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

# Read Excel File
df = pd.read_excel("Medical_Analysis.xlsx", )

# Dropping Index Col
df.drop(columns=["S.no"], inplace=True)

# Data information
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 81 entries, 0 to 80
Data columns (total 57 columns):
 #   Column
Non-Null Count  Dtype
---  ------
--------------  -----
 0   Name
81 non-null     object
 1   Age
81 non-null     int64
 2   Age Group
81 non-null     object
 3   Sex
81 non-null     object
 4   Height
81 non-null     float64
 5   Weight
81 non-null     int64
 6   BMI
81 non-null     float64
 7   Marital status
81 non-null     object
 8   Since how many years are you involved in this occupation?
81 non-null     int64
 9   How many hours are you working per week in this job?
81 non-null     int64
 10  BP
81 non-null     object
 11  Diabetes
81 non-null     object
 12  Thyroid
81 non-null     object
 13  Hernia
81 non-null     object
 14  Are you suffering from musculoskeletal pain / discomfort now ?
```

```
81 non-null     object
 15  Pain Yes/No
81 non-null     object
 16  Muscle pain duration in Years
81 non-null     int64
 17  Wrist and arm score
81 non-null     int64
 18  Neck,Trunk,Leg score
81 non-null     int64
 19  RULA final score
81 non-null     int64
 20  Do the current musculoskeletal symptoms (pain/discomfort)
62 non-null     object
 21  Following any coping strategy (on your own)? If yes, details
81 non-null     object
 22  Do you involve yourself in any recreational activities like
exercise or yoga?                                          81 non-
null     object
 23  Smoking
81 non-null     object
 24  Drinking
81 non-null     object
 25  Neck Pain
81 non-null     object
 26  Shoulder Pain
81 non-null     object
 27  Elbow Pain
81 non-null     object
 28  Wrist / Hand Pain
81 non-null     object
 29  Upper Back
81 non-null     object
 30  Lower Back
81 non-null     object
 31  Both Hip / Thigh Pain
81 non-null     object
 32  Knee Pain
81 non-null     object
 33  Ankle / Feet Pain
81 non-null     object
 34  2. Have You At Anytime During Last 12 Months Been Prevented From
Doing Your Normal Work Because Of Trouble?
Neck Pain?  81 non-null     object
 35  Shoulder Pain?
81 non-null     object
 36  Elbow?
81 non-null     object
 37  Wrist / Hand Pain?
81 non-null     object
```

```
 38  Upper Back?
81 non-null     object
 39  Lower Back?
80 non-null     object
 40  Hip / Thigh Pain?
81 non-null     object
 41  Knee Pain2
81 non-null     object
 42  Ankle / Feet Pain ?
81 non-null     object
 43  3. At The Time Of Initial Onset Of Trouble What Was Your Age ?
81 non-null     object
 44  4. Have You Ever Been Hospitalized Because Of Trouble?
81 non-null     object
 45  5. Have You Ever Had To Change Job / Duties Because Of Trouble?
81 non-null     object
 46  6. Have You Had Trouble At Anytime During Last 12 Months?
81 non-null     object
 47  7. Do You Take Take Medicine Because Of Trouble?
81 non-null     object
 48  8. Have You Taken Sick Leave From Work?
81 non-null     object
 49  Physical Functioning (PF)
81 non-null     float64
 50  Role Physical (RP)
81 non-null     float64
 51  Bodily Pain (BP)
81 non-null     float64
 52  General Health (GH)
81 non-null     float64
 53  Vitality (VT)
81 non-null     float64
 54  Social Functioning (SF)
81 non-null     float64
 55  Role Emotional (RE)
81 non-null     float64
 56  Mental Health (MH)
81 non-null     float64
dtypes: float64(10), int64(8), object(39)
memory usage: 36.2+ KB
```

```
# Statistical Description of data
np.around(df.describe(), 1)
```

```
        Age  Height  Weight   BMI  \
count  81.0    81.0    81.0  81.0
mean   42.1     5.5    73.6  26.1
std     9.3     0.3     8.8   4.4
min    22.0     5.1    55.0  19.9
25%    35.0     5.1    68.0  22.6
```

