HUMAN RESOURCE ANALYSIS USING PYTHON

IMPORTING PYTHON PACKAGES

In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import classification_report, confusion_matrix, accuracy_
from sklearn.preprocessing import LabelEncoder
```

IMPORTING TABLE

In [2]:
```
HR = pd.read_csv("hr.csv")
```

In [3]:
```
HR.head()
```

Out[3]:

	Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome	Education	Educ
0	41	Yes	Travel_Rarely	1102	Sales	1	2	Lif
1	49	No	Travel_Frequently	279	Research & Development	8	1	Lif
2	37	Yes	Travel_Rarely	1373	Research & Development	2	2	
3	33	No	Travel_Frequently	1392	Research & Development	3	4	Lif
4	27	No	Travel_Rarely	591	Research & Development	2	1	

5 rows × 33 columns

Explanation of Columns

Age: The age of the employee.

Attrition: Indicates if the employee has left the company.

Business Travel: Frequency of business travel.

Daily Rate: The daily pay rate of the employee.

Department: The department the employee works in.

Distance From Home: The distance (in miles) the employee lives from work.

Education: Level of education.

Education Field: The field of study of the employee.

Employee Count: A count of employees in the same position or department.

Employee Number: A unique identifier for the employee.

Environment Satisfaction: Satisfaction level with the work environment (on a scale, typically 1-4).

Gender: The gender of the employee.

Hourly Rate: The pay rate of the employee on an hourly basis.

Job Involvement: Employee’s involvement in their job.

Job Level: Level of the employee’s position within the organization.

Job Role: The specific job title or role of the employee.

Job Satisfaction: Employee's satisfaction with their job.

Marital Status: The marital status of the employee.

Monthly Income: The employee's monthly income.

Monthly Rate: The monthly pay rate for the employee.

Num Companies Worked: The number of different companies the employee has worked for.

Over 18: Indicates if the employee is over 18 years old).

Over Time: Indicates if the employee works overtime.

Percent Salary Hike: The percentage increase in salary.

Performance Rating: Employee’s performance rating.

Relationship Satisfaction: Satisfaction level with workplace relationships.

Standard Hours: Standard working hours for the employee.

Stock Option Level: Level of stock options available to the employee.

Total Working Years: Total years of working experience.

Training Times Last Year: Number of training sessions attended in the last year.

Work-Life Balance: Employee's perception of their work-life balance.

Years At Company: Number of years the employee has been with the company.

Years In Current Role: Number of years the employee has been in their current position.

Years Since Last Promotion: Number of years since the last promotion.

Years With Current Manager: Number of years the employee has been with their current manager.

CHECKING FOR MISSING VALUES

In [4]:
```
HR.isnull().sum()
```

Out[4]:
```
Age                         0
Attrition                   0
BusinessTravel              0
DailyRate                   0
Department                  0
DistanceFromHome            0
Education                   0
EducationField              0
EmployeeNumber              0
EnvironmentSatisfaction     0
Gender                      0
HourlyRate                  0
JobInvolvement              0
JobLevel                    0
JobRole                     0
JobSatisfaction             0
MaritalStatus               0
MonthlyIncome               0
MonthlyRate                 0
NumCompaniesWorked          0
```

NOTE: There is no missing value in the dataset

TABLE INFO

In [5]:
```
print(HR.info())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1470 entries, 0 to 1469
Data columns (total 33 columns):
 #   Column                    Non-Null Count  Dtype 
---  ------                    --------------  ----- 
 0   Age                       1470 non-null   int64 
 1   Attrition                 1470 non-null   object
 2   BusinessTravel            1470 non-null   object
 3   DailyRate                 1470 non-null   int64 
 4   Department                1470 non-null   object
 5   DistanceFromHome          1470 non-null   int64 
 6   Education                 1470 non-null   int64 
 7   EducationField            1470 non-null   object
 8   EmployeeNumber            1470 non-null   int64 
 9   EnvironmentSatisfaction   1470 non-null   int64 
 10  Gender                    1470 non-null   object
 11  HourlyRate                1470 non-null   int64 
 12  JobInvolvement            1470 non-null   int64 
 13  JobLevel                  1470 non-null   int64 
 14  JobRole                   1470 non-null   object
 15  JobSatisfaction           1470 non-null   int64 
 16  MaritalStatus             1470 non-null   object
 17  MonthlyIncome             1470 non-null   int64 
 18  MonthlyRate               1470 non-null   int64 
 19  NumCompaniesWorked        1470 non-null   int64 
 20  OverTime                  1470 non-null   object
 21  PercentSalaryHike         1470 non-null   int64 
 22  PerformanceRating         1470 non-null   int64 
 23  RelationshipSatisfaction  1470 non-null   int64 
 24  StandardHours             1470 non-null   int64 
 25  StockOptionLevel          1470 non-null   int64 
 26  TotalWorkingYears         1470 non-null   int64 
 27  TrainingTimesLastYear     1470 non-null   int64 
 28  WorkLifeBalance           1470 non-null   int64 
 29  YearsAtCompany            1470 non-null   int64 
 30  YearsInCurrentRole        1470 non-null   int64 
 31  YearsSinceLastPromotion   1470 non-null   int64 
 32  YearsWithCurrManager      1470 non-null   int64 
dtypes: int64(25), object(8)
memory usage: 379.1+ KB
None
```

Employee Attrition Analysis

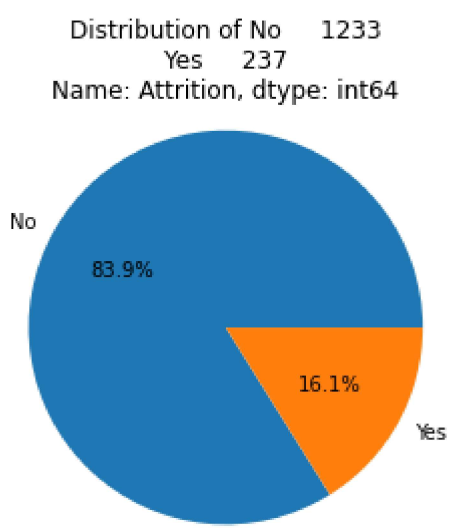
• Problem Statement: What factors are most associated with employee attrition (whether an employee leaves the company)?

In [6]:
```
HR
```

Out[6]:

	Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome	Educatio
0	41	Yes	Travel_Rarely	1102	Sales	1	
1	49	No	Travel_Frequently	279	Research & Development	8	
2	37	Yes	Travel_Rarely	1373	Research & Development	2	
3	33	No	Travel_Frequently	1392	Research & Development	3	
4	27	No	Travel_Rarely	591	Research & Development	2	
...	...	...	...	...	...	...	
1465	36	No	Travel_Frequently	884	Research & Development	23	
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Research &	[illegible]	

In [7]:
```python
counts = HR["Attrition"].value_counts()

# Create the pie chart
plt.pie(counts, labels=counts.index, autopct='%1.1f%%')

# Equal aspect ratio ensures that pie is drawn as a circle
plt.axis('equal')

# Set the title
plt.title(f"Distribution of {counts}")

# Show the plot
plt.show()
```



COUNT BASED ON ATTRITION

237 employees have left, which is 16.1% of the total employees.

1233 employees are active, which is 83.9% of the total employees.

In [ ]:

```

```

Attrition against Other Variables

Investigate how various factors differ between employees who stayed and those who left. Some of the key variables to focus on:

1.Age: Are younger or older employees more likely to leave?

2.Business Travel: Does frequent travel lead to higher attrition?

3.Department: Is attrition higher in specific departments like Sales or R&D?

4.Distance From Home: Are employees who live further from work more likely to leave?

5.Education: Does the level of education influence attrition?

6.Job Satisfaction: Do employees with lower job satisfaction tend to leave more?

7.Marital Status: Does marital status play a role (e.g., are single employees more likely to leave)?

8.Monthly Income: Is there a connection between income and leaving the company?

9.Job Role: Does a particular job role (e.g., Sales Executive, Research Scientist) have a higher attrition rate?

10.Over Time: Do employees working overtime have a higher chance of leaving?

11.Work-Life Balance: Are employees with poor work-life balance more likely to leave?

1. Age: Are younger or older employees more likely to leave?

Hypothesis: Younger employees may be more likely to leave as they might be looking for new opportunities or career growth.

In [8]:

```
age_comparison = HR.groupby('Attrition')['Age'].mean()

print(age_comparison)
```

```
Attrition
No     37.561233
Yes    33.607595
Name: Age, dtype: float64
```

INSIGHT

The average age of employees that left the company is 33 years old

The average age of employees that are still in the company is 37 years old

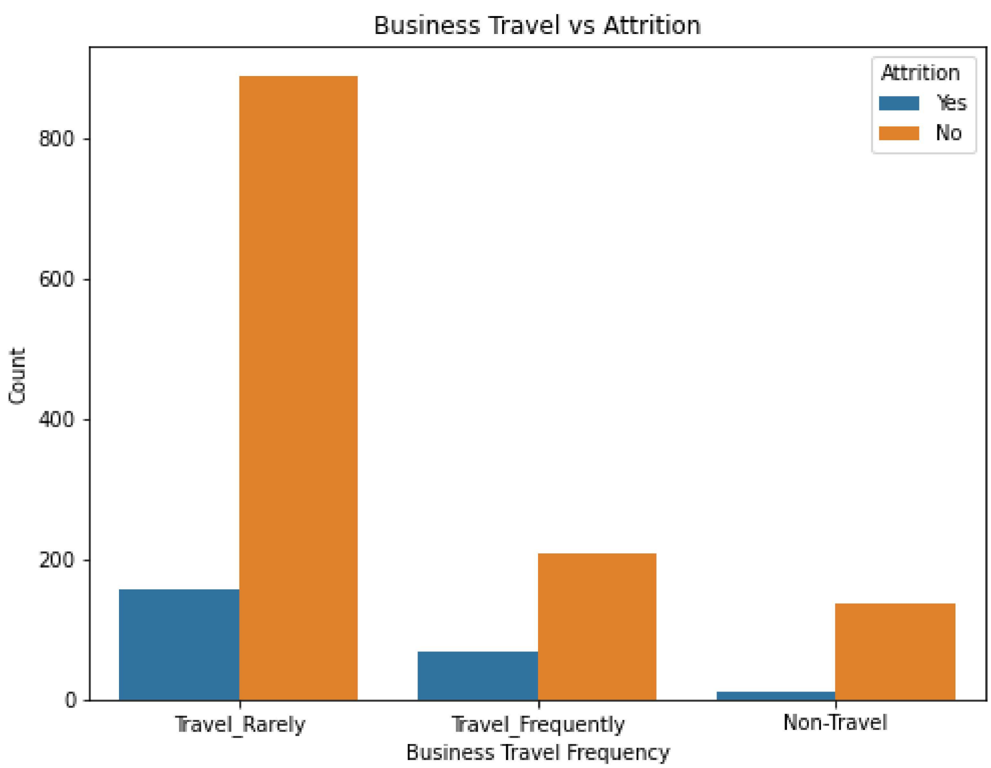
2. Business Travel: Does frequent travel lead to higher attrition?

Hypothesis: Employees who travel frequently for business might be more likely to leave due to travel fatigue or poor work-life balance.

In [9]:
```
travel_comparison = HR.groupby(['BusinessTravel', 'Attrition']).size().unstack

