	6	6	5	5	4	4	4	4
Male	5	5	5	5	5	4	6	4	3	4	4	6	5	5	5	5




Year	Manager	Percentage Difference
1985-1987	No. of male managers > No. of female managers	1.76%
1988-1995	No. of female managers > No. of male managers	Highest Difference: 5.58 % Lowest Difference: 0.54%
1996-2000	No. of male managers > No. of female managers	1.76%

# TASK: 3



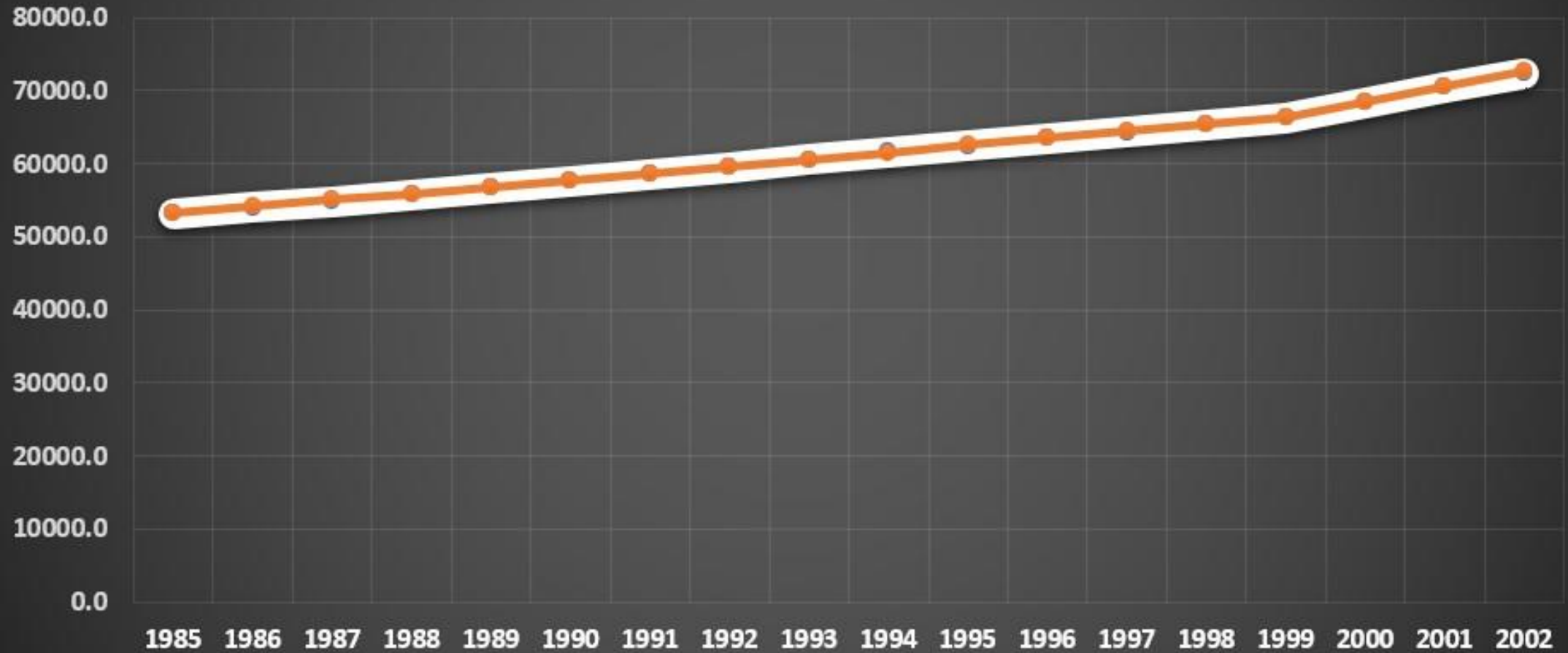
Compare the average salary of  
female employees versus male employees in  
the entire company until year 2002.



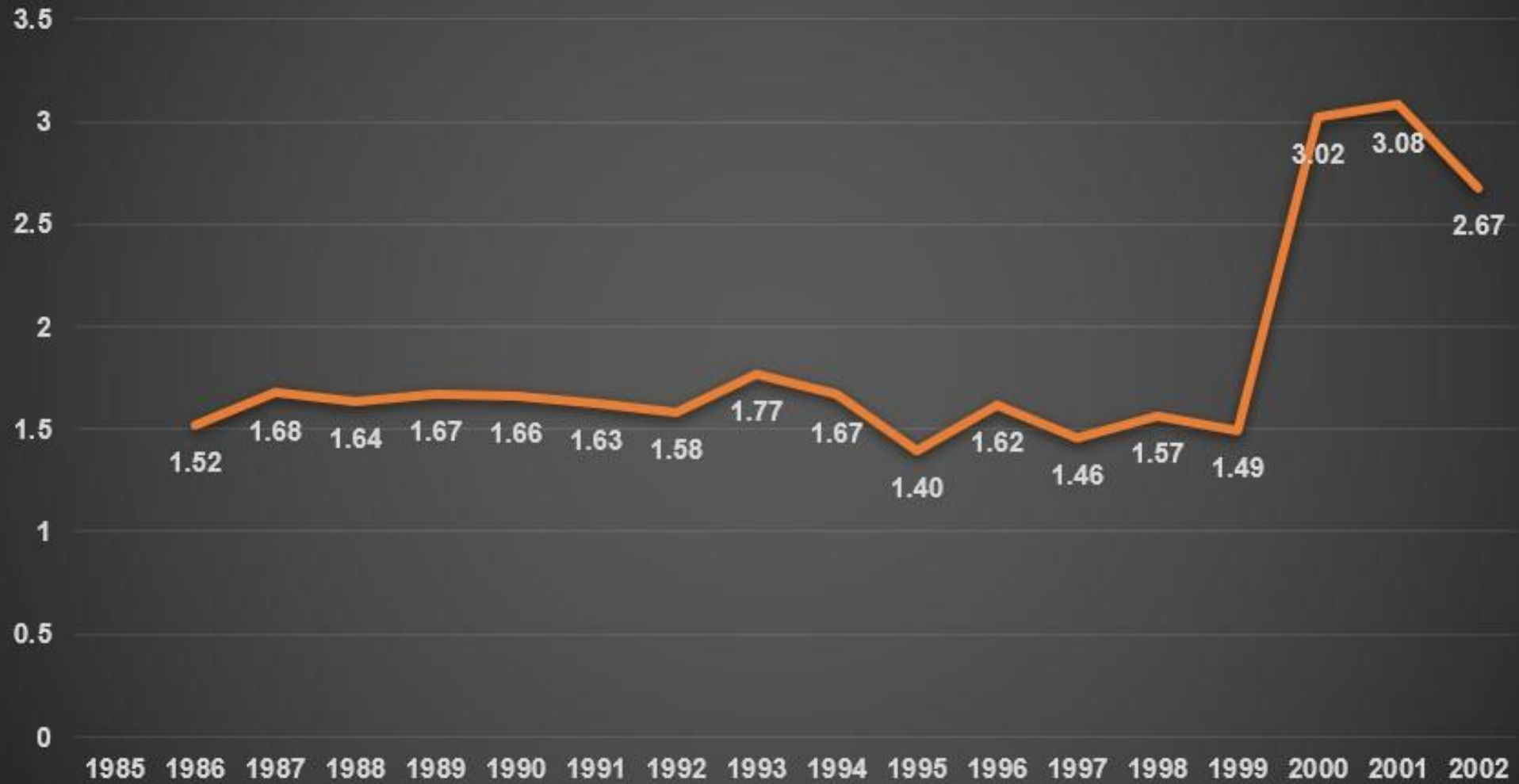
```
SELECT e.gender, d.dept_name, ROUND(AVG(s.salary),2) AS salary, YEAR(s.from_date) AS calendar_year
FROM employees e
JOIN dept_emp de ON de.emp_no = e.emp_no
JOIN departments d ON d.dept_no = de.dept_no
JOIN salaries s ON s.emp_no = e.emp_no
GROUP BY d.dept_no, e.gender, calendar_year
HAVING calendar_year ≤ 2002
ORDER BY d.dept_no;
```

## Average Salary of Gender By Year

F M



## % Increase in Average Salary by Year (Females)



## % Increase in Average Salary by Year(Males)



The average salary has always increased during the years 1985-2002 irrespective of gender.

## FEMALE

Year	Average Salary
1985-1999	Increased by 2% approximately. Minimum Increase Value : 1.40 Maximum Increase Value : 1.77
2000	Sudden increase with 3.02 %
2001	Constant with 3.08%
2002	Declined with 2.67%

## MALE

Year	Average Salary
1985-1999	Increased by 2% approximately. Minimum Increase Value : 1.31 Maximum Increase Value : 1.83
2000	Sudden increase with 3.05 %
2001	Constant with 3.11%
2002	Declined with 2.87%



# TASK: 4



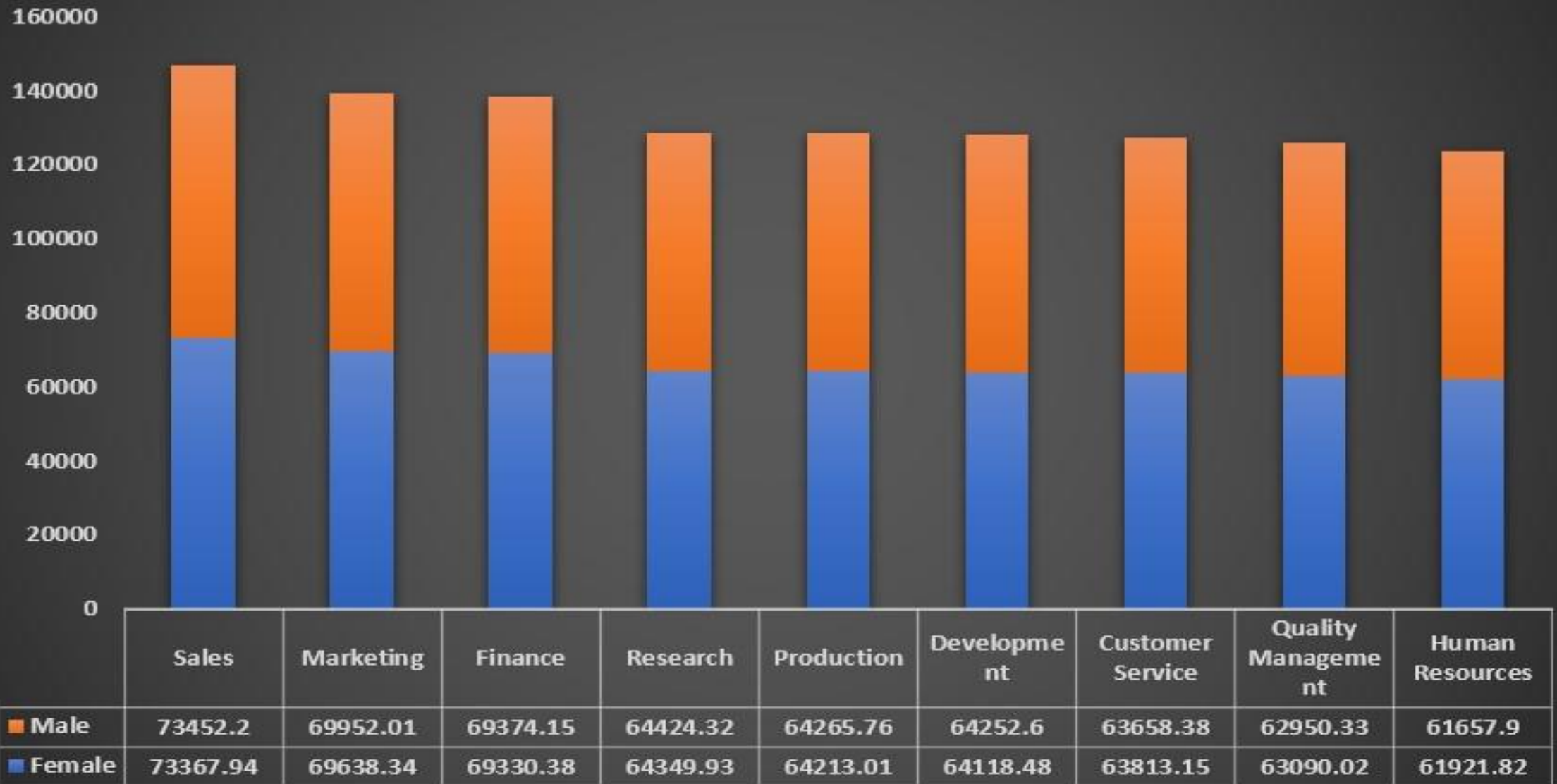
Create a `SQL` stored procedure that will allow you to obtain the average male and female salary per department within a certain salary range.

Let `this` range be defined by two values the user can insert when calling the procedure

```
DELIMITER $$
CREATE PROCEDURE getAverageSalaryMaleVsFemale(IN p_min_salary FLOAT, IN p_max_salary FLOAT)
BEGIN
    SELECT e.gender, d.dept_name, ROUND(AVG(s.salary),2) AS avg_salary
    FROM employees e
    JOIN dept_emp de ON de.emp_no = e.emp_no
    JOIN departments d ON d.dept_no = de.dept_no
    JOIN salaries s ON s.emp_no = e.emp_no
    WHERE s.salary BETWEEN p_min_salary AND p_max_salary
    GROUP BY d.dept_no, e.gender;
END $$
DELIMITER ;

CALL getAverageSalaryMaleVsFemale(50000, 90000);
```

## Average Salary of Employees By Department



1. People employed in the "**Sales**" department earned the highest followed by "**Marketing**" and "**Finance**".

Sales : 73410.07

Marketing : 69795.175

Finance : 69352.265

2. There is a very thin line between males' and females' average salaries with a maximum difference of value 300.