```
#Descriptive
```

```
import pandas as pd
df=pd.read_csv('/content/employee_experience_survey_data.csv')
df.head()
```

	Name	Age Bracket	Gender	Ethnicity	Job Title	Department	Date Survey Completed	Job Satisfaction	Work-Life Balance	Management Support	Team Collaboration	W F
0	John Doe	25-34	Female	Asian	Product Manager	Product Development	2024-10-05	Disagree	Strongly Agree	Neutral	Neutral	
1	Jane Smith	18-24	Female	Middle Eastern	Operations Manager	Sales	2024-10-07	Agree	Strongly Disagree	Strongly Agree	Neutral	
2	Carlos Reyes	45-54	Female	Indian	UX Designer	Consulting	2024-10-08	Neutral	Strongly Disagree	Agree	Disagree	
3	Emily Zhang	35-44	Male	Caucasian	UX Designer	HR	2024-10-07	Neutral	Agree	Agree	Strongly Disagree	
4	Michael Johnson	18-24	Female	Caucasian	UX Designer	Product Development	2024-10-07	Agree	Strongly Agree	Disagree	Neutral	

Next steps: Generate code with df View recommended plots New interactive sheet

```
review={'Strongly Disagree':1,
      'Disagree': 2,
    'Neutral': 3,
    'Agree': 4,
    'Strongly Agree': 5 }
```

```
data=df.copy()
colums=['Job Satisfaction', 'Work-Life Balance', 'Management Support',
    'Team Collaboration', 'Workload Fairness', 'Career Development Opportunities',
    'Workplace Inclusivity', 'Company Communication', 'Compensation Satisfaction',
    'Job Security', 'Overall Engagement'
]
data[colums]=data[colums].replace(review)
data.head()
```

```
<ipython-input-51-15ccba962fba>:7: FutureWarning: Downcasting behavior in `replace` is deprecated and will be removed in
  data[colums]=data[colums].replace(review)
```

	Name	Age Bracket	Gender	Ethnicity	Job Title	Department	Date Survey Completed	Job Satisfaction	Work-Life Balance	Management Support	Team Collaboration	W F
0	John Doe	25-34	Female	Asian	Product Manager	Product Development	2024-10-05	2	5	3	3	
1	Jane Smith	18-24	Female	Middle Eastern	Operations Manager	Sales	2024-10-07	4	1	5	3	
2	Carlos Reyes	45-54	Female	Indian	UX Designer	Consulting	2024-10-08	3	1	4	2	
3	Emily Zhang	35-44	Male	Caucasian	UX Designer	HR	2024-10-07	3	4	4	1	
4	Michael Johnson	18-24	Female	Caucasian	UX Designer	Product Development	2024-10-07	4	5	2	3	

Next steps: Generate code with data View recommended plots New interactive sheet

```
print(data[colums].describe())
```

```
       Job Satisfaction  Work-Life Balance  Management Support  \
count         15.000000          15.000000           15.000000   
mean           3.000000           3.333333            2.800000   
std            1.309307           1.676163            1.521278   
min            1.000000           1.000000            1.000000   
25%            2.000000           1.500000            1.000000   
50%            3.000000           4.000000            3.000000   
75%            4.000000           5.000000            4.000000   
max            5.000000           5.000000            5.000000   

       Team Collaboration  Workload Fairness  \
count           15.000000          15.000000   
mean             2.866667           3.266667   
```

```
std              1.407463           1.533747
min              1.000000           1.000000
25%              1.500000           2.500000
50%              3.000000           3.000000
75%              4.000000           5.000000
max              5.000000           5.000000

       Career Development Opportunities  Workplace Inclusivity  \
count                         15.000000              15.000000   
mean                           3.066667               3.466667   
std                            1.437591               1.060099   
min                            1.000000               2.000000   
25%                            2.000000               3.000000   
50%                            4.000000               3.000000   
75%                            4.000000               4.000000   
max                            5.000000               5.000000   

       Company Communication  Compensation Satisfaction  Job Security  \
count              15.000000                  15.000000     15.000000   
mean                2.933333                   2.733333      3.400000   
std                 1.533747                   1.099784      1.242118   
min                 1.000000                   1.000000      1.000000   
25%                 1.500000                   2.000000      3.000000   
50%                 3.000000                   3.000000      4.000000   
75%                 4.000000                   3.000000      4.000000   
max                 5.000000                   5.000000      5.000000   

       Overall Engagement  
count           15.000000  
mean             3.400000  
std              1.298351  
min              1.000000  
25%              2.500000  
50%              3.000000  
75%              4.500000  
max              5.000000  
```

```
#Related to Job
column1=['Job Satisfaction',
     'Overall Engagement'
]
gender=data.groupby('Gender')[column1].describe()
print(gender)
```

```
       Job Satisfaction                                                \
                  count      mean       std  min   25%  50%  75%  max   
Gender                                                                  
Female              9.0  2.777778  1.301708  1.0  2.00  3.0  4.0  4.0   
Male                6.0  3.333333  1.366260  2.0  2.25  3.0  4.5  5.0   

       Overall Engagement                                           
                    count      mean       std  min  25%  50%   75%  max  
Gender                                                                   
Female                9.0  3.666667  1.414214  2.0  2.0  4.0  5.00  5.0  
Male                  6.0  3.000000  1.095445  1.0  3.0  3.0  3.75  4.0  
```

```
#Related to work
column2=[ 'Work-Life Balance',
     'Workload Fairness',
    'Workplace Inclusivity',
]
gender=data.groupby('Gender')[column2].describe()
print(gender)
```

```
       Work-Life Balance                                                 \
                   count      mean       std  min   25%  50%   75%  max   
Gender                                                                    
Female               9.0  3.111111  1.833333  1.0  1.00  4.0  5.00  5.0   
Male                 6.0  3.666667  1.505545  1.0  3.25  4.0  4.75  5.0   

       Workload Fairness            ...           Workplace Inclusivity  \
                   count      mean  ...  75%  max                 count   
Gender                              ...                                   
Female               9.0  3.111111  ...  3.0  5.0                   9.0   
Male                 6.0  3.500000  ...  5.0  5.0                   6.0   


            mean      std  min  25%  50%   75%  max  
Gender                                               
Female  3.333333  1.00000  2.0  3.0  3.0  4.00  5.0  
Male    3.666667  1.21106  2.0  3.0  3.5  4.75  5.0  

[2 rows x 24 columns]
```

```
#Related to Job
column1=['Job Satisfaction',
     'Overall Engagement'
]
Age=data.groupby('Age Bracket')[column1].describe()
print(Age)
```

```
            Job Satisfaction                                             \
                       count      mean       std  min  25%  50%  75%  max
Age Bracket
18-24                    7.0  3.428571  1.397276  1.0  3.0  4.0  4.0  5.0
25-34                    3.0  3.000000  1.732051  2.0  2.0  2.0  3.5  5.0
35-44                    2.0  2.000000  1.414214  1.0  1.5  2.0  2.5  3.0
45-54                    3.0  2.666667  0.577350  2.0  2.5  3.0  3.0  3.0

            Overall Engagement
                         count      mean       std  min   25%  50%   75%  max
Age Bracket
18-24                      7.0  3.428571  1.511858  1.0  2.50  4.0  4.50  5.0
25-34                      3.0  3.666667  1.527525  2.0  3.00  4.0  4.50  5.0
35-44                      2.0  2.500000  0.707107  2.0  2.25  2.5  2.75  3.0
45-54                      3.0  3.666667  1.154701  3.0  3.00  3.0  4.00  5.0
```

```
#Related to work
column2=[ 'Work-Life Balance',
     'Workload Fairness',
    'Workplace Inclusivity',
]
Age=data.groupby('Age Bracket')[column2].describe()
print(Age)
```

```
            Work-Life Balance                                              \
                        count      mean       std  min   25%  50%   75%  max
Age Bracket
18-24                     7.0  3.142857  1.772811  1.0  1.50  4.0  4.50  5.0
25-34                     3.0  4.666667  0.577350  4.0  4.50  5.0  5.00  5.0
35-44                     2.0  4.500000  0.707107  4.0  4.25  4.5  4.75  5.0
45-54                     3.0  1.666667  1.154701  1.0  1.00  1.0  2.00  3.0

            Workload Fairness           ...          Workplace Inclusivity  \
                        count      mean ...  75%  max                 count
Age Bracket                             ...
18-24                     7.0  2.571429 ...  3.0  5.0                   7.0
25-34                     3.0  3.666667 ...  4.0  5.0                   3.0
35-44                     2.0  3.000000 ...  4.0  5.0                   2.0
45-54                     3.0  4.666667 ...  5.0  5.0                   3.0


                 mean       std  min  25%  50%  75%  max
Age Bracket
18-24        3.000000  1.000000  2.0  2.5  3.0  3.0  5.0
25-34        4.000000  1.000000  3.0  3.5  4.0  4.5  5.0
35-44        4.000000  0.000000  4.0  4.0  4.0  4.0  4.0
45-54        3.666667  1.527525  2.0  3.0  4.0  4.5  5.0

[4 rows x 24 columns]
```

```
#Related to Job
column1=['Job Satisfaction',
     'Overall Engagement'
]
Dept=data.groupby('Department')[column1].describe()
print(Dept)
```

```
                    Job Satisfaction                                       \
                               count      mean       std  min   25%  50%  75%
Department
Consulting                       2.0  3.000000  0.000000  3.0  3.00  3.0  3.0
Design                           1.0  5.000000       NaN  5.0  5.00  5.0  5.0
Finance                          1.0  4.000000       NaN  4.0  4.00  4.0  4.0
HR                               2.0  4.000000  1.414214  3.0  3.50  4.0  4.5
IT                               1.0  1.000000       NaN  1.0  1.00  1.0  1.0
Operations                       1.0  2.000000       NaN  2.0  2.00  2.0  2.0
Product Development              3.0  2.666667  1.154701  2.0  2.00  2.0  3.0
Sales                            4.0  2.750000  1.500000  1.0  1.75  3.0  4.0

                         Overall Engagement                                  \
                    max              count      mean       std  min   25%
Department
Consulting          3.0                2.0  4.000000  1.414214  3.0  3.50
Design              5.0                1.0  1.000000       NaN  1.0  1.00
Finance             4.0                1.0  5.000000       NaN  5.0  5.00
HR                  5.0                2.0  3.500000  0.707107  3.0  3.25
IT                  1.0                1.0  2.000000       NaN  2.0  2.00
Operations          2.0                1.0  3.000000       NaN  3.0  3.00
Product Development 4.0                3.0  3.666667  1.527525  2.0  3.00
Sales               4.0                4.0  3.500000  1.290994  2.0  2.75
```

```
                     50%   75%  max
Department
Consulting           4.0  4.50  5.0
Design               1.0  1.00  1.0
Finance              5.0  5.00  5.0
HR                   3.5  3.75  4.0
IT                   2.0  2.00  2.0
Operations           3.0  3.00  3.0
Product Development  4.0  4.50  5.0
Sales                3.5  4.25  5.0
```

```
#Related to work
column2=[ 'Work-Life Balance',
     'Workload Fairness',
    'Workplace Inclusivity',
]
Dept=data.groupby('Department')[column2].describe()
print(Dept)
```

```
                    Work-Life Balance                                      \
                                count      mean       std  min   25%  50%
Department
Consulting                        2.0  2.000000  1.414214  1.0  1.50  2.0
Design                            1.0  5.000000       NaN  5.0  5.00  5.0
Finance                           1.0  4.000000       NaN  4.0  4.00  4.0
HR                                2.0  4.500000  0.707107  4.0  4.25  4.5
IT                                1.0  5.000000       NaN  5.0  5.00  5.0
Operations                        1.0  1.000000       NaN  1.0  1.00  1.0
Product Development               3.0  4.666667  0.577350  4.0  4.50  5.0
Sales                             4.0  2.000000  1.414214  1.0  1.00  1.5

                               Workload Fairness            ...            \
                      75%  max             count      mean  ...  75%  max
Department                                                  ...
Consulting           2.50  3.0               2.0  5.000000  ...  5.0  5.0
Design               5.00  5.0               1.0  5.000000  ...  5.0  5.0
Finance              4.00  4.0               1.0  2.000000  ...  2.0  2.0
HR                   4.75  5.0               2.0  3.000000  ...  4.0  5.0
IT                   5.00  5.0               1.0  5.000000  ...  5.0  5.0
Operations           1.00  1.0               1.0  4.000000  ...  4.0  4.0
Product Development  5.00  5.0               3.0  1.666667  ...  2.0  3.0
Sales                2.50  4.0               4.0  3.000000  ...  3.0  3.0

                    Workplace Inclusivity                               \
                                    count mean       std  min   25%  50%
Department
Consulting                            2.0  4.5  0.707107  4.0  4.25  4.5
Design                                1.0  3.0       NaN  3.0  3.00  3.0
Finance                               1.0  2.0       NaN  2.0  2.00  2.0
HR                                    2.0  4.5  0.707107  4.0  4.25  4.5
IT                                    1.0  4.0       NaN  4.0  4.00  4.0
Operations                            1.0  2.0       NaN  2.0  2.00  2.0
Product Development                   3.0  3.0  1.000000  2.0  2.50  3.0
Sales                                 4.0  3.5  1.000000  3.0  3.00  3.0


                      75%  max
Department
Consulting           4.75  5.0
Design               3.00  3.0
Finance              2.00  2.0
HR                   4.75  5.0
IT                   4.00  4.0
Operations           2.00  2.0
Product Development  3.50  4.0
Sales                3.50  5.0

[8 rows x 24 columns]
```

```
#Related to Job
column1=['Job Satisfaction',
     'Overall Engagement'
]
Eth=data.groupby('Ethnicity')[column1].describe()
print(Eth)
```

```
                 Job Satisfaction                                        \
                            count      mean       std  min   25%  50%   75%
Ethnicity
African American              2.0  3.500000  2.121320  2.0  2.75  3.5  4.25
Asian                         2.0  2.500000  0.707107  2.0  2.25  2.5  2.75
Caucasian                     3.0  4.000000  1.000000  3.0  3.50  4.0  4.50
Hispanic                      1.0  4.000000       NaN  4.0  4.00  4.0  4.00
Indian                        3.0  1.666667  1.154701  1.0  1.00  1.0  2.00
Middle Eastern                4.0  3.000000  1.154701  2.0  2.00  3.0  4.00

                      Overall Engagement                                   \
                  max               count      mean       std  min   25%  50%
Ethnicity
African American  5.0                 2.0  2.500000  2.121320  1.0  1.75  2.5
Asian             3.0                 2.0  4.000000  1.414214  3.0  3.50  4.0
Caucasian         5.0                 3.0  3.000000  1.000000  2.0  2.50  3.0
Hispanic          4.0                 1.0  5.000000       NaN  5.0  5.00  5.0
Indian            3.0                 3.0  3.666667  1.527525  2.0  3.00  4.0
Middle Eastern    4.0                 4.0  3.250000  1.258306  2.0  2.75  3.0


                   75%  max
Ethnicity
African American  3.25  4.0
Asian             4.50  5.0
Caucasian         3.50  4.0
Hispanic          5.00  5.0
Indian            4.50  5.0
Middle Eastern    3.50  5.0
```

#2 Infrentitial Statistics

```
from scipy import stats
job1 = data[data['Department'] == 'IT']['Job Satisfaction']
job2 = data[data['Department'] == 'HR']['Job Satisfaction']
t_stat, p_value = stats.ttest_ind(job1, job2, nan_policy='omit')
print(f"T-statistic: {t_stat}, P-value: {p_value}")
correlation_coefficient, p_value_corr = stats.pearsonr(data['Work-Life Balance'], data['Overall Engagement'])
print(f"Correlation Coefficient: {correlation_coefficient}, P-value: {p_value_corr}")
```

```
T-statistic: -1.7320508075688774, P-value: 0.33333333333333326
Correlation Coefficient: -0.361040820409166, P-value: 0.18612659993280437
```

```
from scipy import stats
job1 = data[data['Department'] == 'HR']['Job Satisfaction']
job2 = data[data['Department'] == 'Sales']['Job Satisfaction']
t_stat, p_value = stats.ttest_ind(job1, job2, nan_policy='omit')
print(f"T-statistic: {t_stat}, P-value: {p_value}")
correlation_coefficient, p_value_corr = stats.pearsonr(data['Work-Life Balance'], data['Overall Engagement'])
print(f"Correlation Coefficient: {correlation_coefficient}, P-value: {p_value_corr}")
```

```
T-statistic: 0.9759000729485332, P-value: 0.3843727364479507
Correlation Coefficient: -0.361040820409166, P-value: 0.18612659993280437
```

```
from scipy import stats
job1 = data[data['Department'] == 'Consulting']['Job Satisfaction']
job2 = data[data['Department'] == 'Product Development']['Job Satisfaction']
t_stat, p_value = stats.ttest_ind(job1, job2, nan_policy='omit')
print(f"T-statistic: {t_stat}, P-value: {p_value}")
correlation_coefficient, p_value_corr = stats.pearsonr(data['Work-Life Balance'], data['Overall Engagement'])
print(f"Correlation Coefficient: {correlation_coefficient}, P-value: {p_value_corr}")
```

```
T-statistic: 0.38729833462074187, P-value: 0.7243774861131499
Correlation Coefficient: -0.361040820409166, P-value: 0.18612659993280437
```

```
from scipy import stats
job1 = data[data['Gender'] == 'Male']['Job Satisfaction']
job2 = data[data['Gender'] == 'Female']['Job Satisfaction']
t_stat, p_value = stats.ttest_ind(job1, job2, nan_policy='omit')
print(f"T-statistic: {t_stat}, P-value: {p_value}")
correlation_coefficient, p_value_corr = stats.pearsonr(data['Work-Life Balance'], data['Overall Engagement'])
print(f"Correlation Coefficient: {correlation_coefficient}, P-value: {p_value_corr}")
```

```
T-statistic: 0.7943978607505493, P-value: 0.44122530370600777
Correlation Coefficient: -0.361040820409166, P-value: 0.18612659993280437
```