Python Plots

In [1]:
```
# Imports
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import math
from matplotlib.ticker import FuncFormatter
import plotly
import plotly.figure_factory as ff
from pandas.plotting import parallel_coordinates
import numpy as np

%matplotlib inline
```

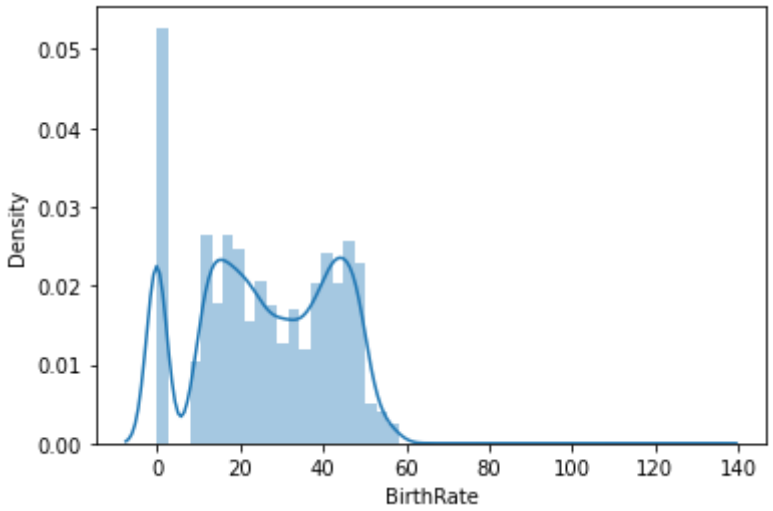
In [2]:
```
education = pd.read_csv('ex6-2/education.csv')
crime = pd.read_csv('ex6-2/crimeratesbystate-formatted.csv')
birthrate = pd.read_csv('ex6-2/birth-rate.csv')

# removing whitespaces
education = education.applymap(lambda x: x.strip() if type(x) is str else x)
crime = crime.applymap(lambda x: x.strip() if type(x) is str else x)
birthrate = birthrate.applymap(lambda x: x.strip() if type(x) is str else x)
```

Python - Histogram

In [3]:
```
birthrate_hist = pd.melt(birthrate, id_vars="Country", var_name="Year", value_na
birthrate_hist["BirthRate"] = birthrate_hist["BirthRate"].apply(lambda x: math.c
sns.distplot( birthrate_hist["BirthRate"] )
```

```
/Users/navavallepalli/opt/anaconda3/lib/python3.9/site-packages/seaborn/distribu
tions.py:2619: FutureWarning: `distplot` is a deprecated function and will be re
moved in a future version. Please adapt your code to use either `displot` (a fig
ure-level function with similar flexibility) or `histplot` (an axes-level functi
on for histograms).
  warnings.warn(msg, FutureWarning)
```

Out[3]:
```
<AxesSubplot:xlabel='BirthRate', ylabel='Density'>
```



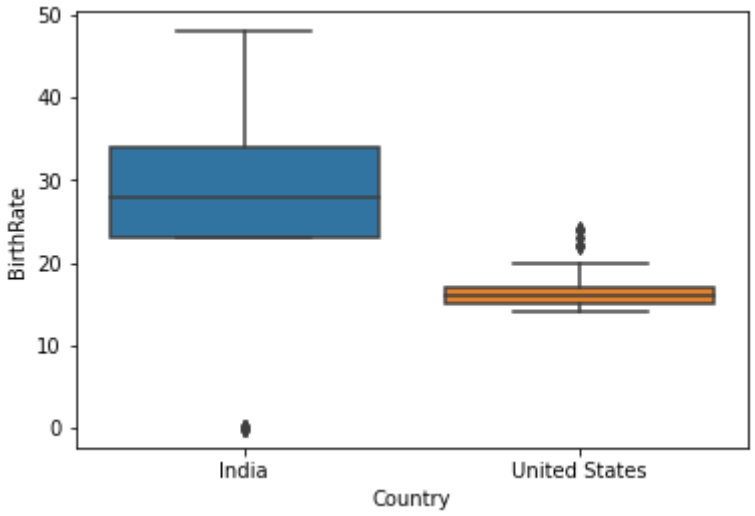
Python - Box plot

In [4]:

```
birthrate_box = birthrate_hist[(birthrate_hist["Country"]=="India") | (birthrate
sns.boxplot(x = birthrate_box["Country"], y=birthrate_box["BirthRate"])
```

