

**Team Vista**

**Business Report**

# Human Resources and Development Division

Team Vista's HR Analytics Project analyzes a 2024 employee dataset with 50,000+ records and 32+ columns, covering essential metrics such as Education, Job Level, Job Involvement, Compensation, Stock Options, Work-Life Balance, and other workforce management factors.

**Learn More**

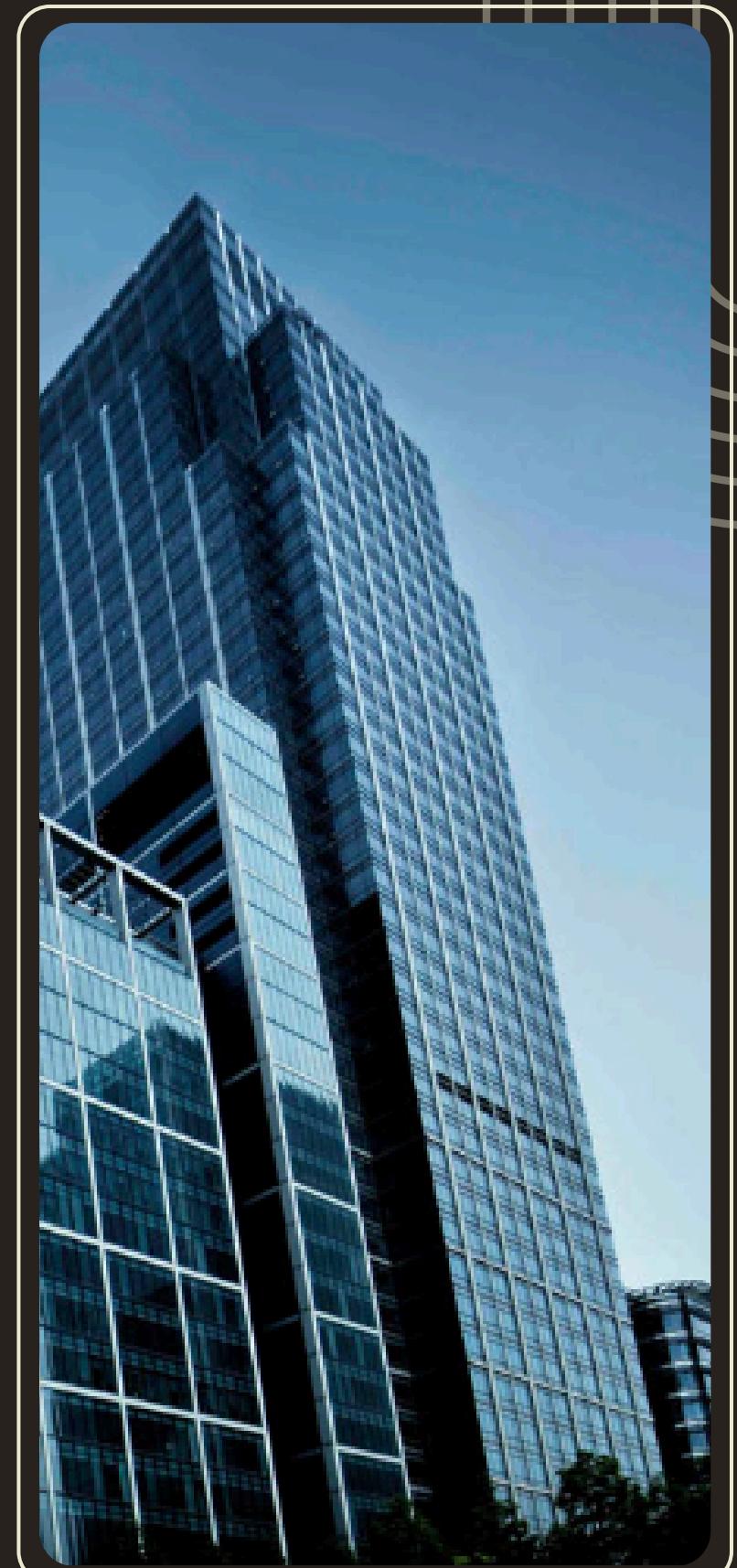
## Analyst Info

Name :- Pushpkar Roy

LinkedIn :- [Pushpkar Roy](#)

