```
In [2]:    import matplotlib.pyplot as plt
           import squarify
           import numpy as np
           import pandas as pd
```

```
In [5]:    df=pd.read_csv("unemployement-rate-1948-2010.csv")
           df
```

Out[5]:

	Series id	Year	Period	Value
0	LNS14000000	1948	M01	3.4
1	LNS14000000	1948	M02	3.8
2	LNS14000000	1948	M03	4.0
3	LNS14000000	1948	M04	3.9
4	LNS14000000	1948	M05	3.5
...	...	...	...	...
741	LNS14000000	2009	M10	10.1
742	LNS14000000	2009	M11	10.0
743	LNS14000000	2009	M12	10.0
744	LNS14000000	2010	M01	9.7
745	LNS14000000	2010	M02	9.7

746 rows × 4 columns

```
In [12]:   df1=pd.read_csv("expenditures.txt", sep='\t')
           df1
```

Out[12]:

	year	category	expenditure	sex
0	2008	Food	6443	1
1	2008	Alcoholic Beverages	444	1
2	2008	Housing	17109	1
3	2008	Apparel	1801	1
4	2008	Transportation	8604	1
...	...	...	...	...
345	1984	Education	303	1
346	1984	Tobacco Products	228	1
347	1984	Miscellaneous	451	1
348	1984	Cash Contributions	706	1
349	1984	Personal Insurance	1897	1

350 rows × 4 columns

In [19]:

```python
n = df1.groupby('category')[['expenditure']].sum()
n
```

Out[19]:

	expenditure
category	
Alcoholic Beverages	8424
Apparel	41833
Cash Contributions	27987
Education	14498
Entertainment	44273
Food	119297
Healthcare	47383
Housing	280256
Miscellaneous	18327
Personal Care	11123
Personal Insurance	84269
Reading	3636
Tobacco Products	6936
Transportation	160694

In [27]:
```python
catg = df1.groupby('category')[['expenditure']].sum().index.get_level_values(
catg
```

Out[27]:
```
['Alcoholic Beverages',
 'Apparel',
 'Cash Contributions',
 'Education',
 'Entertainment',
 'Food',
 'Healthcare',
 'Housing',
 'Miscellaneous',
 'Personal Care',
 'Personal Insurance',
 'Reading',
 'Tobacco Products',
 'Transportation']
```

In [28]:
```python
exp = df1.groupby('category')[['expenditure']].sum().reset_index().expenditur
exp
```

```
Out[28]:  [8424,
           41833,
           27987,
           14498,
           44273,
           119297,
           47383,
           280256,
           18327,
           11123,
           84269,
           3636,
           6936,
           160694]
```

```
In [35]:   plt.figure(figsize=(15, 6))
           squarify.plot(sizes=exp, label=catg)
           plt.axis('off')
           plt.show()
```
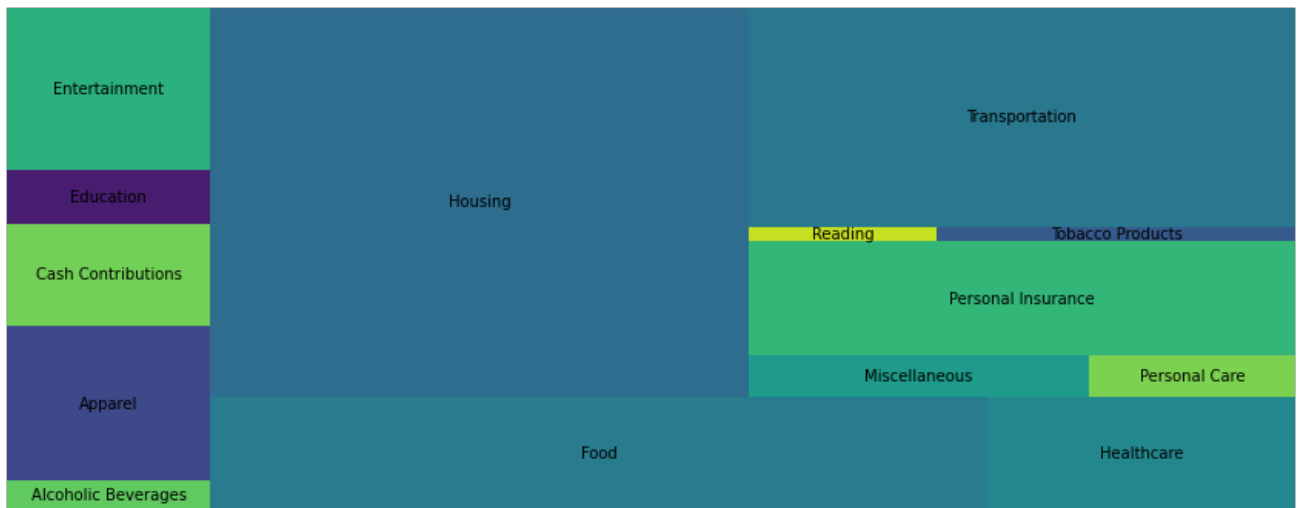


```
In [30]:   l = df.groupby('Year')[['Value']].sum()
           l
```

Out[30]:

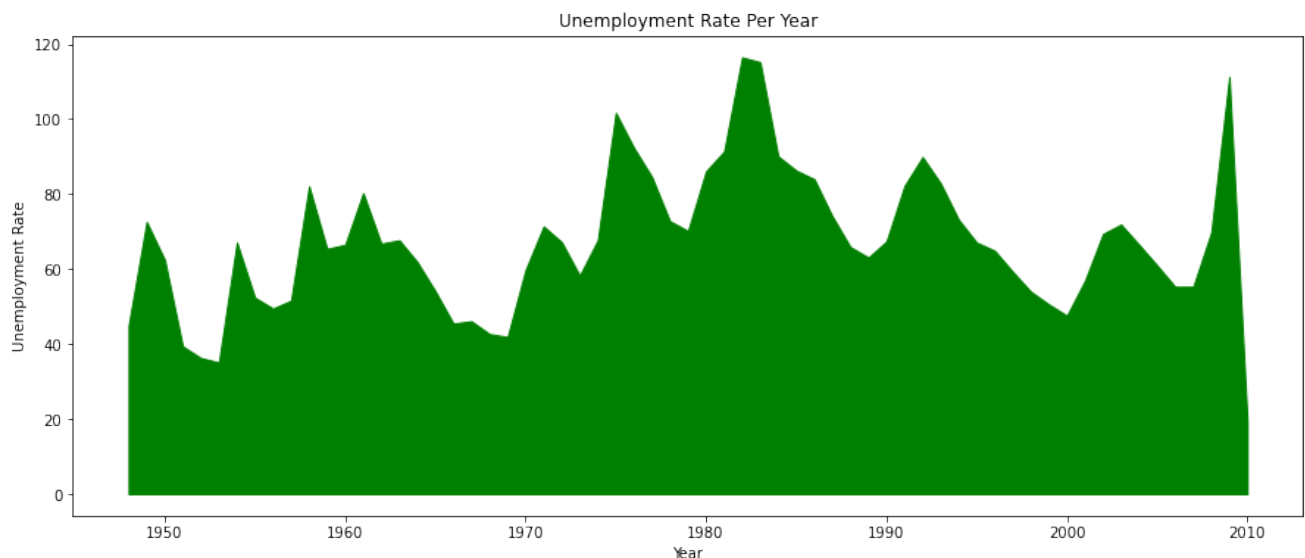|	Value
Year	
1948	45.0
1949	72.6
1950	62.5
1951	39.4
1952	36.3
...	...
2006	55.3
2007	55.3
2008	69.8
2009	111.3
2010	19.4

63 rows × 1 columns

In [34]:

```python
year = df.groupby('Year')[['Value']].sum().index.get_level_values(0).tolist()
val = df.groupby('Year')[['Value']].sum().reset_index().Value.values.tolist()
plt.figure(figsize=(15, 6))
plt.fill_between(year, val, color='green')
plt.title('Unemployment Rate Per Year')
plt.xlabel('Year')
plt.ylabel('Unemployment Rate')
plt.show()
```

In [36]:
```python
df2 = df[(df['Year']>=1984) & (df['Year']<=2008)]
df2
```

Out[36]:

	Series id	Year	Period	Value
432	LNS14000000	1984	M01	8.0
433	LNS14000000	1984	M02	7.8
434	LNS14000000	1984	M03	7.8
435	LNS14000000	1984	M04	7.7
436	LNS14000000	1984	M05	7.4
...	...	...	...	...
727	LNS14000000	2008	M08	6.1
728	LNS14000000	2008	M09	6.2
729	LNS14000000	2008	M10	6.6
730	LNS14000000	2008	M11	6.9
731	LNS14000000	2008	M12	7.4

300 rows × 4 columns

In [38]:
```python
j = df2.groupby('Year')[['Value']].sum()
j
```

Out[38]:

Year	Value
1984	90.1
1985	86.3
1986	84.0
1987	74.1
1988	65.9
1989	63.1
1990	67.4
1991	82.2
1992	89.9
1993	82.9
1994	73.2
1995	67.1
1996	64.9
1997	59.3
1998	54.0
1999	50.6
2000	47.6
2001	56.9
2002	69.4
2003	71.9
2004	66.5
2005	61.0
2006	55.3
2007	55.3
2008	69.8

In [40]:

```python
year = df2.groupby('Year')[['Value']].sum().index.get_level_values(0).tolist(
val = df2.groupby('Year')[['Value']].sum().reset_index().Value.values.tolist(
```

In [41]:
```python
q = df1.groupby('year')[['expenditure']].sum()
q
```

Out[41]:

year	expenditure
1984	21972
1985	23489
1986	23865
1987	24415
1988	25893
1989	27811
1990	28379
1991	29614
1992	30296
1993	30692
1994	31733
1995	32262
1996	33798
1997	34820
1998	35536
1999	36996
2000	38045
2001	39519
2002	40676
2003	40818
2004	43392
2005	46407
2006	48400
2007	49619
2008	50489

In [45]:
```python
exp = df1.groupby('year')[['expenditure']].sum().reset_index().expenditure.va
exp
```
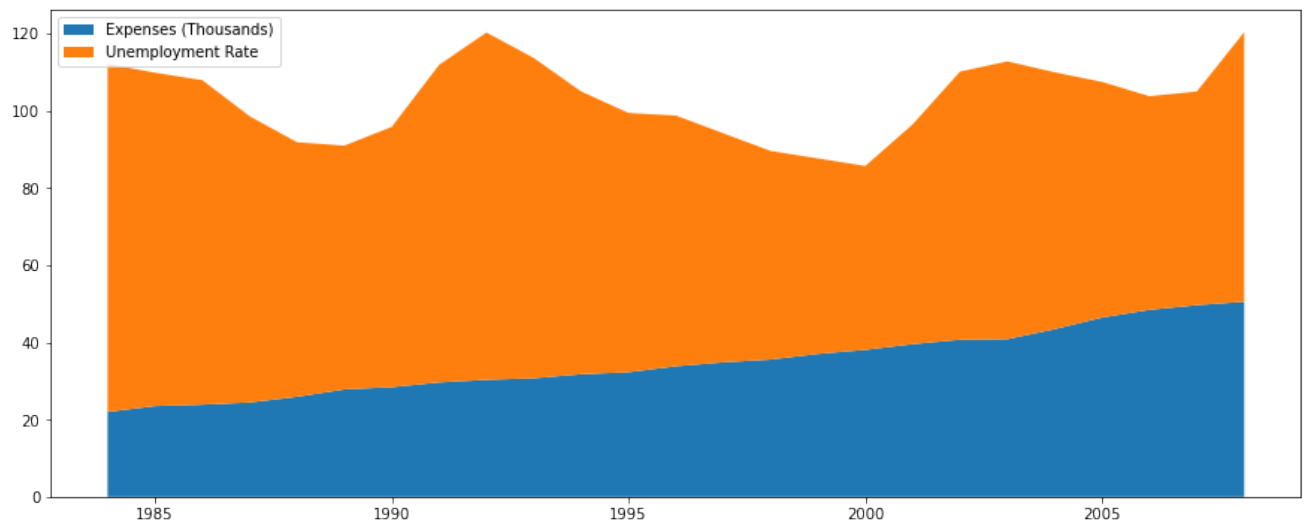
Out[45]:
```
[21972,
 23489,
 23865,
 24415,
 25893,
 27811,
 28379,
 29614,
 30296,
 30692,
 31733,
 32262,
 33798,
 34820,
 35536,
 36996,
 38045,
 39519,
 40676,
 40818,
 43392,
 46407,
 48400,
 49619,
 50489]
```

In [47]:
```python
p = []
for i in exp:
    p.append(i / 1000)
p
```

Out[47]: [21.972,
          23.489,
          23.865,
          24.415,
          25.893,
          27.811,
          28.379,
          29.614,
          30.296,
          30.692,
          31.733,
          32.262,
          33.798,
          34.82,
          35.536,
          36.996,
          38.045,
          39.519,
          40.676,
          40.818,
          43.392,
          46.407,
          48.4,
          49.619,
          50.489]

In [49]:
```python
plt.figure(figsize=(15, 6))
plt.stackplot(year,p, val, labels=['Expenses (Thousands)','Unemployment Rate'
plt.legend(loc='upper left')
```

Out[49]: <matplotlib.legend.Legend at 0x7f82c0e7fa00>

In [ ]: