**PROJECT REPORT**

**Project Topic: SQL and Databases**

**Challenge Statement:**

Given the looming economic recession in the current year, Epoch systems, a Winnipeg based publicly traded company, which specializes in the manufacture of drones has decided to re-evaluate its entire cost structure in a bid to maintain the current level of its share value and meet investor expectations.

After multiple meetings with various stakeholders, management decided that it will embark on a cost cutting exercise shifting focus to become more operationally efficient. Management has decided that a large component of the cost cutting will have to come from layoffs of its workforce.

As a data analyst within Epochs data office - a team under the Development department, you have been instructed by your manager to investigate Epoch’s employee database and retrieve certain information that will enable management to take a more data driven approach to decision making.

The employee database consists of six tables in total. These include the following:

1. Departments table: consists of a department number and department name.
2. Departments-employee table consists of the employee number, department number of employee and the start and end dates of the employees’ time at the department.
3. Departments-manager table consists of the employee number, department number of employee and the start and end dates of the employees’ time at the department for all managers.
4. Employee table: primary employee table which consists of employee number, birth date, first name, last name, gender, and hire date.
5. Salaries table: contains employee number, annual salary, from date and to date.
6. Title table: contains employee number, employee role/title, start date and end date of role.

All current employees have a to\_date column value of ‘9999-01-01’ to indicate that the termination date is unknown.

Based on management discussions, the questions below will enable Epoch’s management to make better decisions about where some trimming of work force can be done.

What is the total head count of all current employees at Epoch systems?

What is the distribution of head count for current employees across the various departments?

What is the average salary of all current employees by department? What is the max salary of all current employees by department?

What is the count of current employees who are aged 70 and above distributed by departments?

What is the average salary for employees grouped by titles?

What is the head count of current employees grouped first by department and then by titles?

Your task is to retrieve the data from the database that answers the questions above. Provide the results in a short report.

**Solutions:**

1. **Total head count of all current employees at Epoch systems**

**Query Used to Extract Data:**

"SELECT COUNT (emp\_no) AS Total\_emplyees FROM dept\_emp WHERE DATE(to\_date) = '9999-01-01'"

**Resulting Table:**

|  |
| --- |
| Total\_emplyees |
| 240124 |

**Graphical Representation of the Result:**

1. **The Distribution of Head Count for Current Employees Across the Various Departments**

**Query Used:**

"SELECT departments.dept\_name as department, COUNT(\*) as count FROM dept\_emp INNER JOIN departments ON dept\_emp.dept\_no = departments.dept\_no WHERE DATE(to\_date) = '9999-01-01' GROUP BY dept\_emp.dept\_no ORDER BY count DESC"

**Resulting Table:**

|  |  |
| --- | --- |
| department | count |
| Development | 61386 |
| Production | 53304 |
| Sales | 37701 |
| Customer Service | 17569 |
| Research | 15441 |
| Marketing | 14842 |
| Quality Management | 14546 |
| Human Resources | 12898 |
| Finance | 12437 |

**Graphical Representation of the Result:**

Chart, bar chart

Description automatically generated

1. **Average Salary of All Current Employees by Department.**

**Query Used to Extract Data:**

"SELECT dept\_name, ROUND(AVG(total\_salary),2) as Average\_salary FROM (SELECT departments.dept\_name, salaries.salary AS total\_salary FROM salaries INNER JOIN dept\_emp ON salaries.emp\_no = dept\_emp.emp\_no INNER JOIN departments ON dept\_emp.dept\_no = departments.dept\_no WHERE DATE(salaries.to\_date) = '9999-01-01' AND DATE(dept\_emp.to\_date) = '9999-01-01') GROUP BY dept\_name ORDER BY Average\_salary DESC"

**Resulting Table:**

|  |  |
| --- | --- |
| dept\_name | Average\_salary |
| Sales | 88852.97 |
| Marketing | 80058.85 |
| Finance | 78559.94 |
| Research | 67913.37 |
| Production | 67843.3 |
| Development | 67657.92 |
| Customer Service | 67285.23 |
| Quality Management | 65441.99 |
| Human Resources | 63921.9 |

**Graphical Representation of the Result:**

Chart, bar chart

Description automatically generated

1. **Maximum Salary of All Current Employees by Department**

**Query Used to Extract Data:**

"SELECT dept\_name, MAX(total\_salary) as Maximum\_salary FROM (SELECT departments.dept\_name, salaries.salary AS total\_salary FROM salaries INNER JOIN dept\_emp ON salaries.emp\_no = dept\_emp.emp\_no INNER JOIN departments ON dept\_emp.dept\_no = departments.dept\_no WHERE DATE(dept\_emp.to\_date) = '9999-01-01') GROUP BY dept\_name ORDER BY Maximum\_salary DESC"

**Resulting Table:**

|  |  |
| --- | --- |
| dept\_name | Maximum\_salary |
| Sales | 158220 |
| Marketing | 145128 |
| Customer Service | 144866 |
| Development | 144434 |
| Finance | 142395 |
| Human Resources | 141953 |
| Production | 138273 |
| Quality Management | 132103 |
| Research | 130211 |

**Graphical Representation of the Result:**

Chart, bar chart

Description automatically generated

1. **Count of Current Employees Who Are Aged 70 and Above Distributed by Departments**

**Query Used to Extract Data:**

"SELECT dept\_name, COUNT(DISTINCT emp\_no) AS Emp\_70\_And\_Above FROM (SELECT DISTINCT departments.dept\_name,employees.emp\_no FROM employees INNER JOIN dept\_emp ON employees.emp\_no = dept\_emp.emp\_no INNER JOIN departments ON dept\_emp.dept\_no = departments.dept\_no WHERE (julianday('now') - julianday(employees.birth\_date)) / 365.25 >=70 AND dept\_emp.to\_date = '9999-01-01') GROUP BY dept\_name ORDER BY Emp\_70\_And\_Above DESC”

**Resulting Table:**

|  |  |
| --- | --- |
| dept\_name | Emp\_70\_And\_Above |
| Development | 5047 |
| Production | 4530 |
| Sales | 3160 |
| Customer Service | 1439 |
| Research | 1332 |
| Quality Management | 1247 |
| Marketing | 1194 |
| Human Resources | 1137 |
| Finance | 1008 |

**Graphical Representation of the Result:**

Chart, bar chart

Description automatically generated

1. **Average Salary of Current Employees Grouped by Titles.**

**Query Used to Extract Data:**

“SELECT title, ROUND(AVG(total\_salary),2) as Average\_salary FROM(SELECT titles.title, salaries.salary AS total\_salary FROM salaries INNER JOIN titles ON salaries.emp\_no = titles.emp\_no WHERE salaries.to\_date = '9999-01-01' AND titles.to\_date='9999-01-01') GROUP BY title ORDER BY Average\_salary DESC"

**Resulting Table:**

|  |  |
| --- | --- |
| title | Average\_salary |
| Senior Staff | 80706.5 |
| Manager | 77723.67 |
| Senior Engineer | 70823.44 |
| Technique Leader | 67506.59 |
| Staff | 67330.67 |
| Engineer | 59602.74 |
| Assistant Engineer | 57317.57 |

**Graphical Representation of the Result:**

Chart, bar chart

Description automatically generated

1. **The Head Count of Current Employees Grouped First by Department and Then by Titles.**

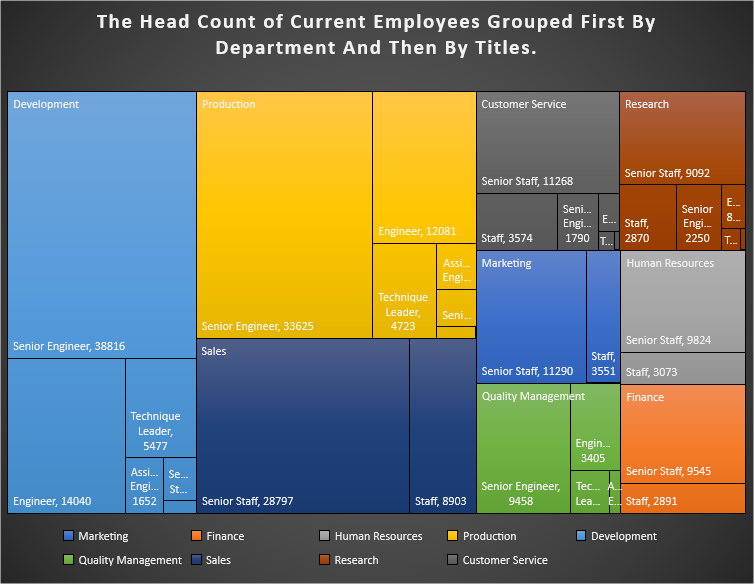
**Query Used to Extract Data:**

"SELECT departments.dept\_name as department, titles.title AS Title, COUNT(\*) AS count FROM dept\_emp INNER JOIN departments ON dept\_emp.dept\_no = departments.dept\_no INNER JOIN TITLES ON titles.emp\_no=dept\_emp.emp\_no WHERE DATE(dept\_emp.to\_date) = '9999-01-01' and DATE(titles.to\_date)= '9999-01-01' GROUP BY dept\_emp.dept\_no , titles.title"

**Resulting Table:**

|  |  |  |
| --- | --- | --- |
| department | Title | count |
| Marketing | Manager | 1 |
| Marketing | Senior Staff | 11290 |
| Marketing | Staff | 3551 |
| Finance | Manager | 1 |
| Finance | Senior Staff | 9545 |
| Finance | Staff | 2891 |
| Human Resources | Manager | 1 |
| Human Resources | Senior Staff | 9824 |
| Human Resources | Staff | 3073 |
| Production | Assistant Engineer | 1402 |
| Production | Engineer | 12081 |
| Production | Manager | 1 |
| Production | Senior Engineer | 33625 |
| Production | Senior Staff | 1123 |
| Production | Staff | 349 |
| Production | Technique Leader | 4723 |
| Development | Assistant Engineer | 1652 |
| Development | Engineer | 14040 |
| Development | Manager | 1 |
| Development | Senior Engineer | 38816 |
| Development | Senior Staff | 1085 |
| Development | Staff | 315 |
| Development | Technique Leader | 5477 |
| Quality Management | Assistant Engineer | 389 |
| Quality Management | Engineer | 3405 |
| Quality Management | Manager | 1 |
| Quality Management | Senior Engineer | 9458 |
| Quality Management | Technique Leader | 1293 |
| Sales | Manager | 1 |
| Sales | Senior Staff | 28797 |
| Sales | Staff | 8903 |
| Research | Assistant Engineer | 77 |
| Research | Engineer | 830 |
| Research | Manager | 1 |
| Research | Senior Engineer | 2250 |
| Research | Senior Staff | 9092 |
| Research | Staff | 2870 |
| Research | Technique Leader | 321 |
| Customer Service | Assistant Engineer | 68 |
| Customer Service | Engineer | 627 |
| Customer Service | Manager | 1 |
| Customer Service | Senior Engineer | 1790 |
| Customer Service | Senior Staff | 11268 |
| Customer Service | Staff | 3574 |
| Customer Service | Technique Leader | 241 |

**Graphical Representation of the Result:**



**Summary:**

The data and bar graphs present the following information about Epoch Systems:

The total number of employees in Epoch Systems is 240,124. Among the departments, the Development department has the highest number of employees with 61,386, while the Finance department has the lowest with 12,437.

In terms of average salary, the Sales department stands out with the highest average salary for current employees. On the other hand, the Human Resources department has the lowest average salary. The Sales department also has the highest maximum salary compared to other departments, while the differences in salary levels among the other departments are not significant.

When considering the age distribution, the Development department has the highest number of employees aged 70 and above. Conversely, the Finance department has the lowest number of employees in this age group.

In terms of job titles, Senior Staff members receive the highest salaries, while Assistant Engineers have the lowest salaries.

Within the Development department, there are 38,816 Senior Engineers and 14,040 Engineers. The Sales department has a count of 28,797 Senior Staff members.

These findings provide an overview of the employee data and highlight certain patterns and variations across different departments within Epoch Systems.

**Recommendations:**

The company, Epoch systems, aims to enhance operational efficiency and reduce costs through workforce layoffs. To achieve this, they can consider the following recommendations:

* Assess departments with high employee numbers, such as Development, Production, and Sales, which have employee counts of 61,386, 53,304, and 37,701 respectively. Explore the possibility of eliminating positions or combining similar roles to streamline operations.
* Evaluate departments with higher salary averages. For instance, Sales has an average salary of $88,852.97, while Marketing and Finance have averages of $80,058.85 and $78,559.94 respectively. Determine if these salary levels align with department performance.
* Assess the relevance and necessity of departments such as Marketing, Human Resources, and Quality Management, which have employee counts of 14,842, 12,898, and 14,546 respectively. Consider the possibility of restructuring positions or reallocating responsibilities within these departments.
* Identify employees nearing retirement age, particularly those aged 70 and above, in departments like Sales, Marketing, and Customer Service. For example, Sales has employees earning the maximum salary of $158,220 in this age group. Offer early retirement options to reduce costs.
* Evaluate the number of managers in each department. For instance, Marketing, Finance, and Human Resources each have one Manager. Assess if the current number of managers is necessary or if some managerial positions can be eliminated or combined.
* Analyze the average salaries across different job titles. For example, Senior Staff has an average salary of $80,706.50, while Managers have an average of $77,723.67. Compare these salaries to industry standards to identify areas where adjustments can be made.
* Assess employees who have multiple roles or work across different departments. In Production, for instance, there are employees with titles like Senior Engineer and Technique Leader who also have responsibilities in other departments. Determine if these cross-departmental roles can be streamlined to reduce costs.
* Consider employee performance and productivity across departments. Identify underperforming individuals or teams and assess their impact on the company's goals. Use this evaluation to determine areas where layoffs may be necessary to improve efficiency and save costs.

There is a Python file included in the project. This file incorporates SQL queries along with the pandas and matplotlib libraries for data representation and performing different operations. By running the file, you can observe the charts generated and analyze the data.