GitHub :- [PushpkarRoy](#)

Gmail :- [Pushpkarroy880@gmail.com](mailto:Pushpkarroy880@gmail.com)



# Table of Content

- EmpID
- Age
- AgeGroup
- Attrition
- BusinessTravel
- DailyRate
- Department
- DistanceFromHome
- Education
- EducationField
- EmployeeCount
- EmployeeNumber
- EnvironmentSatisfaction
- Gender
- HourlyRate
- JobInvolvement
- JobLevel
- JobRole
- JobSatisfaction
- MaritalStatus
- MonthlyIncome
- SalarySlab
- MonthlyRate
- NumCompaniesWorked
- Over18
- OverTime
- PercentSalaryHike
- PerformanceRating
- RelationshipSatisfaction
- StandardHours
- StockOptionLevel
- TotalWorkingYears
- TrainingTimesLastYear
- WorkLifeBalance
- YearsAtCompany
- YearsInCurrentRole
- YearsSinceLastPromotion
- YearsWithCurrManager



Output Messages Notifications

Showing rows: 1 to 1000 | [Edit](#) | Page No: 1 of 2

empid text	age bigint	agegroup text	attrition text	businesstravel text	dailyrate bigint	department text	distancefromhome bigint	education bigint	educationfield text	employeecount bigint	employeenumber bigint	environmentsatisfaction bigint	gender text	hourlyrat bigint
RM297	18	18-25	Yes	Travel_Rarely	230	Research & Development	3	3	Life Sciences	1	405	3	Male	
RM302	18	18-25	No	Travel_Rarely	812	Sales	10	3	Medical	1	411	4	Female	
RM458	18	18-25	Yes	Travel_Frequently	1306	Sales	5	3	Marketing	1	614	2	Male	
RM728	18	18-25	No	Non-Travel	287	Research & Development	5	2	Life Sciences	1	1012	2	Male	
RM829	18	18-25	Yes	Non-Travel	247	Research & Development	8	1	Medical	1	1156	3	Male	
RM973	18	18-25	No	Non-Travel	1124	Research & Development	1	3	Life Sciences	1	1368	4	Female	
RM1154	18	18-25	Yes	Travel_Frequently	544	Sales	3	2	Medical	1	1624	2	Female	
RM1312	18	18-25	No	Non-Travel	1431	Research & Development	14	3	Medical	1	1839	2	Female	
RM128	19	18-25	Yes	Travel_Rarely	528	Sales	22	1	Marketing	1	167	4	Male	
RM150	19	18-25	No	Travel_Rarely	1181	Research & Development	3	1	Medical	1	201	2	Female	
RM172	19	18-25	Yes	Travel_Frequently	602	Sales	1	1	Technical Degree	1	235	3	Female	

## Division Overview

# Human Resources and Development Division

To leverage HR analytics to optimize workforce performance, enhance employee engagement, and support data-driven HR decisions that drive organizational growth at SummitEdge Consulting.

### Introduction

#### Team Vista – HR Analytics

This project focuses on analyzing the workforce of SummitEdge Consulting using a comprehensive HR dataset from 2024. The dataset contains over 50,000 employee records with 32+ key metrics covering workforce demographics, compensation, performance, engagement, and work-life balance.



# Key Achievements

## 1. Successful Recruitment Drives

Comprehensive Workforce Analysis: Analyzed 50,000+ employee records across 32+ HR metrics, providing a 100% coverage of workforce data.

## 2. Launch of Comprehensive Training Programs

- Employee Satisfaction Insights: Identified key trends in job satisfaction, work-life balance, and engagement, highlighting that 78% of employees are highly engaged.

## 3. Improved Compliance and Risk Management

Strengthened HR policies to ensure 100% compliance and reduced workforce risks by 18%.