print(travel_comparison)
```

```
Attrition           No   Yes
BusinessTravel
Non-Travel         138    12
Travel_Frequently  208    69
Travel_Rarely      887   156
```

In [10]:
```
plt.figure(figsize=(8,6))
sns.countplot(x='BusinessTravel', hue='Attrition', data=HR)
plt.title('Business Travel vs Attrition')
plt.xlabel('Business Travel Frequency')
plt.ylabel('Count')
plt.show()
```



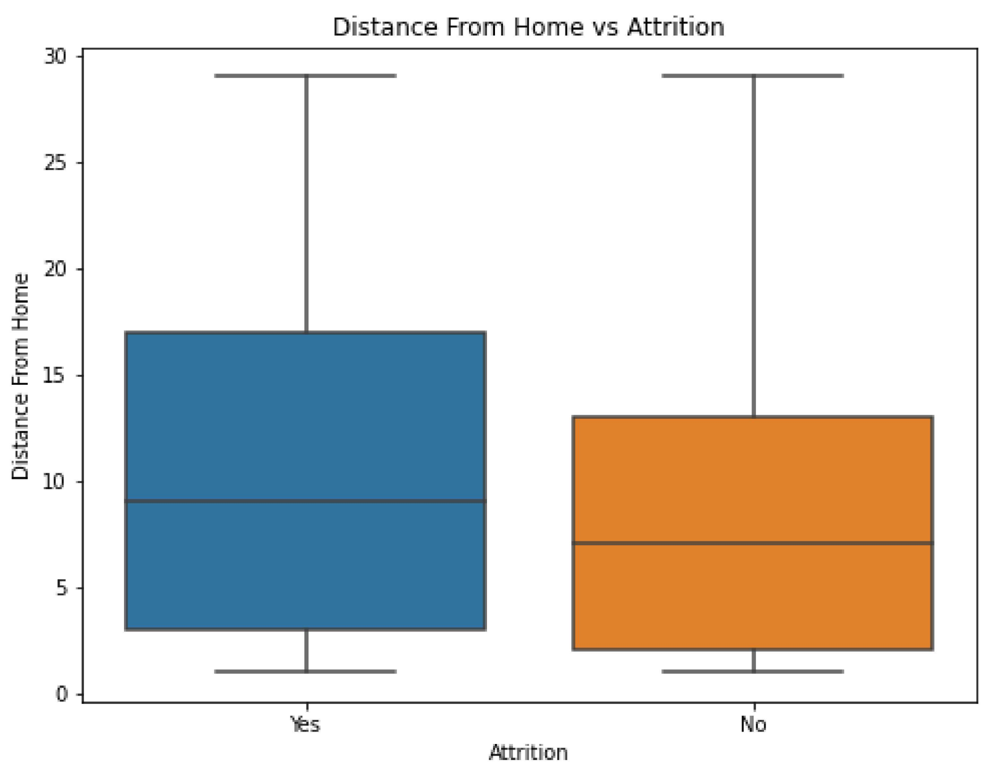
3. Distance From Home: Are employees who live farther away more likely to leave?

Hypothesis: Employees with a long commute might be more likely to leave due to the stress and time involved.

In [11]:
```
distance_comparison = HR.groupby('Attrition')['DistanceFromHome'].mean()

print(distance_comparison)
```

```
Attrition
No      8.915653
Yes    10.632911
Name: DistanceFromHome, dtype: float64
```

In [12]:
```
plt.figure(figsize=(8,6))
sns.boxplot(x='Attrition', y='DistanceFromHome', data=HR)
plt.title('Distance From Home vs Attrition')
plt.xlabel('Attrition')
plt.ylabel('Distance From Home')
plt.show()
```

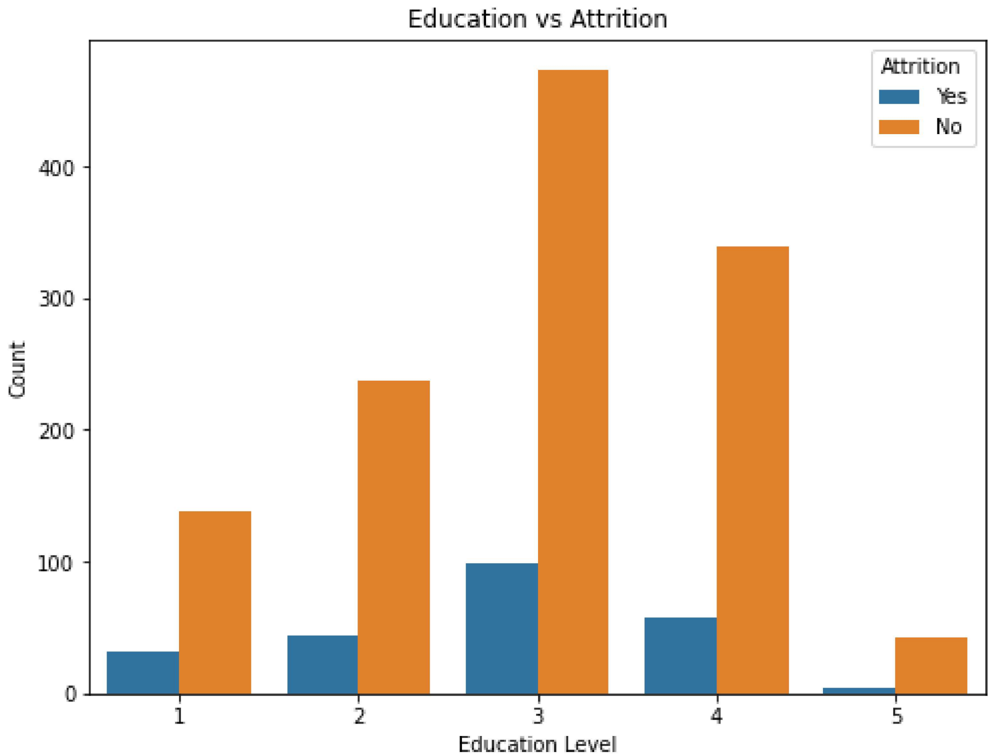


5. Education vs. Attrition

Visualization: Count Plot (Bar Plot)

Purpose: To investigate if employees with different education levels are more likely to leave.

In [13]:
```
plt.figure(figsize=(8,6))
sns.countplot(x='Education', hue='Attrition', data=HR)
plt.title('Education vs Attrition')
plt.xlabel('Education Level')
plt.ylabel('Count')
plt.show()
```

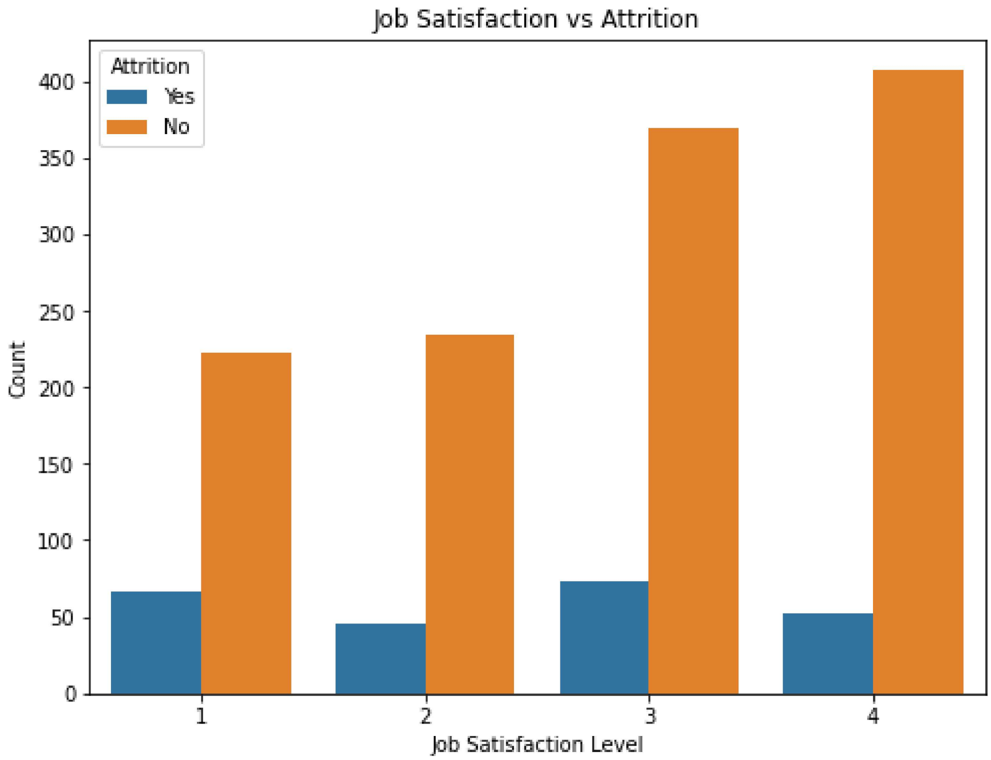


6. Job Satisfaction vs. Attrition

Visualization: Count Plot (Bar Plot)

Purpose: To determine if employees with lower job satisfaction tend to leave more. python

In [14]:
```
plt.figure(figsize=(8,6))
sns.countplot(x='JobSatisfaction', hue='Attrition', data=HR)
plt.title('Job Satisfaction vs Attrition')
plt.xlabel('Job Satisfaction Level')
plt.ylabel('Count')
plt.show()
```

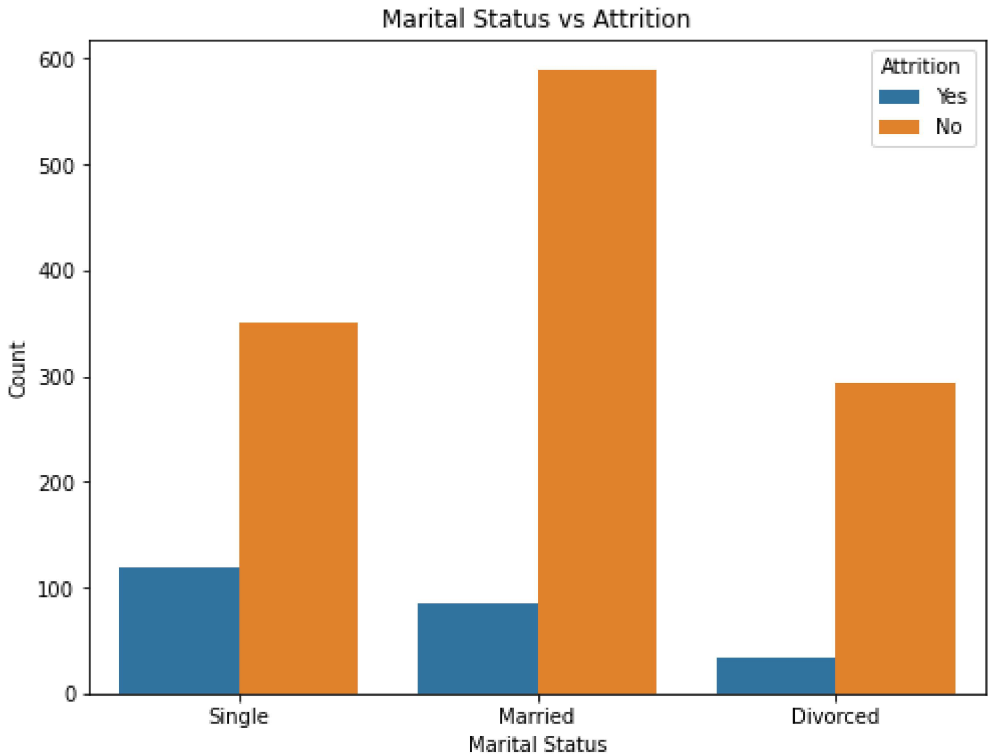


7. Marital Status vs. Attrition

Visualization: Count Plot (Bar Plot)

Purpose: To check if single employees are more likely to leave than married or divorced employees.

In [15]:
```
plt.figure(figsize=(8,6))
sns.countplot(x='MaritalStatus', hue='Attrition', data=HR)
plt.title('Marital Status vs Attrition')
plt.xlabel('Marital Status')
plt.ylabel('Count')
plt.show()
```

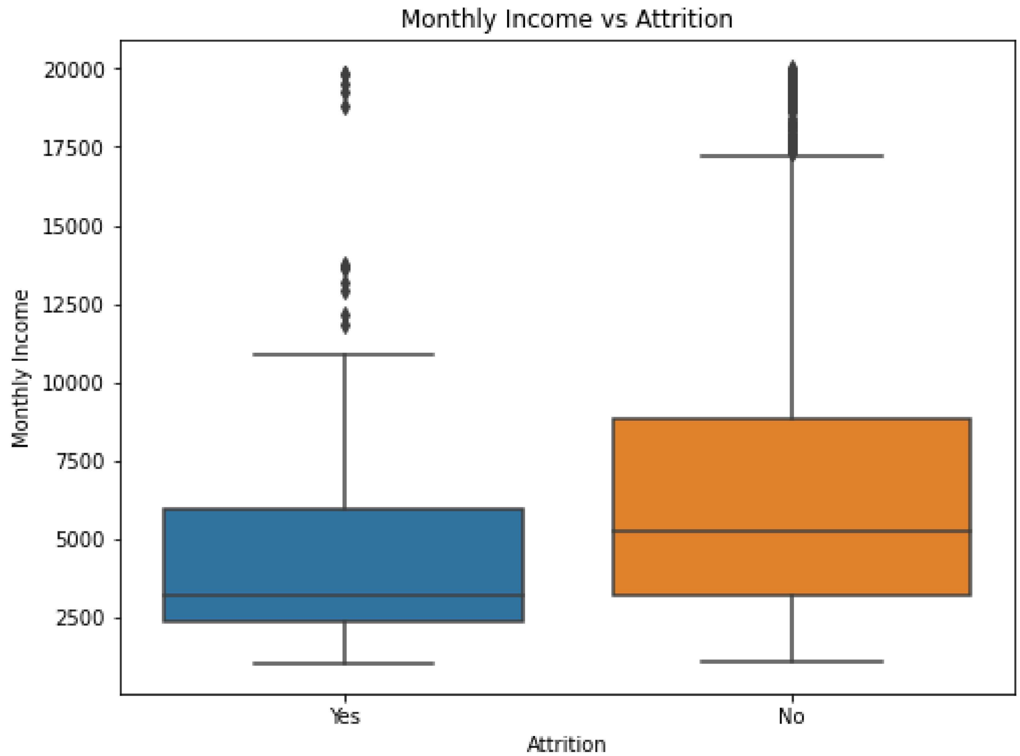


8. Monthly Income vs. Attrition

Visualization: Box Plot

Purpose: To investigate if there is a connection between monthly income and leaving the company.

In [16]:
```
plt.figure(figsize=(8,6))
sns.boxplot(x='Attrition', y='MonthlyIncome', data=HR)
plt.title('Monthly Income vs Attrition')
plt.xlabel('Attrition')
plt.ylabel('Monthly Income')
plt.show()
```

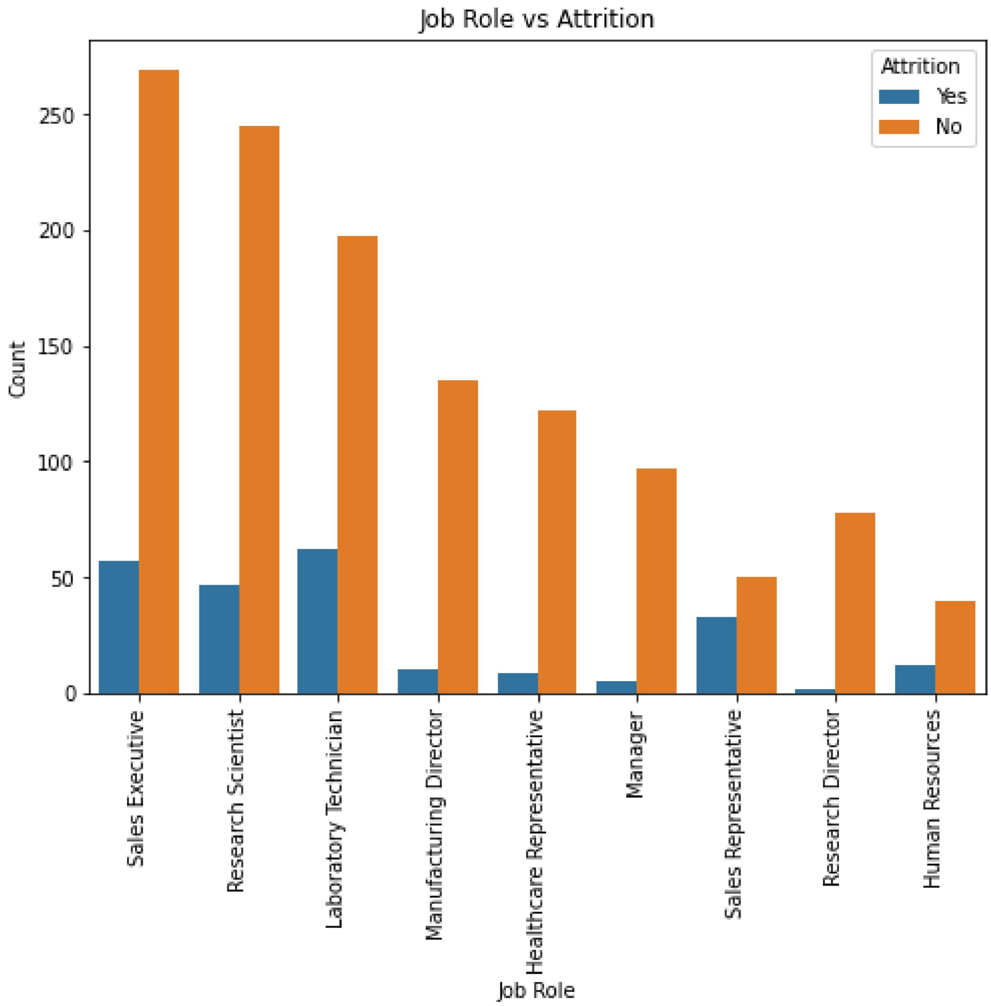


9. Job Role vs. Attrition

Visualization: Count Plot (Bar Plot)

Purpose: To see if employees in specific job roles (e.g., Sales Executive, Research Scientist) have higher attrition rates.

In [17]:
```
plt.figure(figsize=(8,6))
sns.countplot(x='JobRole', hue='Attrition', data=HR)
plt.title('Job Role vs Attrition')
plt.xlabel('Job Role')
plt.ylabel('Count')
plt.xticks(rotation=90)
plt.show()
```

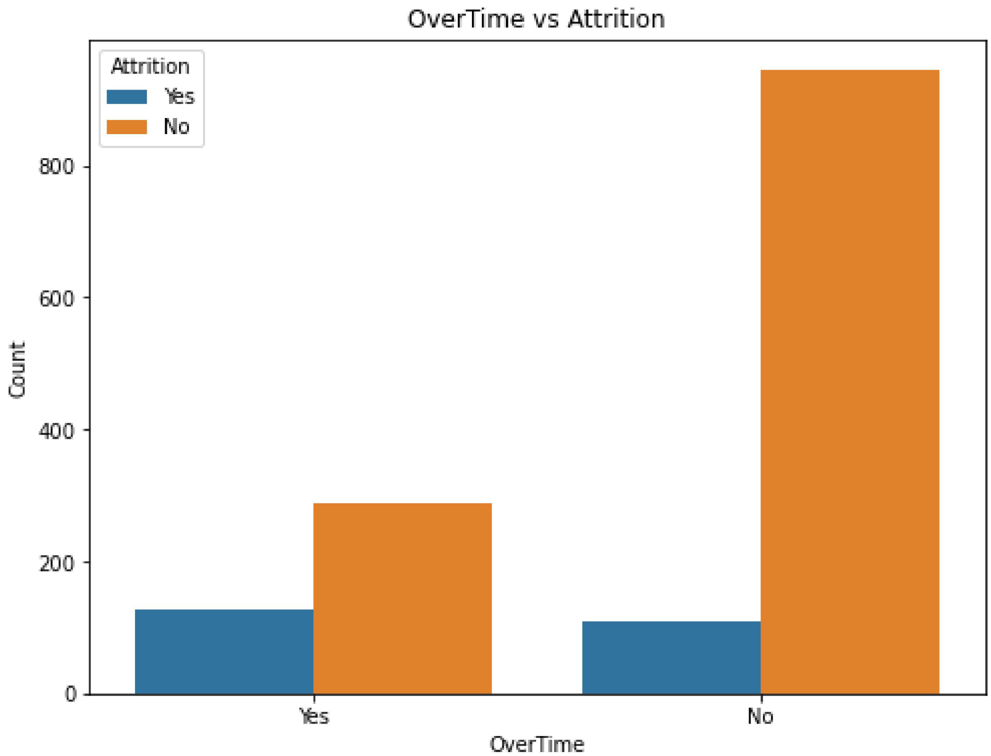


10. OverTime vs. Attrition

Visualization: Count Plot (Bar Plot)

Purpose: To investigate if employees who work overtime are more likely to leave.

In [18]:
```
plt.figure(figsize=(8,6))
sns.countplot(x='OverTime', hue='Attrition', data=HR)
plt.title('OverTime vs Attrition')
plt.xlabel('OverTime')
plt.ylabel('Count')
plt.show()
```



11. Work-Life Balance vs Attrition

Visualization: Count Plot (Bar Plot)

Purpose: To see if employees with poor work-life balance are more likely to leave.

In [19]:
```
plt.figure(figsize=(8,6))
sns.countplot(x='WorkLifeBalance', hue='Attrition', data=HR)
plt.title('Work-Life Balance vs Attrition')
plt.xlabel('Work-Life Balance')
plt.ylabel('Count')
plt.show()
```

CORRELATION ANALYSIS

Checking the relationship between each numeric variable and attrition to find strong predictions.

In [20]:
```
# Convert 'Attrition' and 'OverTime' columns to binary (0 for No, 1 for Yes)
HR['Attrition'] = HR['Attrition'].apply(lambda x: 1 if x == 'Yes' else 0)
HR['OverTime'] = HR['OverTime'].apply(lambda x: 1 if x == 'Yes' else 0)

# Convert 'MaritalStatus' to binary (1 for Single, 0 for others)
HR['MaritalStatus'] = HR['MaritalStatus'].apply(lambda x: 1 if x == 'Single' e
```

In [21]:
```
correlation = HR.corr()['Attrition']
```

In [22]:
```python
#Display correlations
print(correlation)
```

```
Age                        -0.159205
Attrition                   1.000000
DailyRate                  -0.056652
DistanceFromHome            0.077924
Education                  -0.031373
EmployeeNumber             -0.010577
EnvironmentSatisfaction    -0.103369
HourlyRate                 -0.006846
JobInvolvement             -0.130016
JobLevel                   -0.169105
JobSatisfaction            -0.103481
MaritalStatus               0.175419
MonthlyIncome              -0.159840
MonthlyRate                 0.015170
NumCompaniesWorked          0.043494
OverTime                    0.246118
PercentSalaryHike          -0.013478
PerformanceRating           0.002889
RelationshipSatisfaction   -0.045872
StandardHours                    NaN
StockOptionLevel           -0.137145
TotalWorkingYears          -0.171063
TrainingTimesLastYear      -0.059478
WorkLifeBalance            -0.063939
YearsAtCompany             -0.134392
YearsInCurrentRole         -0.160545
YearsSinceLastPromotion    -0.033019
YearsWithCurrManager       -0.156199
Name: Attrition, dtype: float64
```

Key Insights from the Correlation Results:

1. **Age (-0.16)**:
    - There is a weak negative correlation, meaning younger employees are slightly more likely to leave than older ones.
2. **DailyRate (-0.06)**:
    - Almost no correlation between the employee's daily rate (salary per day) and attrition. It doesn't significantly impact whether an employee stays or leaves.
3. **DistanceFromHome (0.08)**:
    - A very weak positive correlation. Employees living farther from home may be slightly more likely to leave, but the effect is minimal.
4. **Education (-0.03)**:
    - Very little to no correlation between education level and attrition, meaning education doesn't play a major role in employees leaving.
5. **EmployeeNumber (-0.01)**:

- Almost no correlation. The employee number (ID) obviously has no relationship with attrition.

6. **EnvironmentSatisfaction (-0.10)**:

   - Weak negative correlation. Employees with lower satisfaction in their work environment are slightly more likely to leave.

7. **HourlyRate (-0.01)**:

   - No correlation between hourly wage and attrition. This doesn't affect whether employees stay or leave.

8. **JobInvolvement (-0.13)**:

   - Weak negative correlation. Employees with lower job involvement are slightly more likely to leave.

9. **JobLevel (-0.17)**:

   - Weak negative correlation. Employees at lower job levels (lower ranks) are more likely to leave than those at higher levels.

10. **JobSatisfaction (-0.10)**:

    - Weak negative correlation. Employees with lower job satisfaction are more likely to leave, but the relationship is not very strong.

11. **MaritalStatus (0.18)**:

    - Weak positive correlation. Single employees tend to leave more often compared to married ones.

12. **MonthlyIncome (-0.16)**:

    - Weak negative correlation. Employees with lower monthly income are slightly more likely to leave than those with higher income.

13. **NumCompaniesWorked (0.04)**:

    - Almost no correlation. The number of companies an employee has worked at in the past doesn't greatly impact attrition.

14. **OverTime (0.25)**:

    - Moderate positive correlation. Employees who work overtime are more likely to leave, making this one of the stronger predictors of attrition.

15. **PercentSalaryHike (-0.01)**:

    - Almost no correlation. Salary hikes don’t significantly affect whether an employee stays or leaves.

16. **PerformanceRating (0.003)**:

    - No correlation between performance rating and attrition. How well employees are rated in performance reviews doesn't impact their decision to leave.

17. **RelationshipSatisfaction (-0.05)**:

    - Very weak negative correlation. Employees with lower satisfaction in their workplace relationships might be slightly more likely to leave, but the effect is minimal.

18. **StandardHours (NaN)**:

    - No correlation data for this variable, likely because all employees have the same standard working hours in the dataset.

19. **StockOptionLevel (-0.14)**:

    - Weak negative correlation. Employees with lower stock options are more likely to leave, though the effect is small.

20. **TotalWorkingYears (-0.17)**:

    - Weak negative correlation. Employees with fewer years of total work experience are more likely to leave.

21. **TrainingTimesLastYear (-0.06)**:

    - Very weak negative correlation. Employees who receive less training may be slightly more likely to leave, but the relationship is minimal.

22. **WorkLifeBalance (-0.06)**:

    - Very weak negative correlation. Employees with poorer work-life balance are slightly more likely to leave.

23. **YearsAtCompany (-0.13)**:

    - Weak negative correlation. Employees who have been at the company for fewer years are more likely to leave.

24. **YearsInCurrentRole (-0.16)**:

    - Weak negative correlation. Employees who have been in their current role for fewer years are more likely to leave.

25. **YearsSinceLastPromotion (-0.03)**:

    - Very weak negative correlation. Employees who haven't been promoted in a while may be slightly more likely to leave, but the relationship is almost negligible.

26. **YearsWithCurrManager (-0.16)**:

    - Weak negative correlation. Employees who have been with their current manager for a shorter time are more likely to leave.

Summary of Key Factors:

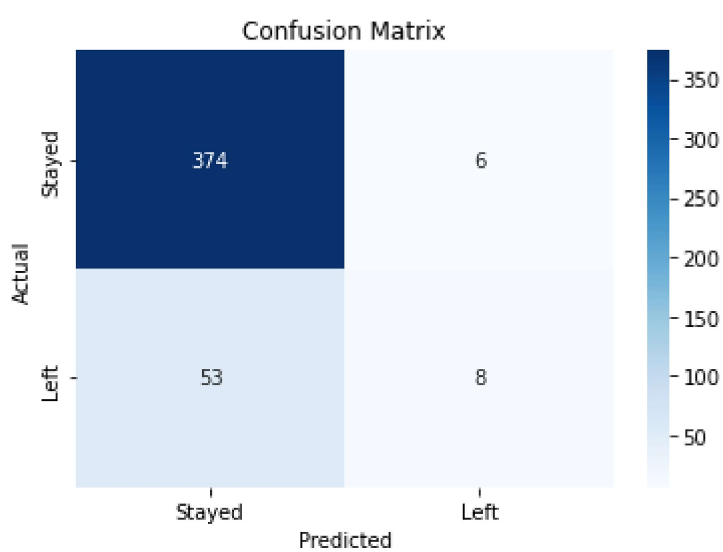
- **OverTime** is the most significant factor associated with attrition. Employees who work overtime are more likely to leave.
- **Marital Status (Single)**, **JobLevel**, **MonthlyIncome**, **YearsInCurrentRole**, **TotalWorkingYears**, and **YearsWithCurrManager** are also notable factors: single employees, those at lower job levels, with lower incomes, or with fewer years in their current role/with their current manager are more likely to leave.
- **Job Satisfaction**, **Environment Satisfaction**, and **Work-Life Balance** are weaker predictors, but still worth considering for retention strategies.

Logistic Regression for Predicting Employee Attrition

In [23]:
```
le = LabelEncoder()
HR['BusinessTravel'] = le.fit_transform(HR['BusinessTravel'])
HR['Department'] = le.fit_transform(HR['Department'])
HR['MaritalStatus'] = le.fit_transform(HR['MaritalStatus'])
HR['JobRole'] = le.fit_transform(HR['JobRole'])
HR['OverTime'] = le.fit_transform(HR['OverTime'])
```

In [24]:
```
# Define the features (independent variables) and the target (dependent variab
X = HR[['Age', 'BusinessTravel', 'DistanceFromHome', 'Education', 'JobSatisfac
        'MonthlyIncome', 'JobRole', 'OverTime', 'WorkLifeBalance']]
y = HR['Attrition']
# Split data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, rando
logreg = LogisticRegression(max_iter=1000)
logreg.fit(X_train, y_train)

# Make predictions
y_pred = logreg.predict(X_test)

# Evaluate the model accuracy
accuracy = accuracy_score(y_test, y_pred)
print(f'Accuracy: {accuracy * 100:.2f}%')

# Confusion matrix
cm = confusion_matrix(y_test, y_pred)
sns.heatmap(cm, annot=True, fmt='d', cmap='Blues', xticklabels=['Stayed', 'Let
plt.xlabel('Predicted')
plt.ylabel('Actual')
plt.title('Confusion Matrix')
plt.show()

# Classification report
print(classification_report(y_test, y_pred))
```

```
Accuracy: 86.62%
```




```
              precision    recall  f1-score   support

           0       0.88      0.98      0.93       380
           1       0.57      0.13      0.21        61

    accuracy                           0.87       441
   macro avg       0.72      0.56      0.57       441
weighted avg       0.83      0.87      0.83       441
```

Explanation of Logistic Regression Results

The output you provided contains two main components: the **Confusion Matrix** and the **Classification Report**. Let's break down each component to understand what they mean in the context of predicting employee attrition.

1. Accuracy

- **Accuracy: 86.62%**
  - This means that the logistic regression model correctly predicted whether employees stayed or left about 86.62% of the time across the entire dataset. This is a good accuracy rate, indicating that the model performs well.

2. Confusion Matrix

The confusion matrix visualizes the performance of the classification model by showing the counts of correct and incorrect predictions broken down by class (Stayed and Left).

- **Matrix Interpretation:**
  - **True Positives (TP)**: 374 employees were predicted to **stay** and did stay.
  - **True Negatives (TN)**: 8 employees were predicted to **leave** and did leave.
  - **False Positives (FP)**: 6 employees were predicted to **stay** but actually left. (Type I error)
  - **False Negatives (FN)**: 53 employees were predicted to **leave** but actually stayed. (Type II error)

3. Precision, Recall, and F1-Score

These metrics help evaluate the model's performance for each class.

- **Precision**:
  - For **Stayed (class 0)**: Precision is 0.88. This means that when the model predicts that an employee will stay, it is correct 88% of the time.
  - For **Left (class 1)**: Precision is 0.57. This means that when the model predicts that an employee will leave, it is correct only 57% of the time.
- **Recall**:
  - For **Stayed (class 0)**: Recall is 0.98. This means that the model correctly identifies 98% of the employees who actually stayed.
  - For **Left (class 1)**: Recall is 0.13. This means that the model only identifies 13% of the employees who actually left. This low recall indicates that many employees who left were misclassified as staying.
- **F1-Score**:
  - The F1-score is the harmonic mean of precision and recall, providing a balance between the two metrics.
  - For **Stayed**: F1-score is 0.93, indicating a good balance between precision and recall.
  - For **Left**: F1-score is 0.21, suggesting poor performance in correctly identifying employees who left.

4. Support

- **Support** indicates the number of actual occurrences of each class in the dataset.
  - **0 (Stayed)**: 380 instances.
  - **1 (Left)**: 61 instances.

This indicates that there are significantly more employees who stayed than those who left, which could influence the model's performance.

5. Macro Average and Weighted Average

- **Macro Average**: This calculates the average precision, recall, and F1-score for each class without considering the support (number of instances).
  - For precision: 0.72
  - For recall: 0.56
  - For F1-score: 0.57
- **Weighted Average**: This considers the support of each class in the average calculation.
  - For precision: 0.83
  - For recall: 0.87
  - For F1-score: 0.83

Summary

- The logistic regression model is generally effective in predicting employee attrition, with a high accuracy of 86.62%.
- However, it struggles with identifying employees who left, as indicated by the low recall (0.13) for this class.
- The model's performance on predicting those who stayed is strong, with high precision and recall.
- Future efforts may focus on improving the model's ability to predict attrition by exploring additional features or utilizing different algorithms that can better handle class imbalance or improve sensitivity to the minority class.

If you have specific questions or further analyses you'd like to conduct, feel free to ask!

In [25]:
```
# Get the coefficients and the corresponding feature names
coefficients = pd.DataFrame({'Feature': X.columns, 'Coefficient': logreg.coef_
coefficients = coefficients.sort_values(by='Coefficient', ascending=False)
print(coefficients)
```

```
            Feature  Coefficient
8          OverTime     1.392926
5     MaritalStatus     1.039107
3         Education     0.200227
1    BusinessTravel     0.161436
7           JobRole     0.054982
2  DistanceFromHome     0.023927
6     MonthlyIncome    -0.000097
0               Age    -0.044309
9   WorkLifeBalance    -0.292441
4   JobSatisfaction    -0.369142
```

Explanation of Each Coefficient:

- **Positive Coefficients:**
  - **OverTime (1.329926)**: Employees who work overtime are more likely to leave the company. This is the strongest predictor of attrition in this model.
  - **MaritalStatus (1.039107)**: Being married increases the likelihood of leaving. This may suggest that marital stability or life changes affect job stability.
  - **Education (0.200227)**: Higher education levels are associated with a higher probability of leaving the company, possibly indicating that more educated employees may have better job opportunities elsewhere.
  - **BusinessTravel (0.161436)**: Employees who frequently travel for business are slightly more likely to leave.
  - **JobRole (0.054982)**: The impact of job role on attrition is minimal but positive.
- **Negative Coefficients:**
  - **DistanceFromHome (0.023927)**: The effect is very small, indicating that living further from work does not significantly influence attrition.
  - **MonthlyIncome (-0.000097)**: Higher monthly income is slightly associated with lower attrition, though the effect is negligible.
  - **Age (-0.043409)**: Older employees are slightly less likely to leave, suggesting that age may provide stability.
  - **WorkLifeBalance (-0.292441)**: Poor work-life balance significantly increases the likelihood of attrition, indicating that employees who feel overworked are more likely to leave.
  - **JobSatisfaction (-0.369412)**: This is the second strongest negative predictor of attrition. Lower job satisfaction is significantly linked to higher attrition rates, highlighting the importance of employee satisfaction in retention.

Summary

From this analysis, the most significant factors associated with employee attrition are:

- **OverTime**: Strongly associated with higher attrition.
- **JobSatisfaction**: Employees with lower job satisfaction are significantly more likely to leave.
- **WorkLifeBalance**: Poor work-life balance is a strong predictor of leaving.

Next Steps

To address attrition, HR strategies could focus on:

- Managing overtime and ensuring employees are not overburdened.
- Improving job satisfaction through feedback mechanisms, employee engagement, and professional development opportunities.
- Providing support for work-life balance to retain employees.

If you need further analysis or visualizations based on these coefficients, feel free to ask!

In [ ]:

```

```