databricks Employee Attrition Project

(https://databricks.com/)

```
%sql
select * from emp_attrition
```

Table

	Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome	Education
1	41	Yes	Travel_Rarely	1102	Sales	1	2
2	49	No	Travel_Frequently	279	Research & Development	8	1
3	37	Yes	Travel_Rarely	1373	Research & Development	2	2
4	33	No	Travel_Frequently	1392	Research & Development	3	4
5	27	No	Travel_Rarely	591	Research & Development	2	1
6	32	No	Travel_Frequently	1005	Research & Development	2	2
7	59	No	Travel_Rarely	1324	Research & Development	3	3

1,470 rows

Total Employee Count

```
%sql
select sum(EmployeeCount) from emp_attrition
```

Table

	sum(EmployeeCount)
1	1470

1 row

FIND OUR ATTRITION DIVISION

```
%sql
SELECT SUM(EmployeeCount) FROM emp_attrition GROUP BY Attrition
```

Table

	sum(EmployeeCount)
1	1233
2	237

2 rows

AGE Analysis -- lets find out which particular age attrition is high (18-24,25-31,32-38,39-45,46-52,52+)

```
%sql
SELECT SUM(EmployeeCount),
case when age BETWEEN 20 and 25 then '20-25' when age BETWEEN 26 and 32 then '26-32' when age BETWEEN 33 and 40 then
'33-40' else '40+' end age_group
FROM emp_attrition where Attrition = 'Yes' group by 2
```

TableVisualization 1

	sum(EmployeeCount)	age_group
1	34	20-25
2	56	33-40
3	85	26-32
4	62	40+

4 rows

most of the people are leaving the org between 26-32

FIND OUT Attrition by Department

```
select count(EmployeeCount),department from emp_attrition where Attrition = 'Yes' group by Department;
```

TableVisualization 1

	count(EmployeeCount)	department
1	92	Sales
2	133	Research & Development
3	12	Human Resources

3 rows

Attrition by education (1-below college, 2-college, 3-bachelor, 4-masters,5-doctor)

```
select count(EmployeeCount),
case when Education = 1 then 'below college' when Education=2 then 'college' when Education=3 then 'bachelor' when
Education=4 then 'master' else 'doctor' end ed_group
 from  emp_attrition where Attrition = 'Yes' group by Education
```

TableVisualization 1

	count(EmployeeCount)	ed_group
1	31	below college
2	99	bachelor
3	5	doctor
4	58	master
5	44	college

5 rows

Attrition by Environment Satisfaction (1-low, 2-medium, 3-high,4-higly satisfied)

```
select count(EmployeeCount),
EnvironmentSatisfaction
from  emp_attrition where Attrition = 'Yes' group by EnvironmentSatisfaction
```

TableVisualization 1

	count(EmployeeCount)	EnvironmentSatisfaction
1	72	1
2	62	3
3	60	4
4	43	2

4 rows

Attrition by business travel

```
select count(EmployeeCount), BusinessTravel from emp_attrition where Attrition = 'Yes' group by 2
```

TableVisualization 1

	count(EmployeeCount)	BusinessTravel
1	69	Travel_Frequently
2	12	Non-Travel
3	156	Travel_Rarely

3 rows

Insights we got as of now:

1.employee age between 26-32 leaving organization more. 2.research and devlopment department left the organization. 3.41.8% people having bachelors degree left org. 4.72 employee not satisfied with environment location left org. 5.156 employee travelling rarely left org.