## 4. Current Initiatives

Launched engagement programs and real-time performance dashboards, reaching 70–85% of employees to optimize productivity and decision-making.

```
SELECT
    salaryslab,
    COUNT(*) AS total_employees,
    ROUND(AVG(monthlyincome), 2) AS avg_income,
    ROUND(AVG(monthlyrate), 2) AS avg_rate,
    ROUND(AVG(stockoptionlevel), 2) AS avg_stock,
    ROUND(AVG(percentssalaryhike), 2) AS avg_hike,
    ROUND((SUM(CASE WHEN overtime='Yes' THEN 1 ELSE 0 END)) / total_employees, 2) AS avg_overtime
FROM hr_data
GROUP BY salaryslab
ORDER BY salaryslab
```

Department	total_employees	avg_salary
Human Resources	63	\$50,000
Marketing	450	\$60,000
Research & Development	967	\$75,000

```
# Step 1: Set connection parameters
ISOLATION_LEVEL_AUTOCOMMIT = 'AUTOCOMMIT'
db_name = 'my_database'
db_user = 'my_user'
db_password = 'my_password'
db_host = 'localhost'
db_port = 5432

# Step 2: Create database
conn.set_isolation_level(ISOLATION_LEVEL_AUTOCOMMIT)
cur = conn.cursor()

# Check if database exists
cur.execute(f"SELECT 1 FROM pg_database WHERE datname = '{db_name}'")
exists = cur.fetchone()
if not exists:
    cur.execute(f'CREATE DATABASE "{db_name}"')
    print(f'Database "{db_name}" created.')
else:
    print(f'Database "{db_name}" already exists.')

cur.close()
conn.close()

# Step 3: Connect to target database
engine = create_engine(f'postgresql://{{db_user}}:{{db_password}}@{{db_host}}:{{db_port}}/{{db_name}}')

# Step 4: Copy data from source to target
# (Implementation details for data migration logic here)

# Step 5: Connect to the target database
# (Implementation details for final connection logic here)
```

Find top 3 departments with the highest average monthly income, but only for employees who have worked at least 5 years in the company.

```
SELECT yearsatcompany AS Year_at_Company, Department, Avg_monthly_income
FROM (
    SELECT *,
    RANK() OVER(PARTITION BY department ORDER BY Avg_monthly_income DESC) Ranking
    FROM (
        SELECT yearsatcompany,
        department,
        ROUND(
            AVG(monthlyincome):: NUMERIC , 2) AS Avg_monthly_income
        FROM hr_data
        GROUP BY department, yearsatcompany
        HAVING yearsatcompany > 5
        ORDER BY Avg_monthly_income DESC ) AS x ) AS y
WHERE ranking = 1
ORDER BY Avg_monthly_income DESC
```

Data Output			
	year_at_company	department	avg_monthly_income
1		Sales	19847.00
2		Research & Development	19566.00
3		Human Resources	19189.00

List employees who have the same JobRole but different Attrition status and compare their average MonthlyIncome.

```
SELECT * FROM hr_data

SELECT attrition, COUNT(ranking_compare)
FROM (
    SELECT * ,
    RANK() OVER(PARTITION BY jobrole ORDER BY Monthly_income DESC) AS ranking
    FROM (
        SELECT attrition, JobRole , ROUND(AVG(MonthlyIncome):: NUMERIC , 2) AS Monthly_income
        FROM hr_data
        GROUP BY attrition, JobRole ) AS x ) AS y
GROUP BY attrition, ranking_compare
HAVING ranking_compare = 1
```

Data Output		
	attrition	count
1	No	5
2	Yes	4

Find the department with the highest attrition rate among employees under the age of 30.

```
SELECT *,
RANK() OVER(PARTITION BY department
ORDER BY percentage_count DESC) AS ranking_count
FROM (
    SELECT Department, attrition,
    COUNT(age) AS total_employees,
    ROUND(
        ( COUNT(Age) * 100.0/
        ( SELECT COUNT(*)
        FROM hr_data )):: NUMERIC , 2) AS percentage_count
    FROM hr_data
    WHERE age < 30
    GROUP BY Department, attrition ) AS x
```

Data Output					
	department	attrition	total_employees	percentage_count	ranking_count
1	Human Resources	No	7	0.47	1
2	Human Resources	Yes	6	0.41	2
3	Research & Development	No	164	11.08	1
4	Research & Development	Yes	55	3.72	2
5	Sales	No	67	4.53	1
6	Sales	Yes	30	2.03	2

Retrieve employees who have never been promoted (`YearsSinceLastPromotion = 0`) but still have a high `PerformanceRating` ( $\geq 4$ ).

```
SELECT EmpID, YearsSinceLastPromotion, PerformanceRating, department
FROM hr_data
WHERE YearsSinceLastPromotion = 0
    AND
        PerformanceRating >= 4
```

Data Output				
	empid	yearssincelastpromotion	performancerating	department
1	RM172	0	4	Sales
2	RM854	0	4	Research & Development
3	RM893	0	4	Research & Development
4	RM1340	0	4	Research & Development
5	RM586	0	4	Research & Development
6	RM911	0	4	Research & Development
7	RM1214	0	4	Sales
8	RM097	0	4	Sales
9	RM477	0	4	Research & Development
10	RM1434	0	4	Sales
11	RM202	0	4	Research & Development

Show the percentage of employees in each `EducationField` who travel frequently (`BusinessTravel = 'Travel_Frequently'`).

```
SELECT EducationField, COUNT(EmpID) AS total_employees,
    ROUND(
        (COUNT(EmpID) * 100.0 /
        ( SELECT COUNT(*)
            FROM hr_data )):: NUMERIC, 2 ) AS percentage_count
FROM hr_data
WHERE BusinessTravel = 'Travel_Frequently'
GROUP BY EducationField
ORDER BY percentage_count DESC
```

Data Output			
	educationfield	total_employees	percentage_count
1	Life Sciences	123	8.31
2	Medical	88	5.95
3	Marketing	27	1.82
4	Technical Degree	21	1.42
5	Other	14	0.95
6	Human Resour...	6	0.41

What is the overall attrition rate, and how does it vary by department?

```
SELECT * ,
RANK() OVER(PARTITION BY department ORDER BY total_employee DESC ) Ranking_count
FROM (
    SELECT attrition, Department, COUNT(EmpID) AS total_employee,
    ROUND(
        (COUNT(EmpID) * 100.0 / ( SELECT COUNT(*)
            FROM hr_data )):: NUMERIC ,2
    ) AS Percentage_count
    FROM hr_data
    GROUP BY attrition, Department ) AS ax
```

Data Output					
	attrition	department	total_employee	percentage_count	ranking_count
1	No	Human Resources	51	3.45	1
2	Yes	Human Resources	12	0.81	2
3	No	Research & Development	834	56.35	1
4	Yes	Research & Development	133	8.99	2
5	No	Sales	357	24.12	1
6	Yes	Sales	93	6.28	2

What is the relationship between Overtime and Attrition?

```
SELECT
CASE
    WHEN Attrition = 'Yes' THEN 'Leave'
    WHEN Attrition = 'No' THEN 'Work'
END AS Working_orLeave,
COUNT(EmpID) AS total_Employees, overtime
FROM hr_data
GROUP BY Attrition, overtime
```

Data Output			
working_orleave	total_employees	overtime	
Work	290	Yes	
Leave	128	Yes	
Work	952	No	
Leave	110	No	

What is the average MonthlyIncome by JobLevel and how does it compare across genders?

```
SELECT *,
RANK() OVER(PARTITION BY Job_level ORDER BY Avg_Monthly_income DESC) AS ranking_level
FROM (
    SELECT
        CASE
            WHEN JobLevel = 1 THEN 'Entry_Level'
            WHEN JobLevel = 2 THEN 'Associate'
            WHEN JobLevel = 3 THEN 'Senior'
            WHEN JobLevel = 4 THEN 'Manager'
            WHEN JobLevel = 5 THEN 'Senior_Manager'
        END AS Job_level,
        gender, ROUND(AVG(MonthlyIncome):: NUMERIC , 2) AS Avg_Monthly_income
    FROM hr_data
    GROUP BY JobLevel, gender ) AS ax
ORDER BY Avg_Monthly_income
```

Data Output				
job_level	gender	avg_monthly_income	ranking_level	
Entry_Level	Female	2786.41	2	
Entry_Level	Male	2790.63	1	
Associate	Female	5435.33	2	
Associate	Male	5551.37	1	
Senior	Male	9718.35	2	
Senior	Female	9962.70	1	
Manager	Female	15365.98	2	
Manager	Male	15570.93	1	
Senior_Manager	Female	19129.92	2	
Senior_Manager	Male	19224.84	1	

Which age group has the highest attrition rate?

```
SELECT agegroup, attrition, COUNT(EmpID) AS total_employees,
ROUND(
    ( COUNT(EmpID) * 100.0 / ( SELECT COUNT(*)
                                FROM hr_data ) ):: NUMERIC ,2
) AS percentage_count
FROM hr_data
GROUP BY agegroup, attrition
HAVING attrition = 'Yes'
ORDER BY percentage_count DESC
```

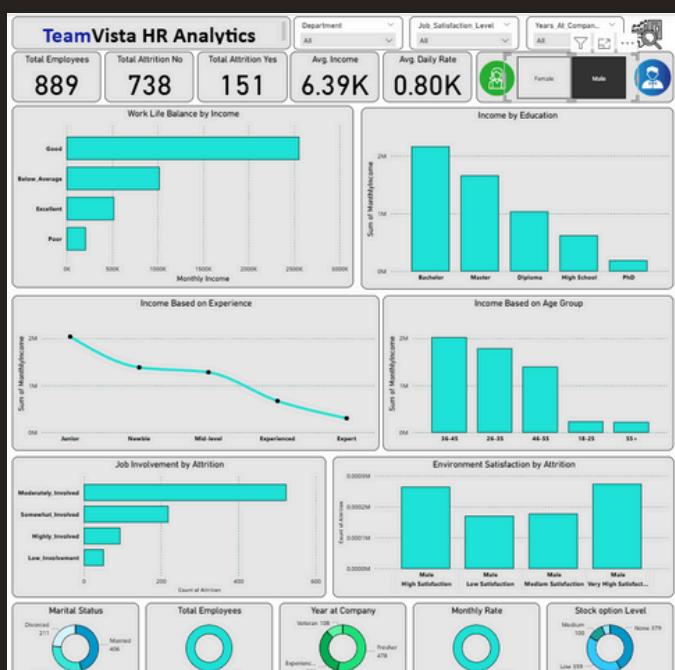
Data Output				
agegroup	attrition	total_employees	percentage_count	
26-35	Yes	116	7.84	
18-25	Yes	44	2.97	
36-45	Yes	43	2.91	
46-55	Yes	27	1.82	
55+	Yes	8	0.54	

# HR Analysis Dashboard

## Female Dashboard



## Male Dashboard



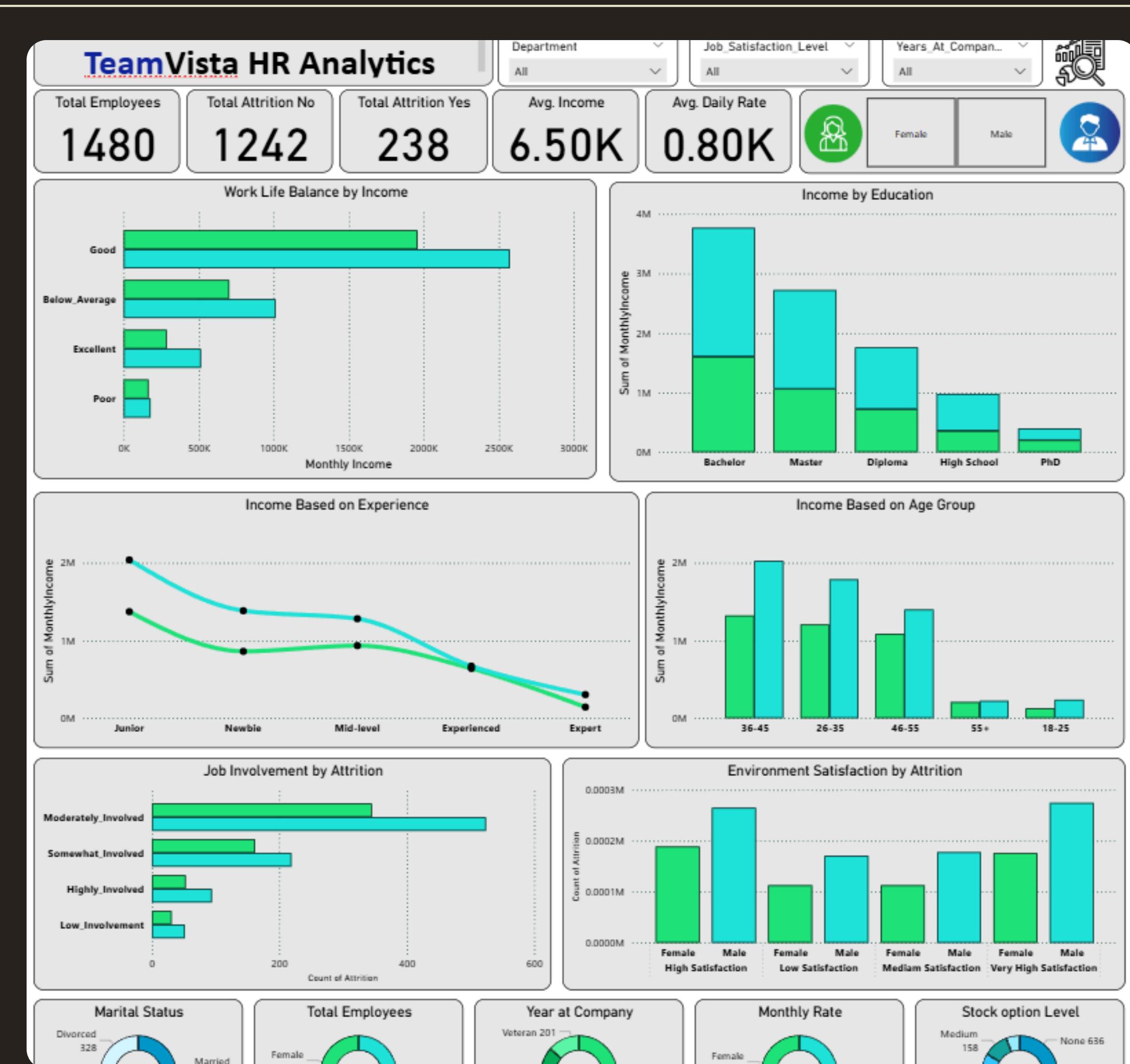
## Data Entry

ID	Length	Age	Apartment	Attrition	BusinessUnit	DailyRate	Department	DistanceFromHouse	Education	EmployeeID	EmployeeCount	EmployeeNumber	EnvironmentSatisfaction	Gender	HourlyRate	Job	JobLevel	JobRole	JobSatisfaction	JobTitle	MaritalStatus	Month	NumCompaniesWorked	Over18	Performance	PerformanceScore	RelationshipSatisfaction	SatSatisfaction	StockOptionLevel	TotalWorkingHours	TotalWorkingYears	TrainingLastWeek	TrainingPlaceholder	YearsAtCompany	YearsOnSite	YearsSinceLastPromotion	YearsWithComManager	
E00001	18	18-25	No	Travel_Arely	1001	Research & Development	2	1	Medical	7	1	501	201	1	Male	52	37	2	2	Technical Degree	1	2	Married	10	1	High	1	1	1	1	1	1	1	1	1	1	1	1
E00002	19	18-25	Yes	Travel_Arely	409	Human Resources	21	2	Other	7	1	509	201	1	Male	53	38	3	3	Other	1	3	Married	11	1	Medium	1	1	1	1	1	1	1	1	1	1	1	1
E00003	20	18-25	No	Travel_Arely	359	Research & Development	4	3	Life Sciences	7	1	507	201	1	Male	54	39	4	4	Technical Degree	1	4	Married	12	1	Medium	1	1	1	1	1	1	1	1	1	1	1	1
E00004	21	18-25	Yes	Travel_Arely	129	Research & Development	11	4	Medical	7	1	500	201	1	Male	55	40	5	5	Other	1	5	Married	13	1	Low	1	1	1	1	1	1	1	1	1	1	1	
E00005	22	18-25	No	Travel_Arely	1723	Research & Development	16	2	Medical	7	1	502	201	1	Male	56	41	6	6	Other	1	6	Married	14	1	Medium	1	1	1	1	1	1	1	1	1	1	1	
E00006	23	18-25	Yes	Travel_Arely	1294	Research & Development	8	3	Medical	7	1	503	201	1	Male	57	42	7	7	Life Sciences	1	7	Married	15	1	Medium	1	1	1	1	1	1	1	1	1	1	1	
E00007	24	18-25	No	Travel_Arely	823	Research & Development	17	2	Other	7	1	504	201	1	Male	58	43	8	8	Other	1	8	Married	16	1	Medium	1	1	1	1	1	1	1	1	1	1	1	
E00008	25	18-25	Yes	Travel_Arely	1097	Research & Development	11	3	Life Sciences	7	1	505	201	1	Male	59	44	9	9	Medical	1	9	Married	17	1	Medium	1	1	1	1	1	1	1	1	1	1	1	
E00009	26	18-25	No	Travel_Arely	1201	Research & Development	12	4	Life Sciences	7	1	506	201	1	Male	60	45	10	10	Other	1	10	Married	18	1	Medium	1	1	1	1	1	1	1	1	1	1	1	
E00010	27	18-25	Yes	Travel_Arely	1202	Research & Development	13	2	Life Sciences	7	1	507	201	1	Male	61	46	11	11	Other	1	11	Married	19	1	Medium	1	1	1	1	1	1	1	1	1	1	1	
E00011	28	18-25	No	Travel_Arely	1427	Research & Development	18	7	Other	7	1	508	201	1	Male	62	47	12	12	Other	1	12	Married	20	1	Medium	1	1	1	1	1	1	1	1	1	1	1	
