

```
# Week 3 Homework
import pandas as pd
import numpy as np

import seaborn as sns

data = {
    'Rows' : ['Entire Sample', 'RExposed', 'RNExposed', 'BExposed', 'BNExposed', 'WExposed', 'WN
    'Coronary Heart Disease': [79, 50, 29, 36, 43, 35, 44],
    'Stroke': [54, 40, 14, 30, 24, 29, 25],
    'Congestive Heart Failure': [52, 38, 14, 28, 24, 25, 27]
}

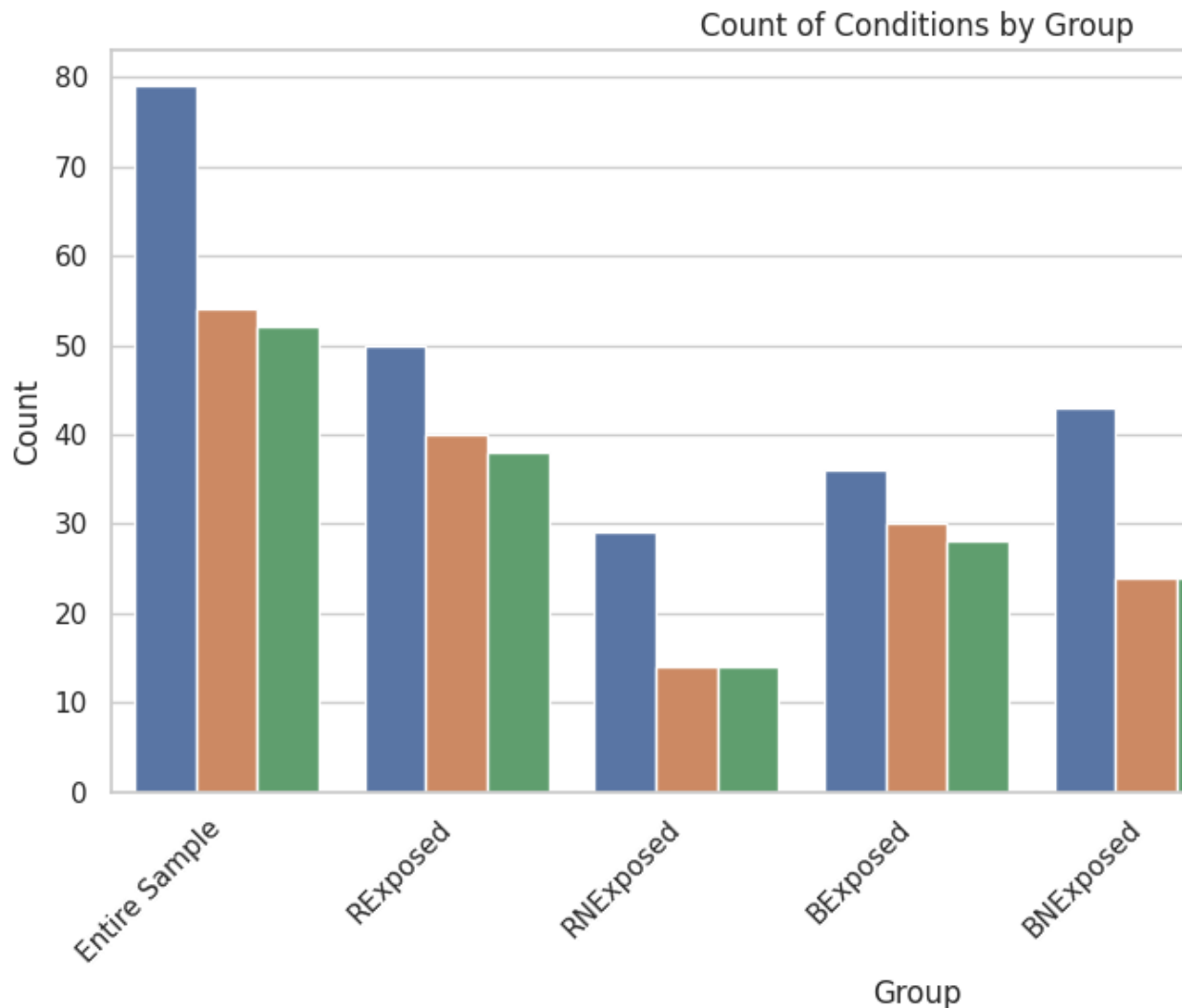
df = pd.DataFrame(data)

df_melted = df.melt(id_vars='Rows', var_name='Condition', value_name='Count')

sns.set(style="whitegrid")

import matplotlib.pyplot as plt
plt.figure(figsize=(10, 6))
bar_plot = sns.barplot(data=df_melted, x='Rows', y='Count', hue='Condition')
bar_plot.set_xticklabels(bar_plot.get_xticklabels(), rotation=45, horizontalalignment='rig
plt.title('Count of Conditions by Group')
plt.xlabel('Group')
plt.ylabel('Count')
plt.tight_layout()
plt.show()
```

```
<ipython-input-15-6464f70620d3>:4: UserWarning: FixedFormatter should only be used with
bar_plot.set_xticklabels(bar_plot.get_xticklabels(), rotation=45, horizontal
```



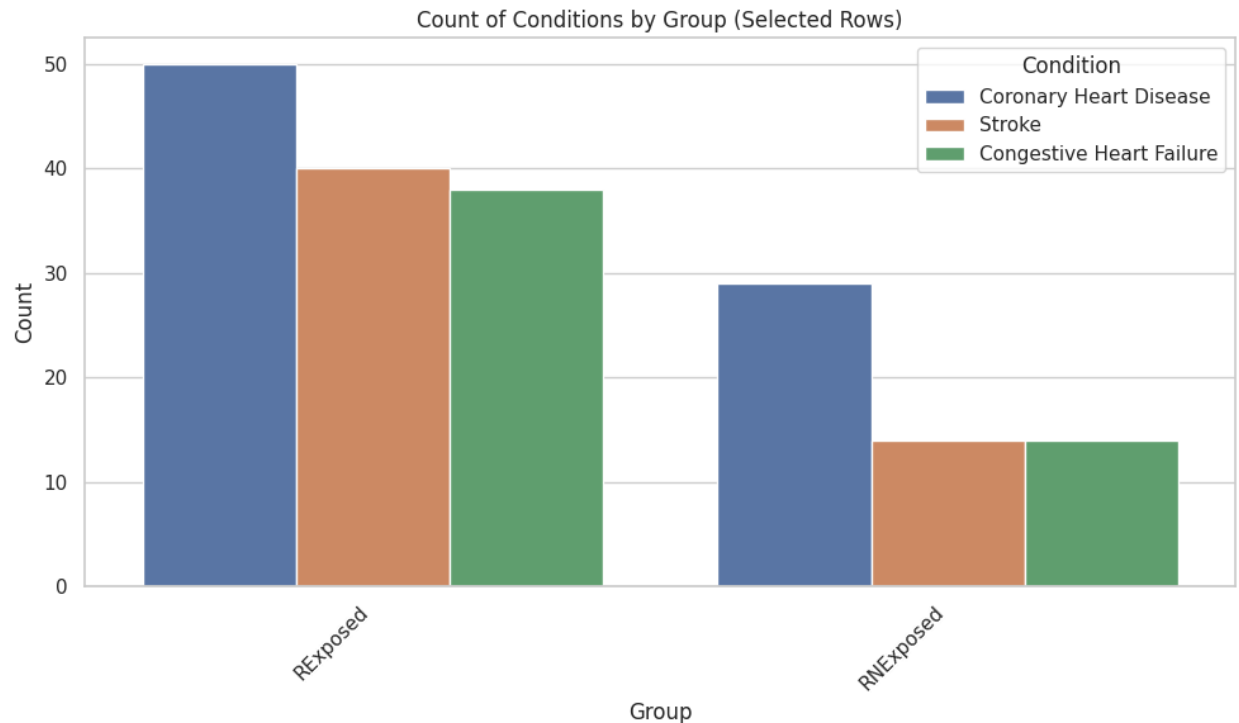
From this plot we can see how each different business type both exposed and not exposed to cigarette smoke stack up against each other. Included as well is the entire sample.

```
selected_rows = ['REExposed', 'RNExposed']
df_filtered = df[df['Rows'].isin(selected_rows)]
```

```
df_filtered_melted = df_filtered.melt(id_vars='Rows', var_name='Condition', value_name='Co
```

```
plt.figure(figsize=(10, 6))
grouped_bar_plot = sns.barplot(data=df_filtered_melted, x='Rows', y='Count', hue='Condition')
grouped_bar_plot.set_xticklabels(grouped_bar_plot.get_xticklabels(), rotation=45, horizontal=True)
plt.title('Count of Conditions by Group (Selected Rows)')
plt.xlabel('Group')
plt.ylabel('Count')
plt.tight_layout()
plt.show()
```

<ipython-input-19-c7c5f5f1b9d5>:3: UserWarning: FixedFormatter should only be used with `grouped_bar_plot.set_xticklabels(grouped_bar_plot.get_xticklabels(), rotation=`



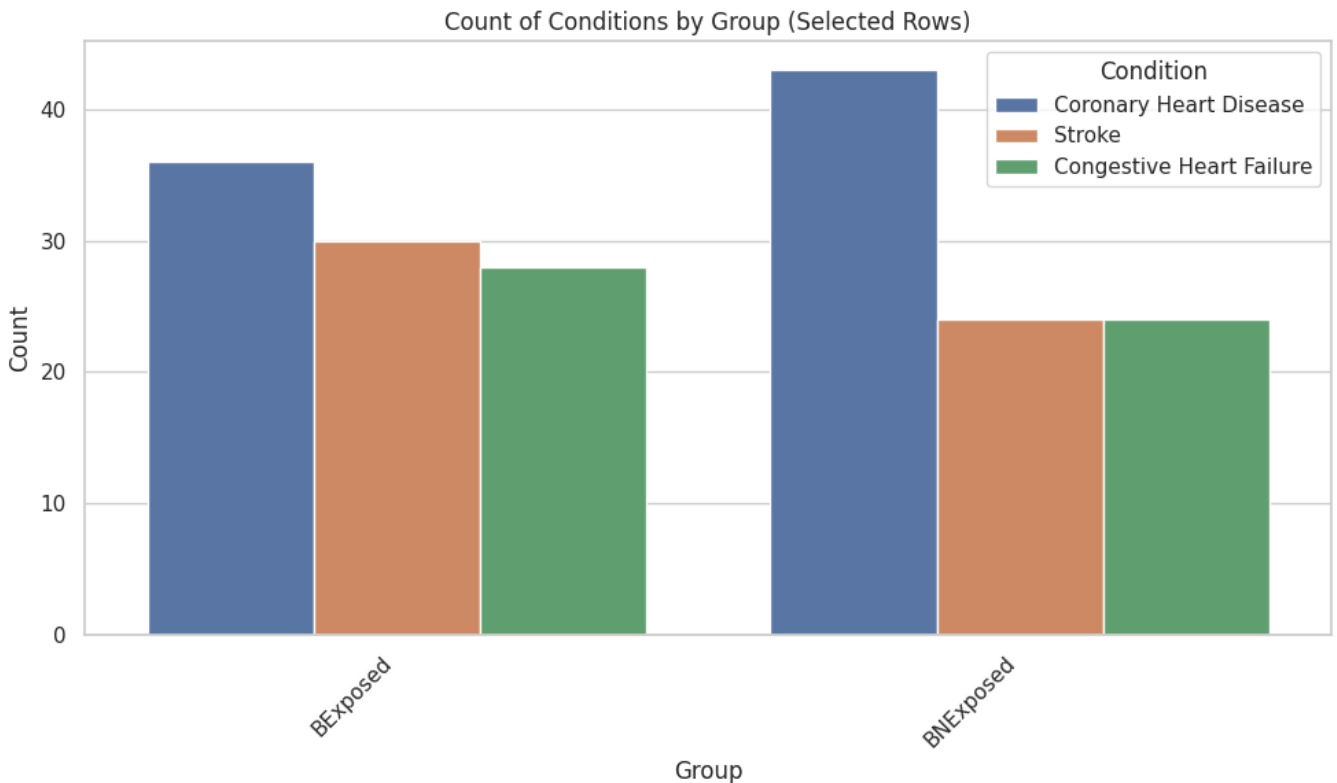
Looking at this, we can see for restaurants with exposed conditions have a higher count of heart disease, strokes and heart failures. This is what we would expect to see as smoking is known to cause all of these things.

```
selected_rows1 = ['BExposed', 'BNExposed']
df_filtered1 = df[df['Rows'].isin(selected_rows1)]
```

```
df_filtered_melted1 = df_filtered1.melt(id_vars='Rows', var_name='Condition', value_name='')
```

```
plt.figure(figsize=(10, 6))
grouped_bar_plot = sns.barplot(data=df_filtered_melted1, x='Rows', y='Count', hue='Condition')
grouped_bar_plot.set_xticklabels(grouped_bar_plot.get_xticklabels(), rotation=45, horizontal=True)
plt.title('Count of Conditions by Group (Selected Rows)')
plt.xlabel('Group')
plt.ylabel('Count')
plt.tight_layout()
plt.show()
```

<ipython-input-28-44262b3216fc>:3: UserWarning: FixedFormatter should only be used to group
 grouped_bar_plot.set_xticklabels(grouped_bar_plot.get_xticklabels(), rotation=45, ho



When looking at bars, the results are a bit different as people at bars who have never been exposed to smoking environment had a higher count of heart disease and only slightly less counts of strokes and heart failure.

```
selected_rows2 = ['WExposed', 'WNEExposed']
df_filtered2 = df[df['Rows'].isin(selected_rows2)]
```

```
df_filtered_melted2 = df_filtered2.melt(id_vars='Rows', var_name='Condition', value_name='Count')
```

```
plt.figure(figsize=(10, 6))
grouped_bar_plot = sns.barplot(data=df_filtered_melted2, x='Rows', y='Count', hue='Condition')
grouped_bar_plot.set_xticklabels(grouped_bar_plot.get_xticklabels(), rotation=45, horizontal=True)
plt.title('Count of Conditions by Group (Selected Rows)')
plt.xlabel('Group')
plt.ylabel('Count')
plt.tight_layout()
plt.show()
```

<ipython-input-35-e0c3efbf3f6a>:3: UserWarning: FixedFormatter should only be used to group
grouped_bar_plot.set_xticklabels(grouped_bar_plot.get_xticklabels(), rotation=45, ho

