```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from scipy import stats

df = pd.read_csv("employee_experience_survey_data (1).csv")

df
```

	Name	Age Bracket	Gender	Ethnicity
0	John Doe	25-34	Female	Asian
1	Jane Smith	18-24	Female	Middle Eastern
2	Carlos Reyes	45-54	Female	Indian
3	Emily Zhang	35-44	Male	Caucasian
4	Michael Johnson	18-24	Female	Caucasian
5	Sara Ahmed	45-54	Male	Middle Eastern
6	Tom Davis	25-34	Male	Caucasian
7	Linda Lopez	18-24	Male	African American
8	Raj Patel	35-44	Female	Indian
9	Amara Njeri	18-24	Male	African American
10	Helen Martin	45-54	Male	Asian
11	Peter Brown	18-24	Female	Indian
12	Nina Roberts	25-34	Female	Middle Eastern
13	Jake Williams	18-24	Female	Hispanic
14	Aisha Thompson	18-24	Female	Middle Eastern

	Job Title	Department	Date Survey Completed
0	Product Manager	Product Development	2024-10-05
1	Operations Manager	Sales	2024-10-07
2	UX Designer	Consulting	2024-10-08
3	UX Designer	HR	2024-10-07
4	UX Designer	Product Development	2024-10-07
5	Business Consultant	Operations	2024-10-09
6	UX Designer	HR	2024-10-08
7	Customer Support	Product Development	2024-10-10
8	Product Manager	IT	2024-10-07
9	HR Specialist	Design	2024-10-10
10	HR Specialist	Consulting	2024-10-05
11	HR Specialist	Sales	2024-10-12
12	Sales Executive	Sales	2024-10-11
13	Business Consultant	Sales	2024-10-08
14	Data Analyst	Finance	2024-10-12

	Job Satisfaction	Work-Life Balance	Management Support
0	Disagree	Strongly Agree	Neutral
1	Agree	Strongly Disagree	Strongly Agree
2	Neutral	Strongly Disagree	Agree
3	Neutral	Agree	Agree
4	Agree	Strongly Agree	Disagree

```
5            Disagree  Strongly Disagree            Neutral
6      Strongly Agree     Strongly Agree  Strongly Disagree
7            Disagree              Agree            Neutral
8   Strongly Disagree     Strongly Agree  Strongly Disagree
9      Strongly Agree     Strongly Agree              Agree
10           Neutral             Neutral  Strongly Disagree
11  Strongly Disagree           Disagree              Agree
12           Disagree              Agree  Strongly Disagree
13              Agree  Strongly Disagree     Strongly Agree
14              Agree              Agree  Strongly Disagree

   Team Collaboration  Workload Fairness Career Development
Opportunities  \
0             Neutral            Neutral
Disagree
1             Neutral            Neutral
Disagree
2            Disagree     Strongly Agree
Agree
3   Strongly Disagree  Strongly Disagree                   Strongly
Agree
4             Neutral  Strongly Disagree
Disagree
5             Neutral              Agree
Agree
6      Strongly Agree     Strongly Agree                Strongly
Disagree
7   Strongly Disagree  Strongly Disagree
Agree
8               Agree     Strongly Agree
Agree
9               Agree     Strongly Agree                Strongly
Disagree
10              Agree     Strongly Agree                   Strongly
Agree
11     Strongly Agree            Neutral
Agree
12  Strongly Disagree            Neutral
Agree
13            Neutral            Neutral
Neutral
14  Strongly Disagree           Disagree                Strongly
Disagree

   Workplace Inclusivity Company Communication Compensation
Satisfaction  \
0                  Agree        Strongly Agree            Strongly
Agree
1                Neutral               Neutral
```

```
Neutral
2                  Agree     Strongly Disagree
Neutral
3                  Agree     Strongly Disagree          Strongly
Disagree
4               Disagree              Disagree          Strongly
Disagree
5               Disagree               Neutral
Agree
6         Strongly Agree     Strongly Disagree
Neutral
7                Neutral                 Agree
Disagree
8                  Agree               Neutral
Agree
9                Neutral                 Agree
Neutral
10        Strongly Agree     Strongly Disagree
Disagree
11               Neutral        Strongly Agree
Neutral
12               Neutral                 Agree
Disagree
13        Strongly Agree              Disagree
Neutral
14              Disagree        Strongly Agree
Disagree

         Job Security Overall Engagement
0               Agree     Strongly Agree
1               Agree            Neutral
2      Strongly Agree     Strongly Agree
3               Agree            Neutral
4             Neutral           Disagree
5   Strongly Disagree            Neutral
6             Neutral              Agree
7            Disagree              Agree
8      Strongly Agree           Disagree
9             Neutral  Strongly Disagree
10  Strongly Disagree            Neutral
11              Agree              Agree
12              Agree           Disagree
13              Agree     Strongly Agree
14              Agree     Strongly Agree

df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 15 entries, 0 to 14
Data columns (total 18 columns):
```

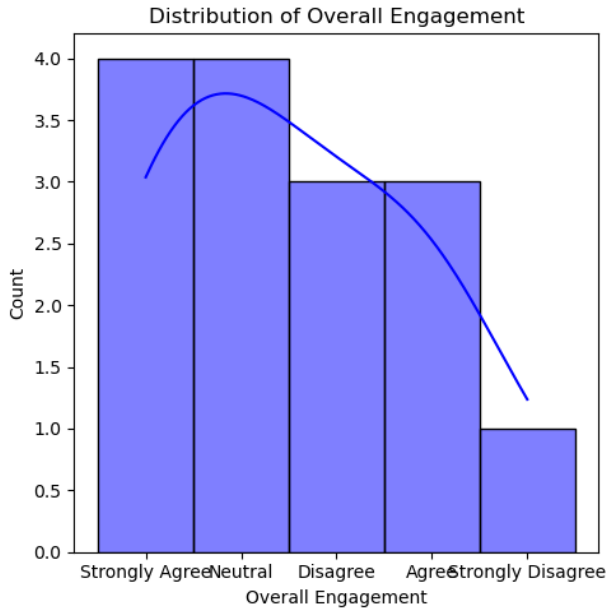
```
 #   Column                                Non-Null Count  Dtype
---  ------                                --------------  -----
 0   Name                                  15 non-null     object
 1   Age Bracket                           15 non-null     object
 2   Gender                                15 non-null     object
 3   Ethnicity                             15 non-null     object
 4   Job Title                             15 non-null     object
 5   Department                            15 non-null     object
 6   Date Survey Completed                 15 non-null     object
 7   Job Satisfaction                      15 non-null     object
 8   Work-Life Balance                     15 non-null     object
 9   Management Support                    15 non-null     object
 10  Team Collaboration                    15 non-null     object
 11  Workload Fairness                     15 non-null     object
 12  Career Development Opportunities      15 non-null     object
 13  Workplace Inclusivity                 15 non-null     object
 14  Company Communication                 15 non-null     object
 15  Compensation Satisfaction             15 non-null     object
 16  Job Security                          15 non-null     object
 17  Overall Engagement                    15 non-null     object
dtypes: object(18)
memory usage: 2.2+ KB

df.describe()

            Name Age Bracket  Gender       Ethnicity    Job Title
Department  \
count         15           15      15              15           15
15
unique        15            4       2               6            8
8
top     John Doe       18-24  Female  Middle Eastern  UX Designer
Sales
freq           1            7       9               4            4
4

       Date Survey Completed Job Satisfaction Work-Life Balance  \
count                     15               15                15
unique                     7                5                 5
top               2024-10-07         Disagree    Strongly Agree
freq                       4                4                 5

       Management Support Team Collaboration Workload Fairness  \
count                  15                 15                15
unique                  5                  5                 5
top     Strongly Disagree            Neutral           Neutral
freq                    5                  5                 5

       Career Development Opportunities Workplace Inclusivity  \
count                                15                    15
```

```
unique                                  5                     4
top                                 Agree               Neutral
freq                                    6                     5

       Company Communication Compensation Satisfaction Job Security  \
count                      15                         15           15
unique                      5                          5            5
top         Strongly Disagree                    Neutral        Agree
freq                        4                          6            7

       Overall Engagement
count                  15
unique                  5
top         Strongly Agree
freq                    4
```

