This data set is gotten from the internet and used for Practise purpose
will explore and pull out the ncessary infomation that would be necessary.

However what this information was meant for is best know to those that posted it

- My find will be at the conclusion section
- Will tag this data set as Company Q

In [1]:
```
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline
import pandas as pd
```

Reading the csv file

In [27]:
```
staffs=pd.read_csv("Staffs_Ifo_data_Clean.csv")
```

Checking the setdata

In [3]:
```
staffs.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1000 entries, 0 to 999
Data columns (total 13 columns):
 #   Column           Non-Null Count  Dtype
---  ------           --------------  -----
 0   Full Name        1000 non-null   object
 1   State of Origin  1000 non-null   object
 2   Department       1000 non-null   object
 3   Unit             1000 non-null   object
 4   Bank             1000 non-null   object
 5   Basic            1000 non-null   object
 6   Housing          1000 non-null   object
 7   Transport        1000 non-null   object
 8   Meal             1000 non-null   object
 9   Utility          1000 non-null   object
 10  Entertainment    1000 non-null   object
 11  Leave            1000 non-null   object
 12  13th Month       1000 non-null   object
dtypes: object(13)
memory usage: 101.7+ KB
```

Checking the head of the data set

In [4]:
```
staffs.head()
```

Out[4]:

	Full Name	State of Origin	Department	Unit	Bank	Basic	Housing	Transport	Meal	Utility	Entertainment	Leave	13th Month
0	Olubunmi Adeyemi	Bauchi	Finance	Account Receivables	Sub-Standard Bank	3,268,505	150,000	20,000	5,000	10,000	6,000	326,851	227,181
1	Albert Ornams	Borno	Human Resources	Training & Development	Distressed Bank	2,787,420	150,000	20,000	5,000	10,000	6,000	278,742	195,430
2	Ibi-Kunle Ayodeji	Rivers	Admin.	MD's Office	Oasis Bank	725,915	150,000	20,000	5,000	10,000	6,000	72,592	60,493
3	Alihu Sarkin	Benue	Internal Control	Risk Management	Distressed Bank	469,436	150,000	20,000	5,000	10,000	6,000	46,944	42,443
4	Offiong Bassey	Rivers	IT	Desktop Support	Distressed Bank	3,338,673	150,000	20,000	5,000	10,000	6,000	333,867	231,812

Checking the tail of data set

In [5]:
```
staffs.tail(5)
```

Out[5]:

	Full Name	State of Origin	Department	Unit	Bank	Basic	Housing	Transport	Meal	Utility	Entertainment	Leave	13th Month
995	Rudayat Akinola	Rivers	Finance	Account Payables	Sub-Standard Bank	1,552,342	150,000	20,000	5,000	10,000	6,000	155,234	113,915
996	Onyemaechi Eyitayo	Ekiti	IT	Connectivity	Atlantic Bank	3,806,904	150,000	20,000	5,000	10,000	6,000	380,690	262,716
997	Olusola Ajobena	Ondo	IT	Business Application Support	Atlantic Bank	2,389,391	150,000	20,000	5,000	10,000	6,000	238,939	169,160
998	Adebukola Areo	Enugu	Admin.	Expenditure	Distressed Bank	795,035	150,000	20,000	5,000	10,000	6,000	79,504	63,932
999	Amena Ogunranti	Imo	Finance	Account Receivables	Last Bank Of Nigeria	2,122,051	150,000	20,000	5,000	10,000	6,000	212,205	151,515

Checking for missing

In [6]:
```
staffs.isnull().any()
```

```
Full Name          False
State of Origin    False
Department         False
Unit               False
Bank               False
Basic              False
Housing            False
Transport          False
Meal               False
Utility            False
Entertainment      False
Leave              False
13th Month         False
dtype: bool
```

In [7]:
```
staffs.isnull().sum()
```

```
Full Name          0
State of Origin    0
Department         0
Unit               0
Bank               0
Basic              0
Housing            0
Transport          0
Meal               0
Utility            0
Entertainment      0
Leave              0
13th Month         0
dtype: int64
```

