Exit Analysis

Table of contents

Introduction

In XYZ company, many employees leave each month. The stackholders want to know facts about employees who leave, so the recrutment will be ready and will keep an enough supply for hot jobs that employees leave it, this is on one hand. On the other hand, they want to know what rules have less lifespan more than others. In this analysis, we will explore the date from the employees' service certificates.

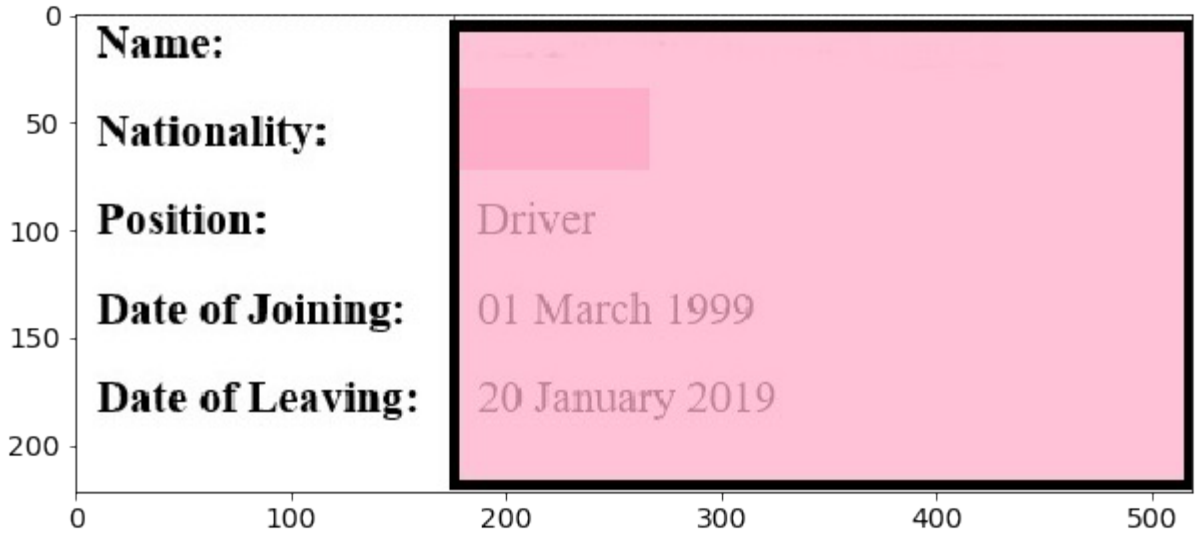
```
Imported Necessary libs
```

DATA Collection process

A script was prepared to collect data from word file forms (Service Certificates) and gathers the data in excel file

The information in the pink box were extracted for the prupose of this Analysis.

Out[6]: <matplotlib.image.AxesImage at 0x230f4247978>



After Extracting the data, the data was cleaned and prepared for the Analysis.

Explanatory Data Analysis

Out[9]:

	nat	job	join	empno	dept	service leave
2019-09-01	Nepalese	Merchandiser	2017-10-15	2344	General Products Department	1.879452
2019-09-01	Indian	Van Salesman	2014-09-13	1927	Consumer Products Department	4.969863
2019-09-01	Nepalese	Warehouse Asst	2012-05-12	1587	Logistics Department	7.309589
2019-09-01	Sri-Lankan	Sales Supervisor	2006-12-16	1107	Consumer Electronics & Home Appliances Department	12.717808
2019-08-28	Nepalese	Driver	2006-11-30	1093	Logistics Department	12.750685

Questions & Answers

What are the minimum, maximum and average

service period for employee?

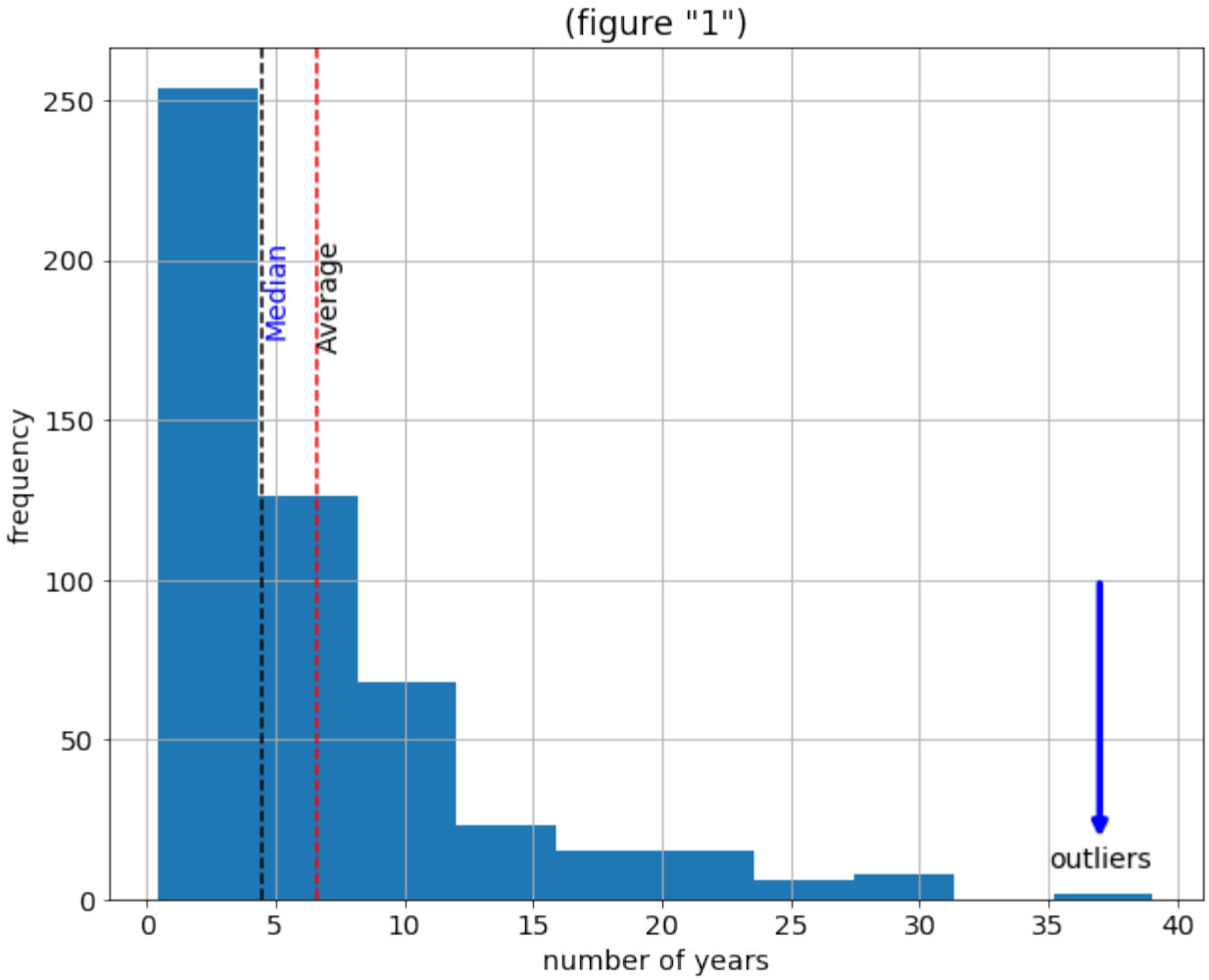
Out[12]:

	service
count	517.000000
mean	6.616942
std	6.319235
min	0.449315
25%	2.345205
50%	4.495890
75%	8.663014
max	39.057534

Explaining the summary

- The number of records is **517** service certificate.
- We can see that the average service period for employee is **6.6** years.
- We can see that the standard deviation is **6.3** years > which is considered high and means that data includes wide range of values.
- The minimum service period is **0.4** year which means 5 month.
- The first quartile is **2.35** years.
- The Second quartile - Median is **4.4** years.
- The third quartile is **8.6** years.
- The interquartile range (IQR) is 8.66 - 2.35 = **6.31** years.
- The maximum service period is **39** years.

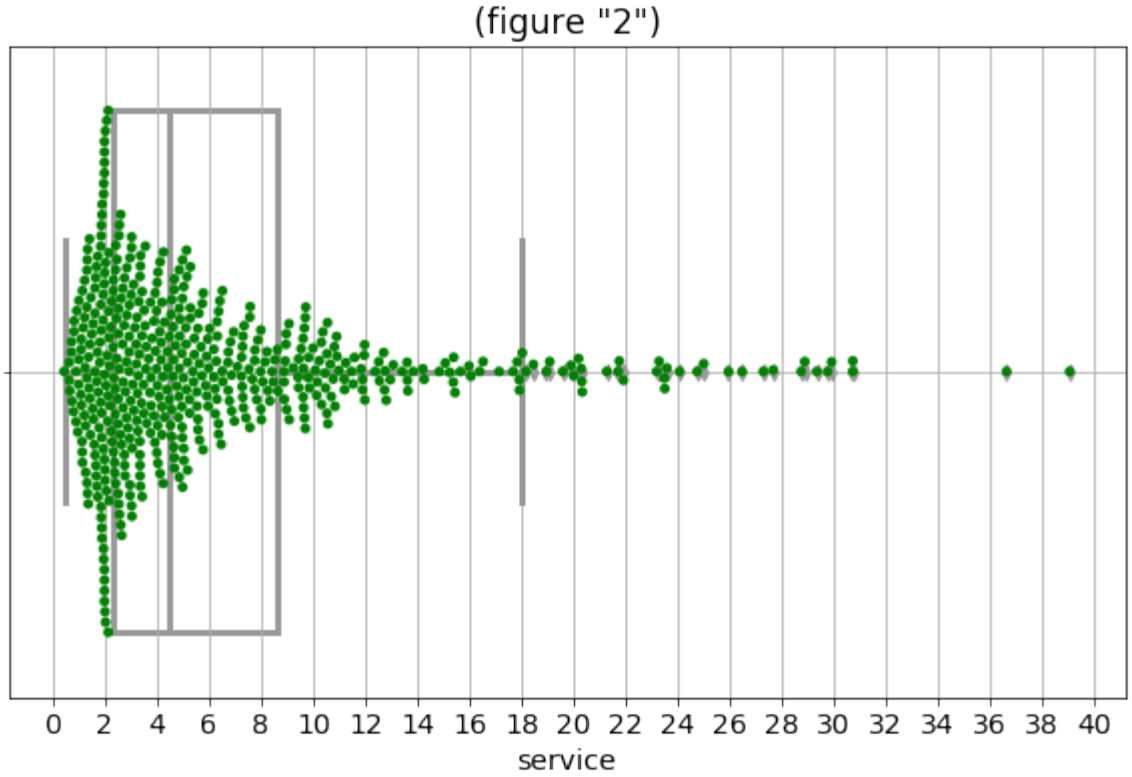
Out[13]: `Text(6.616941787446015, 200, 'Average')`



explaining histogram (figure 1)

- We notice that the sample has positive skewness.
- More than 250 employees did not compelete 5 years at least in the company
- we can see on the right side, there is a few number of employees they worked for more than 30 years, they are outliers. #### As we have outliers, using The Average (mean) ~6.6~ will be misleading, it is better to use median **4.4** as a less misleading value for the average of service period for the employee

Out[14]: `Text(0.5, 1.0, '(figure "2")')`



Explaining Box plot (figure 2)

We notice that most of population is between 0.4 and 8.5 years.

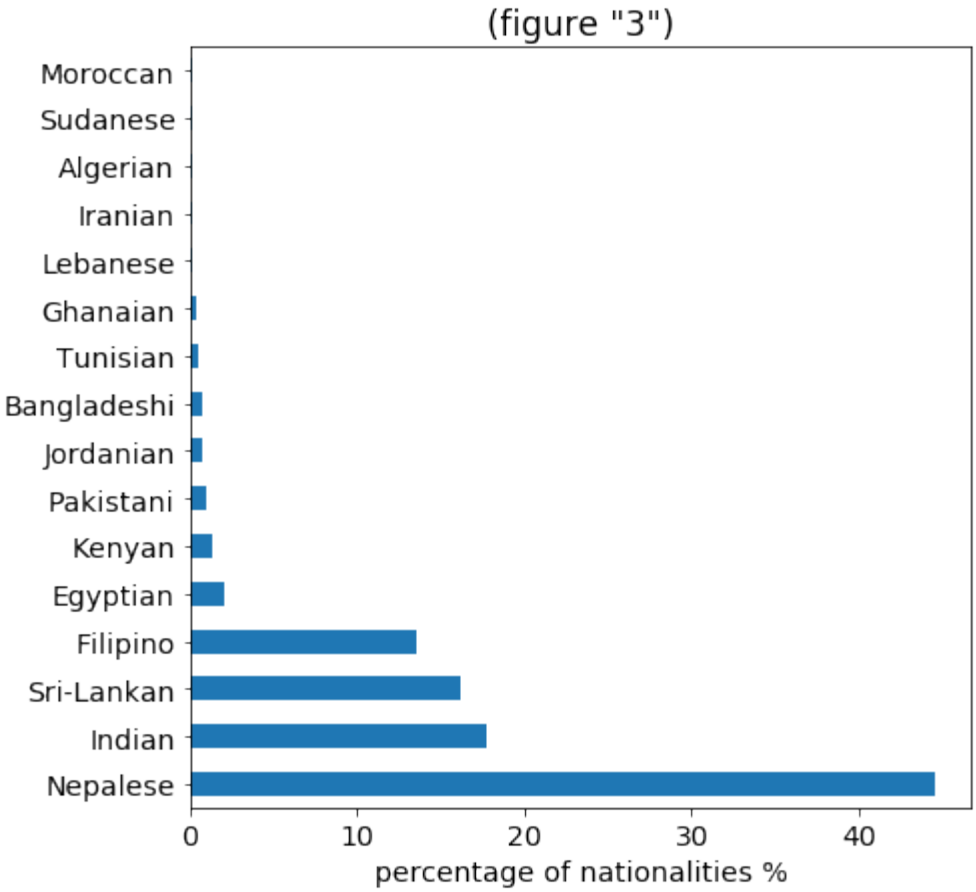
Employees who spent more than 18 years are considered outliers (too far from the average employees).

What are the nationalities that tend to leave the company more than the others?

