

```
In [1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
file_path=r"C:\Users\LENOVO\OneDrive\Desktop\my_folder\Visadataset.csv"
visa_df=pd.read_csv(file_path)
```

```
In [9]: wage_data=visa_df['prevailing_wage']
```

```
In [11]: #apply MinMax Scale or Normalization on wage data based on formulae
wage_min=wage_data.min()
wage_max=wage_data.max()
normalized_wages= (wage_data - wage_min) / (wage_max - wage_min)
print(normalized_wages)
```

```
0      0.001849
1      0.261345
2      0.385312
3      0.261371
4      0.469616
```

...

```
25475  0.241505
25476  0.874579
25477  0.458311
25478  0.269895
25479  0.222033
```

Name: prevailing_wage, Length: 25480, dtype: float64

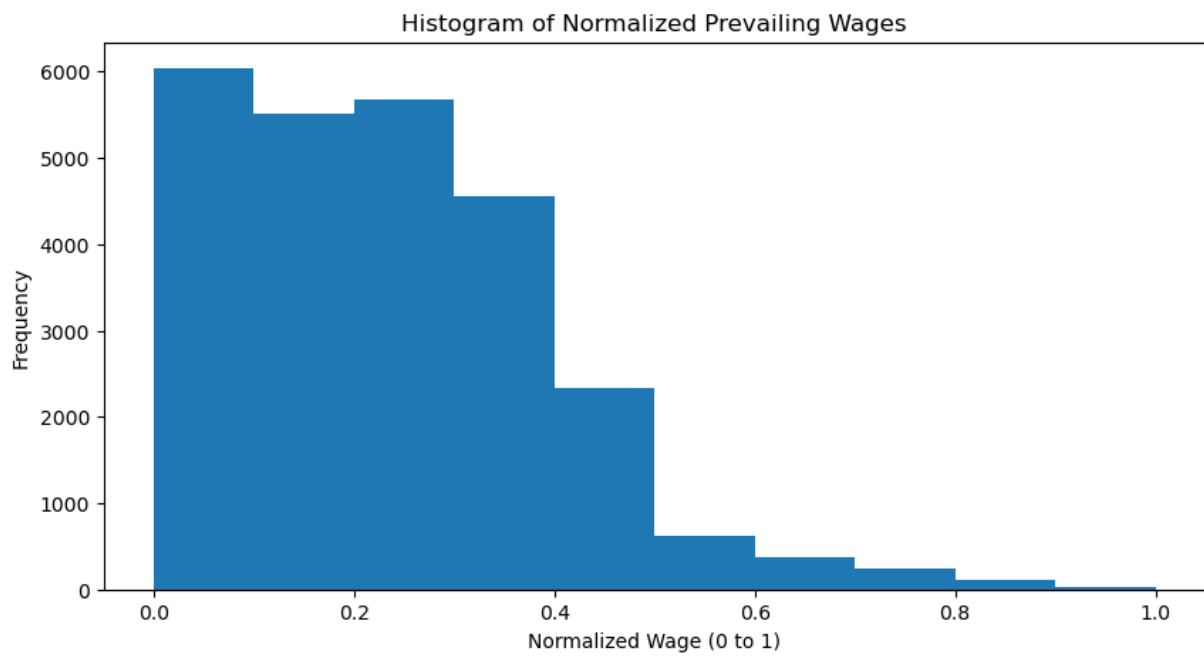
```
In [15]: #apply MinMax Scale or Normalization on wage data based on package
from sklearn.preprocessing import MinMaxScaler
wage_data=visa_df[['prevailing_wage']]
ms=MinMaxScaler() #save the method
wage_ms=ms.fit_transform(wage_data)# apply the fit transform
```

```
In [17]: wage_ms
```

```
Out[17]: array([[0.00184853],
               [0.2613452 ],
               [0.385312  ],
               ...,
               [0.45831136],
               [0.26989486],
               [0.22203311]])
```

```
In [23]: plt.figure(figsize=(10,5))
plt.hist(wage_ms)
plt.title('Histogram of Normalized Prevailing Wages')
plt.xlabel('Normalized Wage (0 to 1)')
plt.ylabel('Frequency')
```

```
Out[23]: Text(0, 0.5, 'Frequency')
```



In []: