HR ANALYTICS PROJECT SQL REPORT

BY GAURAV CHANDRA

```
--##Data Cleaning
--## Here ,i have added a column to the table which has value 1 or 0 depending if attrition is yes or no
,
--## as by summing and counting this column, we would get better insights about attrition

alter table dbo.HR_Analytics
add attrition_value int

update dbo.HR_Analytics set attrition_value =
case
when attrition = 'Yes' then 1
else 0
end from dbo.HR_Analytics

select * from dbo.HR_Analytics

--as we can see that attrition_value column have been added at the end of the table
```

ıy	YearsInCurrentRole	YearsSinceLastPromotion	YearsWithCurrManager	attrition_value
	0	0	0	0
	9	2	8	0
	10	1	8	0
	10	15	6	1
	0	0	0	0
	3	1	2	0
	15	13	8	1
	0	0	0	1
	0	0	4	0
	0	0	0	1

H\SQLEXPRESS ... | LAPTOP-6U8M0FEH\chand ... | HR analytics | 00:00:00 | 1,480 rows

```
--##We had a extra column in our table which has blanks values and column isn't of much use for us for
visualizaton purposes.

alter table dbo.HR_Analytics drop column YearsWithCurrManager


--##Lets check if we have any duplicate rows in our table.
select empid,ROW_NUMBER() over (partition by empid order by empid) as num_rows
from dbo.HR_Analytics
order by 2 desc
```

	empid	num_rows
1	RM1470	2
2	RM1461	2
3	RM1462	2
4	RM1463	2
5	RM1464	2
6	RM1465	2
7	RM1466	2
8	RM1467	2
9	RM1468	2
10	RM1469	2
11	RM147	1

```sql
--##So the data has some duplicate rows in it, so I decided to first import all distinct rows in another table and then
--##I truncated our original table and then added the distinct rows into the table.
--##NOTE:--one can easily do all these steps within excel and power bi , but I still chose to do all the work via SQL and just
--## use these results in power bi to make visualizations ,just to get most out of SQL.

select distinct * into dbo.HR2
from dbo.HR_Analytics

truncate table dbo.HR_analytics

insert into dbo.HR_Analytics select * from dbo.HR2;

drop table dbo.HR2

select empid,ROW_NUMBER() over (partition by empid order by empid) as num_rows
from dbo.HR_Analytics
order by 2 desc
---##We can see that duplicate rows are now removed, empid with 2as rownumber have different values for column
```

	empid	rn
1	RM001	1
2	RM002	1
3	RM003	1
4	RM004	1
5	RM005	1
6	RM006	1
7	RM007	1
8	RM008	1
9	RM009	1
10	RM010	1
11	RM011	1

```sql
--##Here, the data has some Travel_Rarely and TravelRarely as two different values in the Buisness Travel column,so I decided to update the values

select BusinessTravel  from dbo.HR_Analytics
group  by BusinessTravel
```

Results Messages

	BusinessTravel
1	TravelRarely
2	Non-Travel
3	Travel_Rarely
4	Travel_Frequently

```sql
update dbo.HR_Analytics
set BusinessTravel='Travel_Rarely'
where BusinessTravel='TravelRarely'

select BusinessTravel  from dbo.HR_Analytics
group  by BusinessTravel
```

	BusinessTravel
1	Non-Travel
2	Travel_Rarely
3	Travel_Frequently

```
--##Our Data is now cleaned,The Next step is to do some analysis on data ,
--##Here ,I have created different views from our data which includes Attrition by
Age,Department,Education Field,Job Role,Salary,Years at company ,Gender and overtime.
--##We can further import these saved views in power bi, and do the analysis from there.
--##NOTE:--If the following view are already created one first needs to run the drop view command and
then the following queries.

--Table for Attrition vs Age
drop view if exists AttVSAge

create view AttVSAge as(
select AgeGroup,sum(Attrition_value) as [number of employee attrited] ,count(agegroup)as [Employee
count],(convert(float,sum(Attrition_value))/convert(float,count(AgeGroup)) )*100 as Attrition_percentage
from dbo.HR_Analytics
group by AgeGroup)

select * from AttVSAge
```

Results Messages

	AgeGroup	number of employee attrited	Employee count	Attrition_percentage
1	36-45	43	468	9.18803418803419
2	18-25	44	123	35.7723577235772
3	46-55	26	226	11.5044247787611
4	26-35	116	606	19.1419141914191
5	55+	8	47	17.0212765957447

```
--Table for Attrition vs Department
drop view if exists AttVSDep

create view AttVSDep as
select Department,sum(Attrition_value) as [number of employee attrited] ,count(Department)as
TotalEmployeePerDepartment,(convert(float,sum(Attrition_value))/convert(float,count(Department)) )*100 as
Attrition_percentage
from dbo.HR_Analytics
group by Department

select * from AttVSDep
```

Results Messages

	Department	number of employee attrited	TotalEmployeePerDepartment	Attrition_percentage
1	Sales	92	446	20.627802690583
2	Research & Development	133	961	13.8397502601457
3	Human Resources	12	63	19.047619047619

```
--Table for Attrition vs Educational Field
drop view if exists AttVSEdu

create view AttVSEdu as
select EducationField,sum(Attrition_value) as [number of employee attrited] ,count(EducationField)as
TotalEmployeePerEducation
,(convert(float,sum(Attrition_value))/convert(float,count(EducationField)) )*100 as Attrition_percentage
from dbo.HR_Analytics
group by EducationField

select * from AttVSEdu
```

