DSC 640: WEEK 5-6

Author: Yousof Rahimian

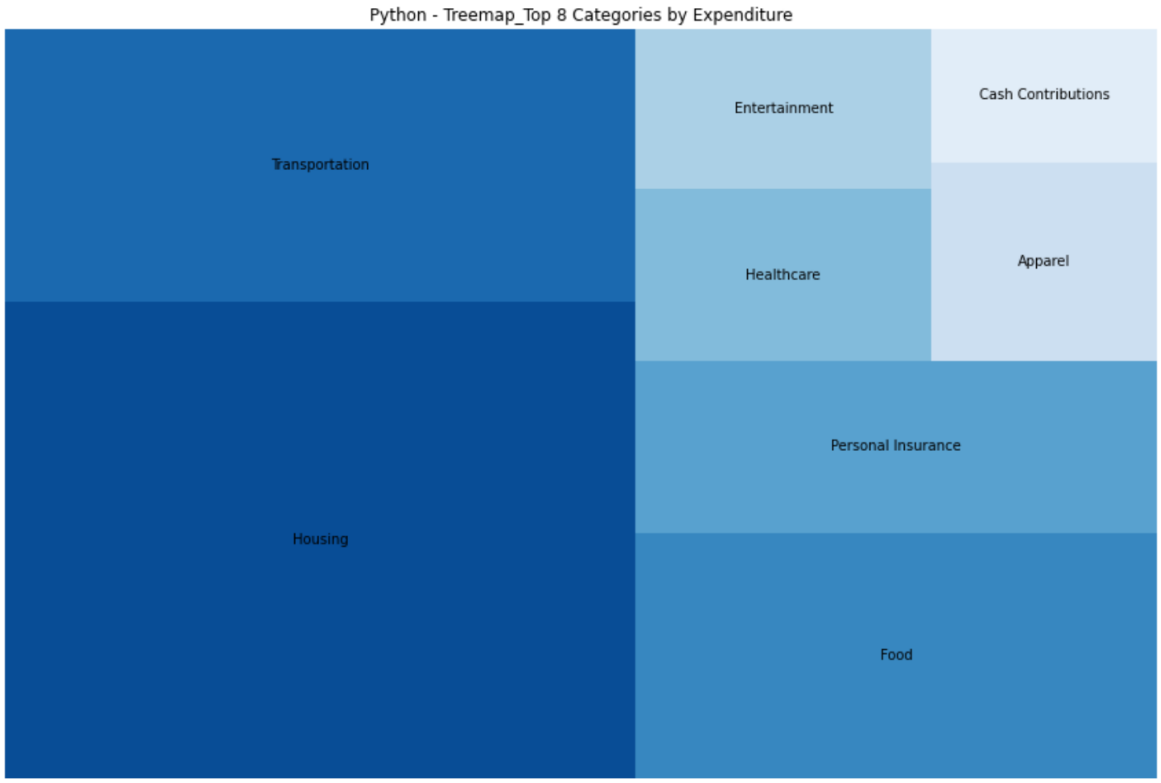
Date: 1/20/23

3.2 Exercises

1 - Tree map

Python

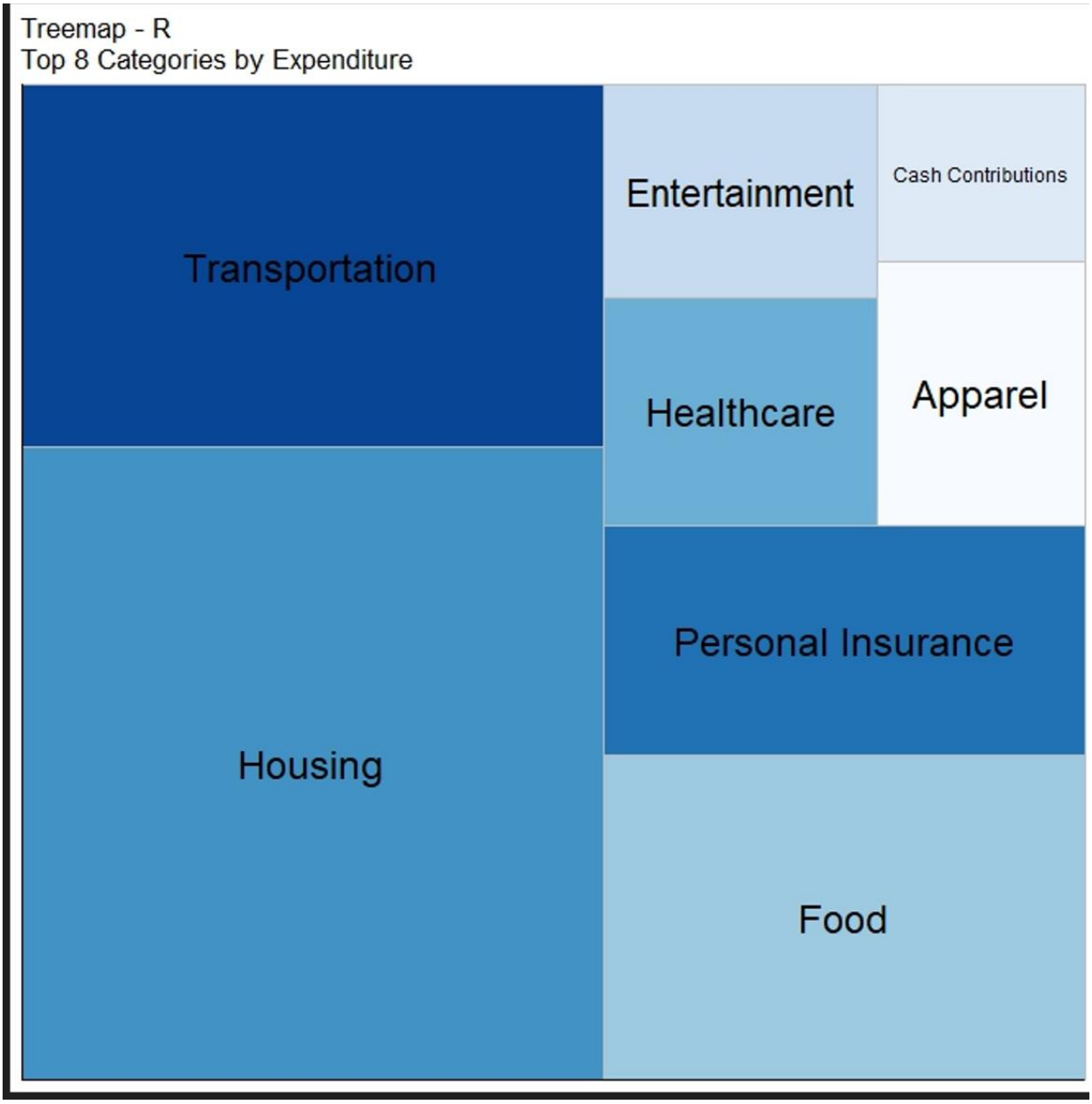
```
# Display the Treemap
fig, ax = plt.subplots(figsize = (15, 10))
squarify.plot(sizes = Ctg_year.values, label = Ctg_year.index,
              color = sb.color_palette("Blues_r", len(Ctg_year.values)))
plt.title('Python - Treemap_Top 8 Categories by Expenditure')
ax.axis('off')

plt.show()
```



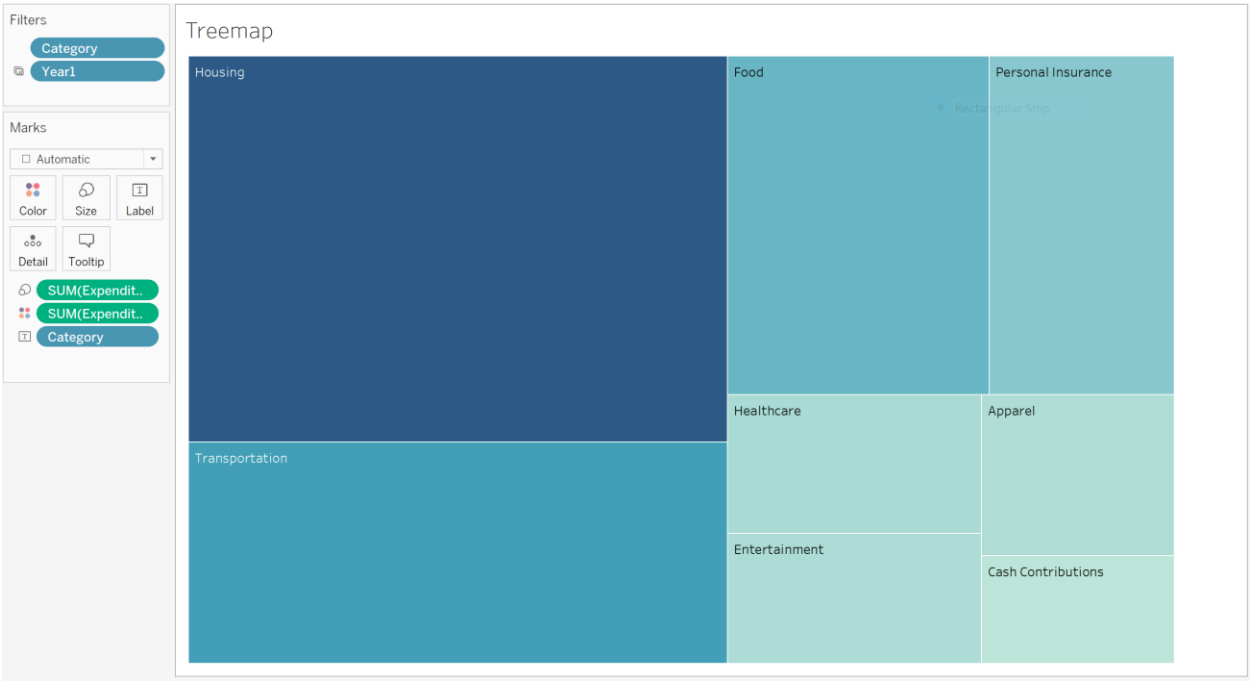
R

```r
#| label: topcategories
category_year <- expenditures_df %>%
  group_by(category) %>%
  summarise(top = sum(expenditure)) %>%
  arrange(desc(top)) %>%
  top_n(8)
category_year
```

```r
#| label: treemap
#| fig-width: 6.5
fig <- ggplot(category_year, aes(area = top, fill = category, label = category)) +
  geom_treemap() +
  geom_treemap_text(colour = "black", min.size = 0.3,
                    place = "centre", padding.x = grid::unit(3, "mm"), padding.y = grid::unit(3, "mm")) +
  theme(legend.position = "none") +
  scale_fill_brewer(palette = "Blues") +
  ggtitle("Treemap - R \nTop 8 Categories by Expenditure")
fig
```


Treemap - R
Top 8 Categories by Expenditure
Transportation
Entertainment
Cash Contributions
Healthcare
Apparel
Personal Insurance
Housing
Food

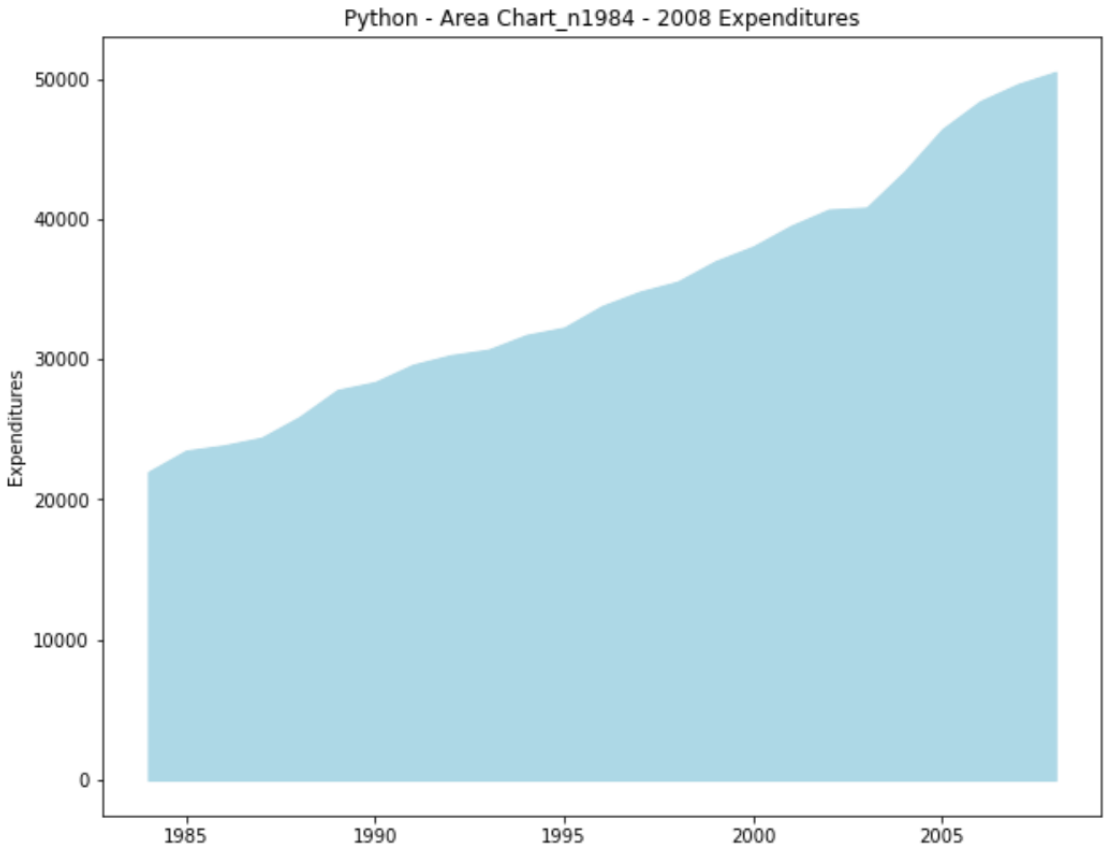
Tableau



2 – Area Chart

Python

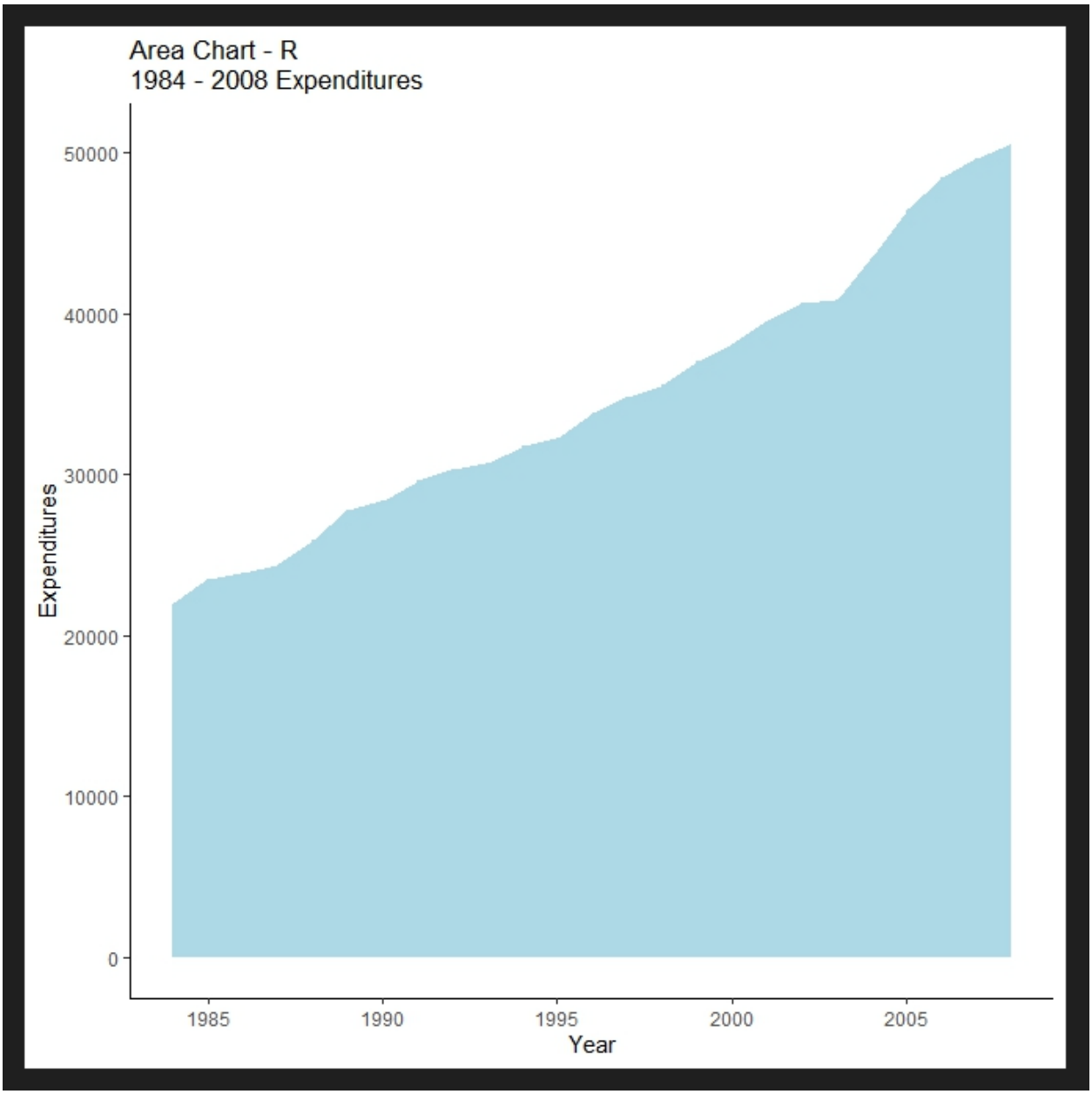
```
# Display the Area Chart
fig, ax = plt.subplots(figsize=(10,8))

ax.fill_between(expenditures_sex1.keys(), expenditures_sex1.values, color='lightblue')

plt.title('Python - Area Chart_n1984 - 2008 Expenditures')
plt.ylabel('Expenditures')
plt.xlabel('')

plt.show()
```



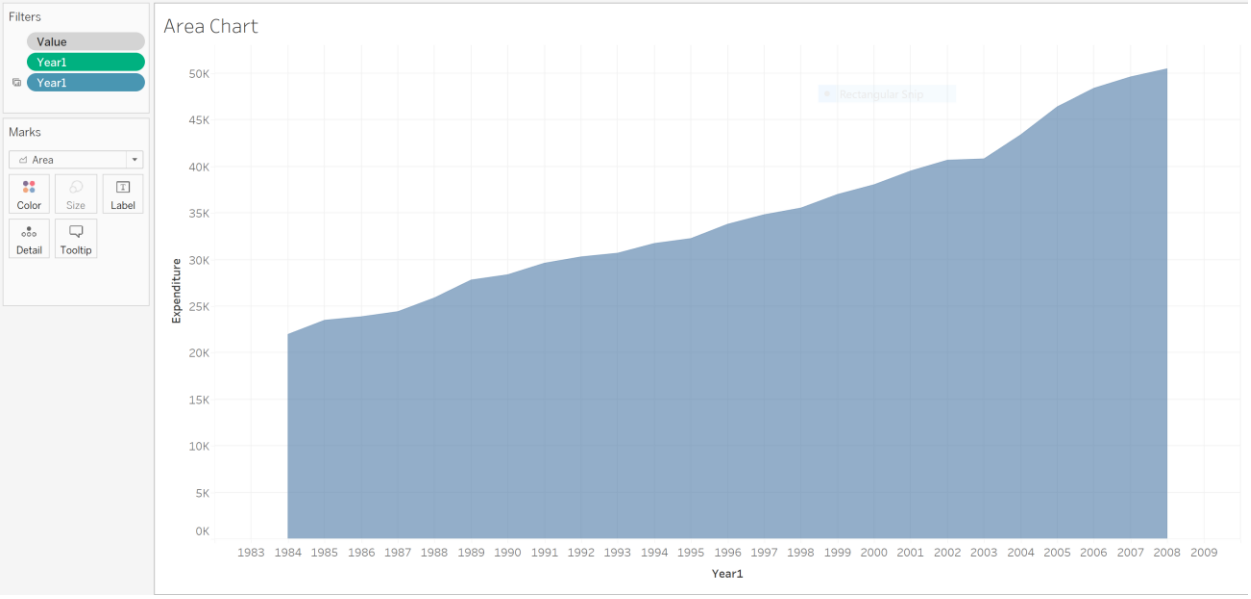
R

```
### 1. Area Chart

`{r}
#| label: expenditure per year for sex = 1
expenditures_per_year <- expenditures_df %>%
  group_by(year) %>%
  summarise_at(vars(expenditure), sum)
expenditures_per_year
```

```
`{r}
#| label: Areachart
#| fig-width: 6.5
fig <- ggplot(expenditures_per_year, aes(x = year, y = expenditure)) +
  geom_area(fill = "lightblue") +
  ggtitle("Area Chart - R \n1984 - 2008 Expenditures") +
  xlab('Year') +
  ylab("Expenditures") +
  scale_x_continuous(breaks = c(1985, 1990, 1995, 2000, 2005))
fig
```


Area Chart - R
1984 - 2008 Expenditures
50000
40000
30000
20000
10000
0
Expenditures
1985
1990
1995
2000
2005
Year

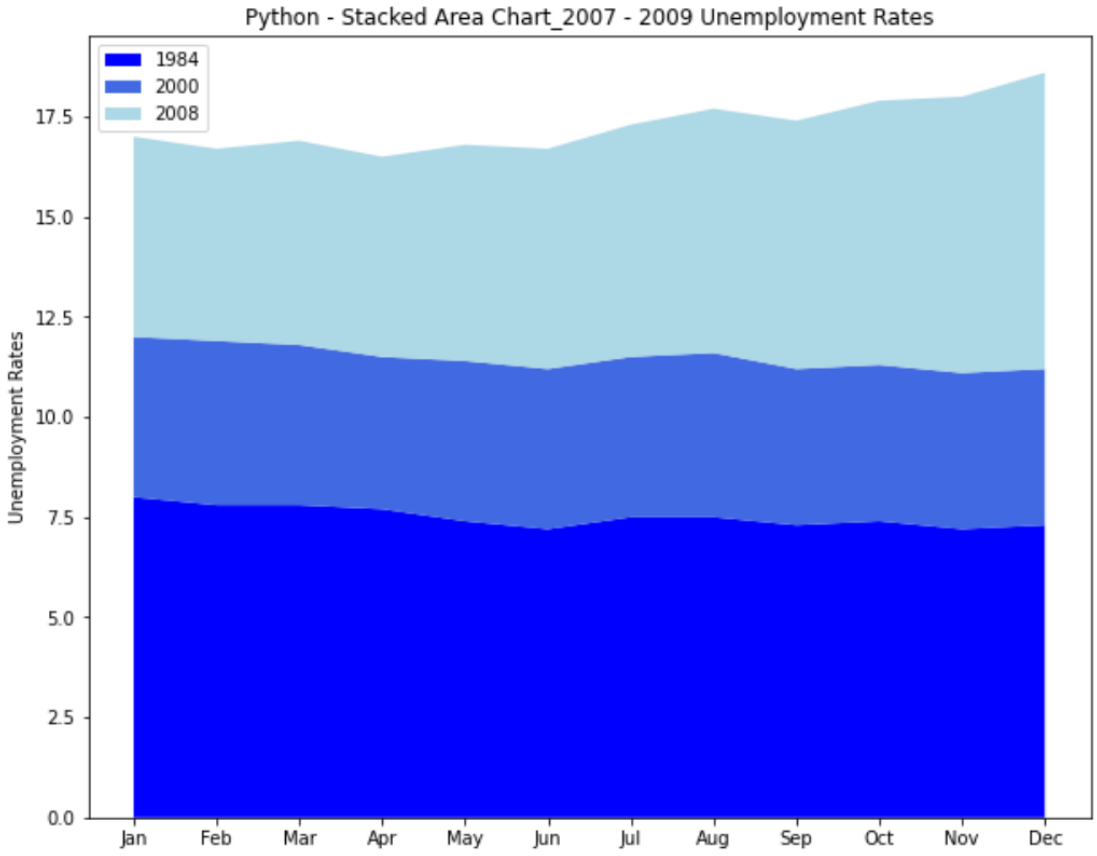
Tableau



3 – Stacked Area Chart

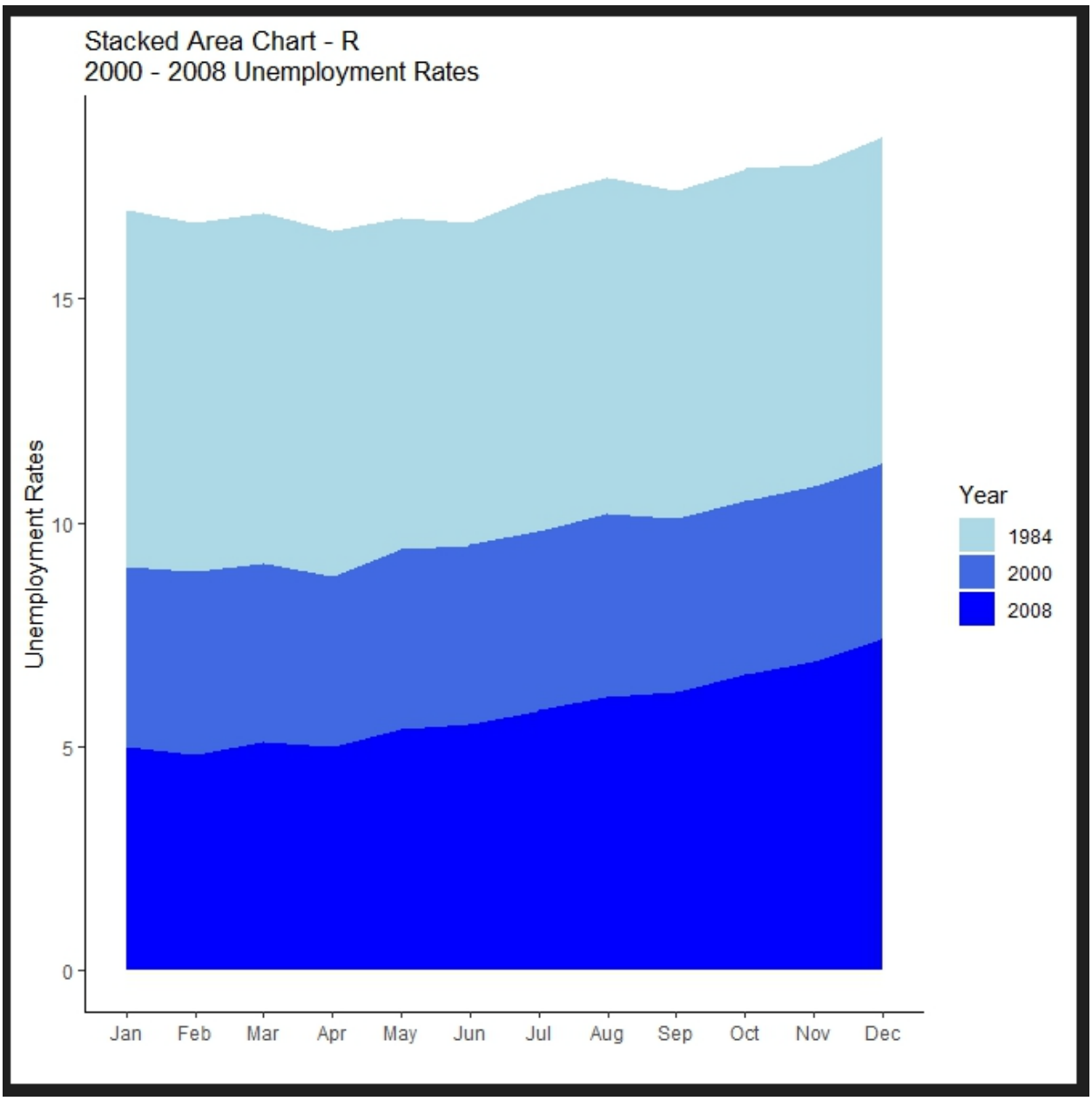
Python

```
# display the Stack Area Chart
fig, ax = plt.subplots(figsize=(10,8))

ax.stackplot(x, y_values.values(), labels = y_values.keys(), colors = ['blue', 'royalblue', 'lightblue'])
plt.title('Python - Stacked Area Chart_2007 - 2009 Unemployment Rates')
plt.ylabel('Unemployment Rates')
plt.xlabel('')

# legend
ax.legend(loc = 'upper left')

plt.show()
```

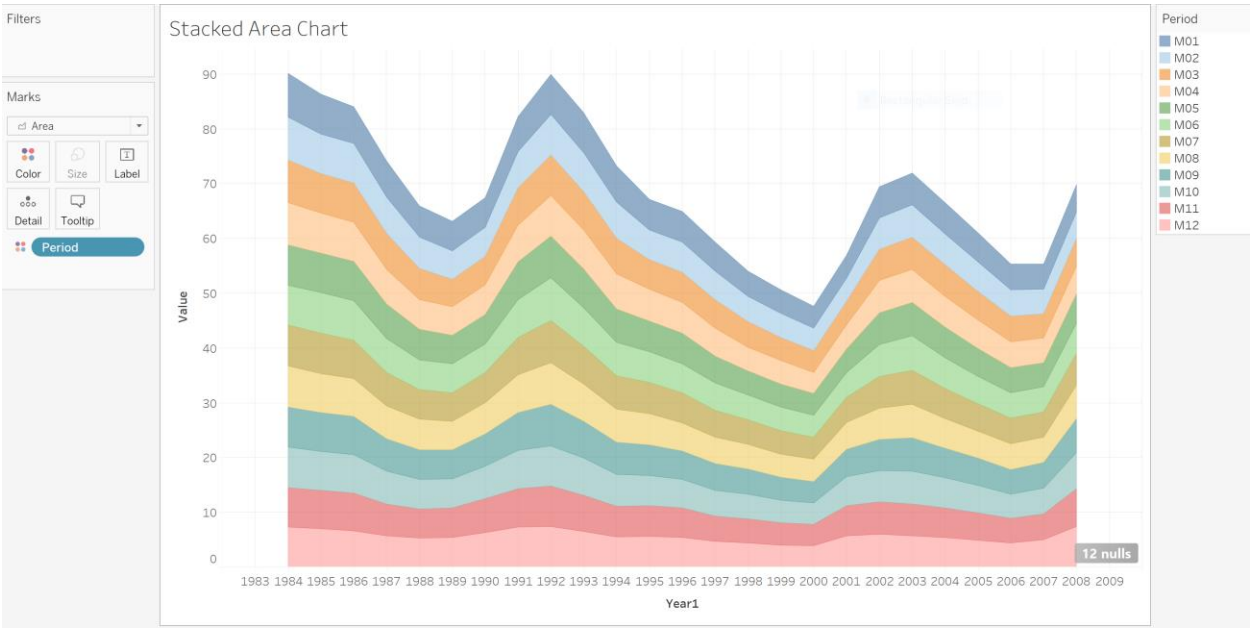


R

```r
#| label: stackedareachart
#| fig-width: 6.5
fig <- ggplot(three_years, aes(x = Period, y = Value, fill = Year, group = Year)) +
  geom_area(position = 'stack') +
  ggtitle("Stacked Area Chart - R \n2000 - 2008 Unemployment Rates") +
  xlab('') +
  scale_x_discrete(labels=c('Jan', 'Feb', 'Mar', 'Apr', 'May', 'Jun', 'Jul',
                            'Aug', 'Sep', 'Oct', 'Nov', 'Dec')) +
  ylab("Unemployment Rates") +
  scale_fill_manual(values = c('lightblue', 'royalblue', 'blue'))
fig
```


Stacked Area Chart - R
2000 - 2008 Unemployment Rates
Unemployment Rates
15
10
5
0
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Year
1984
2000
2008

Tableau


Filters
Marks
Area
Color
Size
Label
Detail
Tooltip
Period
Stacked Area Chart
Value
90
80
70
60
50
40
30
20
10
0
1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009
Year1
12 nulls
Period
M01
M02
M03
M04
M05
M06
M07
M08
M09
M10
M11
M12