Out[4]:

```
<AxesSubplot:xlabel='Country', ylabel='BirthRate'>
```



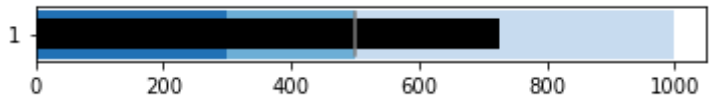
Python - Bullet chart

In [5]:

```
crime_bullet = crime[crime["state"]=="United States"][["state","burglary"]]
crime_bullet['target'] = 500
crime_bullet_tuple = [tuple(x) for x in crime_bullet.values][0]
crime_bullet_tuple

limits = [300, 500, 1000]
palette = sns.color_palette("Blues_r", len(limits))
fig, ax = plt.subplots()
ax.set_aspect('equal')
ax.set_yticks([1])

prev_limit = 0
for idx, lim in enumerate(limits):
```

```
    ax.barh([1], lim-prev_limit, left=prev_limit, height=75, color=palette[idx])
    prev_limit = lim

ax.barh([1], crime_bullet_tuple[1], color='black', height=45)

ax.axvline(crime_bullet_tuple[2], color="gray", ymin=0.10, ymax=0.9)
```

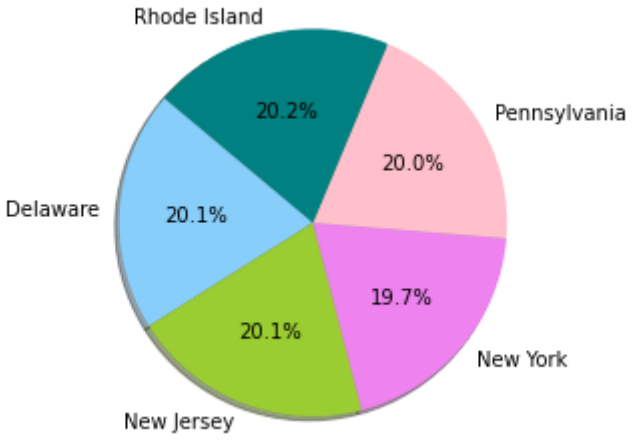
Out[5]: <matplotlib.lines.Line2D at 0x7f820ad9f7c0>



Python - Pie chart

In [6]:

```
# Pie chart: Comparison of reading numbers between 5 states
education_parallel = education[education['state'].isin(['New York','New Jersey',
education_pie = education_parallel[['state','reading']]

# set colors
colors = ['lightskyblue', 'yellowgreen', 'violet', 'pink','teal']

# plot
plt.pie(education_pie['reading'], labels=education_pie['state'], colors=colors,
autopct='%1.1f%%', shadow=True, startangle=140)

plt.axis('equal')
plt.show()
```



R plots

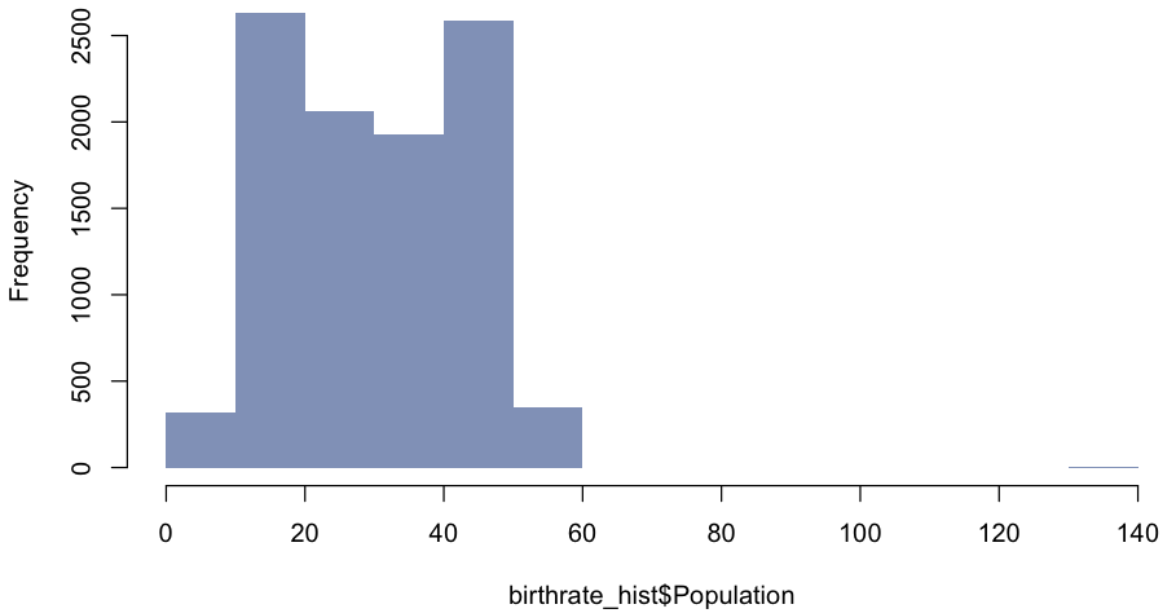
In [1]:
```
# install.packages("reshape2")
```