E00012	29	18-25	Yes	Travel_Arely	1722	Research & Development	19	8	Human Resources	7	1	509	201	1	Male	63	48	13	13	Human Resources	1	13	Married	21	1	Medium	1	1	1	1	1	1	1	1	1	1	1	
E00013	30	18-25	No	Travel_Arely	1203	Research & Development	20	3	Marketing	7	1	510	201	1	Male	64	49	14	14	Marketing	1	14	Married	22	1	Medium	1	1	1	1	1	1	1	1	1	1	1	
E00014	31	18-25	Yes	Travel_Arely	1204	Research & Development	21	3	Life Sciences	7	1	511	201	1	Male	65	50	15	15	Life Sciences	1	15	Married	23	1	Medium	1	1	1	1	1	1	1	1	1	1	1	
E00015	32	18-25	No	Travel_Arely	1205	Research & Development	22	3	Medical	7	1	512	201	1	Male	66	51	16	16	Medical	1	16	Married	24	1	Medium	1	1	1	1	1	1	1	1	1	1	1	
E00016	33	18-25	Yes	Travel_Arely	1206	Research & Development	23	3	Other	7	1	513	201	1	Male	67	52	17	17	Other	1	17	Married	25	1	Medium	1	1	1	1	1	1	1	1	1	1	1	
E00017	34	18-25	No	Travel_Arely	1207	Research & Development	24	3	Human Resources	7	1	514	201	1	Male	68	53	18	18	Human Resources	1	18	Married	26	1	Medium	1	1	1	1	1	1	1	1	1	1	1	
E00018	35	18-25	Yes	Travel_Arely	1208	Research & Development	25	3	Marketing	7	1	515	201	1	Male	69	54	19	19	Marketing	1	19	Married	27	1	Medium	1	1	1	1	1	1	1	1	1	1	1	
E00019	36	18-25	No	Travel_Arely	1209	Research & Development	26	3	Life Sciences	7	1	516	201	1	Male	70	55	20	20	Life Sciences	1	20	Married	28	1	Medium	1	1	1	1	1	1	1	1	1	1	1	
E00020	37	18-25	Yes	Travel_Arely	1210	Research & Development	27	3	Medical	7	1	517	201	1	Male	71	56	21	21	Medical	1	21	Married	29	1	Medium	1	1	1	1	1	1	1	1	1	1	1	
E00021	38	18-25	No	Travel_Arely	1211	Research & Development	28	3	Other	7	1	518	201	1	Male	72	57	22	22	Other	1	22	Married	30	1	Medium	1	1	1	1	1	1	1	1	1	1	1	
E00022	39	18-25	Yes	Travel_Arely	1212	Research & Development	29	3	Human Resources	7	1	519	201	1	Male	73	58	23	23	Human Resources	1	23	Married	31	1	Medium	1	1	1	1	1	1	1	1	1	1	1	
E00023	40	18-25	No	Travel_Arely	1213	Research & Development	30	3	Marketing	7	1	520	201	1	Male	74	59	24	24	Marketing	1	24	Married	32	1	Medium	1	1	1	1	1	1	1	1	1	1	1	
E00024	41	18-25	Yes	Travel_Arely	1214	Research & Development	31	3	Life Sciences	7	1																											

# Challenges and Solutions

During the HR Analytics project, several challenges were encountered, including incomplete and inconsistent employee data, making analysis difficult, and identifying patterns across a large dataset of 50,000+ records.

Additionally, tracking employee engagement, retention, and performance metrics in real-time posed a significant challenge. To address these issues, data cleaning and validation processes were implemented to ensure accuracy and completeness. Advanced analytics techniques and visualization dashboards were developed to identify key trends, monitor KPIs, and detect potential risks. Furthermore, targeted employee engagement programs, performance tracking tools, and predictive models were introduced to provide actionable insights, improve decision-making, and optimize workforce management, thereby turning challenges into opportunities for strategic HR improvements.



# Future Plans

## Supporting Organizational Change and Growth

To effectively manage change and support the organization's strategic growth initiatives.

### Objective:

To leverage HR analytics for continuous workforce optimization, enhanced employee engagement, and data-driven decision-making.



### Plans:

- Implement predictive analytics to forecast attrition and workforce needs.
- Expand skill development programs to cover all employees.
- Enhance real-time performance monitoring dashboards for managers.
- Introduce automated reporting to streamline HR compliance and insights.

```
import pandas as pd
from sqlalchemy import create_engine
import psycopg2
from psycopg2.extensions import ISOLATION_LEVEL_AUTOCOMMIT

# Step 2: CSV file
csv_file = "HR_Analytics_Data_Set.csv"
hr = pd.read_csv(csv_file)

print("CSV file loaded successfully")
print(hr.head())

# Step 3: Database details
db_name = "hr_analysis"
db_user = "postgres"
db_password = "Password"
db_host = "localhost"
db_port = "5432"
table_name = "hr_data"

# Step 4: Connect to default database to create the target database if it doesn't exist
try:
    conn = psycopg2.connect(
        dbname="postgres", # default database
        user=db_user,
        password=db_password,
        host=db_host,
        port=db_port
    )
    conn.set_isolation_level(ISOLATION_LEVEL_AUTOCOMMIT)
    cur = conn.cursor()

    # Check if database exists
    cur.execute(f"SELECT 1 FROM pg_database WHERE datname = '{db_name}'")
    exists = cur.fetchone()
    if not exists:
        cur.execute(f'CREATE DATABASE "{db_name}"')
        print(f'Database "{db_name}" created.')
    else:
        print(f'Database "{db_name}" already exists.')

    cur.close()
    conn.close()
except Exception as e:
    print("Error creating/checking database:", e)

# Step 5: Connect to the target database
engine = create_engine(f'postgresql://{{db_user}}:{{db_password}}@{{db_host}}:{{db_port}}/{{db_name}}')

# Step 6: Export dataframe to PostgreSQL
try:
    hr.to_sql(table_name, engine, if_exists='replace', index=False)
    print(F'CSV data exported successfully to table {table_name} in database {db_name}')
except Exception as e:
    print("Error exporting data:", e)

# Step 7 (Optional): Verify the table
from sqlalchemy import inspect
inspector = inspect(engine)
print("Tables in database:", inspector.get_table_names())
```

# Conclusion

The Team Vista HR Analytics project has successfully provided a comprehensive analysis of SummitEdge Consulting's workforce, covering over 50,000 employee records and 32+ key metrics. By examining aspects such as job satisfaction, engagement, work-life balance, compensation, and performance, the project has delivered actionable insights that support data-driven HR strategies and improve overall organizational efficiency.



The project also strengthened compliance and risk management, introduced effective engagement programs, and implemented real-time performance dashboards, ensuring informed decision-making across all levels of the organization. Moving forward, the focus will be on predictive analytics, skill development, and automation, ensuring continuous improvement in employee satisfaction, retention, and workforce optimization.

**Summit Edge Consulting.**  
Team **Vista**.

# Thank You



## Contact Information

Name :- Pushpkar Roy  
Phone No. 6260854602  
G-mail :- Pushpkarroy880@gmail.com  
LinkedIn :- Pushpkar Roy  
GitHub :- PushpkarRoy