Out[15]:

	percentages	counts
Nepalese	44.487427	230
Indian	17.794971	92
Sri-Lankan	16.247582	84
Filipino	13.539652	70
Egyptian	2.127660	11
Kenyan	1.353965	7
Pakistani	0.967118	5
Jordanian	0.773694	4
Bangladeshi	0.773694	4

Tunisian	0.580271	3
Ghanaian	0.386847	2
Lebanese	0.193424	1
Iranian	0.193424	1
Algerian	0.193424	1
Sudanese	0.193424	1
Moroccan	0.193424	1



Nationalities column statistics (figure 3)

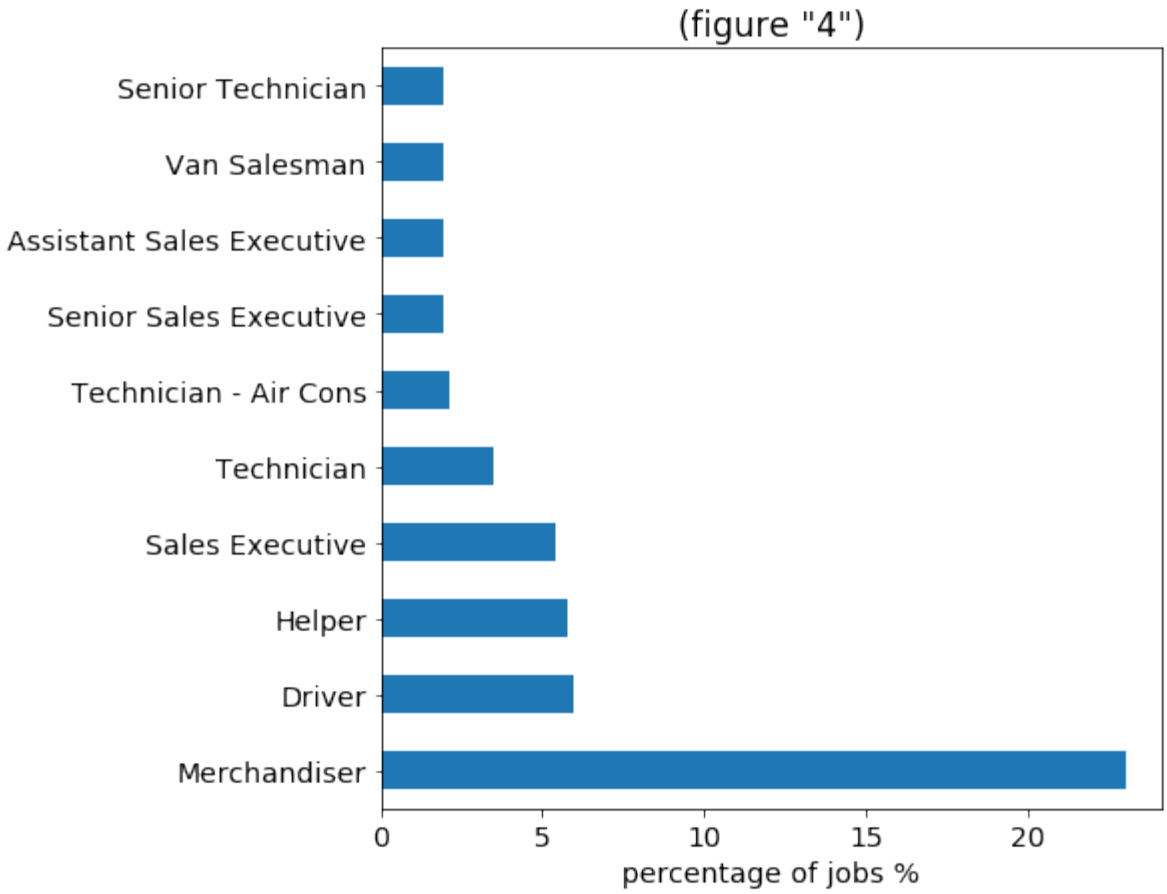
We notice that the dominant nationalities of the sample are Nepalese, Indians, Sri-Lankans and Filipinos (As they represent the most of the company's population)

What are the top 10 positions that they tend to leave the company more than others?

Out[16]:

	percentages	counts
Merchandiser	23.017408	119

Driver	5.996132	31
Helper	5.802708	30
Sales Executive	5.415861	28
Technician	3.481625	18
Technician - Air Cons	2.127660	11
Senior Sales Executive	1.934236	10
Assistant Sales Executive	1.934236	10
Van Salesman	1.934236	10
Senior Technician	1.934236	10



Positions/jobs column statistics (figure 4)

There are too many jobs, however we choose to show only the top 10 jobs in order to display them properly

We notice that most of percentage of the sample is Merchanidisers.

Knowing the hot jobs that they require recruting on fast pace will make the company always ready to face these changes without interrupting the workflow.