In [2]:
```
library('magrittr')
library("reshape2")
source("BulletGraph.R", local=TRUE)
```

In [3]:
```
birthrate <- read.csv('ex6-2/birth-rate.csv')
crime <- read.csv('ex6-2/crimeratesbystate-formatted.csv')
education <- read.csv('ex6-2/education.csv')
```

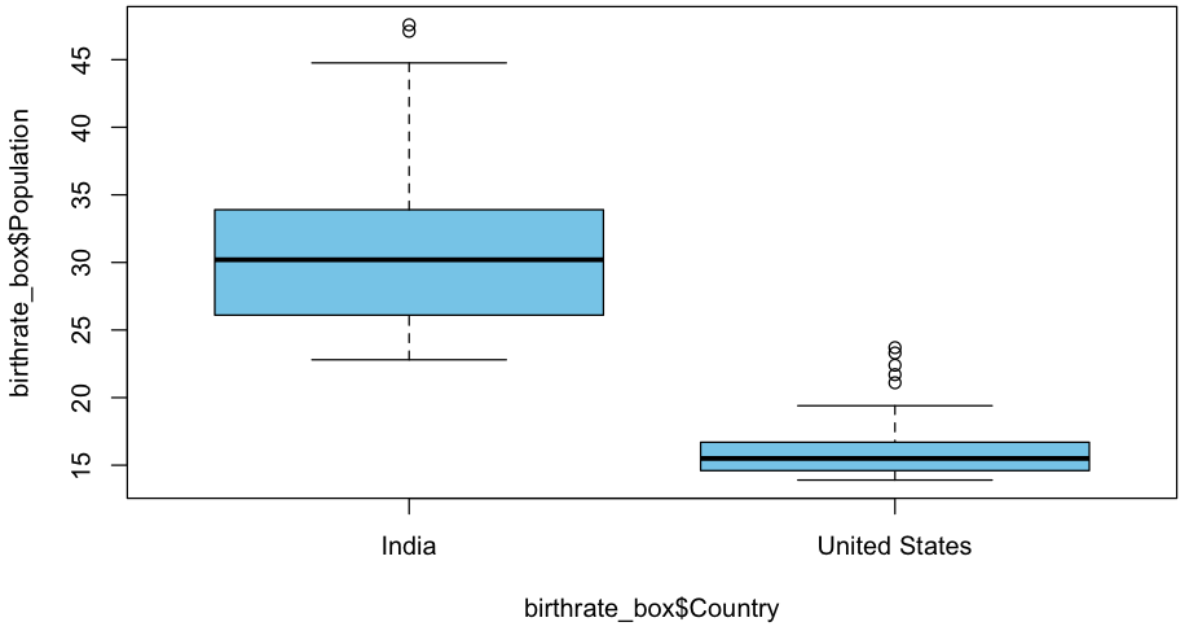
R - Histogram

In [4]:
```
options(repr.plot.width = 8, repr.plot.height = 5)

birthrate_hist <- reshape2::melt(birthrate, id=c("Country")) %>%
  dplyr::mutate("Country" = as.character(Country),
                "Year" = as.character(variable),
                "Population" = value,
                "Population_int" = ceiling(value)) %>%
  dplyr::select(c("Country","Year","Population", "Population_int"))

hist(birthrate_hist$Population, col=rgb(0,0.2,0.5,0.5) , border=F , main="")
```



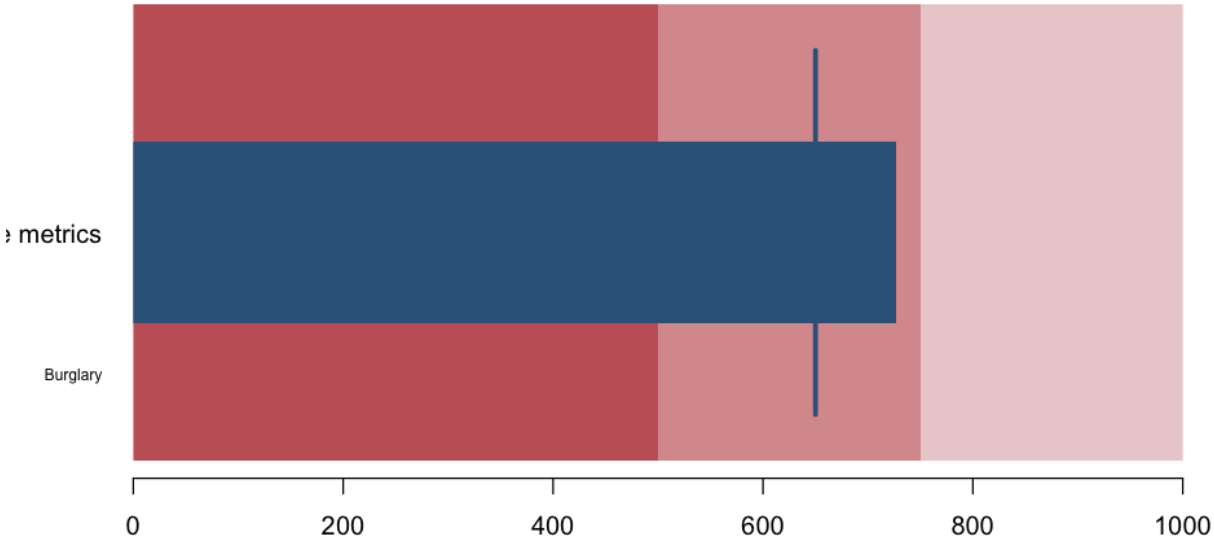
R - Box plot

In [5]:
```
birthrate_box <- birthrate_hist %>%
  dplyr::filter(Country %in% c("United States", "India"))

boxplot(birthrate_box$Population ~ birthrate_box$Country , col="skyblue")
```



R - Bullet graph

In [6]:
```
crime_bullet <- crime %>%
  dplyr::filter(stringr::str_trim(state, "both") == "United States") %>%
  dplyr::select(c(state, burglary))
bulletgraph(x=crime_bullet$burglary,ref=650,limits=c(0,500,750,1000),
            name= "USA Crime metrics",subname="Burglary",
            col="steelblue4",shades="firebrick")
```

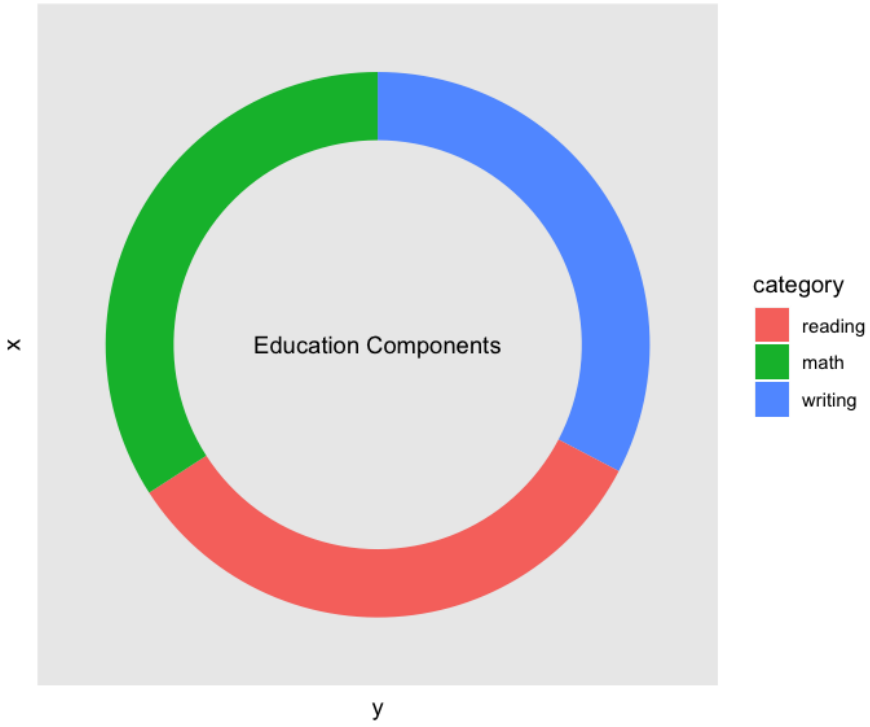


R - Donut chart

In [7]:

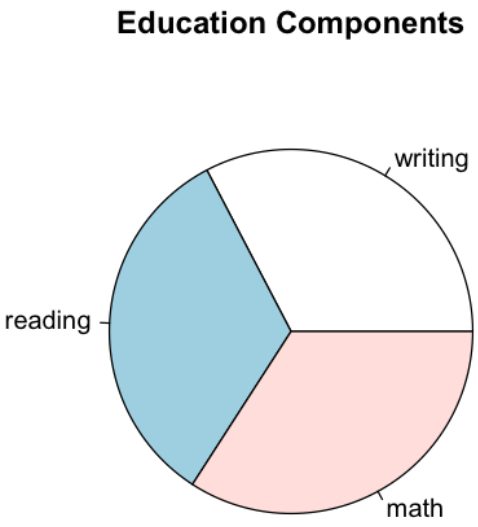
```r
education_donut <- education %>%
  dplyr::filter(stringr::str_trim(state, "both") == "United States") %>%
  reshape2::melt(id=c("state")) %>%
  dplyr::rename("category" = variable) %>%
  dplyr::filter(category %in% c("reading","math","writing")) %>%
  dplyr::select(-state)

# add addition columns, needed for drawing with geom_rect
education_donut$fraction = education_donut$value / sum(education_donut$value)
education_donut = education_donut[order(education_donut$fraction), ]
education_donut$ymax = cumsum(education_donut$fraction)
education_donut$ymin = c(0, head(education_donut$ymax, n=-1))

# make the plot
ggplot2::ggplot(education_donut, ggplot2::aes(fill=category, ymax=ymax, ymin=ym
  ggplot2::geom_rect() +
  ggplot2::coord_polar(theta="y") +
  ggplot2::xlim(c(0, 4)) +
  ggplot2::theme(panel.grid=ggplot2::element_blank()) +
  ggplot2::theme(axis.text=ggplot2::element_blank()) +
  ggplot2::theme(axis.ticks=ggplot2::element_blank()) +
  ggplot2::annotate("text", x = 0, y = 0, label = "Education Components") +
  ggplot2::labs(title="")
```

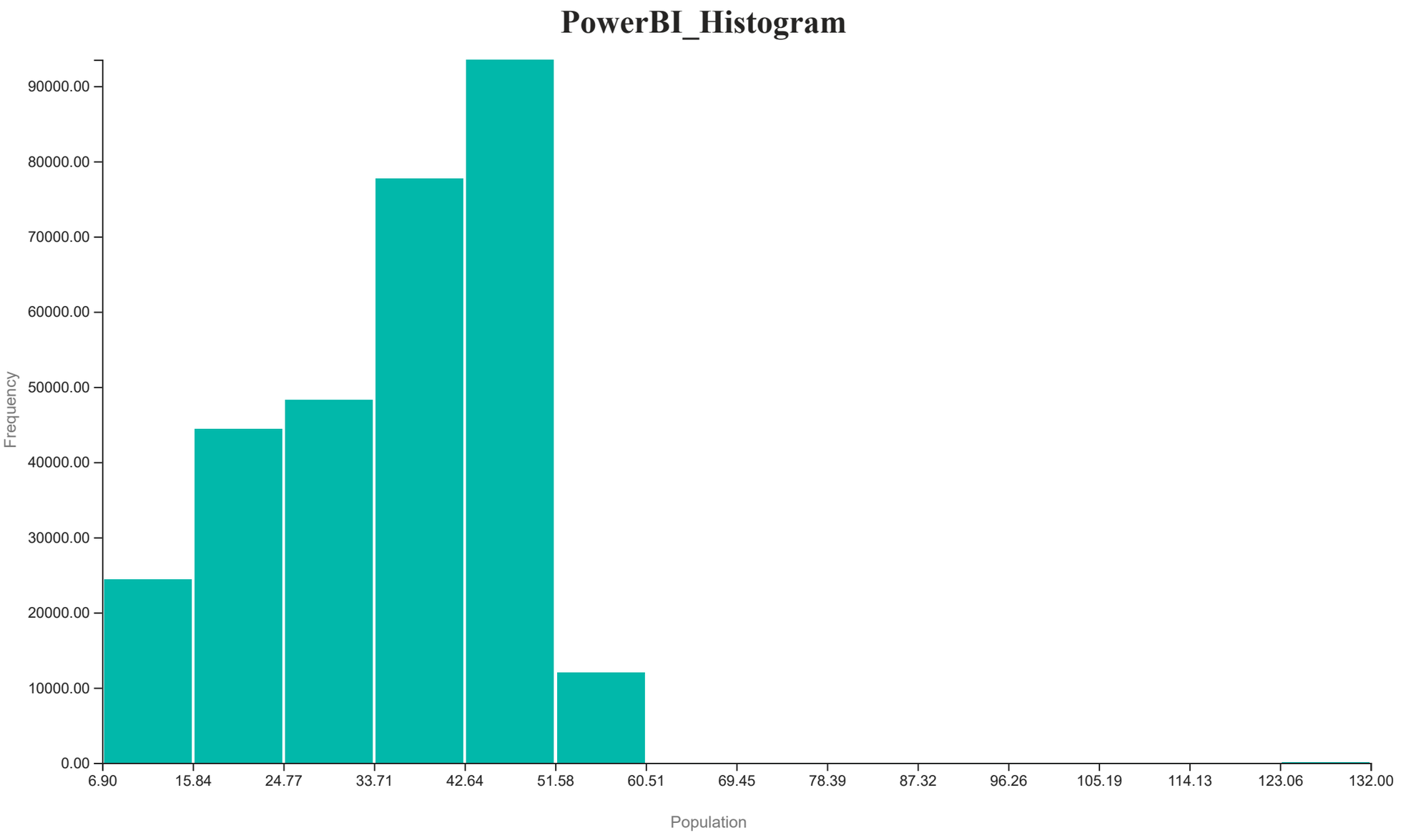
```
Registered S3 methods overwritten by 'ggplot2':
  method         from
  [.quosures     rlang
  c.quosures     rlang
  print.quosures rlang
```



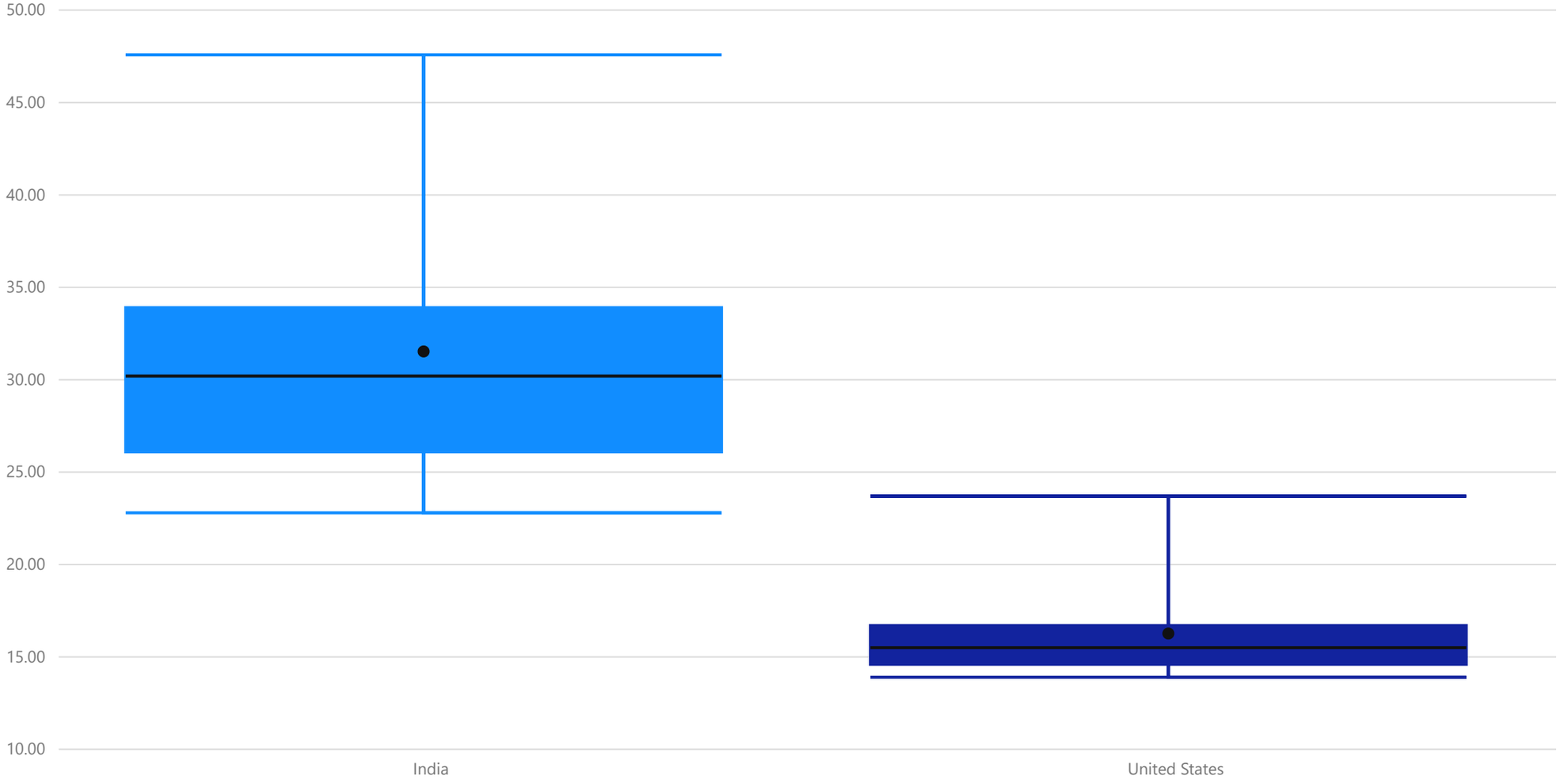
R - Pie chart

In [8]:
```r
# pie chart
slices <- education_donut$value
lbls <- education_donut$category
pie(slices, labels = lbls, main="Education Components")
```

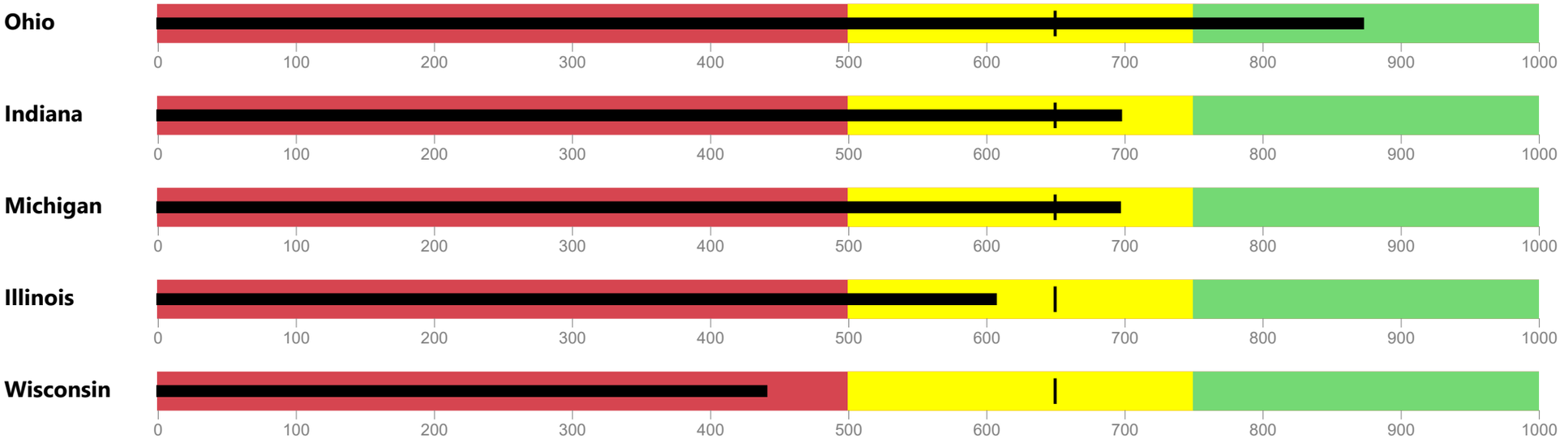



PowerBI_Histogram
90000.00
80000.00
70000.00
60000.00
50000.00
40000.00
30000.00
20000.00
10000.00
0.00
Frequency
6.90
15.84
24.77
33.71
42.64
51.58
60.51
69.45
78.39
87.32
96.26
105.19
114.13
123.06
132.00
Population

PowerBI-Box Plot



PowerBI-Bullet chart



PowerBI-Ribbon chart