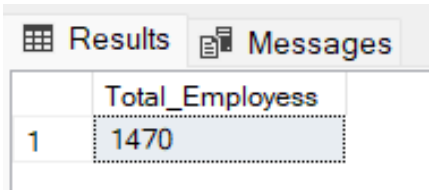
DATA CLEANING PROJECT

1.Total Employees

```
SELECT COUNT(*) AS Total_Employess FROM HR;
```

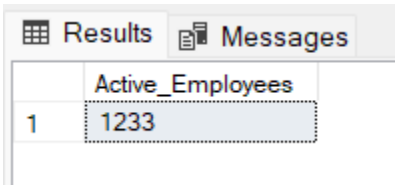


Results | Messages

	Total_Employess
1	1470

2.Active Employees

```
SELECT COUNT(*) AS Active_Employees FROM HR WHERE Attrition = 0;
```

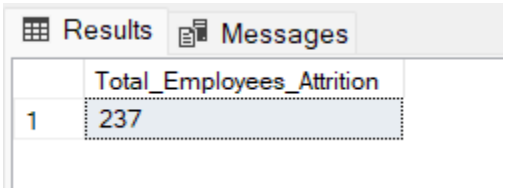


Results | Messages

	Active_Employees
1	1233

3.Attrition Employees

```
SELECT COUNT(*) AS Total_Employees_Attrition FROM HR WHERE Attrition = 1;
```

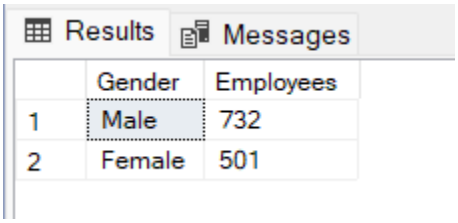


Results | Messages

	Total_Employees_Attrition
1	237

4.Gender Distribution of Active Employees

```
SELECT Gender, COUNT(*) AS Employees FROM HR WHERE Attrition = 0 GROUP BY Gender;
```



Results | Messages

	Gender	Employees
1	Male	732
2	Female	501

5.Age distribution of Employees

```sql
SELECT
    CASE
        WHEN Age BETWEEN 18 AND 23 THEN '18-23'
        WHEN Age BETWEEN 24 AND 29 THEN '24-29'
        WHEN Age BETWEEN 30 AND 35 THEN '30-35'
        WHEN Age BETWEEN 36 AND 41 THEN '36-41'
        WHEN Age BETWEEN 42 AND 47 THEN '42-47'
        WHEN Age BETWEEN 48 AND 60 THEN '48-60'
        ELSE '60+'
    END AS Age_group,COUNT(Age) AS Employees_Count
FROM HR
GROUP BY
    CASE
        WHEN Age BETWEEN 18 AND 23 THEN '18-23'
        WHEN Age BETWEEN 24 AND 29 THEN '24-29'
        WHEN Age BETWEEN 30 AND 35 THEN '30-35'
        WHEN Age BETWEEN 36 AND 41 THEN '36-41'
        WHEN Age BETWEEN 42 AND 47 THEN '42-47'
        WHEN Age BETWEEN 48 AND 60 THEN '48-60'
        ELSE '60+'
    END
ORDER BY Age_group;
```

Results Messages

	Age_group	Employees_Count
1	18-23	71
2	24-29	255
3	30-35	403
4	36-41	316
5	42-47	209
6	48-60	216

6.Attrition trends by Age Group

```sql
SELECT
    CASE
        WHEN Age BETWEEN 18 AND 23 THEN '18-23'
        WHEN Age BETWEEN 24 AND 29 THEN '24-29'
        WHEN Age BETWEEN 30 AND 35 THEN '30-35'
        WHEN Age BETWEEN 36 AND 41 THEN '36-41'
        WHEN Age BETWEEN 42 AND 47 THEN '42-47'
        WHEN Age BETWEEN 48 AND 60 THEN '48-60'
        ELSE '60+'
    END AS Age_Group,

    COUNT(CASE WHEN Attrition = 1 THEN 'Yes' END) AS Attrition_Yes,
    COUNT(CASE WHEN Attrition = 0 THEN 'No' END) AS Attrition_No

FROM HR
GROUP BY
    CASE
```

```
        WHEN Age BETWEEN 18 AND 23 THEN '18-23'
        WHEN Age BETWEEN 24 AND 29 THEN '24-29'
        WHEN Age BETWEEN 30 AND 35 THEN '30-35'
        WHEN Age BETWEEN 36 AND 41 THEN '36-41'
        WHEN Age BETWEEN 42 AND 47 THEN '42-47'
        WHEN Age BETWEEN 48 AND 60 THEN '48-60'
        ELSE '60+'
    END
ORDER BY Age_Group;
```

Results | Messages

	Age_Group	Attrition_Yes	Attrition_No
1	18-23	31	40
2	24-29	60	195
3	30-35	69	334
4	36-41	31	285
5	42-47	19	190
6	48-60	27	189

7.Gender Attrition Analysis

```
SELECT Gender, COUNT(*) AS Employees FROM HR WHERE Attrition = 1 GROUP BY Gender ORDER BY
COUNT(*) DESC;
```

Results | Messages

	Gender	Employees
1	Male	150
2	Female	87

8.Employees Active and Attrition by Department

```
SELECT Department,
    COUNT(CASE WHEN Attrition = 1 THEN 'Yes' END) AS Attrition_Yes,
    COUNT(CASE WHEN Attrition = 0 THEN 'No' END) AS Attrition_No
FROM HR
GROUP BY Department
ORDER BY COUNT(Attrition);
```

Results | Messages

	Department	Attrition_Yes	Attrition_No
1	Human Resources	12	51
2	Sales	92	354
3	Research & Development	133	828

9.Education Field by Employees

```
SELECT EducationField, COUNT(Attrition) AS Employees

FROM HR

GROUP BY EducationField

ORDER BY COUNT(Attrition);
```

Results Messages

	EducationField	Employees
1	Human Resources	27
2	Other	82
3	Technical Degree	132
4	Marketing	159
5	Medical	464
6	Life Sciences	606

10.Impact of Education on Employees Turnover

```
SELECT EducationField,
    COUNT(CASE WHEN Attrition = 1 THEN 'Yes' END) AS Attrition_Yes,
    COUNT(CASE WHEN Attrition = 0 THEN 'No' END) AS Attrition_No
FROM HR
GROUP BY EducationField
ORDER BY COUNT(Attrition);
```

Results Messages

	EducationField	Attrition_Yes	Attrition_No
1	Human Resources	7	20
2	Other	11	71
3	Technical Degree	32	100
4	Marketing	35	124
5	Medical	63	401
6	Life Sciences	89	517

11.Employees by Job Level

```
SELECT JobLevel,COUNT(Attrition) AS Employees
FROM HR
GROUP BY JobLevel
ORDER BY COUNT(Attrition);
```

Results Messages

	JobLevel	Employees
1	Executive Level	69
2	Lead Level	106
3	Senior Level	218
4	Mid Level	534
5	Entry Level	543

12.Attrition Trends Across Different job levels

```
SELECT JobLevel,
    COUNT(CASE WHEN Attrition = 1 THEN 'Yes' END) AS Attrition_Yes,
    COUNT(CASE WHEN Attrition = 0 THEN 'No' END) AS Attrition_No
FROM HR
GROUP BY JobLevel
ORDER BY COUNT(Attrition);
```

Results Messages

	JobLevel	Attrition_Yes	Attrition_No
1	Executive Level	5	64
2	Lead Level	5	101
3	Senior Level	32	186
4	Mid Level	52	482
5	Entry Level	143	400

13.Employees by Job Roles

```
SELECT JobRole,COUNT(Attrition) AS Employees
FROM HR
GROUP BY JobRole
ORDER BY COUNT(Attrition);
```

Results Messages

	JobRole	Employees
1	Human Resources	52
2	Research Director	80
3	Sales Representative	83
4	Manager	102
5	Healthcare Representative	131
6	Manufacturing Director	145
7	Laboratory Technician	259
8	Research Scientist	292
9	Sales Executive	326

14.Analyzing Departures by Job Roles

```
SELECT JobRole,
 COUNT(CASE WHEN Attrition = 1 THEN 'Yes' END) AS Attrition_Yes,
 COUNT(CASE WHEN Attrition = 0 THEN 'No' END) AS Attrition_No
FROM HR
GROUP BY JobRole
ORDER BY COUNT(Attrition);
```

Results Messages

	JobRole	Attrition_Yes	Attrition_No
1	Human Resources	12	40
2	Research Director	2	78
3	Sales Representative	33	50
4	Manager	5	97
5	Healthcare Representative	9	122
6	Manufacturing Director	10	135
7	Laboratory Technician	62	197
8	Research Scientist	47	245
9	Sales Executive	57	269

15.Marital Status Insights

```
SELECT MaritalStatus,
 COUNT(CASE WHEN Attrition = 1 THEN 'Yes' END) AS Attrition_Yes,
 COUNT(CASE WHEN Attrition = 0 THEN 'No' END) AS Attrition_No
FROM HR
GROUP BY MaritalStatus
ORDER BY COUNT(Attrition);
```

Results Messages

	MaritalStatus	Attrition_Yes	Attrition_No
1	Divorced	33	294
2	Single	120	350
3	Married	84	589

16.Attrition by Years of Service

```
SELECT
    CASE
        WHEN YearsAtCompany BETWEEN 0 AND 7 THEN '0-7'
        WHEN YearsAtCompany BETWEEN 8 AND 15 THEN '8-15'
        WHEN YearsAtCompany BETWEEN 16 AND 23 THEN '16-23'
        WHEN YearsAtCompany BETWEEN 24 AND 31 THEN '24-31'
        ELSE '32-40'
    END AS Years_Group,
          COUNT(CASE WHEN Attrition = 1 THEN 'Yes' END) AS Attrition_Yes,
        COUNT(CASE WHEN Attrition = 0 THEN 'No' END) AS Attrition_No
FROM HR
GROUP BY CASE
        WHEN YearsAtCompany BETWEEN 0 AND 7 THEN '0-7'
        WHEN YearsAtCompany BETWEEN 8 AND 15 THEN '8-15'
        WHEN YearsAtCompany BETWEEN 16 AND 23 THEN '16-23'
        WHEN YearsAtCompany BETWEEN 24 AND 31 THEN '24-31'
        ELSE '32-40'
    END
ORDER BY Years_Group;
```

Results Messages

	Years_Group	Attrition_Yes	Attrition_No
1	0-7	182	760
2	16-23	8	95
3	24-31	2	20
4	32-40	3	10
5	8-15	42	348