(EDA) Exploratory Data Analysis

In [8]:
```
staffs.describe ()
```

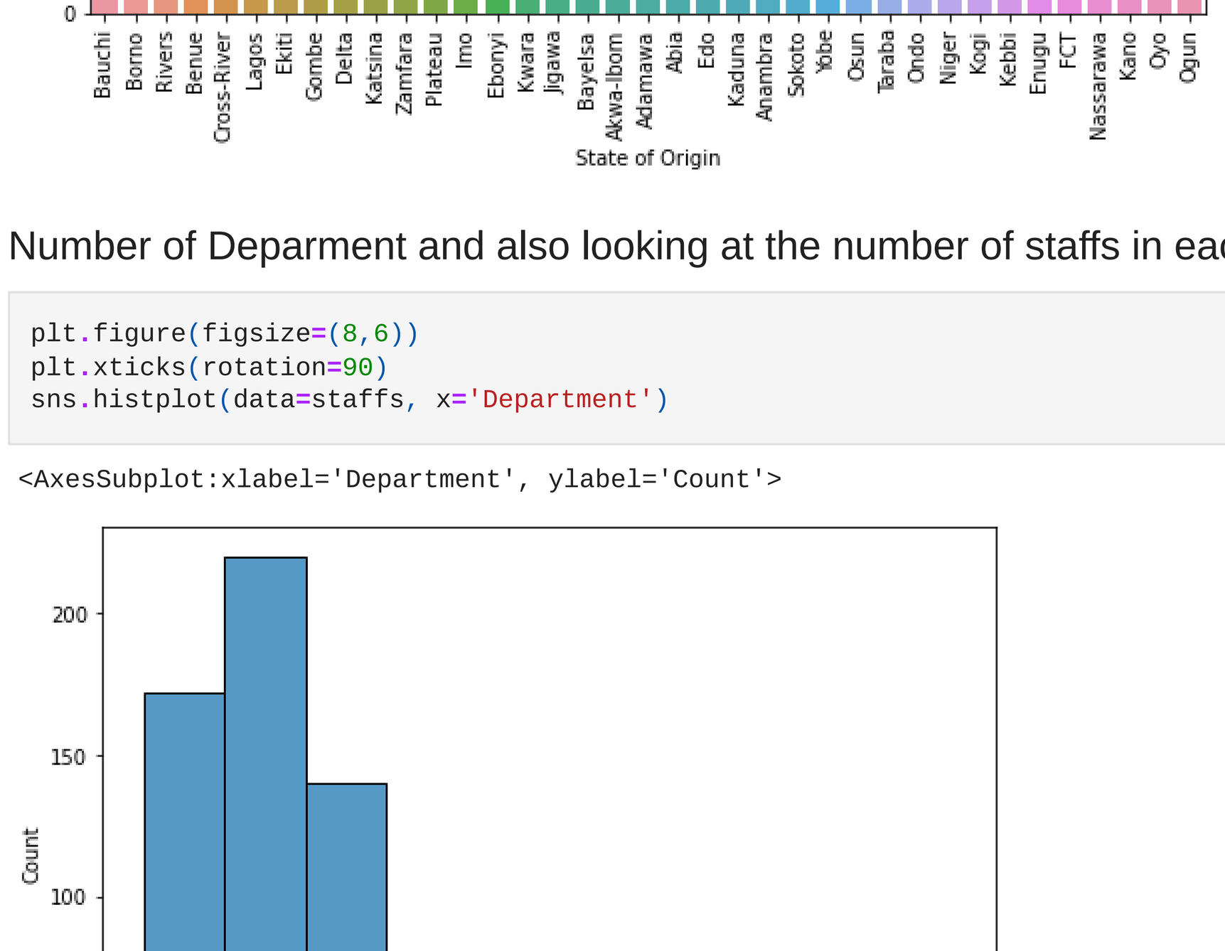
Out[8]:

	Full Name	State of Origin	Department	Unit	Bank	Basic	Housing	Transport	Meal	Utility	Entertainment	Leave	13th Month
count	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
unique	998	37	10	28	6	999	1	1	1	1	1	996	997
top	Mary Onwuachuke	Bauchi	Human Resources	Training & Development	Sub-Standard Bank	742,817	150,000	20,000	5,000	10,000	6,000	318,852	61,901
freq	2	86	220	74	268	2	1000	1000	1000	1000	1000	2	2