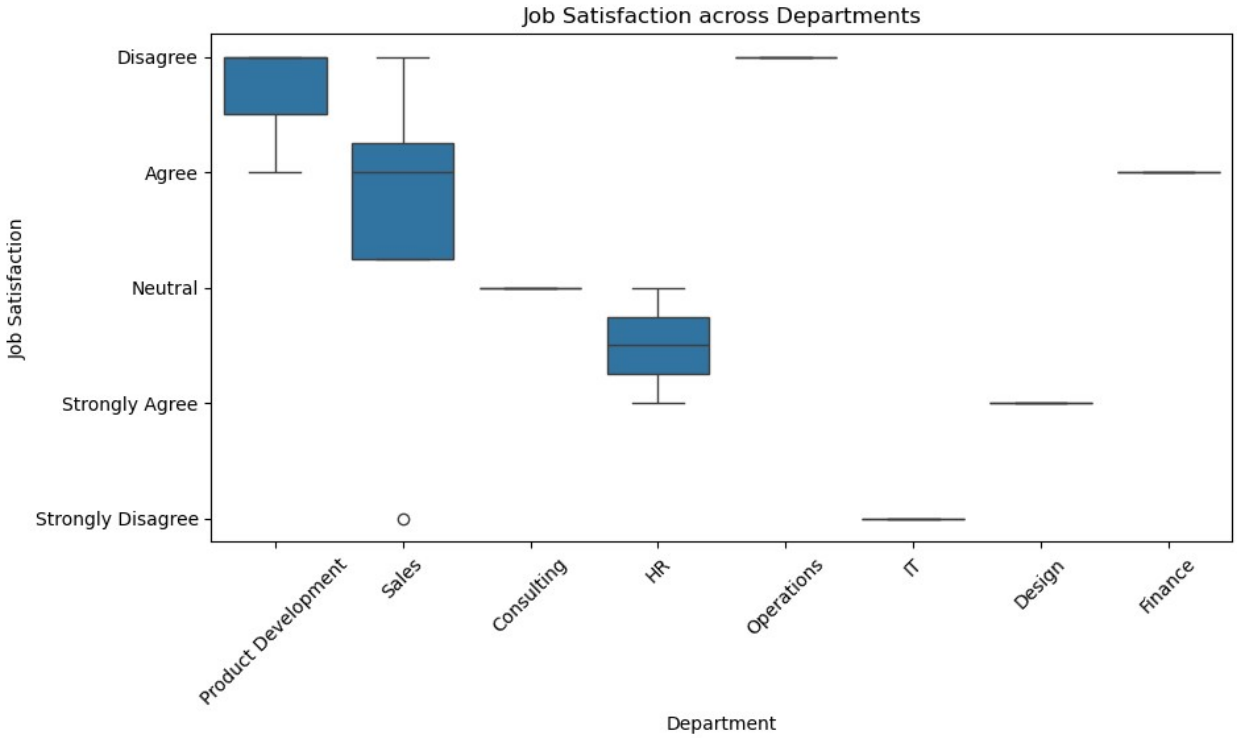
```
# Descriptive Statistics: Summary for 'Overall Engagement' and 'Job
Satisfaction'
print("\nDescriptive statistics for Overall Engagement and Job
Satisfaction:")
print(df[['Overall Engagement', 'Job Satisfaction']].describe())
```

```
Descriptive statistics for Overall Engagement and Job Satisfaction:
       Overall Engagement Job Satisfaction
count                  15               15
unique                  5                5
top         Strongly Agree         Disagree
freq                    4                4
```

```
#Visualization: Distribution of Overall Engagement and Job
Satisfaction
plt.figure(figsize=(10,5))

plt.subplot(1, 2, 1)
sns.histplot(df['Overall Engagement'], bins=10, kde=True,
color='blue')
plt.title('Distribution of Overall Engagement')

plt.subplot(1, 2, 2)
sns.histplot(df['Job Satisfaction'], bins=10, kde=True, color='green')
plt.title('Distribution of Job Satisfaction')

plt.tight_layout()
plt.show()
```




```
# Analysis of Satisfaction across Departments
plt.figure(figsize=(10,5))
sns.boxplot(x='Department', y='Job Satisfaction', data=df)
plt.title('Job Satisfaction across Departments')
plt.xticks(rotation=45)
plt.show()
```



```
# Create a mapping for Likert scale to numeric values
likert_scale_mapping = {
    'Strongly Disagree': 1,
    'Disagree': 2,
    'Neutral': 3,
    'Agree': 4,
    'Strongly Agree': 5
}

# Apply the mapping to the 'Job Satisfaction' column
df['Job Satisfaction'] = df['Job
Satisfaction'].map(likert_scale_mapping)

# Ensure no NaN values remain after the conversion
df = df.dropna(subset=['Job Satisfaction'])

# Now perform the t-test between IT and HR departments after
conversion
dept1 = df[df['Department'] == 'IT']['Job Satisfaction']
dept2 = df[df['Department'] == 'HR']['Job Satisfaction']

# Perform a t-test
t_stat, p_value = stats.ttest_ind(dept1, dept2)

print("\nT-test results between IT and HR:")
print(f"T-statistic: {t_stat}, P-value: {p_value}")
```

```
T-test results between IT and HR:
T-statistic: -1.7320508075688774, P-value: 0.33333333333333326
```

```
# Create a mapping for Likert scale to numeric values
likert_scale_mapping = {
    'Strongly Disagree': 1,
    'Disagree': 2,
    'Neutral': 3,
    'Agree': 4,
    'Strongly Agree': 5
}
# Step 3: Apply the mapping to 'Work-Life Balance' and 'Overall
Engagement'
df['Work-Life Balance'] = df['Work-Life
Balance'].map(likert_scale_mapping)
df['Overall Engagement'] = df['Overall
Engagement'].map(likert_scale_mapping)

# Step 4: Check for missing (NaN) values after mapping
print("\nNumber of NaN values in 'Work-Life Balance':", df['Work-Life
Balance'].isna().sum())
print("Number of NaN values in 'Overall Engagement':", df['Overall
Engagement'].isna().sum())
```

```
# Step 5: Drop rows with NaN values in these columns
df = df.dropna(subset=['Work-Life Balance', 'Overall Engagement'])

# Step 6: Calculate the correlation between Work-Life Balance and
Overall Engagement
correlation = df['Work-Life Balance'].corr(df['Overall Engagement'])
print(f"\nCorrelation between Work-Life Balance and Overall
Engagement: {correlation}")

# Step 7: Visualization of correlation
plt.figure(figsize=(6,6))
sns.scatterplot(x='Work-Life Balance', y='Overall Engagement',
data=df)
plt.title('Work-Life Balance vs Overall Engagement')
plt.show()
```

```
Number of NaN values in 'Work-Life Balance': 0
Number of NaN values in 'Overall Engagement': 0

Correlation between Work-Life Balance and Overall Engagement: -
0.36104082040916596
```


Work-Life Balance vs Overall Engagement
Overall Engagement
Work-Life Balance
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0