```
50%    43.0     5.6    73.0  24.5
75%    50.0     5.8    78.0  29.1
max    57.0     6.1   102.0  40.4

       Since how many years are you involved in this occupation?  \
count                                                  81.0
mean                                                   14.4
std                                                     9.1
min                                                     1.0
25%                                                     8.0
50%                                                    14.0
75%                                                    22.0
max                                                    32.0

       How many hours are you working per week in this job?   \
count                                                  81.0
mean                                                   50.8
std                                                     9.2
min                                                    28.0
25%                                                    50.0
50%                                                    50.0
75%                                                    60.0
max                                                    60.0

       Muscle pain duration in Years  Wrist and arm score   \
count                           81.0                  81.0
mean                             4.0                   5.2
std                              3.6                   0.4
min                              0.0                   4.0
25%                              0.0                   5.0
50%                              4.0                   5.0
75%                              6.0                   5.0
max                             10.0                   6.0

       Neck,Trunk,Leg score  RULA final score  Physical Functioning
(PF)  \
count                 81.0              81.0
81.0
mean                   4.1               5.1
67.7
std                    0.5               0.6
15.5
min                    3.0               4.0
35.0
25%                    4.0               5.0
55.0
50%                    4.0               5.0
75.0
75%                    4.0               5.0
80.0
```

```
max                      6.0                6.0
95.0

       Role Physical (RP)  Bodily Pain (BP)  General Health (GH)  \
count                81.0              81.0                 81.0
mean                 69.3              51.4                 45.5
std                  15.0              39.5                 13.3
min                  33.3               0.0                 20.0
25%                  50.0               0.0                 36.7
50%                  66.7              66.7                 46.7
75%                  83.3             100.0                 53.3
max                  83.3             100.0                 80.0

       Vitality (VT)  Social Functioning (SF)  Role Emotional (RE)  \
count           81.0                     81.0                 81.0
mean            53.5                     67.1                 56.9
std             11.3                     11.7                 11.7
min             10.0                     35.0                 26.7
25%             50.0                     60.0                 53.3
50%             55.0                     70.0                 53.3
75%             60.0                     75.0                 66.7
max             75.0                     90.0                 80.0

       Mental Health (MH)
count                81.0
mean                 41.8
std                   5.0
min                  32.0
25%                  40.0
50%                  44.0
75%                  44.0
max                  56.0
```

```
# Setting Index Value starting from 1
df.index += 1
```

Number of unique Items

```
# Number of Unique Items
df.nunique()
```

```
Name
78
Age
31
Age Group
4
Sex
1
Height
```

12
Weight
26
BMI
60
Marital status
2
Since how many years are you involved in this occupation?
22
How many hours are you working per week in this job?
7
BP
2
Diabetes
2
Thyroid
2
Hernia
2
Are you suffering from musculoskeletal pain / discomfort now ?
43
Pain Yes/No
2
Muscle pain duration in Years
10
Wrist and arm score
3
Neck,Trunk,Leg score
4
RULA final score
3
Do the current musculoskeletal symptoms (pain/discomfort)
3
Following any coping strategy (on your own)? If yes, details
7
Do you involve yourself in any recreational activities like exercise or yoga? 2
Smoking
2
Drinking
2
Neck Pain
2
Shoulder Pain
2
Elbow Pain
2
Wrist / Hand Pain
2

Upper Back
2
Lower Back
2
Both Hip / Thigh Pain
2
Knee Pain
2
Ankle / Feet Pain
1
2. Have You At Anytime During Last 12 Months Been Prevented From Doing Your Normal Work Because Of Trouble?\nNeck Pain? 2
Shoulder Pain?
2
Elbow?
2
Wrist / Hand Pain?
2
Upper Back?
2
Lower Back?
2
Hip / Thigh Pain?
2
Knee Pain2
2
Ankle / Feet Pain ?
1
3. At The Time Of Initial Onset Of Trouble What Was Your Age ?
22
4. Have You Ever Been Hospitalized Because Of Trouble?
2
5. Have You Ever Had To Change Job / Duties Because Of Trouble?
2
6. Have You Had Trouble At Anytime During Last 12 Months?
2
7. Do You Take Take Medicine Because Of Trouble?
2
8. Have You Taken Sick Leave From Work?
2
Physical Functioning (PF)
13
Role Physical (RP)
4
Bodily Pain (BP)
4
General Health (GH)
11
Vitality (VT)

```
11
Social Functioning (SF)
12
Role Emotional (RE)
9
Mental Health (MH)
7
dtype: int64
```

Columns

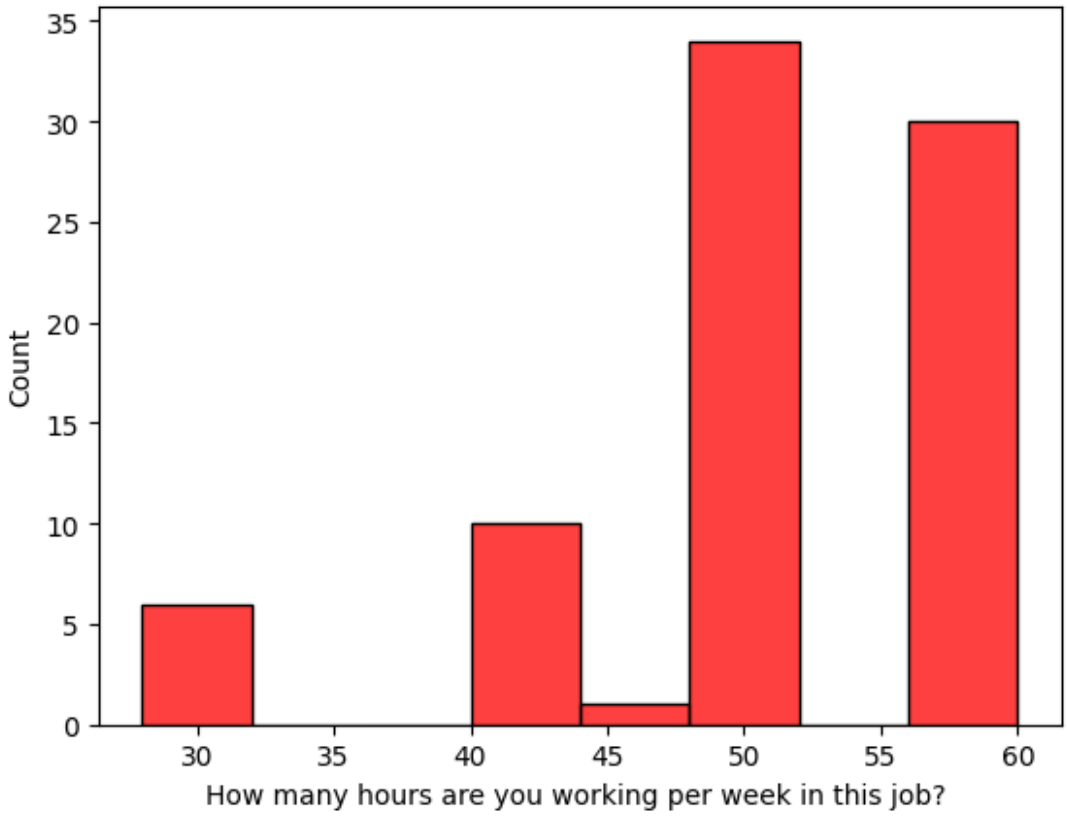
```
# Columns
df.columns.tolist()
```

```
['Name',
 'Age',
 'Age Group',
 'Sex',
 'Height',
 'Weight',
 'BMI',
 'Marital status',
 'Since how many years are you involved in this occupation?',
 'How many hours are you working per week in this job? ',
 'BP',
 'Diabetes',
 'Thyroid',
 'Hernia',
 'Are you suffering from musculoskeletal pain / discomfort now ? ',
 'Pain Yes/No',
 'Muscle pain duration in Years',
 'Wrist and arm score ',
 'Neck,Trunk,Leg score',
 'RULA final score',
 'Do the current musculoskeletal symptoms (pain/discomfort)',
 'Following any coping strategy (on your own)? If yes, details ',
 'Do you involve yourself in any recreational activities like exercise
or yoga? ',
 'Smoking',
 'Drinking',
 'Neck Pain',
 'Shoulder Pain',
 'Elbow Pain',
 'Wrist / Hand Pain',
 'Upper Back',
 'Lower Back',
 'Both Hip / Thigh Pain',
 'Knee Pain',
 'Ankle / Feet Pain',
 '2. Have You At Anytime During Last 12 Months Been Prevented From
```

```
Doing Your Normal Work Because Of Trouble?\nNeck Pain?',
 'Shoulder Pain?',
 'Elbow?',
 'Wrist / Hand Pain?',
 'Upper Back?',
 'Lower Back?',
 'Hip / Thigh Pain?',
 'Knee Pain2',
 'Ankle / Feet Pain ?',
 '3. At The Time Of Initial Onset Of Trouble What Was Your Age ?',
 '4. Have You Ever Been Hospitalized Because Of Trouble?',
 '5. Have You Ever Had To Change Job / Duties Because Of Trouble?',
 '6. Have You Had Trouble At Anytime During Last 12 Months?',
 '7. Do You Take Take Medicine Because Of Trouble?',
 '8. Have You Taken Sick Leave From Work?',
 'Physical Functioning (PF)',
 'Role Physical (RP)',
 'Bodily Pain (BP)',
 'General Health (GH)',
 'Vitality (VT)',
 'Social Functioning (SF)',
 'Role Emotional (RE)',
 'Mental Health (MH)']
```

Hours Working

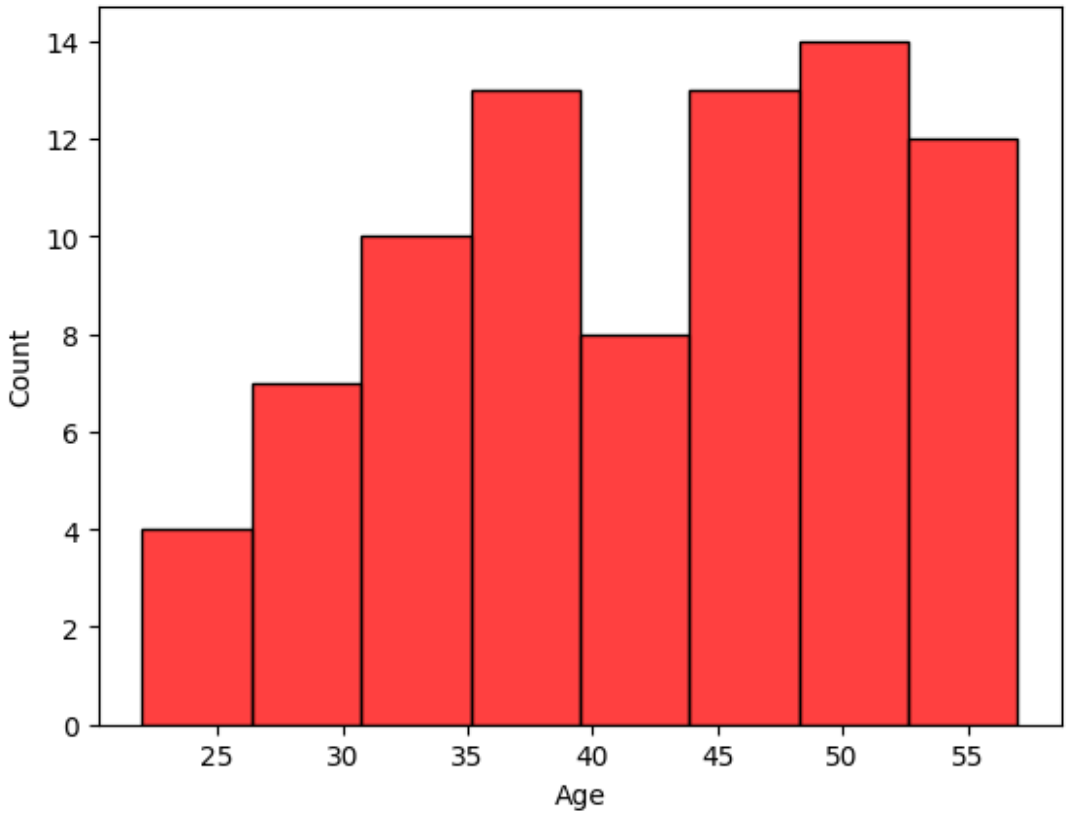
```
sns.histplot(data=df, x="How many hours are you working per week in
this job? ", color="red")
plt.show()
```



Age Histogram Chart

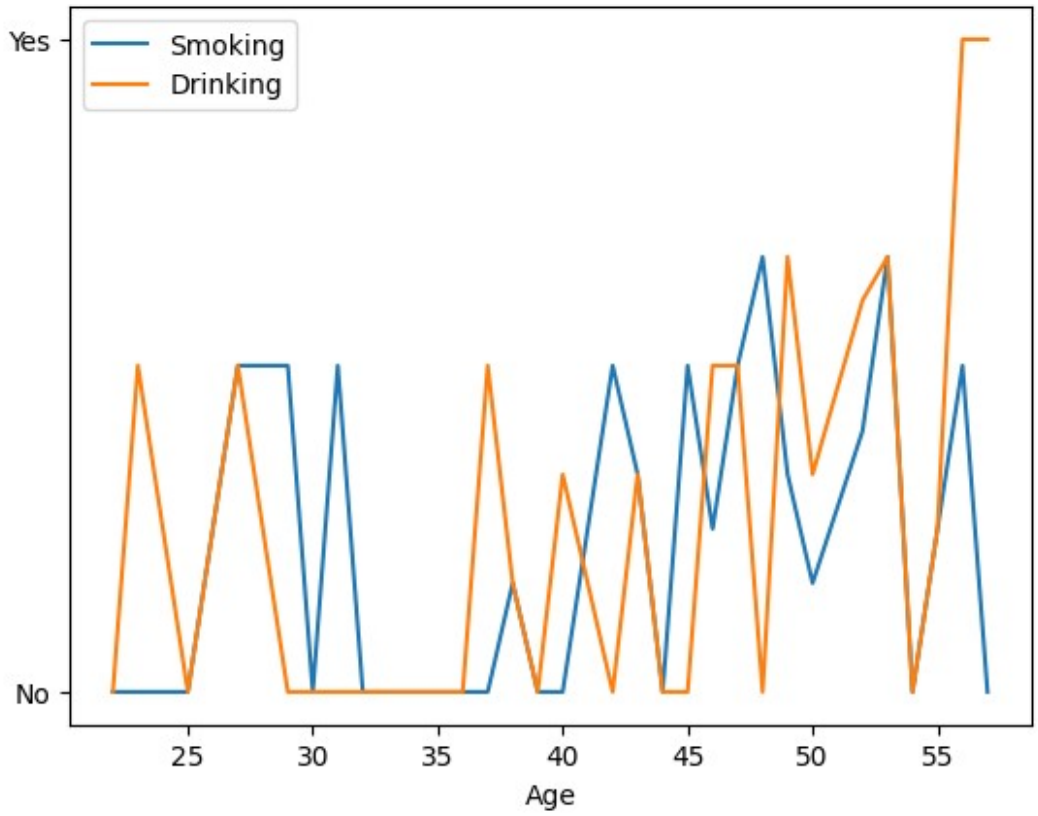
```
# Histogram distribution of Age
sns.histplot(data=df, x="Age", color='red')

plt.show()
```



Smokers and Drinkers vs Age

```
# Line Plots for Smokers and Drinkers in comparison to their age
%matplotlib inline
sns.lineplot(data=df, x="Age", y="Smoking", errorbar=None,
label="Smoking")
sns.lineplot(data=df, x="Age", y="Drinking", errorbar=None,
label="Drinking")
plt.ylabel("")
plt.legend()
plt.show()
```



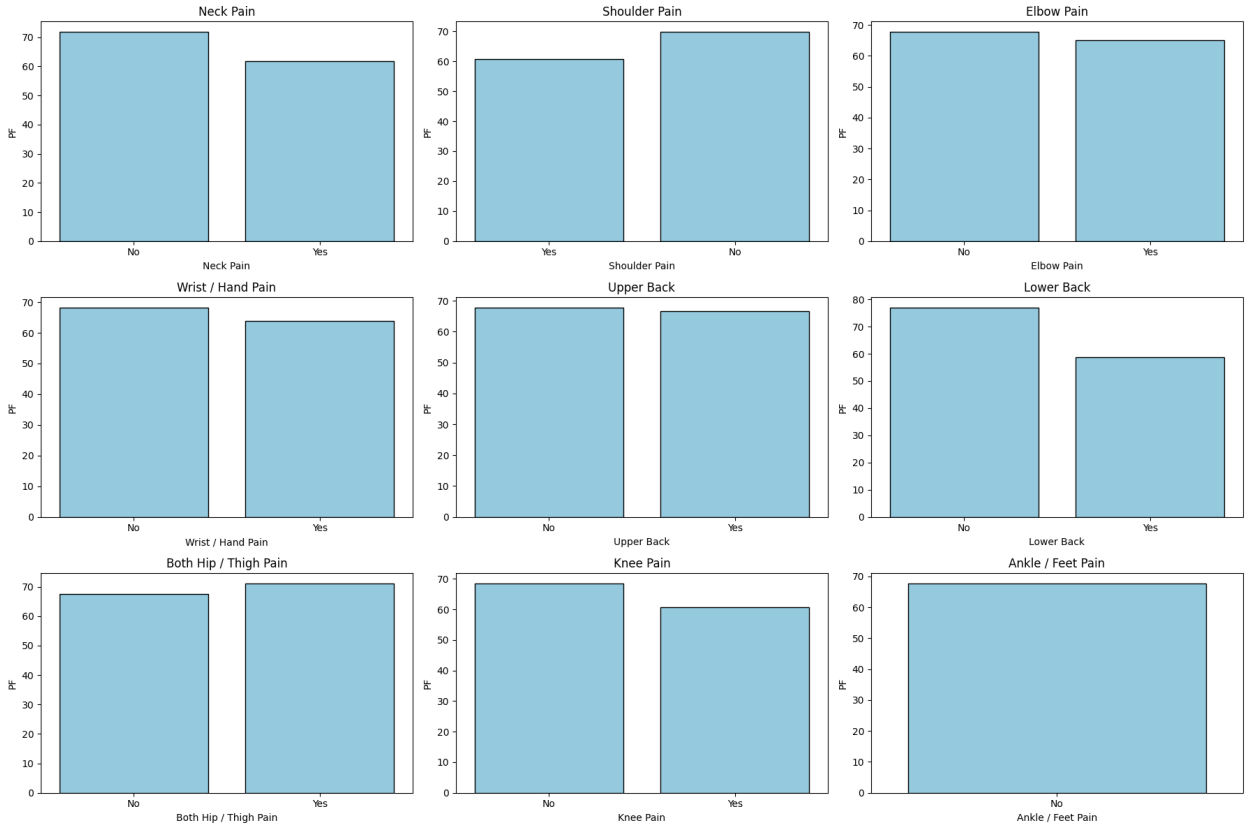
Comparison of Pains vs Physical Functioning

```
# List of pain columns
pain_columns = [
    'Neck Pain', 'Shoulder Pain', 'Elbow Pain', 'Wrist / Hand Pain',
    'Upper Back', 'Lower Back', 'Both Hip / Thigh Pain',
    'Knee Pain', 'Ankle / Feet Pain'
]

# Create a subplot for each pain type
fig, axes = plt.subplots(nrows=3, ncols=3, figsize=(18, 12))
axes = axes.flatten()

for i, pain in enumerate(pain_columns):
    ax = axes[i]
    sns.barplot(data=df, x=pain, y="Physical Functioning (PF)",
errorbar=None, ax=ax, color='skyblue', edgecolor="black")
    ax.set_title(pain)
    ax.set_ylabel("PF")
    ax.set_xlabel(pain)

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



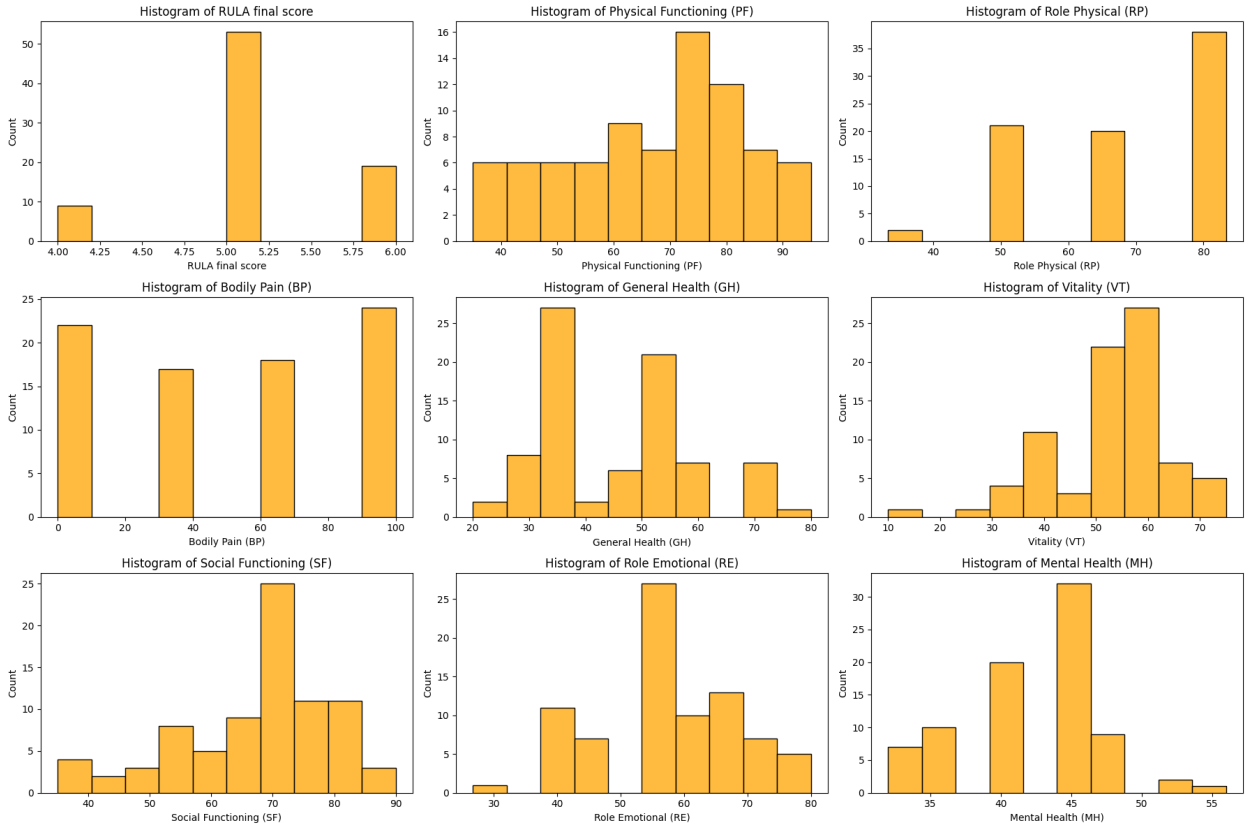
Comparison of RULA Score, SF-36 vs Counts

```
# List of target columns for histogram
columns = [
    'RULA final score',
    'Physical Functioning (PF)',
    'Role Physical (RP)',
    'Bodily Pain (BP)',
    'General Health (GH)',
    'Vitality (VT)',
    'Social Functioning (SF)',
    'Role Emotional (RE)',
    'Mental Health (MH)'
]

# Create subplots (3 rows x 3 columns)
fig, axes = plt.subplots(nrows=3, ncols=3, figsize=(18, 12))
axes = axes.flatten()

# Plot each histogram
for i, col in enumerate(columns):
    sns.histplot(data=df, x=col, ax=axes[i], kde=False, bins=10,
color="orange")
    axes[i].set_title(f"Histogram of {col}")
    axes[i].set_ylabel("Count")
```

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



Comparison of disease vs work hours

```
# Comparison of disease vs work hours
columns = [
    'Since how many years are you involved in this occupation?',
    'How many hours are you working per week in this job? ',
    'BP',
    'Diabetes',
    'Thyroid',
    'Hernia'
]

# Create subplots (2 rows x 3 columns)
fig, axes = plt.subplots(nrows=2, ncols=3, figsize=(18, 10))
axes = axes.flatten()

# Plot histograms
for i, col in enumerate(columns):
    sns.histplot(data=df, x=col, ax=axes[i], kde=False, bins=10,
color='#4c72b0')
```

```
    axes[i].set_title(f"Histogram of {col}")
    axes[i].set_ylabel("Count")

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