Checking which State has more people working for the company Q

In [9]:
```
plt.figure(figsize=(10,6))
plt.xticks(rotation=90)
sns.countplot(data=staffs, x ='State of Origin')
```

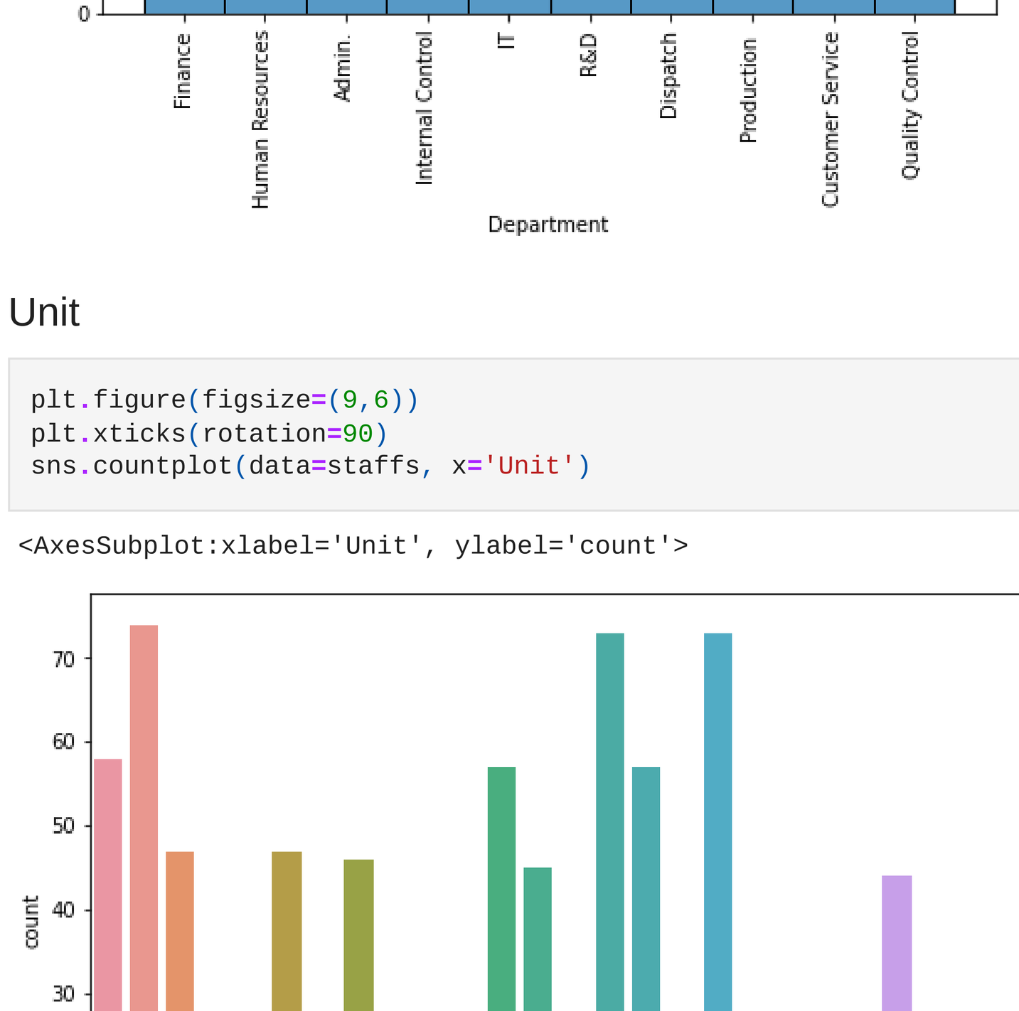
Out[9]: `<AxesSubplot:xlabel='State of Origin', ylabel='count'>`



Number of Deparment and also looking at the number of staffs in each Department

In [10]:
```
plt.figure(figsize=(8,6))
plt.xticks(rotation=90)
sns.histplot(data=staffs, x='Department')
```

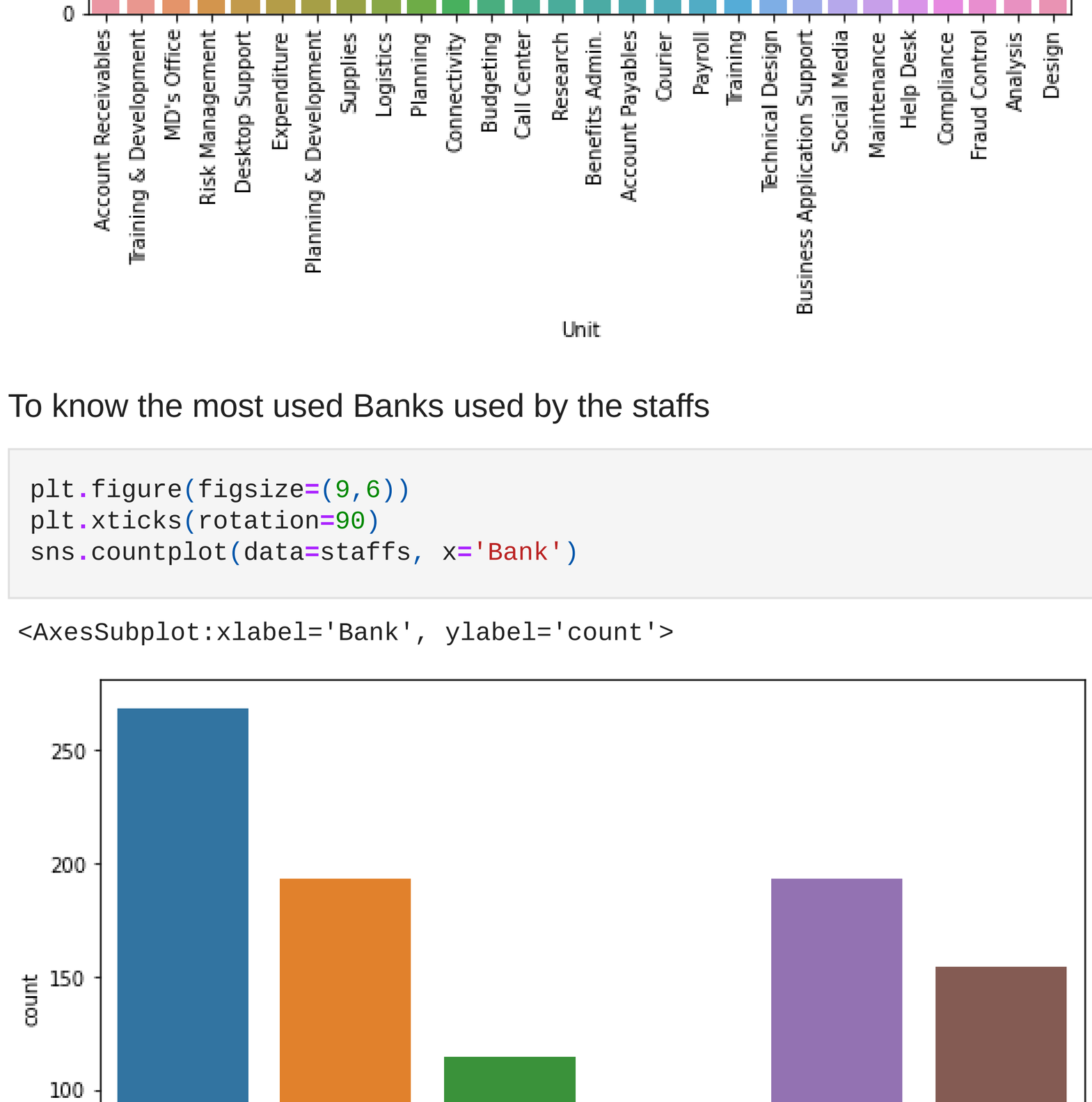
Out[10]: `<AxesSubplot:xlabel='Department', ylabel='Count'>`



Unit

In [11]:
```
plt.figure(figsize=(9,6))
plt.xticks(rotation=90)
sns.countplot(data=staffs, x='Unit')
```

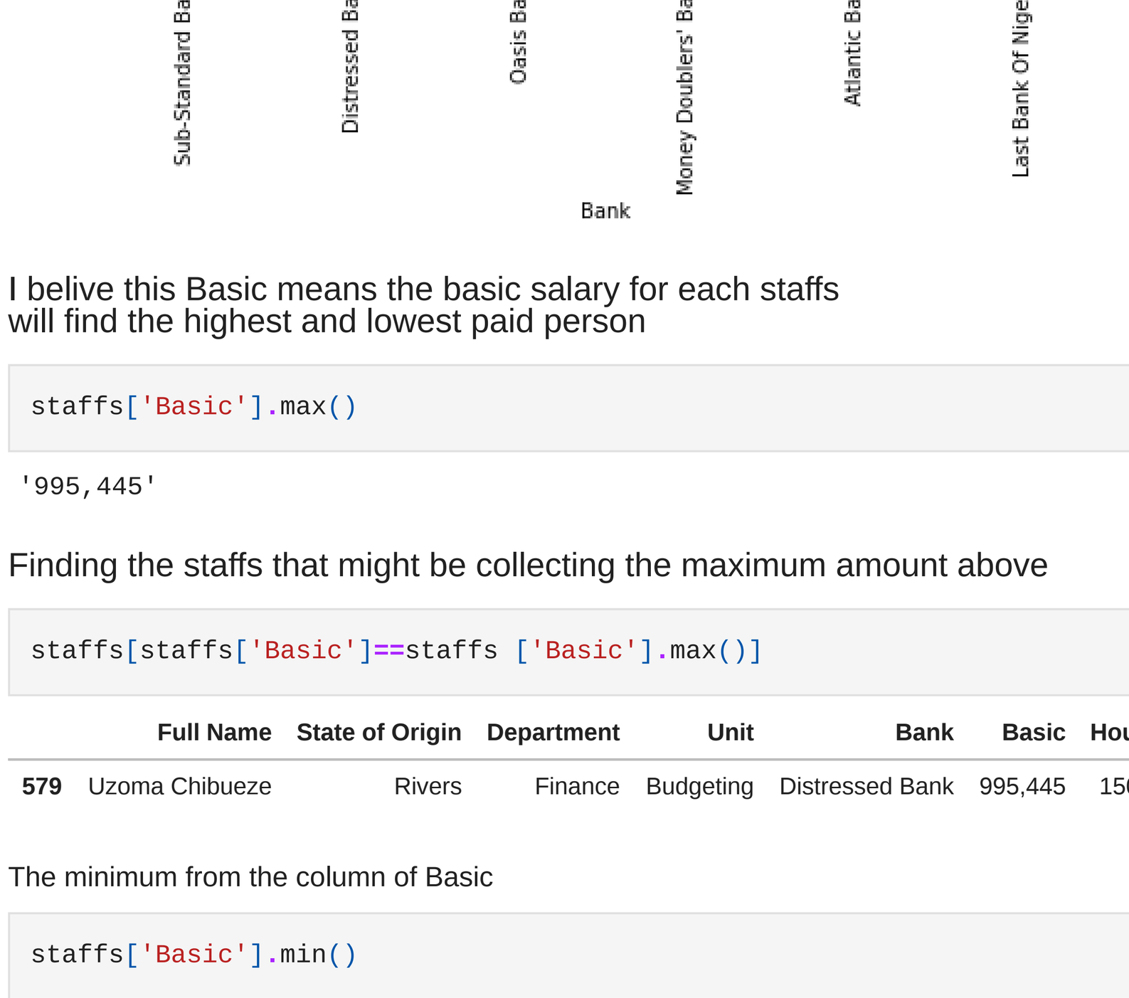
Out[11]: `<AxesSubplot:xlabel='Unit', ylabel='count'>`



To know the most used Banks used by the staffs

In [12]:
```
plt.figure(figsize=(9,6))
plt.xticks(rotation=90)
sns.countplot(data=staffs, x='Bank')
```

Out[12]: `<AxesSubplot:xlabel='Bank', ylabel='count'>`



I belive this Basic means the basic salary for each staffs
will find the highest and lowest paid person

In [13]:
```
staffs['Basic'].max()
```

Out[13]: '995,445'

Finding the staffs that might be collecting the maximum amount above

In [14]:
```
staffs[staffs['Basic']==staffs ['Basic'].max()]
```

Out[14]:

	Full Name	State of Origin	Department	Unit	Bank	Basic	Housing	Transport	Meal	Utility	Entertainment	Leave	13th Month
579	Uzoma Chibueze	Rivers	Finance	Budgeting	Distressed Bank	995,445	150,000	20,000	5,000	10,000	6,000	99,545	77,159

The minimum from the column of Basic

In [15]:
```
staffs['Basic'].min()
```

Out[15]: '1,002,442'

Finding the staffs that might be collecting the minimum amount above

In [16]:
```
staffs [staffs['Basic']==staffs['Basic'].min()]
```

Out[16]:

	Full Name	State of Origin	Department	Unit	Bank	Basic	Housing	Transport	Meal	Utility	Entertainment	Leave	13th Month
398	Akeem Adedokun	Katsina	Admin.	Expenditure	Atlantic Bank	1,002,442	150,000	20,000	5,000	10,000	6,000	100,244	77,621

This varaible below are all fix for all the staffs there is no need to do an EDA 5 varible for this.

In [17]:
```
staffs[['Housing','Transport','Meal','Utility','Entertainment']]
```

Out[17]:

	Housing	Transport	Meal	Utility	Entertainment
0	150,000	20,000	5,000	10,000	6,000
1	150,000	20,000	5,000	10,000	6,000
2	150,000	20,000	5,000	10,000	6,000
3	150,000	20,000	5,000	10,000	6,000
4	150,000	20,000	5,000	10,000	6,000
...	...	...	...	...	...
995	150,000	20,000	5,000	10,000	6,000
996	150,000	20,000	5,000	10,000	6,000
997	150,000	20,000	5,000	10,000	6,000
998	150,000	20,000	5,000	10,000	6,000
999	150,000	20,000	5,000	10,000	6,000

1000 rows × 5 columns

which stasffs got the highest Leave paid

In [18]:
```
staffs[staffs['Leave']== staffs['Leave'].max()]
```

Out[18]:

	Full Name	State of Origin	Department	Unit	Bank	Basic	Housing	Transport	Meal	Utility	Entertainment	Leave	13th Month
579	Uzoma Chibueze	Rivers	Finance	Budgeting	Distressed Bank	995,445	150,000	20,000	5,000	10,000	6,000	99,545	77,159

which staffs got the minimum

In [19]:
```
staffs[staffs['Leave']== staffs['Leave'].min()]
```

Out[19]:

	Full Name	State of Origin	Department	Unit	Bank	Basic	Housing	Transport	Meal	Utility	Entertainment	Leave	13th Month
76	Ibhade Omoyeni	Sokoto	Human Resources	Payroll	Sub-Standard Bank	104,316	150,000	20,000	5,000	10,000	6,000	10,432	18,345

the same code will be run for the 13th month

In [20]:
```
staffs['13th Month'].max()
```

Out[20]: '99,189'

In [21]:
```
staffs[staffs['13th Month'] == staffs['13th Month'].max()]
```

Out[21]:

	Full Name	State of Origin	Department	Unit	Bank	Basic	Housing	Transport	Meal	Utility	Entertainment	Leave	13th Month
279	Kayode Michael	Gombe	Internal Control	Risk Management	Distressed Bank	1,329,230	150,000	20,000	5,000	10,000	6,000	132,923	99,189

Minimum of 13th Month

In [22]:
```
staffs['13th Month'].min()
```

Out[22]: '100,732'

In [31]:
```
staffs[staffs['13th Month']==staffs['13th Month'].min()]
```

Out[31]:

	Full Name	State of Origin	Department	Unit	Bank	Basic	Housing	Transport	Meal	Utility	Entertainment	Leave	13th Month
124	Olaitan Asiegbu	Borno	R&D	Analysis	Atlantic Bank	1,352,608	150,000	20,000	5,000	10,000	6,000	135,261	100,732

Find the correlation coefficient between the

In [40]:
```
staffs['Department']= staffs['Basic'].apply(len)
```

In [43]:
```
staffs[['Department','Basic']].corr()
```

Out[43]:

	Department
Department	1.0

My Conclusion

Going through the data which was gotten from the internter

- It shows all the state are not even distributed but is either here nor there but base on their qualification each person have for the post
- it show there should be a reduction in the unite (offices) and staffs also
- There is a correlation between the Department and the Basic (Basic salary) this implies that there is an increase of salary at each Department

In [ ]: