

```
In [1]: # import libraries
import pandas as pd
import seaborn as sns
import squarify
import matplotlib.pyplot as plt
```

```
In [2]: # Load data
df_ue = pd.read_csv(r'C:\GitHub\DSC640\DSC640\unemployment-rate-1948-2010.csv')
df_usps = pd.read_excel(r'C:\Github\DSC640\DSC640\us-postage.xlsx')
df_pop = pd.read_excel(r'C:\GitHub\DSC640\DSC640\world-population.xlsx')
```

```
In [3]: # view data
print('df_ue: \n', df_ue.head(), '\n')
print('df_ue shape: ', df_ue.shape, '\n')
print('df_usps: \n', df_usps.head(), '\n')
print('df_usps shape: ', df_usps.shape, '\n')
print('df_pop: \n', df_pop.head(), '\n')
print('df_pop shape: ', df_pop.shape)
```

```
df_ue:
   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
```

```
df_ue shape: (746, 4)
```

```
df_usps:
   Year  Price
0  1991   0.29
1  1995   0.32
2  1999   0.33
3  2001   0.34
4  2002   0.37
```

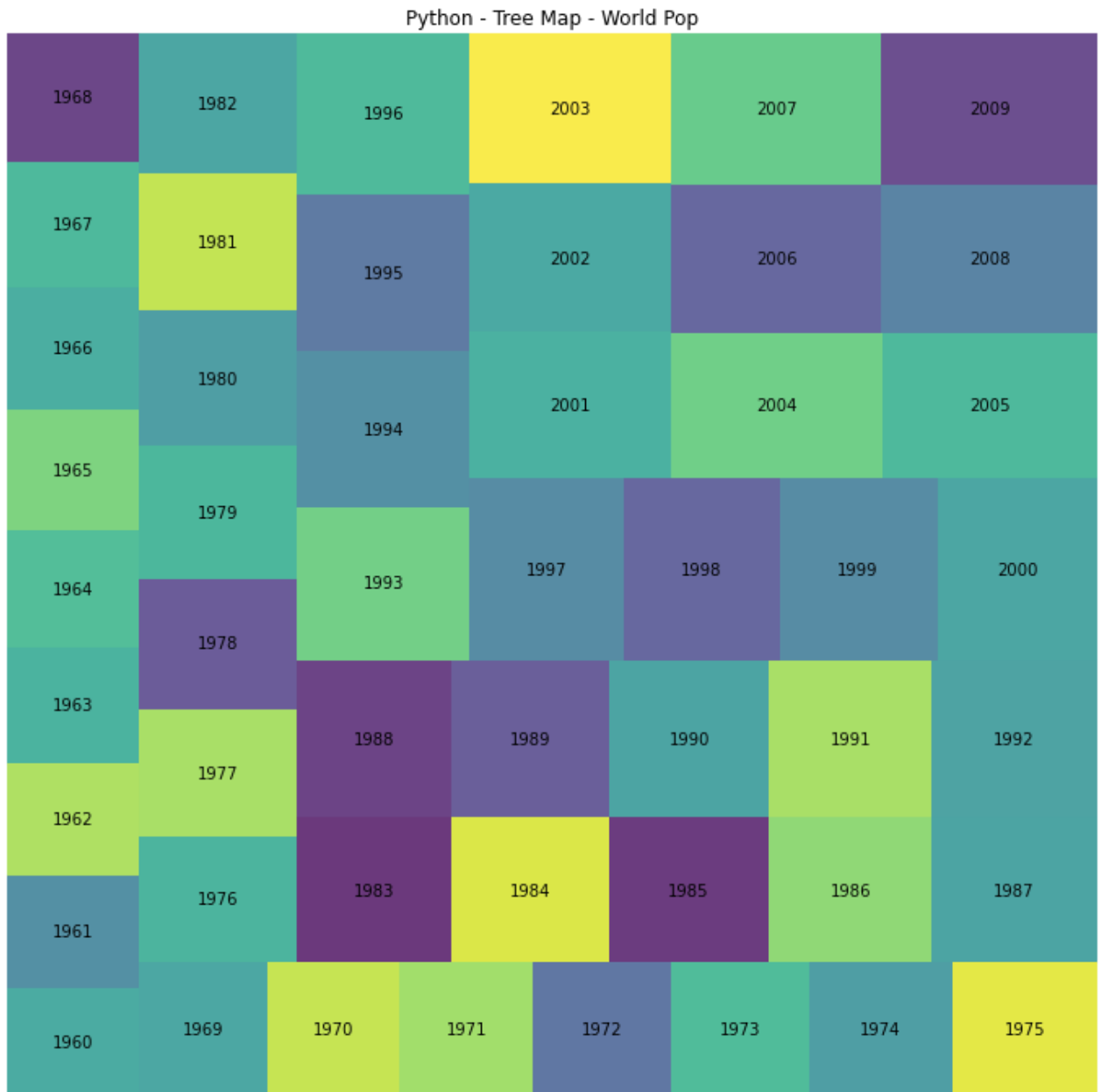
```
df_usps shape: (10, 2)
```

```
df_pop:
   Year  Population
0  1960  3028654024
1  1961  3068356747
2  1962  3121963107
3  1963  3187471383
4  1964  3253112403
```

```
df_pop shape: (50, 2)
```

In [4]:

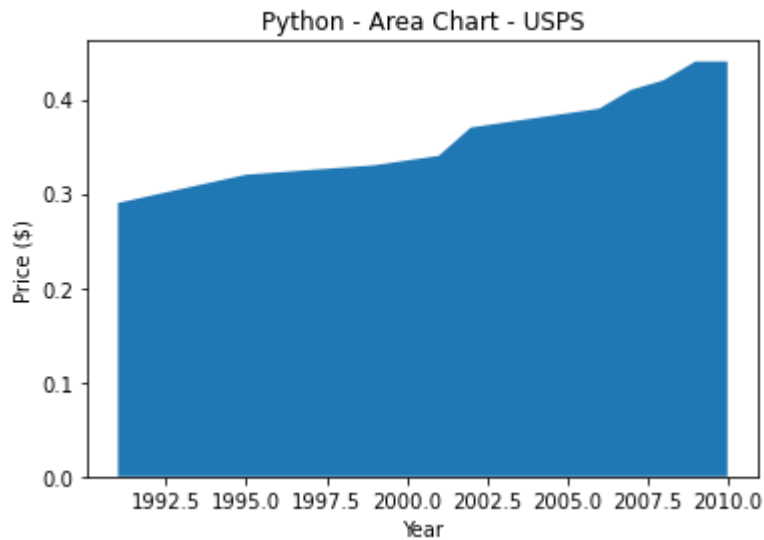
```
# tree map
fig, ax = plt.subplots(1, figsize = (12,12))
squarify.plot(sizes = df_pop['Population'],
              label = df_pop['Year'],
              alpha = 0.8)
plt.axis('off')
plt.title('Python - Tree Map - World Pop')
plt.show()
```



In [5]:

```
# 1 area chart
plt.stackplot(df_usps.Year, df_usps.Price)
plt.title('Python - Area Chart - USPS')
```

```
plt.xlabel('Year')
plt.ylabel('Price ($)')
plt.show()
```



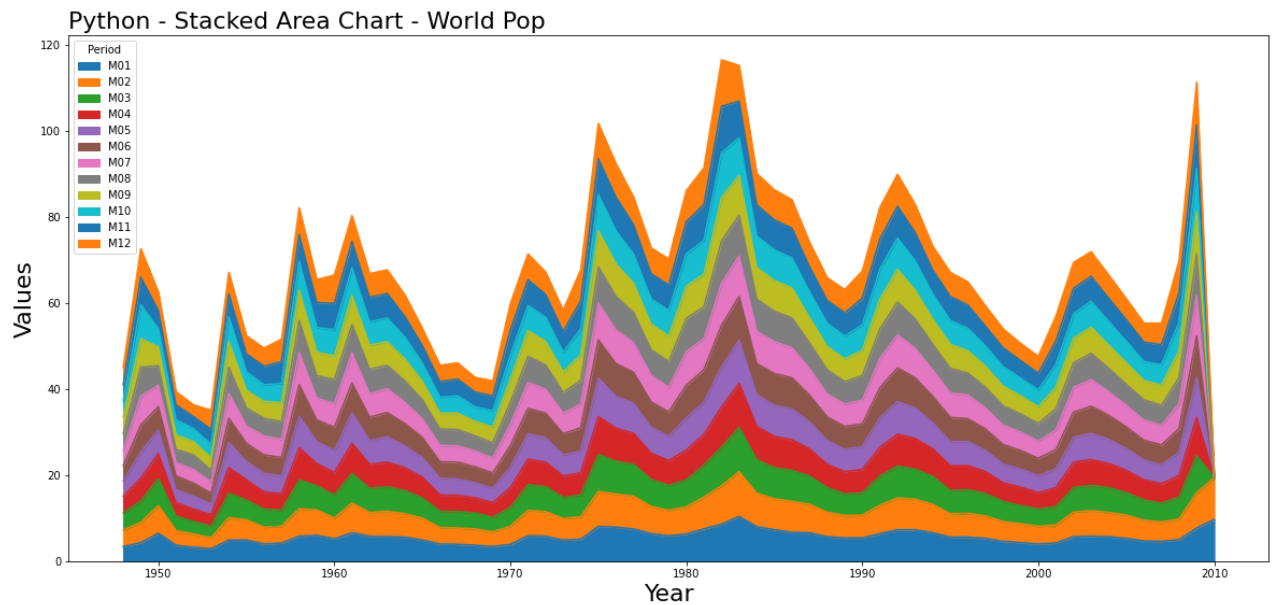
In [6]:

```
# Stacked area chart

# turn df_ue to pivot table
df_ue_pv = df_ue.pivot(index = 'Year',
                        columns = 'Period',
                        values = 'Value')

df_ue_pv.plot.area(figsize = (20,9))
plt.title('Python - Stacked Area Chart - World Pop', fontsize = 22, loc =
'left')
plt.ylabel("Values", fontsize = 22)
plt.xlabel("Year", fontsize = 22)
```

Out[6]: Text(0.5, 0, 'Year')

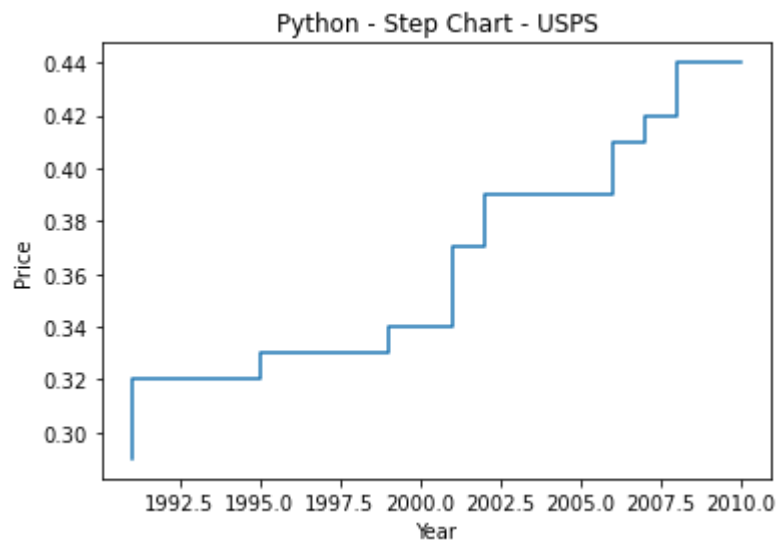


In [7]:

```
# 1 step chart
sns.lineplot(df_usps.Year, df_usps.Price, drawstyle = 'steps-pre')
plt.title('Python - Step Chart - USPS')
plt.show()
```

C:\Users\howla\anaconda3\lib\site-packages\seaborn_decorators.py:36: FutureWarning: Pass the following variables as keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.

warnings.warn(



In []:

Howland_DSC640_wk4

Howland_E

2022-07-01

R Markdown

Load libraries and data

```
# import libraries
library(ggplot2)
library(readxl)
library(treemap)

# load files
setwd("c:/GitHub/DSC640/DSC640")
df_usps <- read_excel("us-postage.xlsm")
df_pop <- read_excel("world-population.xlsm")
df_ue <- read.csv("unemployment-rate-1948-2010.csv")
```

Look at data

```
head(df_usps)

## # A tibble: 6 x 2
##   Year Price
##   <dbl> <dbl>
## 1  1991  0.29
## 2  1995  0.32
## 3  1999  0.33
## 4  2001  0.34
## 5  2002  0.37
## 6  2006  0.39

head(df_pop)

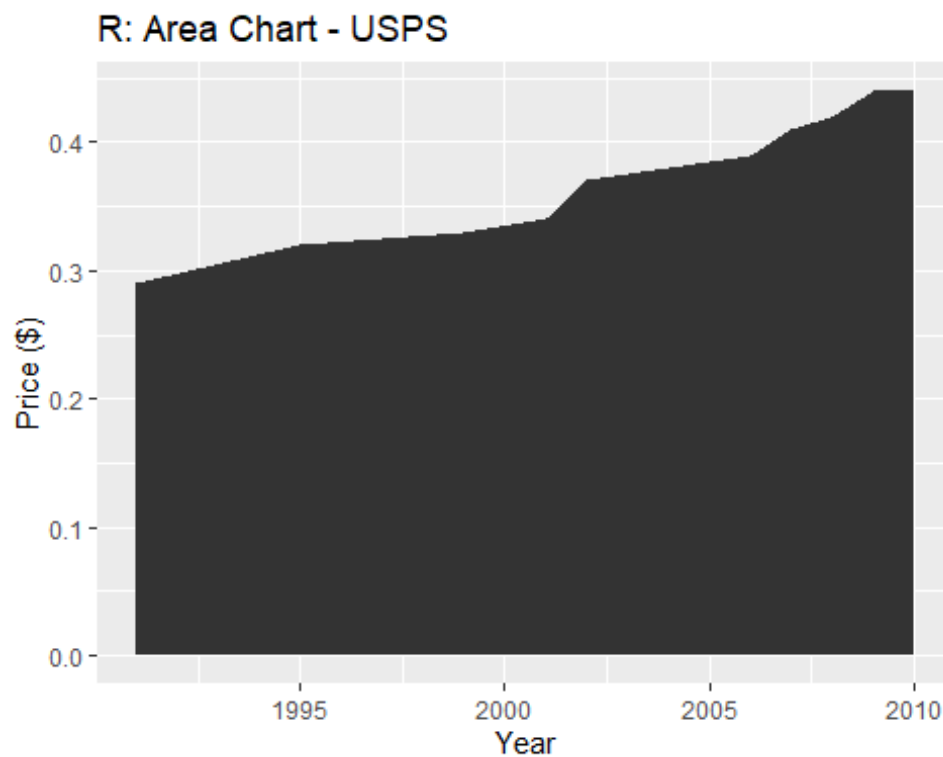
## # A tibble: 6 x 2
##   Year Population
##   <dbl>      <dbl>
## 1  1960 3028654024
## 2  1961 3068356747
## 3  1962 3121963107
## 4  1963 3187471383
## 5  1964 3253112403
## 6  1965 3320396924
```

```
head(df_ue)
```

```
##      Series.id Year Period Value
## 1 LNS14000000 1948   M01   3.4
## 2 LNS14000000 1948   M02   3.8
## 3 LNS14000000 1948   M03   4.0
## 4 LNS14000000 1948   M04   3.9
## 5 LNS14000000 1948   M05   3.5
## 6 LNS14000000 1948   M06   3.6
```

R: Area Chart

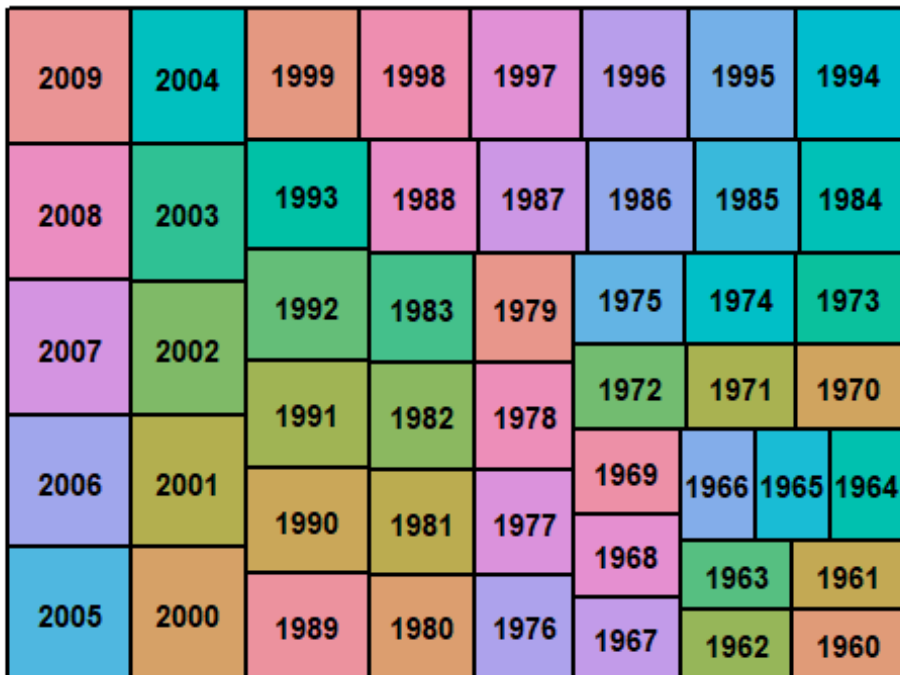
```
ggplot(df_usps, aes(x =Year, y = Price)) +  
  geom_area() +  
  ggtitle("R: Area Chart - USPS") +  
  labs(x = "Year", y = "Price ($)")
```



R: Tree Map

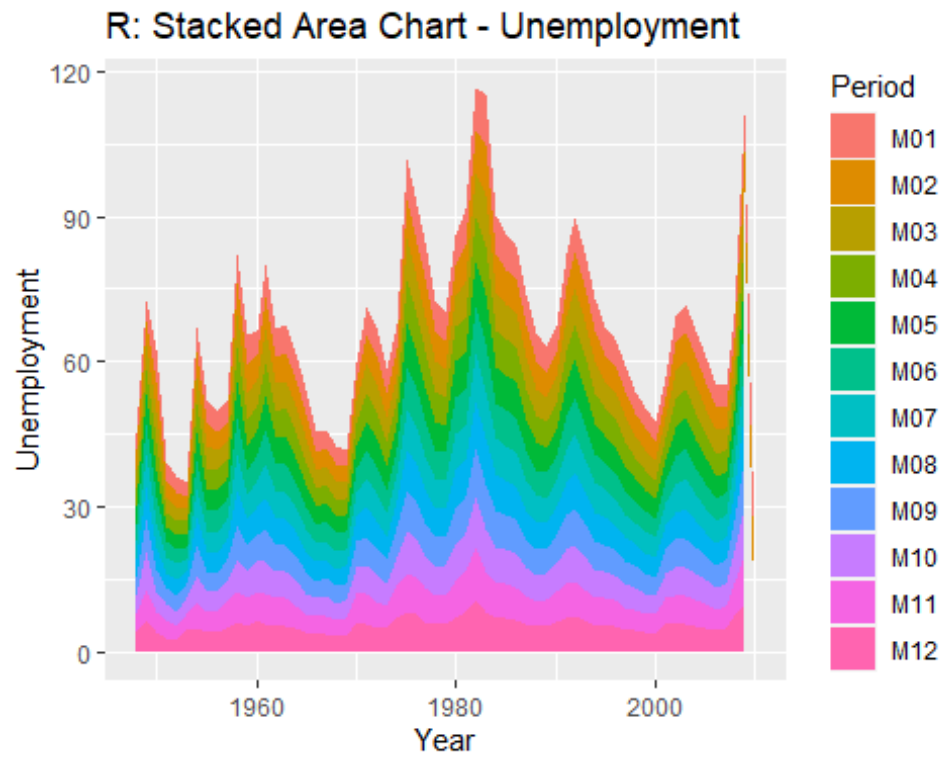
```
treemap(df_pop,  
        index = c("Year"),  
        vSize = "Population",  
        title = "R: Treemap - Population"  
)
```

R: Treemap - Population



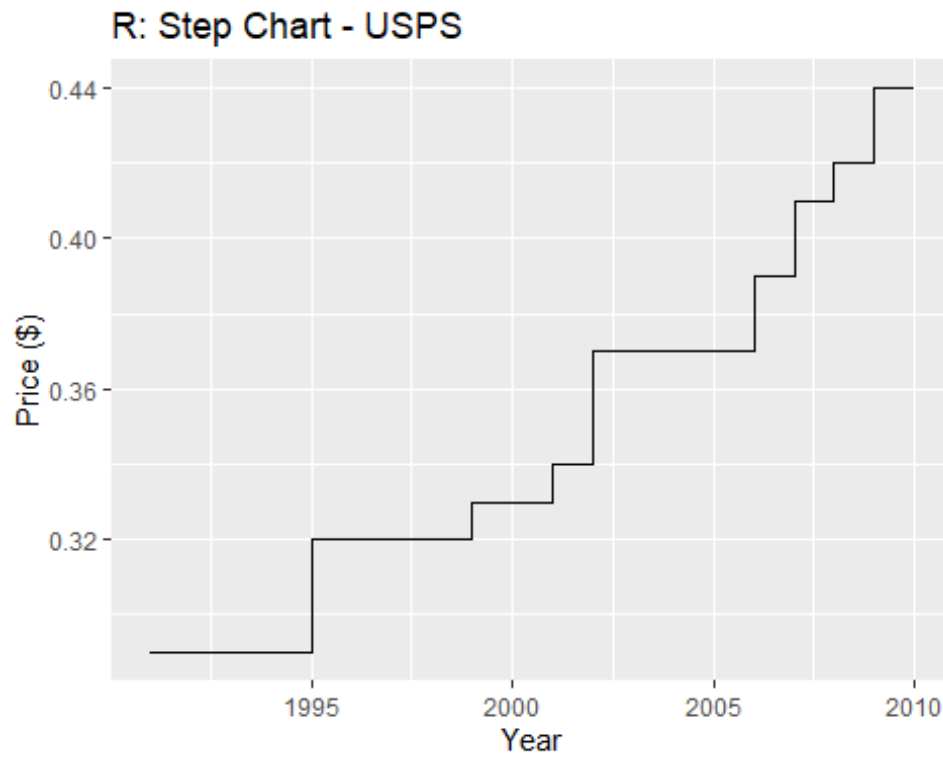
R: Stacked Area Chart

```
ggplot(df_ue, aes(x = Year, y = Value, fill = Period)) +  
  geom_area() +  
  ggtitle("R: Stacked Area Chart - Unemployment") +  
  labs(x = "Year", y = "Unemployment")
```

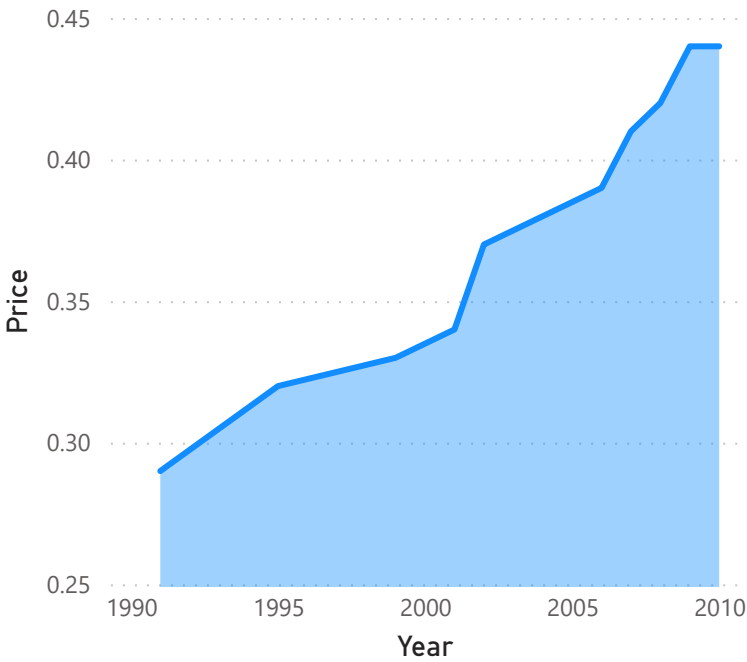


R: Step Chart

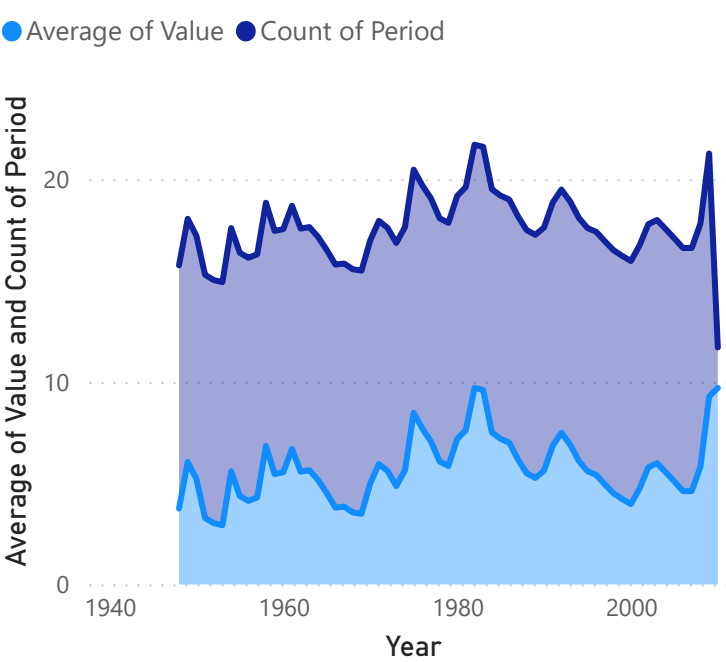
```
ggplot(df_usps, aes(x = Year, y = Price)) +  
  geom_step() +  
  ggtitle("R: Step Chart - USPS") +  
  labs(x = "Year", y = "Price ($)")
```



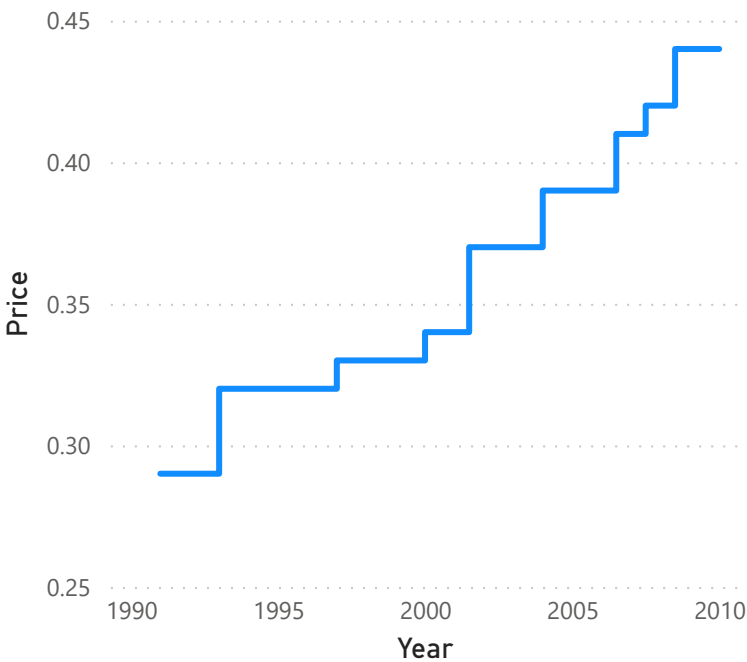
PBI - Area Chart - Price by Year



PBI - Stacked Area Chart



PBI - Step Chart - Price by Year



PBI - Tree map - World Pop