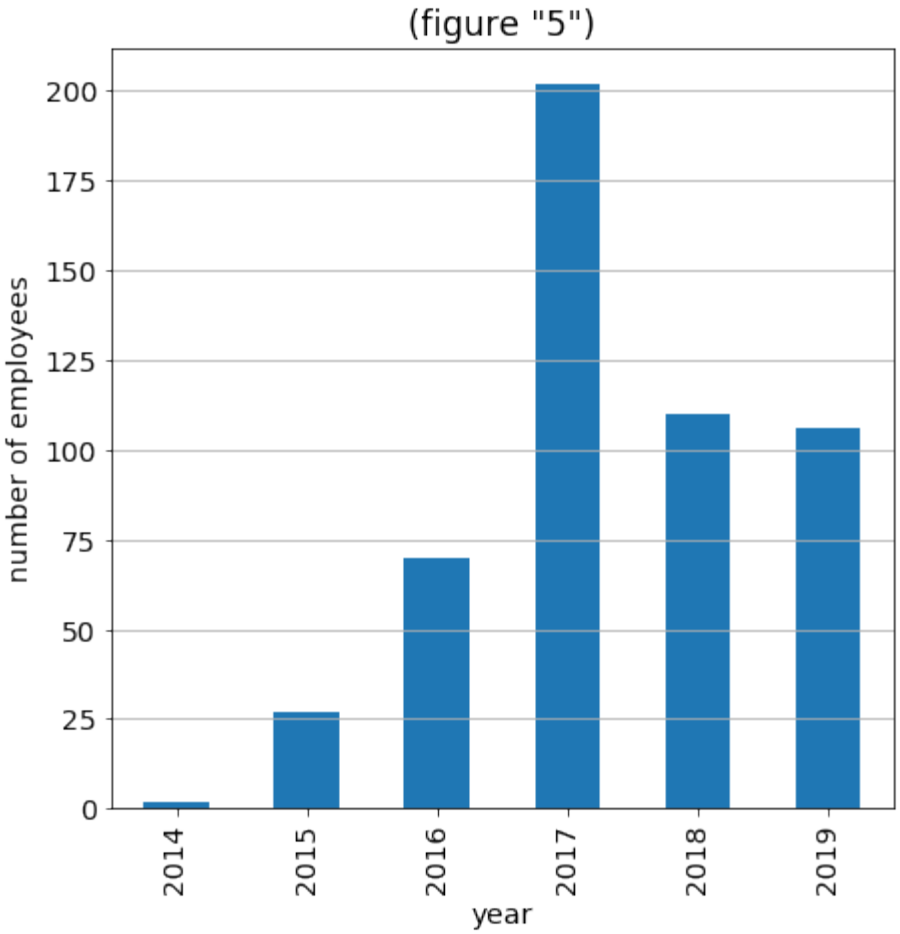
What is the average number of people they are leaving each month/quarter/year?

```
The expected amount of employees to leave per month are: 9.0
per Quarter: 27.0
per half year: 52.0
per year 103.0
```

General year-month Statistics

	nat												Total_per_year
leave	1	2	3	4	5	6	7	8	9	10	11	12	
leave													
2014	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	2
2015	0.0	0.0	0.0	0.0	2.0	0.0	0.0	1.0	11.0	5.0	4.0	4.0	27
2016	3.0	2.0	4.0	5.0	8.0	4.0	8.0	5.0	9.0	3.0	4.0	15.0	70
2017	12.0	12.0	25.0	4.0	31.0	7.0	23.0	30.0	7.0	22.0	16.0	13.0	202
2018	15.0	14.0	9.0	2.0	8.0	7.0	5.0	11.0	6.0	16.0	7.0	10.0	110
2019	23.0	8.0	12.0	9.0	10.0	6.0	9.0	9.0	11.0	4.0	5.0	0.0	106

Out[18]: `Text(0.5, 1.0, '(figure "5")')`



Year statistics (figure 5)

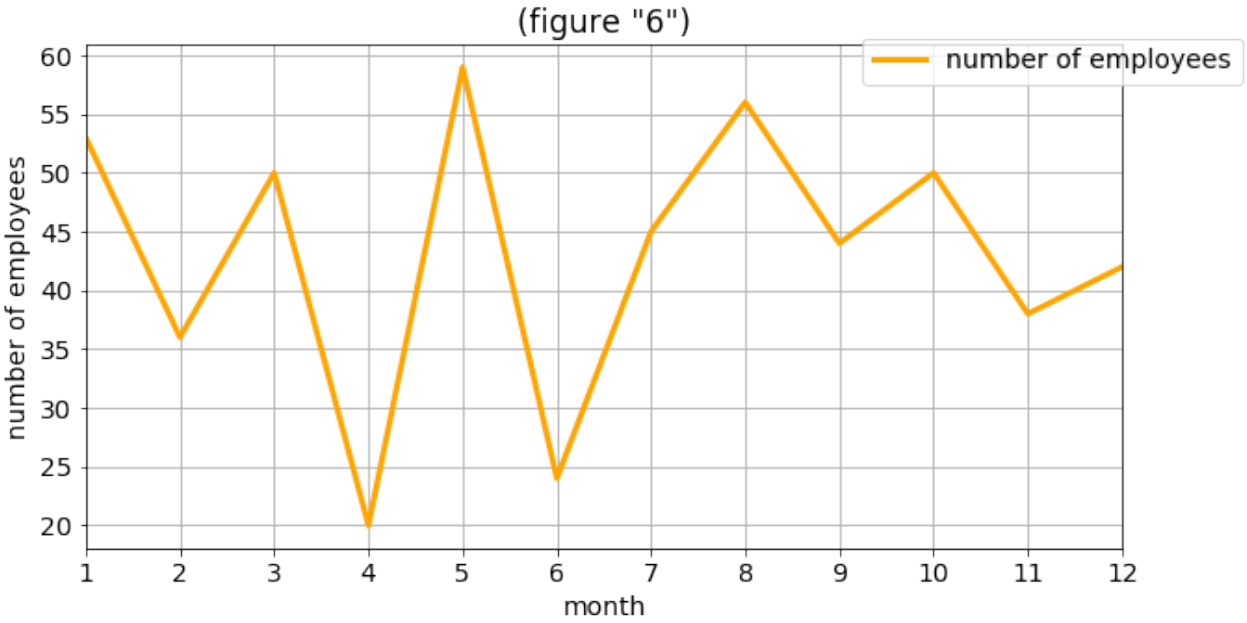
We notice from the table that 2014 the data is not complete and also in 2015.

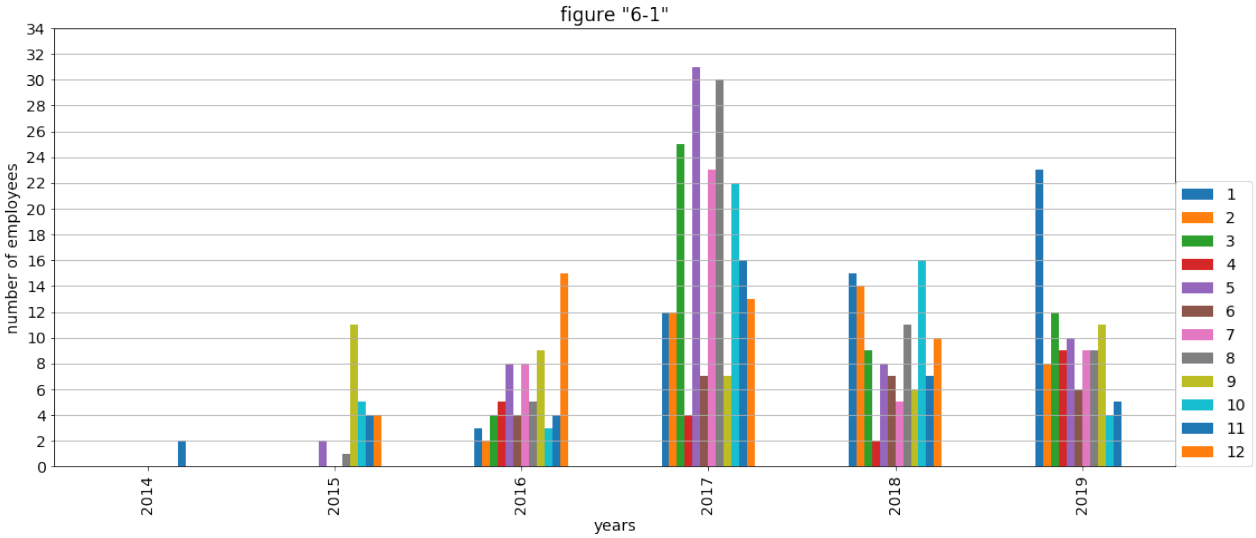
The years that have the full data of all months are 2016, 2017, and 2018.

As this analysis were made by the end of Nov. 2019, December is not completed yet.

2017 has the peak if we compared it to the other full years in 2016, 2018, which may indicate major incident or event affected the exit rate of employees.

Month over years	1	2	3	4	5	6	7	8	9	10	11	12
number of employees	53	36	50	20	59	24	45	56	44	50	38	42





Out[25]:

leave	1	2	3	4	5	6	7	8	9	10	11	12
leave												
2014	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0
2015	0.0	0.0	0.0	0.0	2.0	0.0	0.0	1.0	11.0	5.0	4.0	4.0
2016	3.0	2.0	4.0	5.0	8.0	4.0	8.0	5.0	9.0	3.0	4.0	15.0
2017	12.0	12.0	25.0	4.0	31.0	7.0	23.0	30.0	7.0	22.0	16.0	13.0
2018	15.0	14.0	9.0	2.0	8.0	7.0	5.0	11.0	6.0	16.0	7.0	10.0
2019	23.0	8.0	12.0	9.0	10.0	6.0	9.0	9.0	11.0	4.0	5.0	0.0

Month statistics (figure 6)

We notice from the table that may is highest month at all, however it is highest in the peak

year 2017 which affects the total.

Despite the fact that May is highest ever, in other years it is normal but we notice its always higher than April and June.

August had also higher rates in 2017 affected its performance in overall

Month statistics (figure 6-1)

providing a chart of all months in each year in order to compare them to each other

Out[35]:

job	Assistant Sales Executive	Driver	Helper	Merchandiser	Sales Executive	Senior Sales Executive	Senior Technician	Technician
1	0.0000	5.660	5.660000	22.6400	5.660000	5.660000	1.8900	0.0000
2	2.7800	0.000	2.780000	30.5600	5.560000	2.780000	0.0000	5.5600
3	2.0000	4.000	6.000000	24.0000	6.000000	4.000000	2.0000	0.0000
4	5.0000	10.000	5.000000	15.0000	20.000000	0.000000	0.0000	0.0000
5	6.7800	8.470	5.080000	20.3400	6.780000	0.000000	0.0000	5.0800
6	0.0000	12.500	0.000000	29.1700	4.170000	4.170000	0.0000	4.1700
7	0.0000	8.890	13.330000	15.5600	0.000000	0.000000	4.4400	0.0000
8	1.7900	7.140	1.790000	28.5700	1.790000	1.790000	0.0000	3.5700
9	0.0000	4.550	18.180000	20.4500	2.270000	0.000000	4.5500	4.5500
10	4.0000	8.000	4.000000	20.0000	6.000000	0.000000	6.0000	8.0000
11	0.0000	2.630	2.630000	34.2100	5.260000	2.630000	2.6300	5.2600
12	0.0000	2.380	2.380000	16.6700	9.520000	2.380000	0.0000	4.7600
Minimum	0.0000	0.000	0.000000	15.0000	0.000000	0.000000	0.0000	0.0000
Average	1.8625	6.185	5.569167	23.0975	6.084167	1.950833	1.7925	3.4125
Maximum	6.7800	12.500	18.180000	34.2100	20.000000	5.660000	6.0000	8.0000

Explaining table of top 10 jobs percentages

- **The table shows the percentage of each job that left in each month over all the years.**
- **The minimum and maximum and average are the expected rate for each of these jobs to leave each month**
  #### for example if there is 100 employees leave the company per month, it is expected that the average around 23% of them will be Merchandisers and up to 34% or more of total number of employees will be Merchandisers.
  ### **As we saw above the average employees they leave the company every month**

around 10 employees, we expect 2:3 or more to be Merchandisers.

What are the highest and lowest months?

```
'The lowest month'
```

	nat											
leave	1	2	3	4	5	6	7	8	9	10	11	12
leave												
2014	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2015	NaN	NaN	NaN	NaN	NaN	NaN	NaN	1.0	NaN	NaN	NaN	NaN
2016	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2017	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2018	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2019	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

```
'The highest month'
```

	nat											
leave	1	2	3	4	5	6	7	8	9	10	11	12
leave												
2014	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2015	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2016	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2017	NaN	NaN	NaN	NaN	31.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2018	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2019	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

We see that the lowest month is August 2015, this is because the employee was assigned the task of certificates prepration and the old certificates prepared upon request.

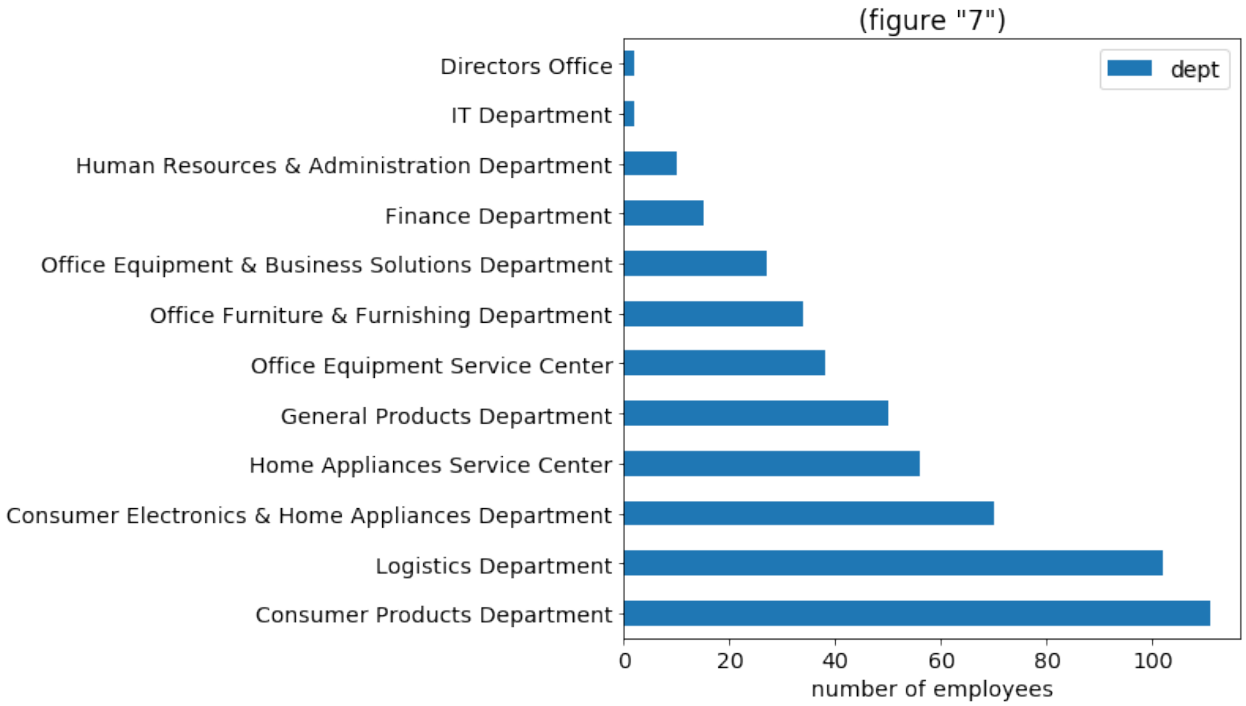
The maximum number was in May 2017.

What is the higest department in employees left and what are the departments that they have less

service periods?

	dept
Consumer Products Department	111
Logistics Department	102
Consumer Electronics & Home Appliances Department	70
Home Appliances Service Center	56
General Products Department	50
Office Equipment Service Center	38
Office Furniture & Furnishing Department	34
Office Equipment & Business Solutions Department	27
Finance Department	15
Human Resources & Administration Department	10
IT Department	2
Directors Office	2

Out[45]: Text(0.5, 1.0, '(figure "7")')



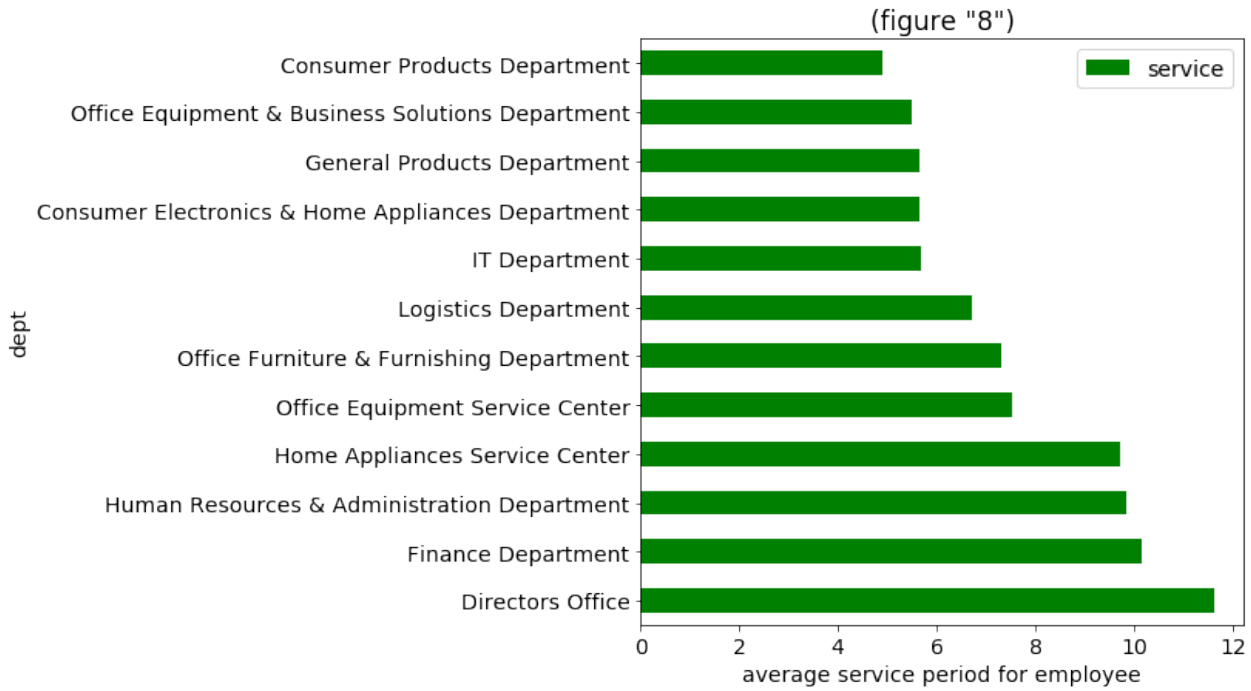
We notice (figure 7) that Consumer Products department has the highest exit rate.

The lowest rate in Directors office and IT as the total number of employees in these departments are extremely low.

We notice also that HR and Finance Departments are stable and have less number of employees.

dept	service
Directors Office	11.636986
Finance Department	10.147580
Human Resources & Administration Department	9.830137
Home Appliances Service Center	9.733953
Office Equipment Service Center	7.514348
Office Furniture & Furnishing Department	7.296454
Logistics Department	6.725813
IT Department	5.669863
Consumer Electronics & Home Appliances Department	5.659765
General Products Department	5.644932
Office Equipment & Business Solutions Department	5.488179
Consumer Products Department	4.905048

Out[46]: Text(0.5, 1.0, '(figure "8")')



In last table and plot (figure 8), if we excluded IT and Directors for their low numbers, we will find that Finance department has higher rates of employee service and stability and on the other hand Consumer Products Department is not stable and it has one of hot positions that requires continuous hiring.

Final Notes:

- The data were collected from word files.

- All the service certificates were prepared at the times of the employee vacation were not in the dataset, another employee was assigned to prepare the Certificates. Thus, there is some missing data.

- In order to improve the data and get more accurate results, it is better to acquire the data from the from the company system as it will be complete without any missing values.