RELATIONAL DATABASES
END OF COURSE PROJECT

BAKERS' DATABASE ANALYSIS FOR HR

Group N°2: Yohan Rossi – Wang Jia – Enhe Dong – Emilien Chollet - Murugesh Manthiramoorthi

ASSUMPTIONS MADE :

- The company founded in 1980s that had been running successfully so far is facing some difficulties in the recent years due to tough competition in the software market.
- The newly appointed HR seek our help to get him/her understand the current company situation.
- We assume the date of analysis as 2002-08-01 because of the following reason. because we notice that in the salaries table, the salaries were recorded by years based on each employees' hire date. But the "to-date" becomes "9999-01-01" since 2001-08-02 though the "from-date" shows still the normal date. We think that these employees with "9999-01-01" are still working for the company. Based on this logic, we assume that the company's employees' demission system update once a year and the update date is on august 2nd. Considering the maximum date in the table on august first 2002, we assume that today is 2002-08-01 because the next day is the date that the system will update "to-date" and the result won't show "9999-01-01" anymore.

We use Tableau Desktop to visualize the results of our queries.

```sql
USE employees;

-- 1. Dealing with to_date issue (9999 as year)

-- 1.a Update the tables to_date column based on the pattern of the table salaries

UPDATE salaries
SET to_date = '2002-08-01'
WHERE to_date = '9999-01-01';

UPDATE dept_emp
SET to_date = '2002-08-01'
WHERE to_date = '9999-01-01';

UPDATE titles
SET to_date = '2002-08-01'
WHERE to_date = '9999-01-01';
```

#QUERY 1

Query:

```
--List of employees with the following information: First name, last name, gender,
age, number of years spent in the company, department, the year they joined the
department, and their current job title.

SELECT
employees.first_name,
employees.last_name,
employees.gender,
DATE_FORMAT (FROM_DAYS(TO_DAYS('2003-08-02')-TO_DAYS(employees.birth_date)), '%y')
AS "age",
DATE_FORMAT (FROM_DAYS(TO_DAYS(MAX(salaries.to_date))-
TO_DAYS(employees.hire_date)), '%y') AS "experience",
departments.dept_name AS 'department',
dept_emp.from_date AS 'dept_join_date',
titles.title AS 'job_title',
MAX(titles.to_date) AS 'title.to_date',
MAX(dept_emp.to_date) AS 'dept.to_date'
FROM employees
JOIN salaries
ON salaries.emp_no = employees.emp_no
JOIN dept_emp
ON dept_emp.emp_no = employees.emp_no
JOIN departments
ON dept_emp.dept_no = departments.dept_no
JOIN titles
ON employees.emp_no = titles.emp_no
GROUP BY employees.emp_no, employees.first_name, employees.last_name,
employees.gender, department, job_title
ORDER BY employees.emp_no;
```

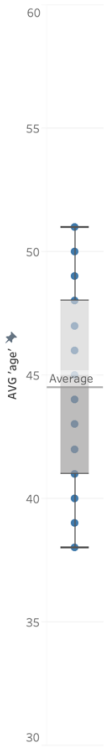
Result:

first_nam	last_nan	gend	age	experienc	department	dept_join_date	job_title	title.to_date	dept.to_date
Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter
Georgi	Facello	M	49	16	Development	1986-06-26	Senior Engineer	2003-06-22	2003-06-22
Bezalel	Simmel	F	39	16	Sales	1996-08-03	Staff	2002-08-02	2002-08-02
Parto	Bamford	M	43	16	Production	1995-12-03	Senior Engineer	2002-12-01	2002-12-01
Chirstian	Koblick	M	49	15	Production	1986-12-01	Engineer	1995-12-01	2002-11-27
Chirstian	Koblick	M	49	15	Production	1986-12-01	Senior Engineer	2002-11-27	2002-11-27
Kyoichi	Maliniak	M	48	12	Human Resources	1989-09-12	Senior Staff	2002-09-09	2002-09-09
Kyoichi	Maliniak	M	48	12	Human Resources	1989-09-12	Staff	1996-09-12	2002-09-09
Anneke	Preusig	F	50	13	Development	1990-08-05	Senior Engineer	2002-08-02	2002-08-02
Tzvetan	Zielinski	F	46	13	Research	1989-02-10	Senior Staff	2003-02-07	2003-02-07
Tzvetan	Zielinski	F	46	13	Research	1989-02-10	Staff	1996-02-11	2003-02-07

Analysis/interpretation:

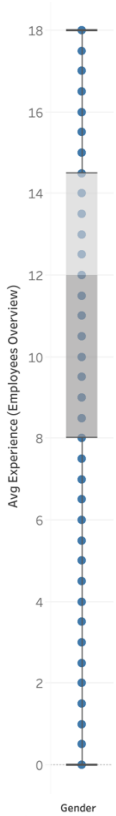
With this query we cannot make a specific analysis because the purpose of this query is to provide an overview of basics information of all employees present in the database.

However, we can make basics statistics on the employees especially on the age, the experience and the gender proportion.

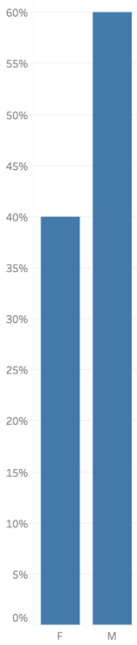


If we focus on the repartition of the age of employees in the company, we can notice the average of the age is 44, and minimum is 38 and maximum is 51.
80% of the employees are above 40 years old and most of them are between 40 and 46 years old.

We can notice this is an very high average, and assume the company may be stop hiring and doesn't want to hire young people according to the HR hiring politic. So we can easily imagine the company doesn't want to hire young people to not spend time on their formation.



Now if we focus on the experience of employees, we can notify the average of experience is around 11 years for an employee and we can see most of employees have an experience between 8 and 15 years, which is a very big sign of a company that seeks to retain its employees over a long term.



By focusing on the gender repartition in the company, we can find a inequality between men and women in each department.

By the way, by coding the query we face some lack of information in the database, especially on information about identifying the employees still working for the company, the last available title for each employee or the last department.

Recommendations:

We offer them to launch a specific program to acquire young talents, in order to train them and make them evolve within the company.
This will allow them to acquire new skills and bring a different point of view, which is essential to develop in the software market.

Based on the gender repartition, we encourage the company to hire more women in order to achieve equality between men and women, particularly for departments whose field is not naturally predominantly male

We also want to provide data base recommendations and encourage the company to insert into the database different information for a easier access to important information, especially on :

- *Active employees*

```sql
ALTER TABLE salaries
ADD in_activity VARCHAR(255);

UPDATE salaries
SET in_activity = 'YES'
WHERE  from_date > '2002-01-01';
```

- *Last available title for each employee*

```sql
-- Marking the currents and pasts job title
ALTER TABLE titles
ADD title_availability VARCHAR (255);

CREATE TABLE last_title
SELECT emp_no, MAX(to_date) AS 'to_date'
FROM titles
GROUP BY emp_no;

UPDATE titles
JOIN last_title
ON titles.emp_no = last_title.emp_no
SET title_availability = 'current'
WHERE titles.to_date = last_title.to_date;


UPDATE titles
SET title_availability = 'past'
WHERE title_availability IS NULL;
```

- Last department for each employee

Same process as for titles table.

#QUERY 2

Query:

```
#The number of employees per department

SELECT COUNT(DISTINCT dept_emp.emp_no) AS 'number_of_employees',
departments.dept_name AS 'department'
FROM dept_emp
JOIN departments
ON departments.dept_no = dept_emp.dept_no
GROUP BY  departments.dept_name
ORDER BY number_of_employees DESC ;
```

Result:

number_of_employees	department
85 707	Development
73 485	Production
52 245	Sales
23 580	Customer Service
21 126	Research
20 211	Marketing
20 117	Quality Management
17 786	Human Resources
17 346	Finance

Analysis/interpretation:

This query provide the total number of employees for the all period of the database. As we can see, the top 3 departments in terms of employees volume are :

- Development
- Production
- Sales

With this information, we can conclude the company have adopted a strategy based on combination between a big production of software and an aggressive commercial strategy based on human agents on the field.

We can also divide the departments of the company in different clusters, based on the total of employees :

- Cluster 1 (production & sales) : development, production, sales (211 437)
- Cluster 2 (marketing & customer relationship, designing & research) : customer service, research, marketing (64 917)
- Cluster 3 (support functions) : quality management, HR & finance (55 249)

Based on the value chain of the company, we can divide the departments of the company in two types of activity : Main activities & Support activites.

Main activities regroup all departments related to the production and sales (cluster 1) & the marketing & research part (cluster 2), but also the quality management from cluster 3
Support activities regroup the departments related to the management (financial & human) so it's elements from cluster 3.

Now we can analyze the typology of offer composition from the company based on the repartition of human resources through the different departments.

Because of the big dominance of the cluster 1 (production & sales), we can assume the company wants to product a lot and sell a big volume of products, but this kind of strategy is usually at the expense of product quality, marketing attributes and customer service quality.

Since a volume strategy requires reducing costs in order to sell a lot at a lower price than the reference product, it is necessary that the company does not end up in a situation called "stuck in the middle", i.e. a flat volume of units sold compared to the market and costs not sufficiently optimized.

Recommendations:

Based on these assumptions, we can recommend that the company revise its initial strategy by improving the functions related to product design (marketing & research), in order to present an apparently more qualitative product with a more attractive perceived value.

The company can also opt for an upward repositioning of its product by increasing the weight of the research and quality departments, as well as by improving customer loyalty through the improvement of the customer relations department.

To support our hypotheses and go further in our analysis, we need to identify the annual evolution of the number of employees per department, in order to identify a possible attempt at strategic repositioning.

We need to :

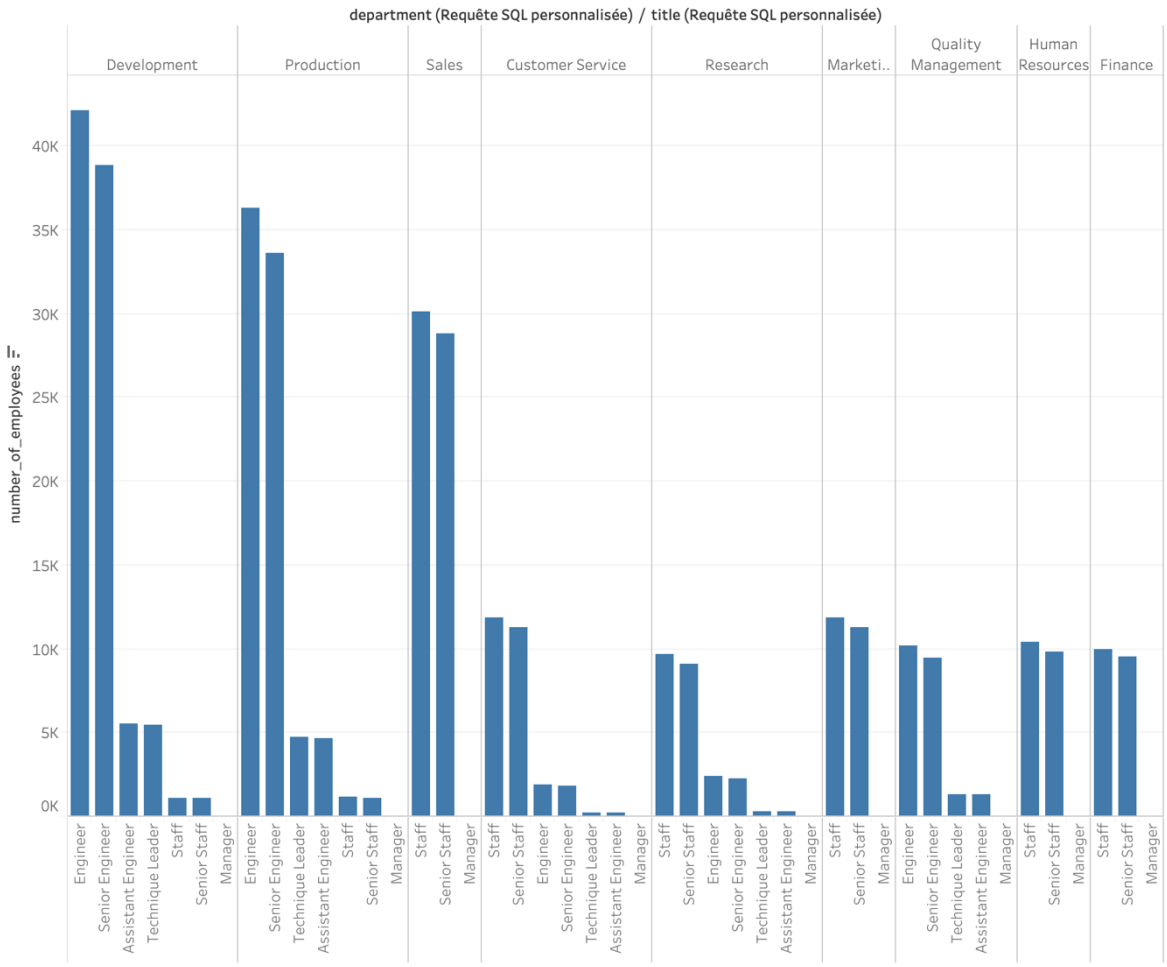
- Prioritize the employees who are still active in order to get an overview of the company's situation at the most recent date in the database.
- Identify the HR policy for acquiring and strengthening skills within the departments by identifying the processes for appointing department heads and managers.
- Identify the connections that exist between the different departments, to identify whether there is a lack of communication or skills in inter-departmental synergies.

Additional Queries based on recommendations

```sql
--The number of employees per department at last date of data base (2003-08-01)
SELECT COUNT(DISTINCT dept_emp.emp_no) AS 'number_of_employees',
departments.dept_name AS 'department'
FROM dept_emp
JOIN departments
ON departments.dept_no = dept_emp.dept_no
WHERE dept_emp.to_date = '2003-08-01'
GROUP BY  departments.dept_name
ORDER BY number_of_employees DESC ;
```

number_of_employees	department
61392	Development
53310	Production
37706	Sales
17572	Customer Service
15442	Research
14845	Marketing
14546	Quality Management
12900	Human Resources
12439	Finance

```sql
-- The number of employees per department & title
SELECT COUNT(DISTINCT dept_emp.emp_no) AS 'number_of_employees',
departments.dept_name AS 'department', titles.title AS 'title'
FROM dept_emp
JOIN departments
ON departments.dept_no = dept_emp.dept_no
JOIN titles
ON dept_emp.emp_no = titles.emp_no
WHERE dept_emp.in_activity = "YES"
GROUP BY  departments.dept_name, title
ORDER BY department, title, COUNT(number_of_employees) DESC ;
```


department (Requête SQL personnalisée) / title (Requête SQL personnalisée)
Development
Production
Sales
Customer Service
Research
Marketi..
Quality Management
Human Resources
Finance
number_of_employees
40K
35K
30K
25K
20K
15K
10K
5K
0K
Engineer
Senior Engineer
Assistant Engineer
Technique Leader
Staff
Senior Staff
Manager

#QUERY 3

Query:

```
-- List of employees per department, their positions, and salaries. Make a separate
list for each department.

select distinct concat(e.first_name ,' ', e.last_name) as 'Name',d.dept_name as
'Departments',t.title as 'Position',s.salary as 'Salary'
from employees e join salaries s on e.emp_no=s.emp_no
join titles t on e.emp_no=t.emp_no
join dept_emp de on e.emp_no=de.emp_no
join departments d on de.dept_no=d.dept_no
join dept_manager dm on d.dept_no=dm.dept_no
where s.to_date='9999-01-01' /* still receive a salary so still in the company*/and
t.to_date = '9999-01-01' /*last title job*/and de.to_date = '9999-01-01'/* last
department*/
and d.dept_no = 'd009'/* department id to be changed depend on the departement to
show*/
order by d.dept_no, e.last_name;
```

Result:

d001 : Marketing

Name	Departments	Position	Salary
Mats Aamodt	Marketing	Senior Staff	114221
Phuoc Aamodt	Marketing	Senior Staff	78772
Alexius Aamodt	Marketing	Senior Staff	58834
Gaetan Aamodt	Marketing	Staff	90226
Hidefumi Aamodt	Marketing	Staff	62885
Sreekrishna Aamodt	Marketing	Senior Staff	59087
Hidefumi Aamodt	Marketing	Senior Staff	116825
Roddy Aamodt	Marketing	Staff	59222
Jongsuk Acton	Marketing	Senior Staff	54466

d002 : Finance

Name	Departments	Position	Salary
Bartek Aamodt	Finance	Senior Staff	76842
Shmuel Aamodt	Finance	Senior Staff	80524
Qingxiang Aamodt	Finance	Senior Staff	49129
Shakhar Aamodt	Finance	Senior Staff	87221
Yakkov Aamodt	Finance	Senior Staff	99713
Mingzeng Aamodt	Finance	Staff	77648
Danco Aamodt	Finance	Staff	61553
Takanari Aamodt	Finance	Senior Staff	110018
Bikash Aamodt	Finance	Senior Staff	49012

d003 : Human Resources

Name	Departments	Position	Salary
Dekang Aamodt	Human Resources	Senior Staff	82224
Rafail Aamodt	Human Resources	Staff	54172
Kensyu Aamodt	Human Resources	Senior Staff	54472
Pramod Aamodt	Human Resources	Senior Staff	75613
Shigeichiro Aamodt	Human Resources	Staff	55170
Younwoo Aamodt	Human Resources	Senior Staff	56364
Alois Aamodt	Human Resources	Staff	66988
Sudharsan Acton	Human Resources	Staff	52876
Gina Acton	Human Resources	Senior Staff	79836

d004 : Production

Name	Departments	Position	Salary
Mokhtar Aamodt	Production	Engineer	54900
Sajjad Aamodt	Production	Senior Staff	72192
JiYoung Aamodt	Production	Senior Engineer	69833
Marie Aamodt	Production	Engineer	48795
Shigeaki Aamodt	Production	Senior Engineer	60313
Valeri Aamodt	Production	Senior Engineer	50585
Leah Aamodt	Production	Senior Engineer	47996
Vugranam Aamodt	Production	Senior Engineer	81932
Tristan Aamodt	Production	Assistant Engineer	101737

d005 : Development

Name	Departments	Position	Salary
Sachem Aamodt	Development	Senior Engineer	62286
Shen Aamodt	Development	Senior Engineer	98486
Huican Aamodt	Development	Senior Engineer	50364
Mani Aamodt	Development	Senior Engineer	55636
Sorina Aamodt	Development	Engineer	47468
Erez Aamodt	Development	Senior Engineer	47606
Aluzio Aamodt	Development	Engineer	53272
Gonzalo Aamodt	Development	Technique Leader	64828
Dmitry Aamodt	Development	Senior Engineer	100059

d006 : Quality Management

Name	Departments	Position	Salary
Sreenivas Aamodt	Quality Management	Senior Engineer	68697
Guther Aamodt	Quality Management	Engineer	50682
Garnet Aamodt	Quality Management	Senior Engineer	67467
Tianruo Aamodt	Quality Management	Senior Engineer	59373
Uinam Aamodt	Quality Management	Engineer	58052
Rafael Aamodt	Quality Management	Senior Engineer	69812
Tadahiko Acton	Quality Management	Senior Engineer	81418
Dipankar Acton	Quality Management	Senior Engineer	79401
Reinhard Acton	Quality Management	Senior Engineer	84773

d007 : Sales

Name	Departments	Position	Salary
Vidar Aamodt	Sales	Senior Staff	91272
Chuanyi Aamodt	Sales	Senior Staff	79724
Narain Aamodt	Sales	Senior Staff	63545
Hairong Aamodt	Sales	Senior Staff	95324
Srinidhi Aamodt	Sales	Senior Staff	86268
Teiji Aamodt	Sales	Senior Staff	129115
Jaideep Aamodt	Sales	Senior Staff	93594
Somnath Aamodt	Sales	Senior Staff	103069
Elzbieta Aamodt	Sales	Senior Staff	130777

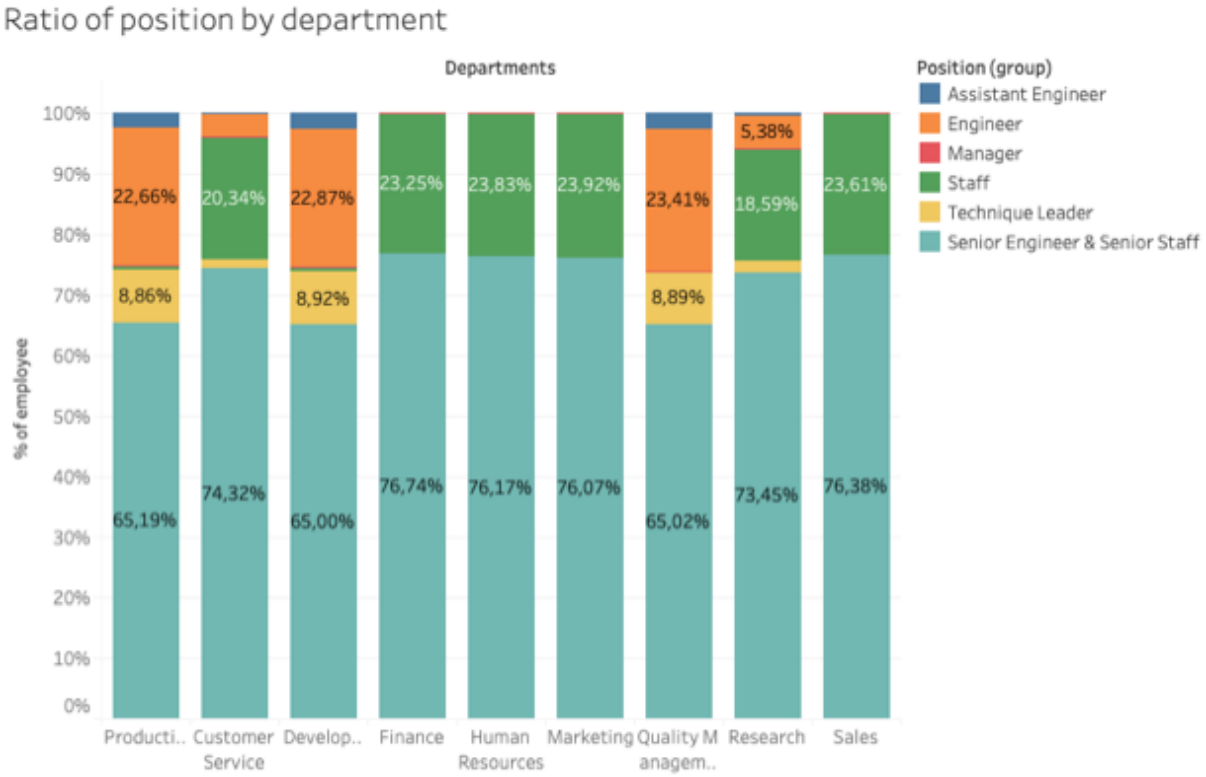
d008 : Research

Name	Departments	Position	Salary
Vasilii Aamodt	Research	Senior Staff	62317
Siddarth Aamodt	Research	Staff	46006
Khatoun Aamodt	Research	Senior Staff	70672
Maha Aamodt	Research	Senior Staff	69746
Hiroyasu Aamodt	Research	Senior Staff	75423
Moto Aamodt	Research	Senior Engineer	89839
Mariangiola Aamodt	Research	Senior Staff	44400
Jeane Aamodt	Research	Senior Engineer	71894
Arlette Aamodt	Research	Senior Staff	60536

d009 : Customer Service

Name	Departments	Position	Salary
Rajmohan Aamodt	Customer Service	Senior Staff	81794
Subhash Aamodt	Customer Service	Senior Engineer	79019
Arumugam Aamodt	Customer Service	Senior Staff	74407
Luigi Aamodt	Customer Service	Staff	58714
Marla Aamodt	Customer Service	Senior Staff	59523
Owen Aamodt	Customer Service	Senior Staff	78890
Youpyo Aamodt	Customer Service	Senior Engineer	71957
Guoxiang Aamodt	Customer Service	Staff	84328
Kamran Aamodt	Customer Service	Senior Engineer	70368

Analysis/interpretation:



We can see on the graph the proportion of each position in the different departments. The two seniors position has been merged to highlight the proportion of senior positions in each department. The senior positions proportion is pretty stable for every department. Around 70% of the employees in each department have a senior position. It means for 1 employee without a senior status, almost 3 senior employees are here to support him. With a senior status comes experience, lot of knowledge but also can go with high salary, routine, old school mind.
We notice there is only one manager in each department.

Recommendations:

Base on this assessment and also the number of employees in each department, we recommend to split the departments into units and to promote some of the senior employees to the management of these units. Each unit can be in charge of one product or one project. It would split the job so it can be done faster and better. The teams would see the result faster and be more engaged to their work. Also, the management would be more effective. We would suggest to take advantage of the senior status employees to hire young employees in order to train them and to transfer their knowledge and experience.

#QUERY 4

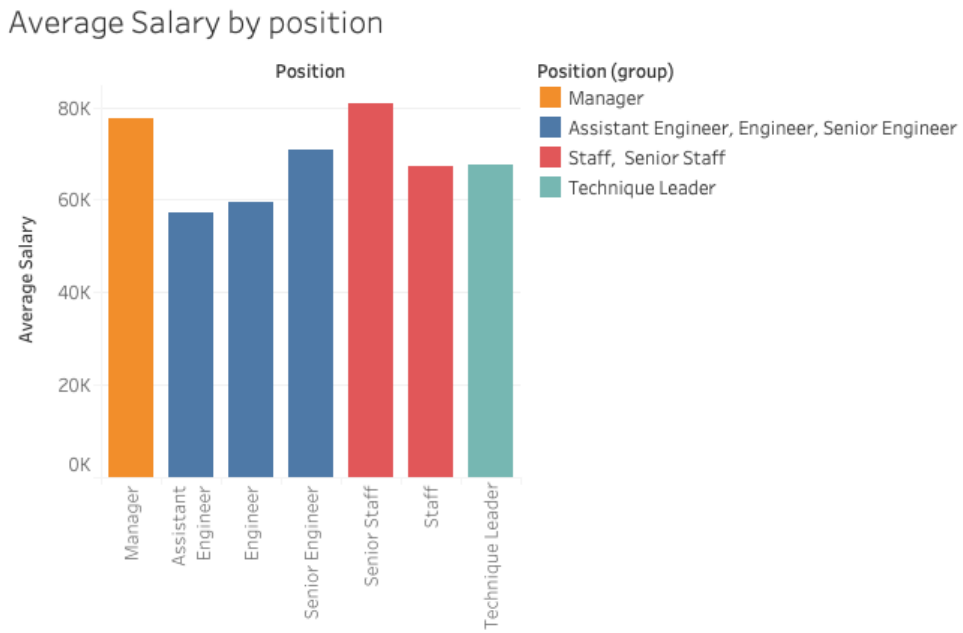
Query:

```
List of the average salary of employees in each job position (title).

select round(AVG(s.salary),0) as 'Average Salary', t.title as 'Position'
from titles t join salaries s on t.emp_no=s.emp_no where s.to_date='9999-01-01' and
t.to_date='9999-01-01'
group by t.title
order by round(AVG(s.salary),0);
```

Result:

Average Salary	Position
57318	Assistant Engineer
59603	Engineer
67331	Staff
67507	Technique Leader
70823	Senior Engineer
77724	Manager
80706	Senior Staff

Analysis/interpretation:



Sum of Average Salary for each Position. Colour shows details about Position (group).

The average salary goes up with the position. A Senior Engineer earns more than an Engineer who earns more than an assistant Engineer. What is striking is the difference of salary between a Senior Engineer and a Staff. The difference is really low. For a software company, the engineers are the spearheads of the company. And in this company, even in a Senior position they can pretend to almost the same salary as the Staff. This salary difference can be really frustrating for the employees. Every employee needs to be recognized by the company if he wants to feel good in his company. To be paid as his fair value is part of this recognition feeling.

Recommendations:

Based on this analysis we recommend to adjust the salaries among all the employees. As it is not possible do decrease the salary of the current employees, we suggest to increase the salaries of the Engineers and Senior Engineers to be grateful to what the bring to the company.

#QUERY 5

Query:

```
-- List of the average age of employees in the company, in each department and each
job position

-- avg age in the company of all current employees
select round(avg((DATE_FORMAT (FROM_DAYS(TO_DAYS('2002-08-01')-
TO_DAYS(birth_date)), '%Y')+0)),2) as age from
(select
e.emp_no,e.first_name,e.last_name,e.birth_date,e.gender,e.hire_date,a.dept_name,d.t
itle,s.salary
from employees as e join departments a join dept_emp b join titles d join salaries
s
on e.emp_no=b.emp_no and b.emp_no=d.emp_no and d.emp_no=s.emp_no and
b.dept_no=a.dept_no
where s.to_date='2002-08-01' and d.to_date='2002-08-01' and b.to_date='2002-08-
01')f ;

-- avg age in each department of current employees
select f.dept_name as department,round(avg((DATE_FORMAT (FROM_DAYS(TO_DAYS('2002-
08-01')-TO_DAYS(birth_date)), '%Y')+0)),2)as age from
(select
e.emp_no,e.first_name,e.last_name,e.birth_date,e.gender,e.hire_date,a.dept_name,d.t
itle,s.salary
from employees as e join departments a join dept_emp b join titles d join salaries
s
on e.emp_no=b.emp_no and b.emp_no=d.emp_no and d.emp_no=s.emp_no and
b.dept_no=a.dept_no
where s.to_date='2002-08-01' and d.to_date='2002-08-01'and b.to_date='2002-08-01')f
group by f.dept_name;

-- avg age in each titles of current employees
select f.title as position,round(avg((DATE_FORMAT (FROM_DAYS(TO_DAYS('2002-08-01')-
TO_DAYS(birth_date)), '%Y')+0)),2) as age from
(select
e.emp_no,e.first_name,e.last_name,e.birth_date,e.gender,e.hire_date,a.dept_name,d.t
itle,s.salary
from employees as e join departments a join dept_emp b join titles d join salaries
s
on e.emp_no=b.emp_no and b.emp_no=d.emp_no and d.emp_no=s.emp_no and
b.dept_no=a.dept_no
where s.to_date='2002-08-01' and d.to_date='2002-08-01' and b.to_date='2002-08-
01')f group by f.title;
```

Result:

age
43.50

department	age
Development	43.49
Sales	43.50
Production	43.52
Human Resources	43.52
Research	43.48
Quality Management	43.53
Marketing	43.48
Customer Service	43.49
Finance	43.51

position	age
Senior Engineer	43.50
Staff	43.49
Senior Staff	43.51
Engineer	43.49
Assistant Engineer	43.48
Technique Leader	43.56
Manager	42.67

The average age of current employees in the firm is 43.5 years old.
There is no big gap between different department about the average age, which is also about 43.5 years old. The gap of oldest department and youngest department is just 0.05 which is even negligible.
The same case for different job position happens. The average age in 7 different job position also wanders at about 43.5 years old. And we notice that an interesting thing is that the manager's average age is the youngest compare to other job position although the difference is very small (less than 1).

Analysis/interpretation:

It demonstrates that the age composition in the company is very stable: the middle age employees plays the most important role in this company. But as a software company and an emerging industry company, the employees average age seems too old because the technological updating is very common in industry and especially for the newest trend: AI. Some old staff they may not grasp the technique that the company wants.

Recommendations:

They need to hire young and high-educated employees who grasp the newest code technology. Besides new employees can bring different working atmosphere to the company, enhance the age composition and bring different site of views.

#QUERY 06

Query:

```
-- What is the ratio of me to women in each department and each job position?

-- men ratio in title of current emmployees
select title as position,
sum(case when gender='M'then 1 else 0 end) as  men ,count(*) as total,
round(sum(case when gender='M' then 1 else 0 end)/count(*),2) as ratio from
(select
e.emp_no,e.first_name,e.last_name,e.birth_date,e.gender,e.hire_date,a.dept_name,d.t
itle,s.salary
from employees as e join departments a join dept_emp b join titles d join salaries
s
on e.emp_no=b.emp_no and b.emp_no=d.emp_no and d.emp_no=s.emp_no and
b.dept_no=a.dept_no
where s.to_date='2002-08-01' and d.to_date='2002-08-01' and b.to_date='2002-08-
01')f group by title;

-- women ratio in  title of current emmployees
select title as position,
sum(case when gender='F'then 1 else 0 end) as women ,count(*) as total,
round(sum(case when gender='F' then 1 else 0 end)/count(*),2) as ratio from
(select
e.emp_no,e.first_name,e.last_name,e.birth_date,e.gender,e.hire_date,a.dept_name,d.t
itle,s.salary
from employees as e join departments a join dept_emp b join titles d join salaries
s
on e.emp_no=b.emp_no and b.emp_no=d.emp_no and d.emp_no=s.emp_no and
b.dept_no=a.dept_no
where s.to_date='2002-08-01' and d.to_date='2002-08-01'and b.to_date='2002-08-01')f
group by title;

-- men ratio in department of current emmployees
select dept_name as department,sum(case when gender='M'then 1 else 0 end) as
men,count(*) as total,
round(sum(case when gender='M'then 1 else 0 end)/count(*),2) as ratio from
(select
e.emp_no,e.first_name,e.last_name,e.birth_date,e.gender,e.hire_date,a.dept_name,d.t
itle,s.salary
from employees as e join departments a join dept_emp b join titles d join salaries
s
on e.emp_no=b.emp_no and b.emp_no=d.emp_no and d.emp_no=s.emp_no and
b.dept_no=a.dept_no
where s.to_date='2002-08-01' and d.to_date='2002-08-01'and b.to_date='2002-08-01')f
group by dept_name;

-- women ratio in department of current emmployees
```

```
select dept_name as department,sum(case when gender='f'then 1 else 0 end) as
women,count(*) as total,
round(sum(case when gender='f'then 1 else 0 end)/count(*),2) as ratio from
(select
e.emp_no,e.first_name,e.last_name,e.birth_date,e.gender,e.hire_date,a.dept_name,d.t
itle,s.salary
from employees as e join departments a join dept_emp b join titles d join salaries
s
on e.emp_no=b.emp_no and b.emp_no=d.emp_no and d.emp_no=s.emp_no and
b.dept_no=a.dept_no
where s.to_date='2002-08-01' and d.to_date='2002-08-01'and b.to_date='2002-08-01')f
group by dept_name;
```

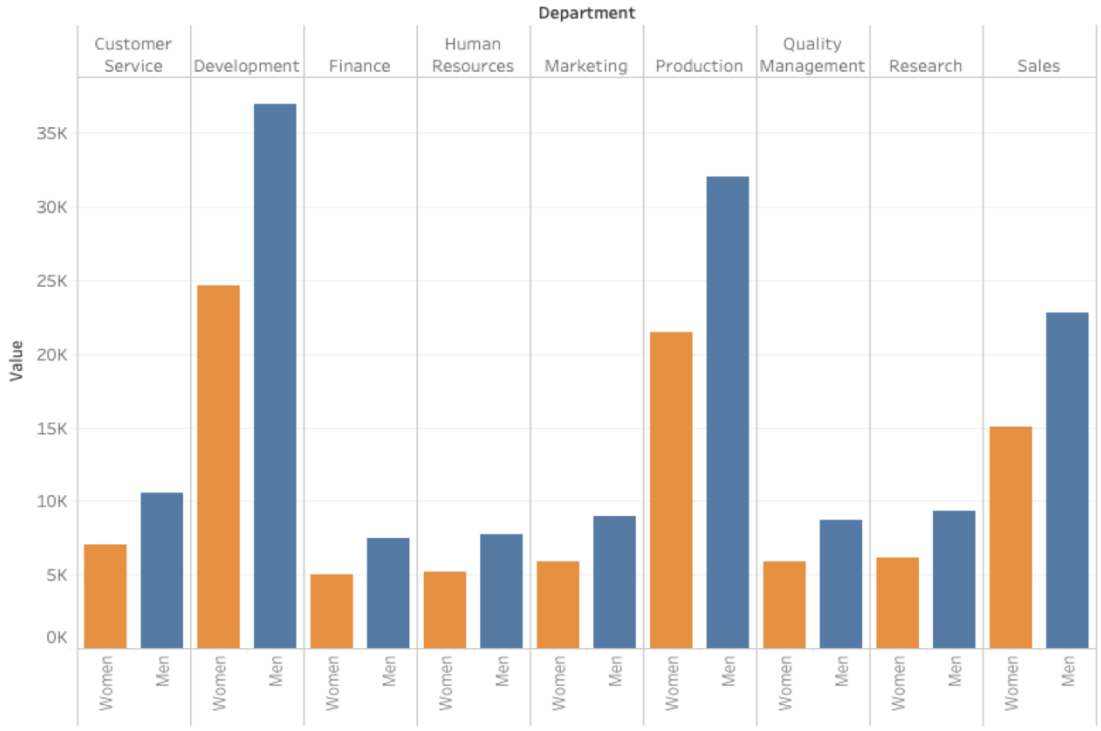
Result:

position	men	total	ratio
Senior Engineer	51666	86170	0.60
Staff	15500	25637	0.60
Senior Staff	49376	82250	0.60
Engineer	18627	31085	0.60
Assistant Engineer	2160	3605	0.60
Technique Leader	7209	12088	0.60
Manager	5	9	0.56

position	women	total	ratio
Senior Engineer	34504	86170	0.40
Staff	10137	25637	0.40
Senior Staff	32874	82250	0.40
Engineer	12458	31085	0.40
Assistant Engineer	1445	3605	0.40
Technique Leader	4879	12088	0.40
Manager	4	9	0.44

department	men	total	ratio
Development	36966	61574	0.60
Sales	22777	37830	0.60
Production	31994	53456	0.60
Human Resources	7772	12938	0.60
Research	9293	15489	0.60
Quality Management	8700	14588	0.60
Marketing	9004	14883	0.60
Customer Service	10592	17617	0.60
Finance	7445	12469	0.60

department	women	total	ratio
Development	24608	61574	0.40
Sales	15053	37830	0.40
Production	21462	53456	0.40
Human Resources	5166	12938	0.40
Research	6196	15489	0.40
Quality Management	5888	14588	0.40
Marketing	5879	14883	0.40
Customer Service	7025	17617	0.40
Finance	5024	12469	0.40



The result about men to women ratio in every department in the whole are nearly the same. The ratio is always about 6:4 in man and woman. As for the job position, the situation is almost the same like the department. The ratio men to women is always 6:4 except that in manager title, the ratio is 56% to 44% of men to women. But considering the very limited number of managers compare to other job titles. This change is negligible.

Analysis/interpretation:

This result shows that this company hires more men than women. Even in some department that women are traditionally seen as dominant like human resource and finance, it is the same situation. As it is a software company, it makes no sense at all to have a comprehensive gender difference in either every department or every job title because it's not a company which focuses on physical labor.

Recommendations:

It is hard to suppose that the strategy of hire employees contains a gender bias in this firm which possibly starts at the beginning of recruitment. But what is sure is that try to avoid a huge gender difference in everywhere of the company is better.

#QUERY 07

Query:

```
-- List of employees with (i) the lowest salary, (ii) the highest salary (iii)
above average salary. Please do not forget to mention the lowest and highest salary
amount. The currency is Euro.


-- (i) the highest salary

SELECT
E.emp_no AS employee_no,
E.first_name AS "First Name",
E.last_name AS "Last Name",
S.salary
FROM employees E JOIN salaries S USING(emp_no)
WHERE to_date="9999-01-01" AND S.salary IN (SELECT MAX(S.salary) FROM salaries S);
SELECT E.emp_no AS employee_no,
E.first_name AS "First Name",
E.last_name AS "Last Name",
S.salary
FROM employees E JOIN salaries S USING(emp_no)
WHERE to_date="9999-01-01" AND S.salary IN (SELECT MIN(S.salary) FROM salaries S);

-- (ii) the lowest salary

SELECT
E.emp_no AS employee_no,
E.first_name AS "First Name",
E.last_name AS "Last Name",
S.salary
FROM employees E JOIN salaries S USING(emp_no)
WHERE to_date="9999-01-01" AND S.salary IN (SELECT MIN(S.salary) FROM salaries S);
```

Results for (i) (ii)

employees_no	First Name	Last Name	salary
43624	Tokuyasu	Pesch	158220

employees_no	First Name	Last Name	salary
253406	Olivera	Baek	38623

```sql
--iii the employees who have above the average salaries

 SELECT AVG(S.salary)
     FROM ((employees E JOIN dept_emp DE USING(emp_no))
     JOIN salaries S USING(emp_no))
     JOIN departments D ON D.dept_no=DE.dept_no
     WHERE S.to_date='9999-01-01' AND DE.to_date='9999-01-01';

    SELECT E.emp_no AS "Employee No", E.first_name AS "First Name", E.last_name AS
"Last Name", S.salary AS "Salary"
    FROM ((employees E JOIN dept_emp DE USING(emp_no))
    JOIN salaries S USING(emp_no))
    JOIN departments D ON D.dept_no=DE.dept_no
    WHERE DE.to_date='9999-01-01' AND S.to_date='9999-01-01' AND S.salary >
72012.2359;
```

Result :

Employee No	First Name	Last Name	Salary
10001	Georgi	Facello	88958
10002	Bezalel	Simmel	72527
10004	Chirstian	Koblick	74057
10005	Kyoichi	Maliniak	94692
10007	Tzvetan	Zielinski	88070

#QUERY 08

Query:

```sql
-- List of employees who joined the company in December.

select
emp_no as employees_no,
first_name,
last_name,
hire_date
from employees
WHERE month(hire_date)=12;
```

Result

employees_no	first_name	last_name	hire_date
10004	Chirstian	Koblick	1986-12-01
10012	Patricio	Bridgland	1992-12-18
10023	Bojan	Montemayor	1989-12-17
10037	Pradeep	Makrucki	1990-12-05
10050	Yinghua	Dredge	1990-12-25
10073	Shir	McClurg	1991-12-01

#QUERY 09

Query:

List of the most experienced employees (by the number of years) in each department and the company as a whole.

```
--(i) Each department

SELECT E.first_name AS "First Name", E.last_name AS "Last Name",
            T.title AS "Position", D.dept_name AS "Department Name",
    YEAR('2002-08-01')-YEAR(E.birth_date) AS "Age",
    YEAR('2002-08-01')-YEAR(E.hire_date) AS "Experience"
            FROM ((employees E JOIN dept_emp DE USING(emp_no))
    JOIN titles T USING(emp_no))
    JOIN departments D ON D.dept_no=DE.dept_no
    WHERE (YEAR('2002-08-01')-YEAR(E.hire_date))=17 AND DE.to_date='9999-01-01' AND
T.to_date='9999-01-01' AND D.dept_name='Development';

--(ii) the company

SELECT E.first_name AS "First Name", E.last_name AS "Last Name", T.title AS
"Position",
            D.dept_name AS "Department Name", YEAR('2002-08-01')-YEAR(E.birth_date)
AS "Age",
            YEAR('2002-08-01')-YEAR(E.hire_date) AS "Experience"
            FROM ((employees E JOIN dept_emp DE USING(emp_no))
    JOIN titles T USING(emp_no))
    JOIN departments D ON D.dept_no=DE.dept_no
    WHERE (YEAR('2002-08-01')-YEAR(E.hire_date))=17 AND T.to_date='9999-01-01' AND
DE.to_date='9999-01-01';
```

Result:

In all the departments, maximum experiences are 17 years as shown in the table below.

Experience	Department Name
17	Development
17	Sales
17	Production
17	Human Resources
17	Research
17	Quality Management
17	Customer Service
17	Marketing
17	Finance

Then, after proceeding further, we obtained the list of the most experience employees in each of the department and the company as a whole. Sample results for the whole company is attached below.

First Name	Last Name	Position	Department Name	Age	Experience
Bezalel	Simmel	Staff	Sales	38	17
Sumant	Peac	Senior Engineer	Quality Management	50	17
Eberhardt	Terkki	Senior Staff	Human Resources	39	17
Otmar	Herbst	Senior Engineer	Quality Management	46	17
Tse	Herber	Senior Staff	Sales	40	17
Reuven	Garigliano	Technique Leader	Research	47	17
Erez	Ritzmann	Senior Engineer	Development	50	17
Premal	Baek	Senior Staff	Human Resources	45	17
Yuichiro	Swick	Staff	Customer Service	39	17
Kayoko	Valtorta	Senior Staff	Customer Service	48	17

(the same query can be used to obtain all the most experienced employees in each of the departments. For sample, I have attached the results of "Development" team.

First Name	Last Name	Position	Department Name	Age	Experience
Erez	Ritzmann	Senior Engineer	Development	50	17
Armond	Peir	Technique Leader	Development	47	17
Nigel	Aloisi	Senior Engineer	Development	48	17
Pranav	Furedi	Senior Engineer	Development	49	17
Carrsten	Schmiedel	Engineer	Development	39	17
Foong	Flasterstein	Engineer	Development	41	17
Basil	Ishibashi	Senior Engineer	Development	47	17
Prodip	Schusler	Assistant Engineer	Development	38	17
Mototsugu	Beilner	Senior Engineer	Development	44	17
Eishiro	Miyakawa	Senior Engineer	Development	46	17

Analysis/interpretation:

As mentioned earlier, 17 is the maximum years of experience of employees in various departments. Total number of employees with this experience in various departments are shown below.

Sl. No.	Department	No. of employees
01	Development	7270
02	Production	6383
03	Sales	4344
04	Customer Service	2021
05	Research	1841
06	Marketing	1773
07	Quality Management	1712
08	Human Resources	1514
09	Finance	1433
	Total	28291

After analyzing the above data of 28291 employees, their age distribution across departments and job roles are shown below.

It is found through further analyses that more aged people are found in Production, Development and Sales departments.

As we have mentioned earlier cluster 1 has most aged people among the most experienced people in the company.

Recommendations:

As more aged people are concentrated in certain departments alone, this leads to an invariability in the distribution of experience across different departments. There might be some misguidance if experienced people are not present in certain departments. We suggest revamping employees' departments according to their interests simultaneously considering their expertise in this company and their previous company. For example, some senior engineer in Quality Management department could be serving for some time in production to share some problems that has been repeating for a long time and causing issues. This will decrease the time to market of the product.

#QUERY 10

Query:

```
-- List of the most recently hired employees (that is, the year the most recent
employee was recruited).

SELECT MAX(YEAR(E.hire_date)) FROM employees E;

SELECT E.emp_no, E.first_name AS "First Name", E.last_name AS "Last Name", T.title
AS "Job Title",
    YEAR('2003-08-01')-YEAR(E.birth_date) AS "Age",
            YEAR('2003-08-01')-YEAR(E.hire_date) AS "Experience"
            FROM employees E JOIN titles T USING(emp_no)
    WHERE YEAR(E.hire_date)=2000;
```

Result:

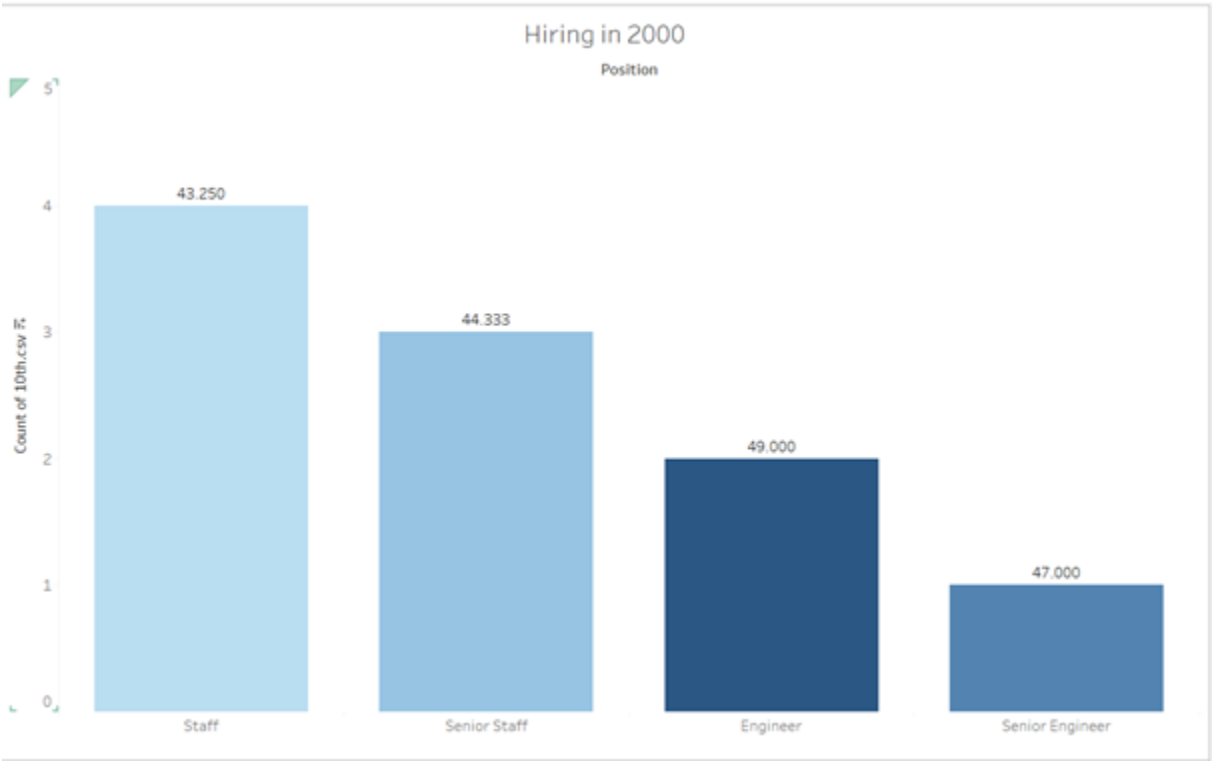
2000 is the last year that an employee has been recruited in the company. Totally, we have 13 employees who were hired in 2000 as shown below.

First Name	Last Name	Position	Age	Gender	Experience
Ulf	Flexer	Staff	42	M	2
Seshu	Rathonyi	Staff	38	F	2
Randi	Luit	Staff	49	F	2
Mariangiola	Boreale	Senior Engineer	47	M	2
Ennio	Alblas	Senior Staff	42	F	2
Volkmar	Perko	Senior Staff	43	F	2
Xuejun	Benzmuller	Staff	44	F	2
Shahab	Demeyer	Senior Staff	48	M	2
Jaana	Verspoor	Engineer	49	F	2
Jeong	Boreale	Engineer	49	M	2
Yucai	Gerlach	Engineer	45	M	2
Bikash	Covnot	Engineer	38	M	2
Hideyuki	Delgrande	Engineer	48	F	2

Analysis/interpretation:

Of all the 13 employees hired in 2000, their distribution of job position along with their average age is shown below. It can be observed that highest number of hires was for “Staff” position. But they have the lowest average age.

It is also interesting to note that higher the number of hires, lower is the average age.



Of all these 13 employees. 4 of them left the company within almost 2 years in the company as shown below.

First Name	Last Name	Position	Age	Relieved date	Months Spent at company
Seshu	Rathonyi	Staff	38	2001-07-24	18
Shahab	Demeyer	Senior Staff	48	2002-03-25	26
Jeong	Boreale	Engineer	49	2002-02-05	25
Bikash	Covnot	Engineer	38	2000-05-19	3

Bikash Covnot left the company within 3 months of joining. Company spends a lot of money on hiring potential employees. If newly recruited employees leave the company before generating much value to the company, it is still a great loss of revenue to the company.

Recommendations:

Efforts must be taken to retain employees in the company. More helpful employee benefits could be given to increase their chances of being retained in the company. The company can also try to sign a contract while joining stating that if the employee stays in the company for a particular number of years, they will be entitled to receive a huge increment and few shares of the company. The company could also create a team to follow up with the employees to check with their level of satisfaction in their roles.

#ADDITIONAL QUERY 01

Query:

```
-- How many employees were hired every year?

SELECT YEAR(E.hire_date) AS "Year" ,COUNT(DISTINCT E.emp_no) AS "No. of employees
hired"
            FROM ((employees E JOIN dept_emp DE USING(emp_no))
    JOIN titles T USING(emp_no))
    JOIN departments D ON D.dept_no=DE.dept_no
    GROUP BY YEAR(E.hire_date);
```

Result:

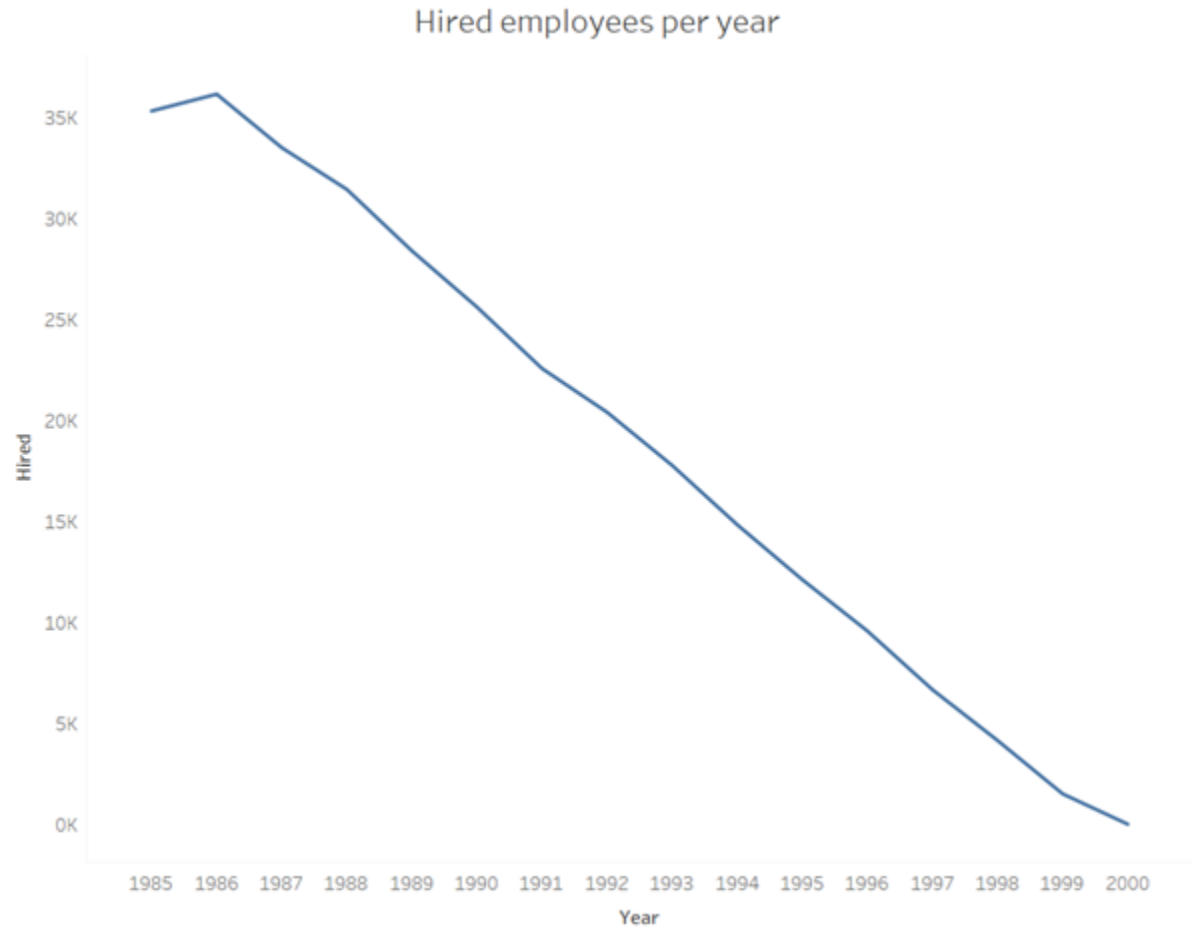
2000 is the last year that an employee has been recruited in the company. Totally, we have 13 employees who were hired in 2000 as shown below.

Year	No. of employees hired
1985	35316
1986	36150
1987	33501
1988	31436
1989	28394
1990	25610
1991	22568
1992	20402
1993	17772
1994	14835
1995	12115
1996	9574
1997	6669
1998	4155
1999	1514
2000	13

The below list shows the total number of employees hired evry year since the beginning of the company.

Analysis/interpretation:

Analyzing the results obtained, we found that the number of employees hired has been decreasing over years.



We assume all the current employees will stay into the company. Starting in 2014 every year around 18000 employees will be retired. So, a huge amount of employee will leave the company.

Recommendations:

Hiring new employees helps the company stay along with its competitors. Decreasing or halting the recruitment process prevent potential employees from joining the company which in turns adversely affects the growth of the company in long run. As time moves, potential employees grow old and eventually leave the company. They also take their knowledge and experience with them. So, we suggest stabilizing the hiring process and brining in new talents to the company thus keeping it sustainable.

#ADDITIONAL QUERY 2(employees rotation rate)

Query:

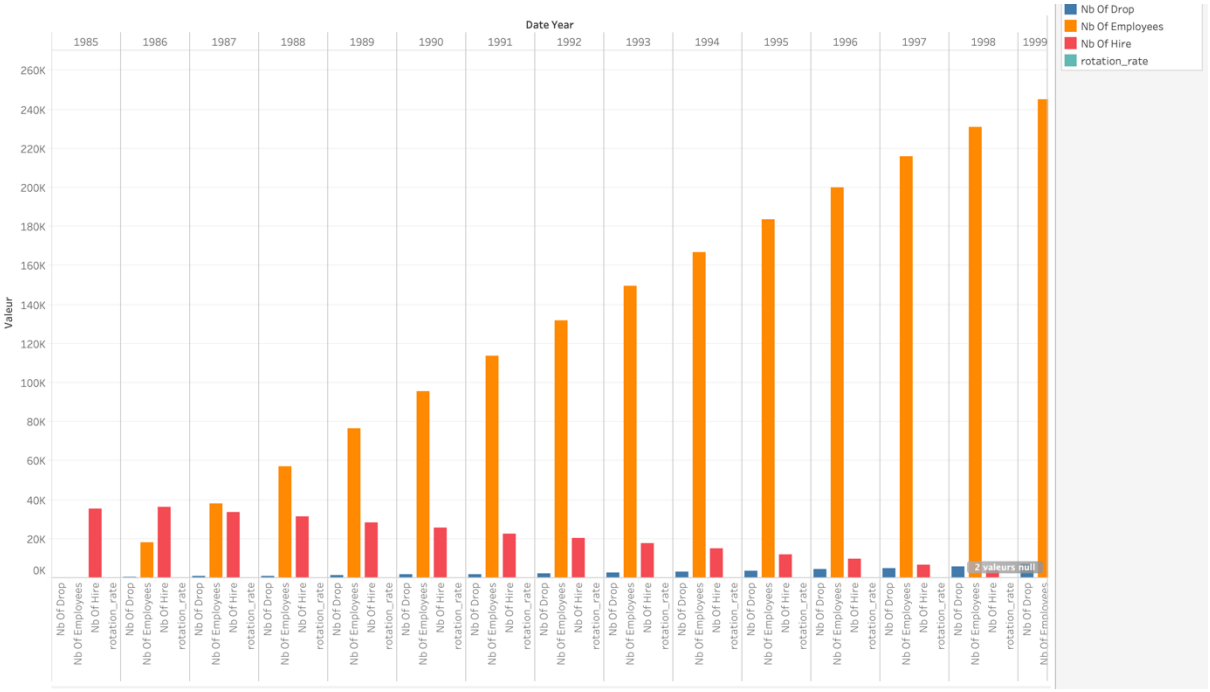
```sql
-- Calculation of rotation rate

-- Identifying number of hire by year
SELECT COUNT(DISTINCT employees.emp_no) AS 'nb_of_hire', DATE_FORMAT(hire_date,
"%Y") AS 'hiredate'
FROM employees
GROUP BY hiredate;


-- Identifying number of employees by year
SELECT COUNT(DISTINCT emp_no) AS 'nb_of_employees', DATE_FORMAT(to_date, "%Y") AS
'todate'
FROM salaries
GROUP BY todate;

-- Creating a table whith number of hire by year
CREATE TABLE employees_rotation
SELECT COUNT(DISTINCT emp_no) AS 'nb_of_hire', DATE_FORMAT(hire_date, "%Y") AS
'hiredate'
FROM employees
GROUP BY hiredate;

-- Creating a table whith number of employees by year
CREATE TABLE employees_rotation2
SELECT COUNT(DISTINCT emp_no) AS 'nb_of_employees', DATE_FORMAT(to_date, "%Y") AS
'todate'
FROM salaries
GROUP BY todate;

-- Identifying the last year for each employee
SELECT emp_no, MAX(from_date),MAX(to_date)
FROM employees.salaries
GROUP BY emp_no;

-- Create a table last_salary
CREATE TABLE last_salary
SELECT emp_no, MAX(from_date) AS 'from_date', MAX(to_date) AS 'to_date'
FROM employees.salaries
GROUP BY emp_no;

-- Marking last year for each employee
ALTER TABLE salaries
ADD last_salary VARCHAR (255);

UPDATE salaries
JOIN last_salary
ON last_salary.emp_no = salaries.emp_no
```

```sql
SET last_salary = 'LAST'
WHERE last_salary.to_date = salaries.to_date;

-- Creating a table with number of employees who leave the company by year
CREATE TABLE employees_rotation3
SELECT COUNT(last_salary) AS 'nb_of_drop', DATE_FORMAT(to_date, "%Y") AS 'todate'
FROM salaries
WHERE last_salary = 'LAST' AND to_date < '2002-08-01'
GROUP BY todate
ORDER BY todate DESC;



-- Merge tables whith number of employees, number of hire and number of number of departures
CREATE TABLE emp_rotation
SELECT nb_of_drop, employees_rotation.nb_of_hire,
employees_rotation2.nb_of_employees, employees_rotation3.todate AS 'date_year'
FROM employees_rotation3
LEFT JOIN employees_rotation
ON employees_rotation3.todate = employees_rotation.hiredate
JOIN employees_rotation2
ON employees_rotation3.todate = employees_rotation2.todate
;

SELECT nb_of_drop/nb_of_employees AS 'rotation_rate'
FROM emp_rotation;

ALTER TABLE emp_rotation
ADD rotation_rate numeric (10, 2);

UPDATE emp_rotation
SET rotation_rate = nb_of_drop/nb_of_employees;
```

Result:

nb_of_dro	nb_of_hire	nb_of_employees	date_yea	rotation_rate
Filter	Filter	Filter	Filter	Filter
89	35316	89	1985	1.00
375	36150	18340	1986	0.02
670	33501	37872	1987	0.02
906	31436	57267	1988	0.02
1292	28394	76385	1989	0.02
1597	25610	95423	1990	0.02
1886	22568	113885	1991	0.02
2326	20402	131907	1992	0.02
2749	17772	149755	1993	0.02
3141	14835	166944	1994	0.02
3752	12115	183770	1995	0.02
4346	9574	200025	1996	0.02
5077	6669	215860	1997	0.02
5955	4155	230973	1998	0.03
6981	1514	245046	1999	0.03
7627	13	257447	2000	0.03
7263	(NULL)	251376	2001	0.03
4011	(NULL)	244135	2002	0.02

Analysis/interpretation:



As we can see on the graph, there is too much hires in the first years of the database but not enough in the last years. It can lead the company to face a problem, too much employees will be retire at the same time and there is not enough rotation of the employees to provide new skills to the company.

Recommendations:

We encourage the company to prepare a HR strategy to use at the best the skills and experience of old employees in order to train an integrate future young employees.

#ADDITIONAL QUERY 3 (HR job title process)

Query:

- Identify employees who changed their title
- Identify number title changes by employee
- Identify rank (Assistant, Junior, Senior)

```sql
-- Identifying employees who get a promotion
SELECT *
FROM titles
WHERE emp_no IN (SELECT emp_no
                    FROM titles
                    GROUP BY emp_no
                    HAVING COUNT(emp_no) > 1);

-- Marking them in the title table
ALTER TABLE titles
ADD promotion VARCHAR (255);

CREATE TABLE titles2
LIKE titles;

INSERT INTO titles2
SELECT * FROM titles;

UPDATE titles
SET promotion = 'promoted'
WHERE emp_no IN (SELECT emp_no
                    FROM titles2
                    GROUP BY emp_no
                    HAVING COUNT(emp_no) > 1);



-- Creating a table whith number of promotion by employee

CREATE TABLE titles3
SELECT titles.emp_no, titles.from_date, titles.to_date, employees.hire_date,
titles.title, titles.Promotion
    FROM titles
    LEFT JOIN employees
    ON titles.emp_no = employees.emp_no;

UPDATE titles3
SET promotion = 'hired'
WHERE from_date = hire_date;

ALTER TABLE titles3
ADD promotions_nb INT;
```

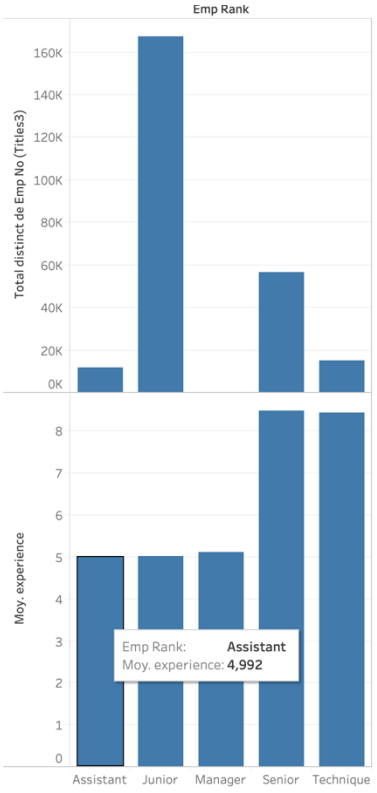
```sql
UPDATE titles3
SET promotions_nb = 1
WHERE Promotion = 'promoted';

UPDATE titles3
SET promotions_nb = 0
WHERE promotions_nb IS NULL;

UPDATE titles3
SET promotion = 'hired'
WHERE  Promotion IS NULL;


-- Identifying differents rank

SELECT DISTINCT title
FROM titles3;


-- Marking rank in the table

ALTER TABLE titles3
ADD emp_rank VARCHAR (255);

UPDATE titles3
SET emp_rank = 'Junior'
WHERE title = 'Engineer' OR title = 'Staff';

UPDATE titles3
SET emp_rank = 'Senior'
WHERE title LIKE '%Senior%';


UPDATE titles3
SET emp_rank = 'Assistant'
WHERE title LIKE '%Assistant%';

UPDATE titles3
SET promotions_nb = 0
WHERE emp_rank = 'Assistant';

UPDATE titles3
SET emp_rank = 'Technique'
WHERE title LIKE '%Technique%';

UPDATE titles3
SET emp_rank = 'Manager'
WHERE title LIKE '%Manager%';


-- Adding experience spend in one title
```

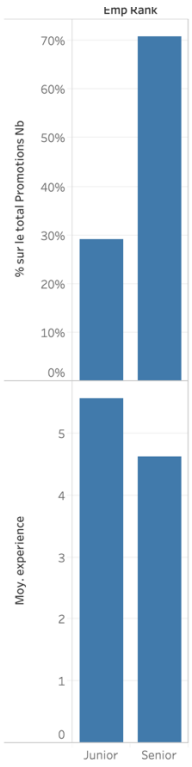
```sql
ALTER TABLE titles3
ADD experience VARCHAR (255);

UPDATE titles3
SET experience = DATE_FORMAT(FROM_DAYS(TO_DAYS(to_date)-TO_DAYS(from_date)), '%y');
```

Result:

emp_	from_date	to_date	hire_date	title	Promoti	promotions_nb	emp_ra	experienc
Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter
10001	1986-06-26	2002-08-01	1986-06-26	Senior Engineer	hired	0	Senior	16
10002	1996-08-03	2002-08-01	1985-11-21	Staff	hired	0	Junior	05
10003	1995-12-03	2002-08-01	1986-08-28	Senior Engineer	hired	0	Senior	06
10004	1986-12-01	1995-12-01	1986-12-01	Engineer	hired	0	Junior	08
10004	1995-12-01	2002-08-01	1986-12-01	Senior Engineer	promoted	1	Senior	06
10005	1996-09-12	2002-08-01	1989-09-12	Senior Staff	promoted	1	Senior	05
10005	1989-09-12	1996-09-12	1989-09-12	Staff	hired	0	Junior	07
10006	1990-08-05	2002-08-01	1989-06-02	Senior Engineer	hired	0	Senior	11

Analysis/interpretation:





As we can see here, the people hired by the company are most hires as a "Junior" position (which correspond to Staff, Engineer). Some of them were directly hired as a senior position.

We can notify the employees hired as a senior position spent more time in the company then employees hired at a Junior position. It's because many people who get hired at a junior position were be promoted at a senior position due to an HR politic which want to keep employees in the company.

Then, if we focus on employees who get promoted we can easily see the difference of experience, because of the time already spent at a previous title.

Recommendations:

Our recommendation is based on the guiding thread of our analysis, hire young people. Here we can encourage the company to switch from the couple Junior > Senior promotion to the couple Assistant to Junior.

In this way, the company will be able to benefit from new talent and skills at a lower cost, by facilitating their integration with the help of its solid base of experienced employees.

#ADDITIONAL QUERY 4

Query:

```
-- What is the average salary among the different department in last year and in
1985 and 2002?

select avg(f.salary) as salary,f.dept_name from
(select
e.emp_no,e.first_name,e.last_name,e.birth_date,e.gender,e.hire_date,a.dept_name,d.t
itle,s.salary
from employees as e join departments a join dept_emp b join titles d join salaries
s
on e.emp_no=b.emp_no and b.emp_no=d.emp_no and d.emp_no=s.emp_no and
b.dept_no=a.dept_no
where s.from_date = '1985-01-01' and d.from_date ='1985-01-01'and b.from_date
='1985-01-01')f group by f.dept_name;

select avg(f.salary) as salary,f.dept_name from
(select
e.emp_no,e.first_name,e.last_name,e.birth_date,e.gender,e.hire_date,a.dept_name,d.t
itle,s.salary
from employees as e join departments a join dept_emp b join titles d join salaries
s
on e.emp_no=b.emp_no and b.emp_no=d.emp_no and d.emp_no=s.emp_no and
b.dept_no=a.dept_no
where s.to_date = '2002-08-01' and d.to_date ='2002-08-01'and b.to_date ='2002-08-
01')f group by f.dept_name;
```

Result:

salary	dept_name
40000.0000	Customer Service
42093.0000	Production
48291.0000	Human Resources
48626.0000	Development
60026.0000	Finance
61357.0000	Quality Management
71166.0000	Marketing
71612.0000	Sales
72446.0000	Research

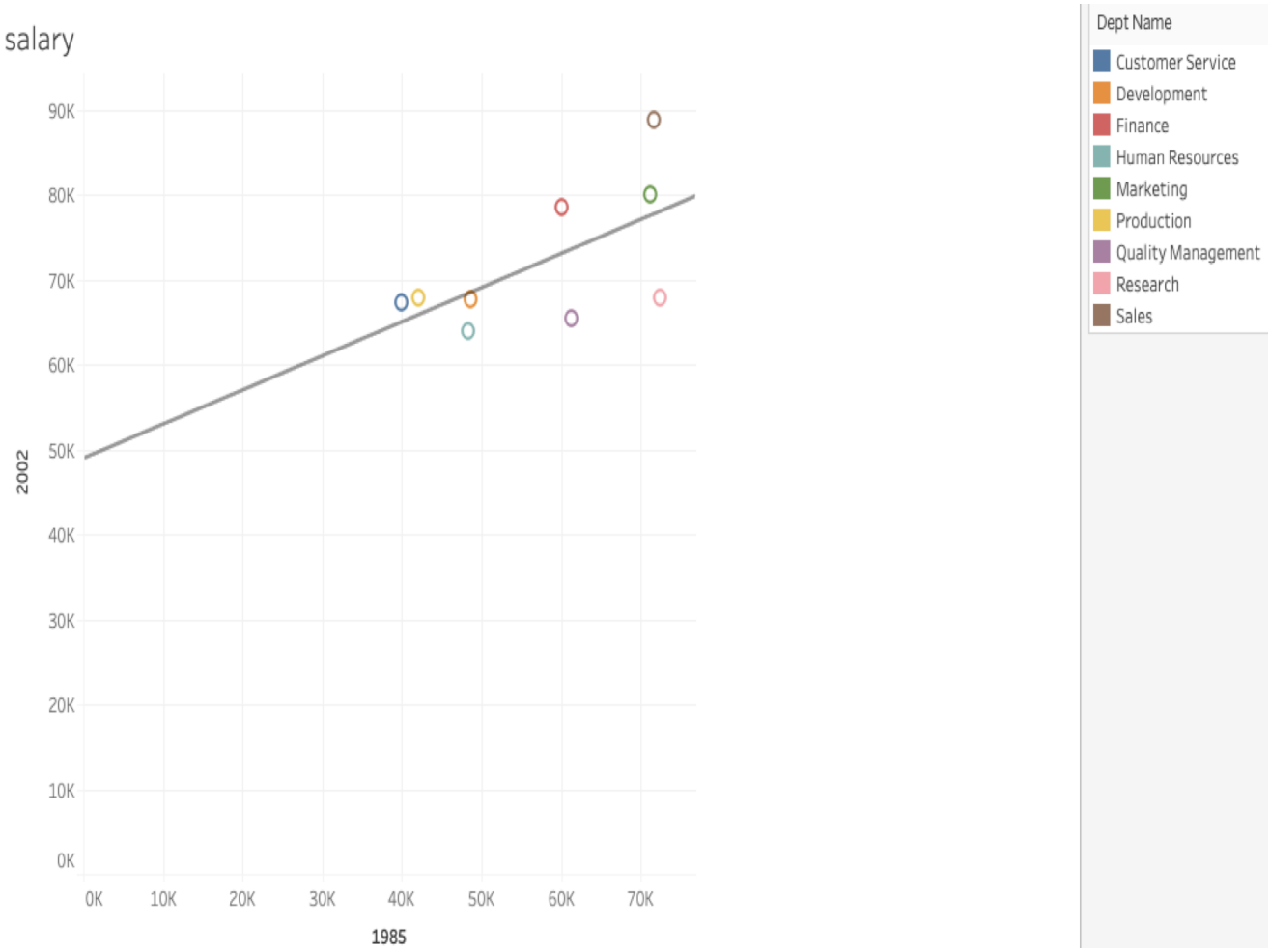
(Average salary in 1985)

salary	dept_name
63934.2343	Human Resources
65442.0659	Quality Management
67289.2184	Customer Service
67659.7426	Development
67838.1020	Production
67915.8959	Research
78555.4644	Finance
80048.0202	Marketing
88833.6909	Sales

(Average salary in 2002)

From the table we can see that in last year the top 3 salary department are sales, marketing and Finance, separately at 88833,80048 and 78555 euros but at the beginning of the establishment of the company, the highest average is in research department. What's more, the research department's average salary last year decrease comparing to 1985, which is the only department which decreases salary among the 9 departments.

Analysis/interpretation:



if we compare the growing range of salary in the 9 departments, we can easily find that sales are always the best department, and finance and marketing also keep a good increase trend. But the research, human resource, quality management and development department increase trend is below the average level of the company. Especially we find that there is a negative increase for the research department. It is not a good sign because as a software company, the technology that they own is always the most important thing because it may lead customers' daily habit and attract new customer. So a high salary for those technicians are important because it can retain the talented employees and attract new talented employees especially technicians. But in this company, what they do is inverse. They pay much more to the "untechnical department" employees.

Recommendations:

Increase the salary in “technical department” like development and research department. It is maybe the best time to hire some new and young knowledge technicians to enrich the code skill and the atmosphere in the company. Pay them more to encourage them to write code and that is a competitive core for a software firm.

#ADDITIONAL QUERY 5

Query:

Compare the average salaries between female and male in different departments and highlights which departments female and male have better talents.

(Because the codes in sql for each department are pretty similar, I will just show the codes in "Human-Resources" and "marketing" domain which have obvious gaps.)

```
-- For ALL departments:

 --MALE
select
employees.gender,
avg(salaries.salary) AS "male_average_salary"
from employees
join salaries
on employees.emp_no=salaries.emp_no
where gender="M";

-- FEMALE
select
employees.gender,
avg(salaries.salary) AS "female_average_salary"
from employees
join salaries
on employees.emp_no=salaries.emp_no
where gender="F";

-- For Marketing department :

--MALE
select
employees.gender,
avg(salaries.salary) as "male-marketing"
from employees
join salaries
on employees.emp_no=salaries.emp_no
join dept_emp
on employees.emp_no=dept_emp.emp_no
join departments
on dept_emp.dept_no=departments.dept_no
where gender="M" and salaries.to_date="9999-01-01" and dept_name="marketing";


--FEMALE
select
employees.gender,
avg(salaries.salary) as "female-Marketing"
```

```sql
from employees
join salaries
on employees.emp_no=salaries.emp_no
join dept_emp
on employees.emp_no=dept_emp.emp_no
join departments
on dept_emp.dept_no=departments.dept_no
where gender="F" and salaries.to_date="9999-01-01" and dept_name="marketing";

-- For Human-Resources department:

--MALE
select
employees.gender,
avg(salaries.salary) as "male-Human Resources"
from employees
join salaries
on employees.emp_no=salaries.emp_no
join dept_emp
on employees.emp_no=dept_emp.emp_no
join departments
on dept_emp.dept_no=departments.dept_no
where gender="M" and salaries.to_date="9999-01-01" and dept_name="Human Resources";

--FEMALE
select
employees.gender,
avg(salaries.salary) as "female-Human Resources"
from employees
join salaries
on employees.emp_no=salaries.emp_no
join dept_emp
on employees.emp_no=dept_emp.emp_no
join departments
on dept_emp.dept_no=departments.dept_no
where gender="F" and salaries.to_date="9999-01-01" and dept_name="Human Resources";
```

Results

gender	male_average_salary
M	63838.1769

gender	female_average_salary
F	63769.6032

gender	male-marketing
M	80244.0238

gender	female-Marketing
F	79664.2762

gender	male-Human resource
M	63641.0794

gender	female-Human resource
F	64025.5980

Analysis/interpretation:

There are 8% employees joining the company in December and it is a pretty reasonable ratio that each month there are some new blood in the company.
Besides the two departments (Human-resources and marketing) which I highlight. For other departments, the average gap between male and female are all within €150.Female does better in customer service, quality management and research domain while male does better job in other domains.

Recommendations:

From query7, we can see that the employees who have the highest salary (€158220) and the lowest salary (€38623) are all coming from the production department. In the same department, there is still a really big gap in salary which may trigger the dissatisfaction and sense of unfair among employees.
To solve this problem, firstly, increase employees' awareness of the value of their positions through open job evaluation methods and let employees understand the rationality and fairness of job evaluation.
Secondly, take multiple indictors into consideration for KPI. It is not comprehensive to evaluate people's value just through how much production they sale, how many clients they have served and how much value they create for company etc., but also their personal contribution for the company, such as some employees create a good team-build day and some employees share their good experiences with others to help them work better. Those employees who have different talents will definitely enrich the culture of the company and create a good atmosphere for the whole company.
Thirdly, the company can private the salaries and compensation of the employees to avoid the sense of unfairness.

As we can see from the additional query that there is a really small gap(€69) in the average salary between male and female. So I go deep into different departments and find that male has much more higher average salary than female which the gap is €580 while apparently female does a better job in human-resources domain.(the gap is €384). I think it is not bad for people in their preference domain and use their specialty, but it would be better that the company try to attract more male HR and marketing woman. That will not only enrich the culture of the whole department, but also give more different perspectives from different genders in different positions.

Conclusion:

Based on the following analysis, we can find that there are 2 main problems for the company. The first is that the composition of employees is not reasonable which shows in age and gender part. As a software company, their age composition is not reasonable. They have too many middle-aged employees, but no young employees and the majority employees' age are close. But young generation is the key for an internet company especially those high educated young people because they have the newest insight and technology skills. However, the employees' work is relatively stable in this firm which can be noticed from the decreasing hired employees every year. So, they don't have enough job position to hire new employees. It causes a vicious circle. What's more, the employees' age is close which may cause a wave of large number of retirements for the company after two decades. The gender quantity imbalance is must also not be overlooked in the company. No matter in different department or different position, there is always much more male employees than female. Even in some traditional female dominant department it is the same case which makes completely no sense.

We suggest that the company optimizes the employees' composition by hiring more young and high educated employees to verify the age composition and bring the newest thought and technology skill to the company. Meantime they should pay attention to the gender balance in the company and try not to have a huge gap between men and women. Besides, because of the huge number of employees in each department, we think it is better to set up more job positions especially in top management to help manager in each department.

The second is that the company doesn't attach enough important to the technical department. The research department's employees are not enough considering they are one of the core department in software company and the average salary in technical department (like development, research and production) is generally lower than untechnical department. So, we advise that the company increases the salary in technical department to encourage the technicians to work harder and it is also a way to attract those talent technicians in other company.