Sai Sriram Uppada

2024-03-11

```
library(readr)
library(tidyverse)
```

```
## ── Attaching core tidyverse packages ──────────────────────── tidyverse
2.0.0 ──
## ✓ dplyr     1.1.4     ✓ purrr     1.0.2
## ✓ forcats   1.0.0     ✓ stringr   1.5.1
## ✓ ggplot2   3.4.4     ✓ tibble    3.2.1
## ✓ lubridate 1.9.3     ✓ tidyr     1.3.1
## ── Conflicts ────────────────────────────────────────
tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all
conflicts to become errors
```

```
library(plotly)
```

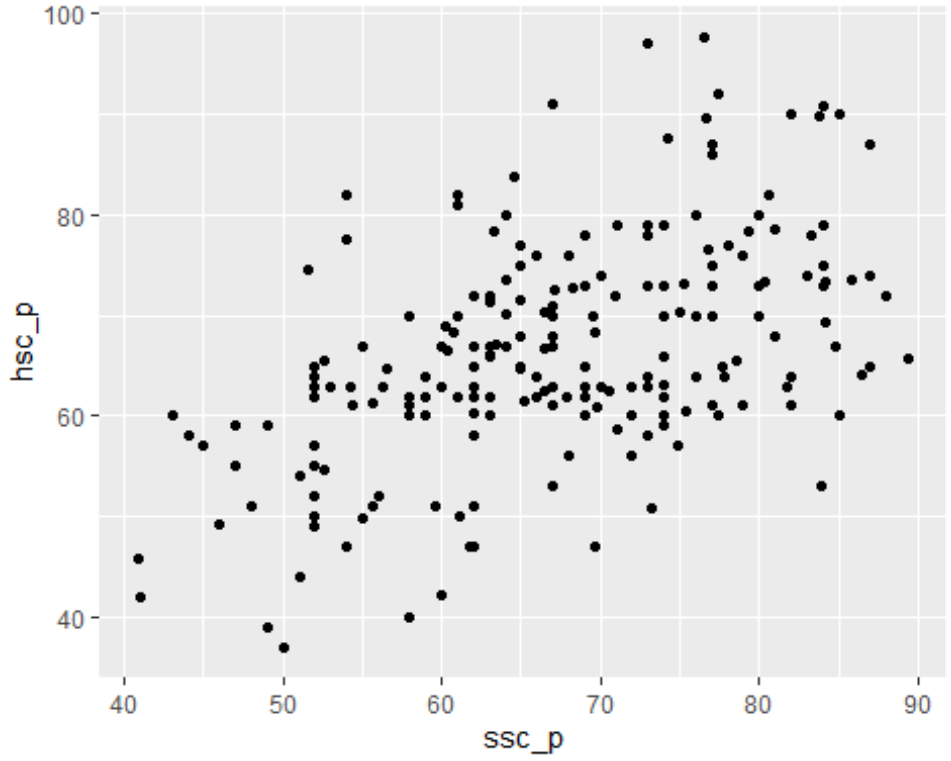
```
## Warning: package 'plotly' was built under R version 4.3.3
```

```
##
## Attaching package: 'plotly'
##
## The following object is masked from 'package:ggplot2':
##
##     last_plot
##
## The following object is masked from 'package:stats':
##
##     filter
##
## The following object is masked from 'package:graphics':
##
##     layout
```

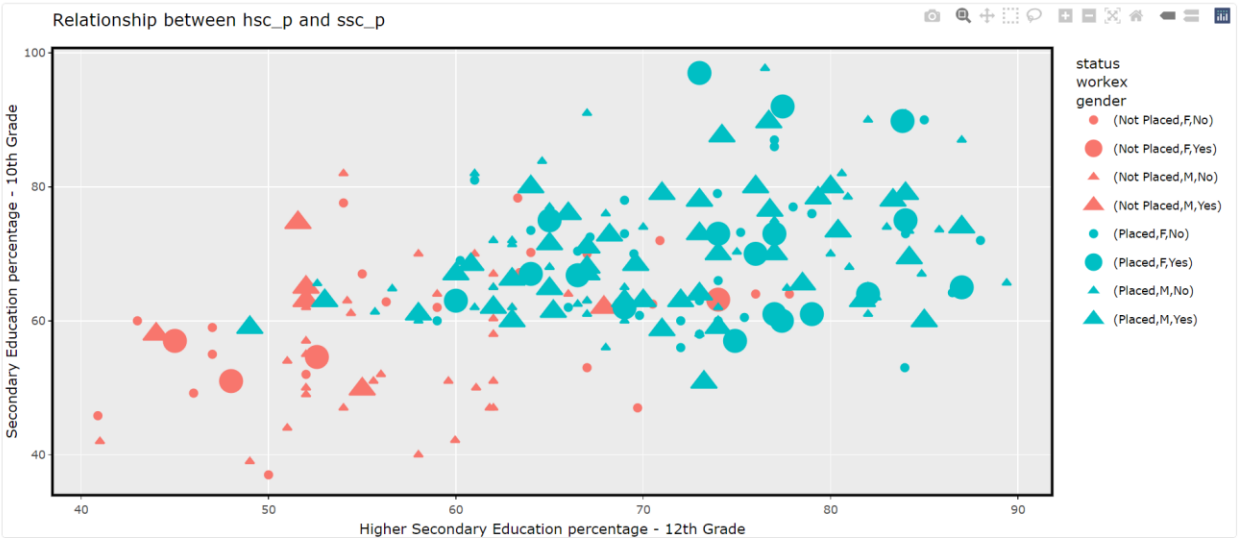
```
data <- read_csv('Placement_Data_Full_Class.csv')
```

```
## Rows: 215 Columns: 15
## ── Column specification
──────────────────────────────────────────────────────
## Delimiter: ","
## chr (8): gender, ssc_b, hsc_b, hsc_s, degree_t, workex, specialisation,
status
## dbl (7): sl_no, ssc_p, hsc_p, degree_p, etest_p, mba_p, salary
##
```

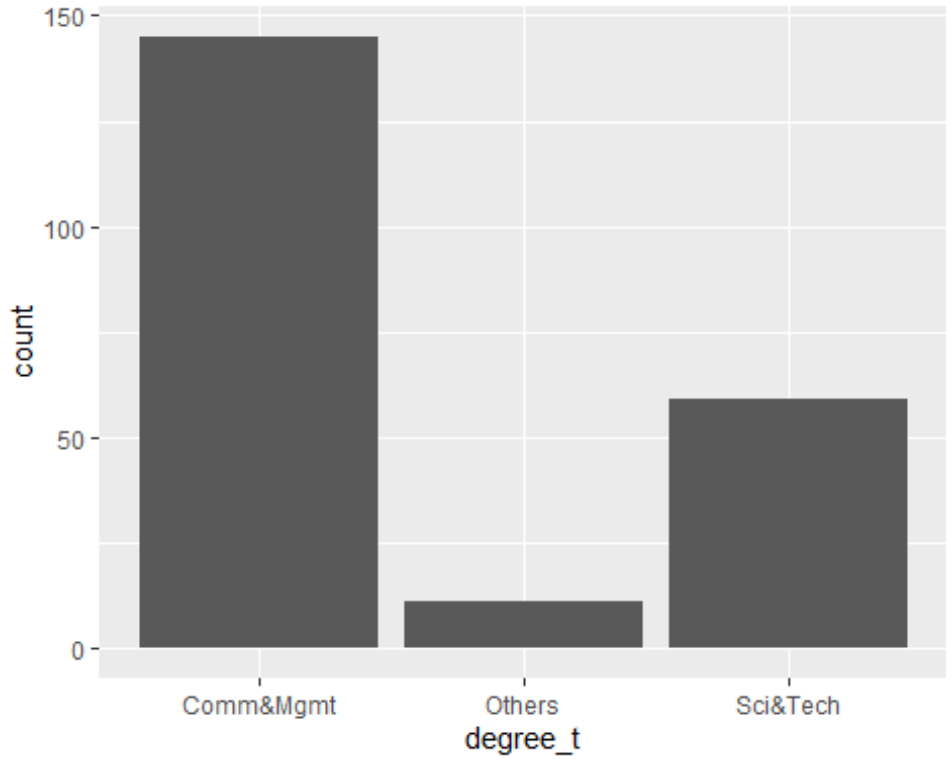
```
## ℹ Use `spec()` to retrieve the full column specification for this data.
## ℹ Specify the column types or set `show_col_types = FALSE` to quiet this
message.

print(data)

## # A tibble: 215 × 15
##    sl_no gender ssc_p ssc_b   hsc_p hsc_b hsc_s degree_p degree_t workex
etest_p
##    <dbl> <chr>  <dbl> <chr>   <dbl> <chr> <chr>    <dbl> <chr>    <chr>
<dbl>
##  1     1 M       67   Others   91   Othe… Comm…     58   Sci&Tech No
55
##  2     2 M       79.3 Central  78.3 Othe… Scie…     77.5 Sci&Tech Yes
86.5
##  3     3 M       65   Central  68   Cent… Arts      64   Comm&Mg… No
75
##  4     4 M       56   Central  52   Cent… Scie…     52   Sci&Tech No
66
##  5     5 M       85.8 Central  73.6 Cent… Comm…     73.3 Comm&Mg… No
96.8
##  6     6 M       55   Others   49.8 Othe… Scie…     67.2 Sci&Tech Yes
55
##  7     7 F       46   Others   49.2 Othe… Comm…     79   Comm&Mg… No
74.3
##  8     8 M       82   Central  64   Cent… Scie…     66   Sci&Tech Yes
67
##  9     9 M       73   Central  79   Cent… Comm…     72   Comm&Mg… No
91.3
## 10    10 M       58   Central  70   Cent… Comm…     61   Comm&Mg… No
54
## # ℹ 205 more rows
## # ℹ 4 more variables: specialisation <chr>, mba_p <dbl>, status <chr>,
## #   salary <dbl>

base1 <- ggplot(data, aes(x= ssc_p, y = hsc_p))
base1 + geom_point()
```



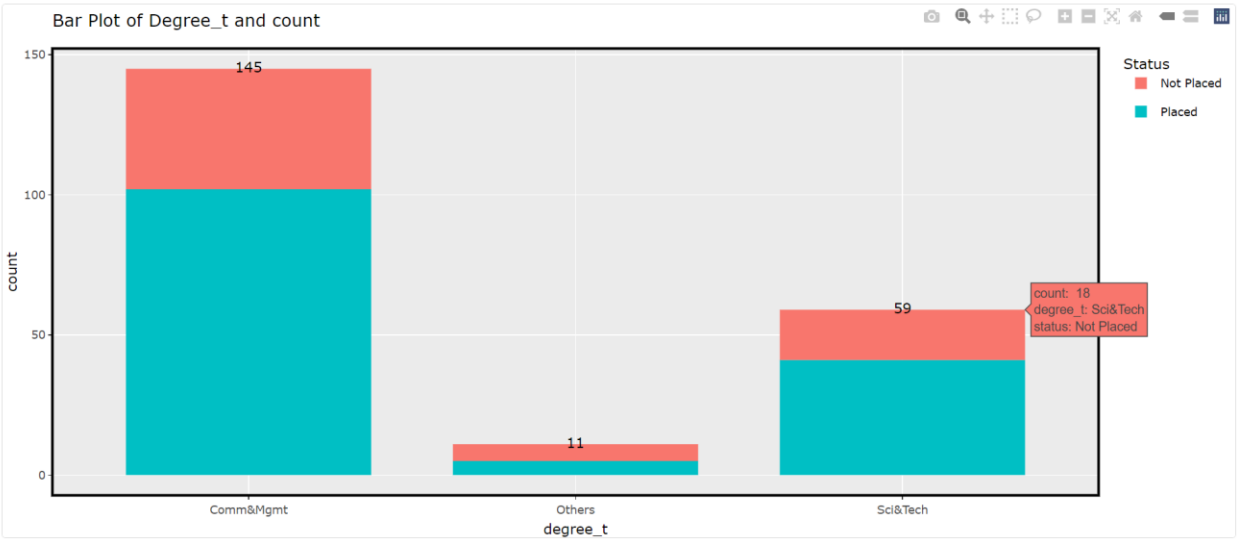

```
b <- base1 +
  geom_point(aes(color = status, shape = gender, size = workex), stroke =
0.5) +
  ggtitle("Relationship between hsc_p and ssc_p") +
  theme(panel.border = element_rect(color = "black", fill = NA, size = 2)) +
  labs(x = "Higher Secondary Education percentage - 12th Grade",
       y = "Secondary Education percentage - 10th Grade")

ggplotly(b)
```



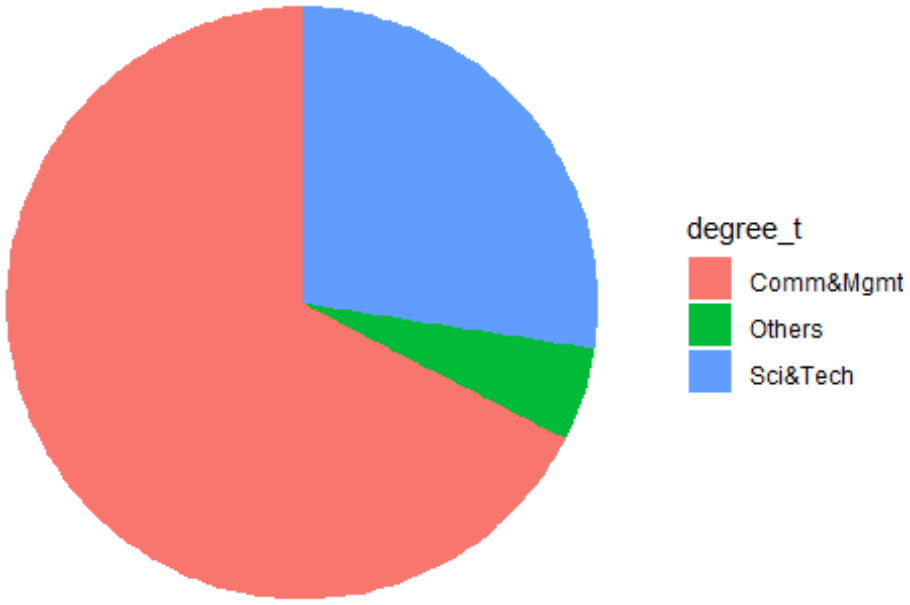
```
base2 <- ggplot(data, aes(degree_t))
base2 +geom_bar()
```



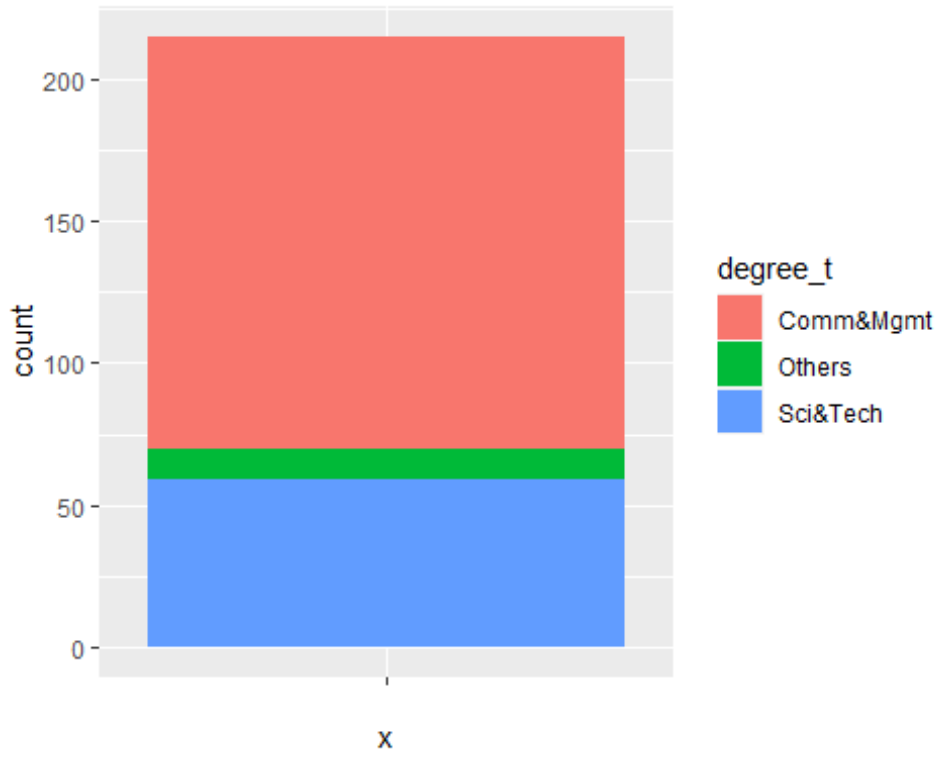

```
base2 <- ggplot(data, aes(x=degree_t,))
p <- base2 +
  geom_bar(aes(fill = status), width = 0.75, stat = "count") +
  geom_text(stat = "count", aes(label=..count..), vjust = -0.5) +
theme(panel.border = element_rect(color = "black", fill = NA, size = 2))+