	EducationField	number of employee attrited	TotalEmployeePerEducation	Attrition_percentage
1	Technical Degree	32	132	24.2424242424242
2	Marketing	35	159	22.0125786163522
3	Life Sciences	89	606	14.6864686468647
4	Medical	63	464	13.5775862068966
5	Human Resources	7	27	25.9259259259259
6	Other	11	82	13.4146341463415

```sql
--Table for Attrition vs Job role
drop view if exists AttVSJob

create view AttVSJob as
select JobRole,sum(Attrition_value) as [number of employee attrited] ,count(JobRole)as
TotalEmployeePerJobrole
,(convert(float,sum(Attrition_value))/convert(float,count(JobRole)) )*100 as Attrition_percentage
from dbo.HR_Analytics
group by JobRole

select * from AttVSJob
```

	JobRole	number of employee attrited	TotalEmployeePerJobrole	Attrition_percentage
1	Sales Representative	33	83	39.7590361445783
2	Manager	5	102	4.90196078431373
3	Healthcare Representative	9	131	6.87022900763359
4	Laboratory Technician	62	259	23.9382239382239
5	Sales Executive	57	326	17.4846625766871
6	Manufacturing Director	10	145	6.89655172413793
7	Human Resources	12	52	23.0769230769231
8	Research Director	2	80	2.5
9	Research Scientist	47	292	16.0958904109589

```sql
--Table for Attrition vs Salary Slab
drop view if exists AttVSSal

create view AttVSSal as
select SalarySlab,sum(Attrition_value) as [number of employee attrited] ,count(SalarySlab)as
TotalEmployeePerSalarySlab
,(convert(float,sum(Attrition_value))/convert(float,count(SalarySlab)) )*100 as Attrition_percentage
from dbo.HR_Analytics
group by SalarySlab

select * from AttVSSal
```

	SalarySlab	number of employee attrited	TotalEmployeePerSalarySlab	Attrition_percentage
1	5k-10k	49	440	11.1363636363636
2	10k-15k	20	148	13.5135135135135
3	15k+	5	133	3.7593984962406
4	Upto 5k	163	749	21.762349799733

```sql
--Table for Attrition vs Years at company
drop view if exists AttVSYrs

create view AttVSYrs as
select YearsAtCompany,sum(Attrition_value) as [number of employee attrited] ,count(YearsAtCompany)as
TotalEmployeePerYearsAtCompany
,(convert(float,sum(Attrition_value))/convert(float,count(YearsAtCompany)) )*100 as Attrition_percentage
from dbo.HR_Analytics
group by YearsAtCompany

select * from AttVSYrs
```

	YearsAtCompany	number of employee attrited	TotalEmployeePerYearsAtCompany	Attrition_percentage
1	0	16	44	36.3636363636364
2	31	1	3	33.3333333333333
3	14	2	18	11.1111111111111
4	30	0	1	0
5	6	9	76	11.8421052631579
6	15	1	20	5
7	29	0	2	0
8	3	20	128	15.625
9	13	2	24	8.33333333333333
10	32	1	3	33.3333333333333
11	16	1	12	8.33333333333333

```sql
--Table for Attrition vs Overtime
drop view if exists AttVSOverT

create view AttVSOverT as
select OverTime,sum(Attrition_value) as [number of employee attrited] ,count(OverTime)as
TotalEmployeePerOverTime
,(convert(float,sum(Attrition_value))/convert(float,count(OverTime)) )*100 as Attrition_percentage
from dbo.HR_Analytics
group by OverTime

select * from AttVSOverT
```

	OverTime	number of employee attrited	TotalEmployeePerOverTime	Attrition_percentage
1	Yes	127	416	30.5288461538462
2	No	110	1054	10.4364326375712

```sql
--Table for Attrition vs gender
drop view if exists AttVSGen

create view AttVSGen as
select gender,sum(Attrition_value) as [number of employee attrited] ,count(gender)as
TotalEmployeePerGender
,(convert(float,sum(Attrition_value))/convert(float,count(Gender)) )*100 as Attrition_percentage
from dbo.HR_Analytics
group by Gender

select * from AttVSGen
```

	gender	number of employee attrited	TotalEmployeePerGender	Attrition_percentage
1	Male	150	882	17.0068027210884
2	Female	87	588	14.7959183673469

```sql
--##one can use the cte here too,

with cte (AgeGroup,emp_attri,count_age)as(
select AgeGroup,sum(Attrition_value) as [emp_attri] ,count(agegroup)
from dbo.HR_Analytics
group by AgeGroup)
select *,(convert(float,emp_attri)/convert(float,count_age))*100
from cte
order by emp_attri desc
```

	AgeGroup	emp_attri	count_age	(No column name)
1	26-35	116	606	19.1419141914191
2	18-25	44	123	35.7723577235772
3	36-45	43	468	9.18803418803419
4	46-55	26	226	11.5044247787611
5	55+	8	47	17.0212765957447