  labs(fill = "Status",title = "Bar Plot of Degree_t and count")
library(plotly)
ggplotly(p)
```



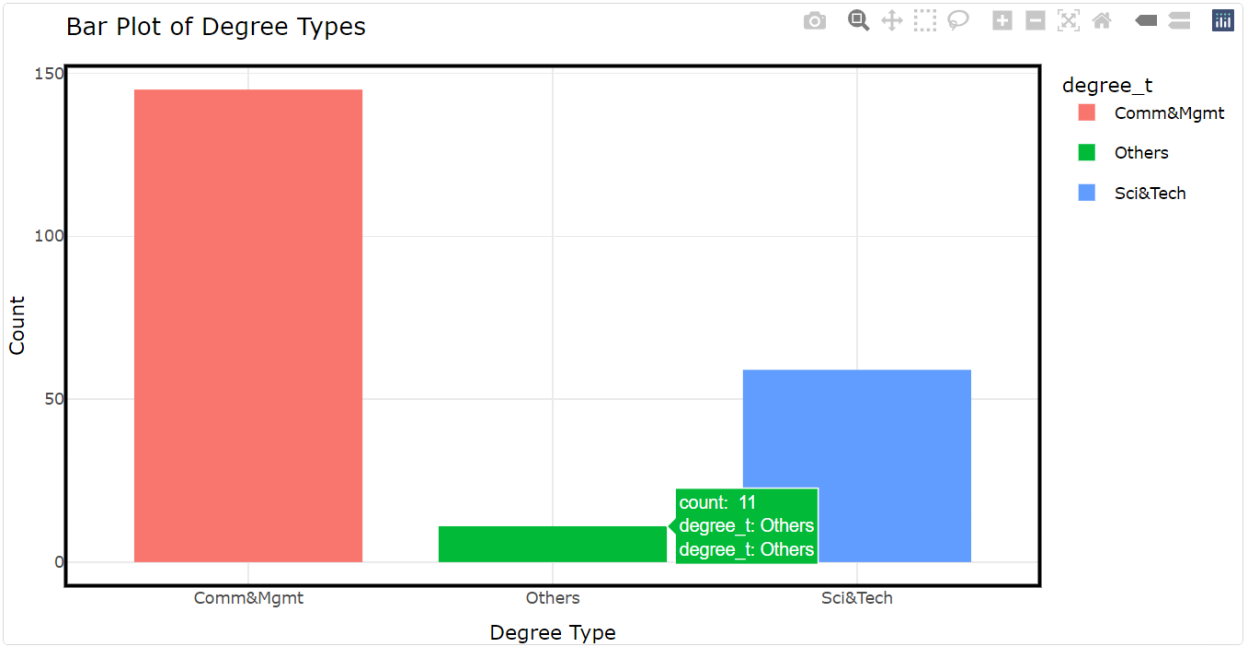

```
piebar <- ggplot(data, aes(x = " ", fill = degree_t))+ geom_bar(width = 1)
piechart <- piebar + coord_polar("y") + theme_void()
piechart
```




```
piebar
```

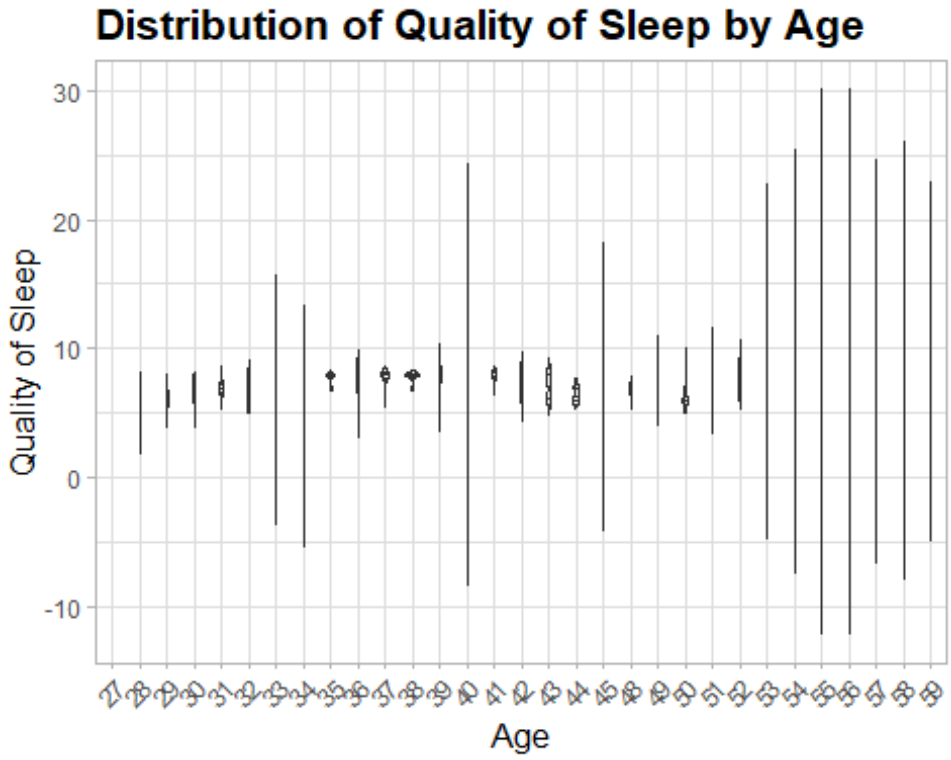



```
bar_plot <- ggplot(data, aes(x = degree_t, fill = degree_t)) +
  geom_bar(stat = "count", width = 0.75) +
  theme_minimal() + theme(panel.border = element_rect(color = "black", fill =
NA, size = 2))+
  labs(title = "Bar Plot of Degree Types",
       x = "Degree Type",
       y = "Count")
ggplotly(bar_plot)
```



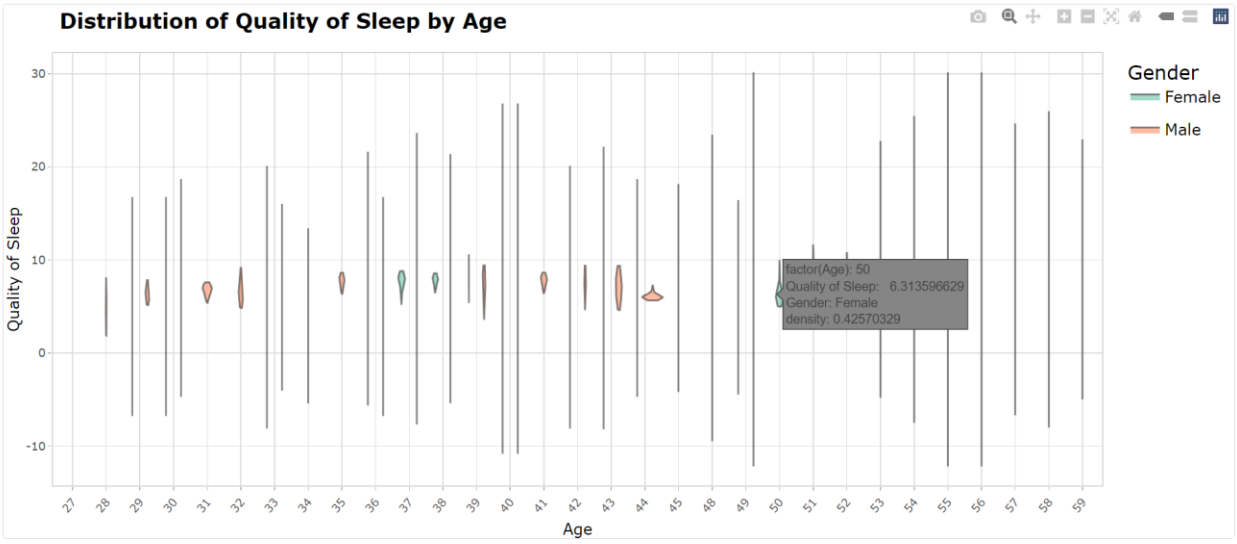

```
data1 <- read_csv('Sleep_health_and_lifestyle_dataset.csv')

data1$`Sleep Duration` <- factor(data1$`Sleep Duration`)
head(data1)

## # A tibble: 6 × 13
##   `Person ID` Gender   Age Occupation        `Sleep Duration` `Quality of
Sleep`
##         <dbl> <chr>  <dbl> <chr>             <fct>
<dbl>
## 1           1 Male      27 Software Engineer 6.1
6
## 2           2 Male      28 Doctor            6.2
6
## 3           3 Male      28 Doctor            6.2
6
## 4           4 Male      28 Sales Representa… 5.9
4
## 5           5 Male      28 Sales Representa… 5.9
4
## 6           6 Male      28 Software Engineer 5.9
4
## # ℹ 7 more variables: `Physical Activity Level` <dbl>, `Stress Level`
<dbl>,
## #   `BMI Category` <chr>, `Blood Pressure` <chr>, `Heart Rate` <dbl>,
## #   `Daily Steps` <dbl>, `Sleep Disorder` <chr>

library(ggplot2)
color_palette <- RColorBrewer::brewer.pal(name = "Set2", n =
length(unique(data1$Gender)))
```

```r
library(ggplot2)
library(RColorBrewer)
color_palette <- brewer.pal(n = 8, name = "Set2")
p <- ggplot(data1, aes(x=factor(Age), y=`Quality of Sleep`)) +
  geom_violin(trim=FALSE, scale="count", adjust=1.5, draw_quantiles=c(0.25,
0.5, 0.75), alpha=0.6) +
  scale_fill_manual(values=color_palette) +
  labs(title="Distribution of Quality of Sleep by Age",
       x="Age", y="Quality of Sleep") +
  theme_light() +
  theme(axis.text.x = element_text(angle=45, hjust=1),
        legend.title = element_text(size=25),
        legend.text = element_text(size=10),
        plot.title = element_text(size=16, face="bold"),
        axis.title.x = element_text(size=12),
        axis.title.y = element_text(size=12))
print(p)
```

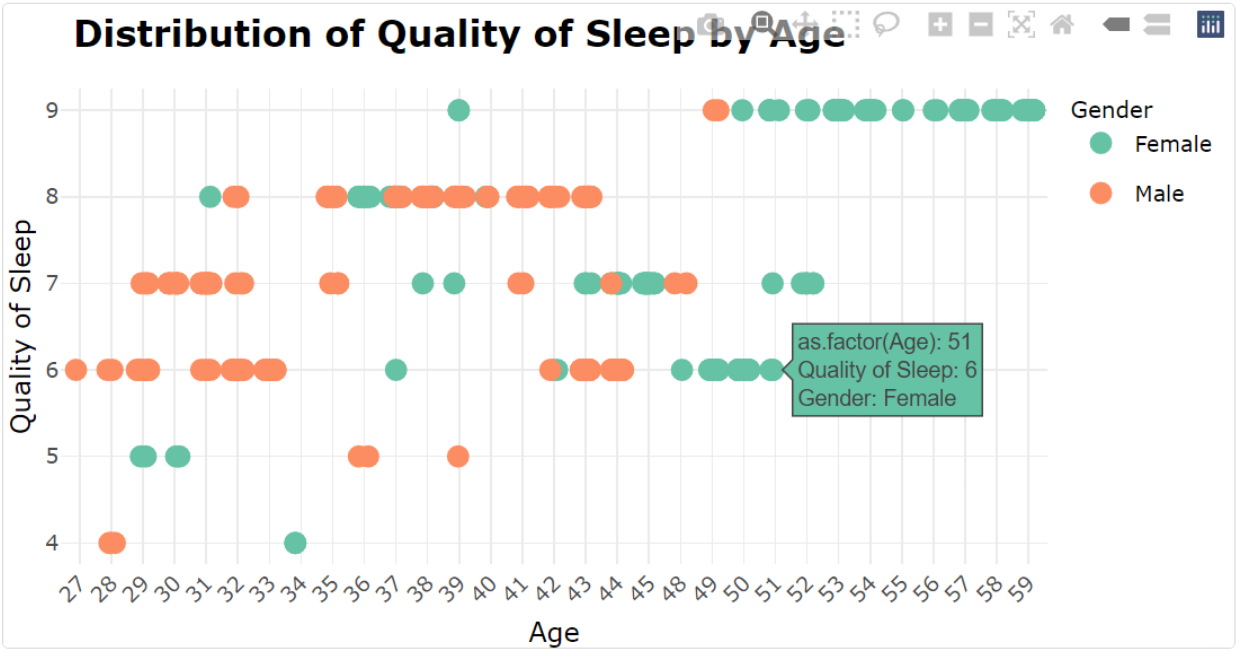



```r
color_palette <- brewer.pal(n = 8, name = "Set2")
p <- ggplot(data1, aes(x=factor(Age), y=`Quality of Sleep`, fill= Gender)) +
  geom_violin(trim=FALSE, scale="count", adjust=2.5, draw_quantiles=c(0.25,
0.75, 0.25), alpha=0.6) +
  scale_fill_manual(values=color_palette) +
  labs(title="Distribution of Quality of Sleep by Age",
       x="Age", y="Quality of Sleep") +
  theme_light() +
  theme(axis.text.x = element_text(angle=50, hjust=1),
```

```
        legend.title = element_text(size=15),
        legend.text = element_text(size=12),
        plot.title = element_text(size=16, face="bold"),
        axis.title.x = element_text(size=12),
        axis.title.y = element_text(size=12))
ggplotly(p)
```




```
a<-ggplot(data1, aes(x = as.factor(Age), y = `Quality of Sleep`, color =
Gender)) +
  geom_point(position = position_jitter(width = 0.2, height = 0), size = 3) +
  scale_color_manual(values = color_palette)  +
  labs(title = "Distribution of Quality of Sleep by Age",
       x = "Age",
       y = "Quality of Sleep",
       color = "Gender") +
  theme_minimal(base_size = 12) +
  theme(legend.title = element_text(size = 10),
        axis.text.x = element_text(angle = 45, hjust = 1),
        plot.title = element_text(size = 16, face = "bold"),
        axis.title = element_text(size = 12)) +
  guides(color = guide_legend(title.position = "top"))
ggplotly(a)
```


Distribution of Quality of Sleep by Age
Quality of Sleep
Age
Gender
Female
Male
as.factor(Age): 51
Quality of Sleep: 6
Gender: Female
9
8
7
6
5
4
27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 48 49 50 51 52 53 54 55 56 57 58 59