DSC640_Exercise_6-2_JakeMeyer

Jake Meyer

2023-02-25

Assignment Instructions

Submit 1 histogram, 1 box plot, 1 bullet chart, and 1 additional chart with R.

Show Working directory

```
getwd()
```

```
## [1] "C:/Users/jkmey/Documents/Github/DSC640_Course_Assignments/DSC640_Repository/Weeks11&12"
```

```
dir()
```

```
##  [1] "~$DSC640_Exercise_6-2_JakeMeyer.pptx"
##  [2] "birth-rate.csv"
##  [3] "crimeratesbystate-formatted.csv"
##  [4] "DSC640_Exercise_6-2_JakeMeyer.ipynb"
##  [5] "DSC640_Exercise_6-2_JakeMeyer.pbix"
##  [6] "DSC640_Exercise_6-2_JakeMeyer.pptx"
##  [7] "DSC640_Exercise_6-2_JakeMeyer_Power_BI_Images.pdf"
##  [8] "DSC640_Exercise_6-2_JakeMeyerPythonCode.pdf"
##  [9] "DSC640_Exercise_6-2_JakeMeyerRCode.Rmd"
## [10] "education.csv"
## [11] "education_revised.csv"
## [12] "tabn084.xls"
## [13] "tabn106.xls"
## [14] "tabn146.xls"
```

```
# setwd("C:/Users/jkmey/Documents/Github/DSC640_Course_Assignments/DSC640_Repository")
```

Import the necessary libraries

```
library(readxl)
library(ggplot2)
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.2.1      v stringr 1.5.0
## v readr   2.1.3      v forcats 0.5.2
## v purrr   0.3.5
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(dplyr)
library(scales)
```

```
##
## Attaching package: 'scales'
##
## The following object is masked from 'package:purrr':
##
##     discard
##
## The following object is masked from 'package:readr':
##
##     col_factor
```

```
library(plotly)
```

```
##
## 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
```
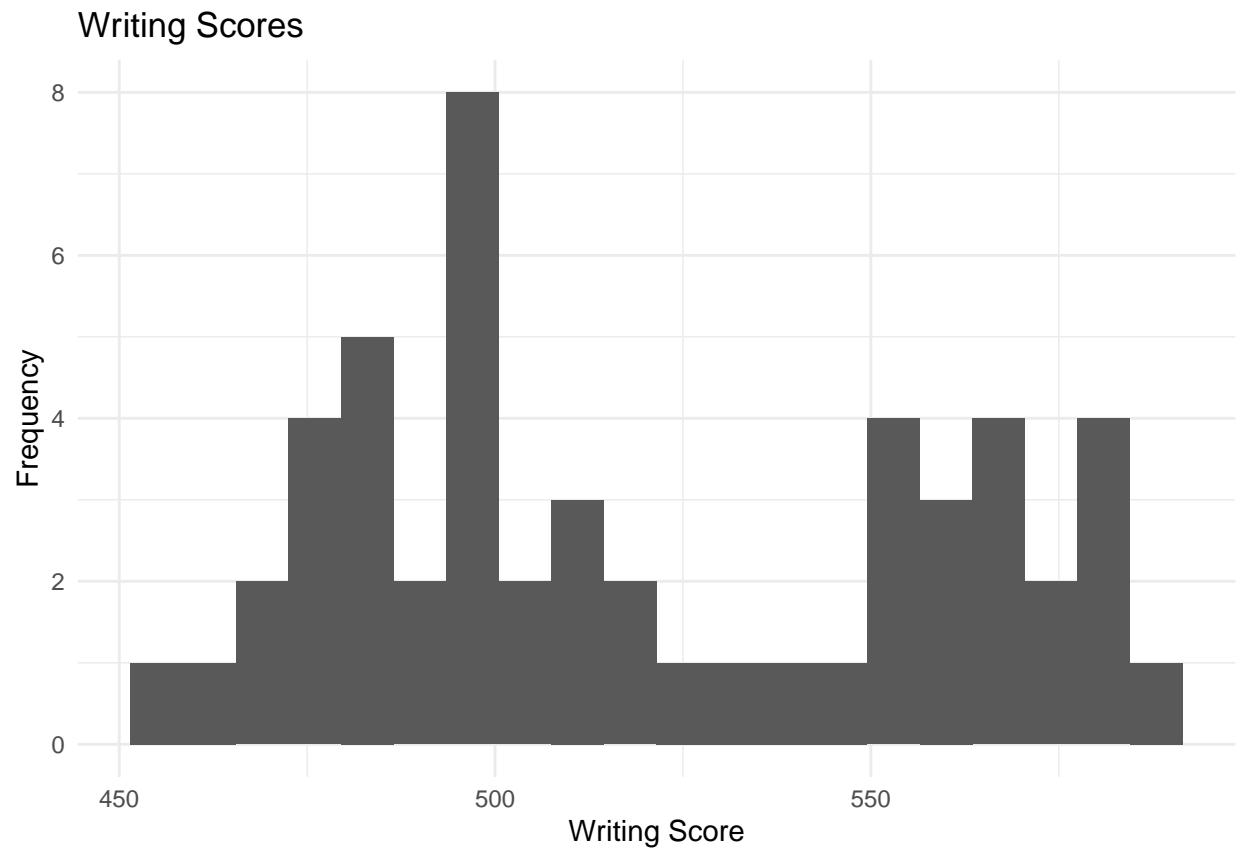
```
theme_set(theme_minimal())
```

Import the data

```
df1 <- read.csv("education.csv", header = TRUE, sep = ",")
df2 <- read.csv("education_revised.csv", header = TRUE, sep = ",")
```

Histogram.

Generate a histogram using the education data.

```
ggplot(df1, aes(writing)) + geom_histogram(binwidth = 7) +
  ggtitle('Writing Scores') + labs(x='Writing Score', y = 'Frequency')
```
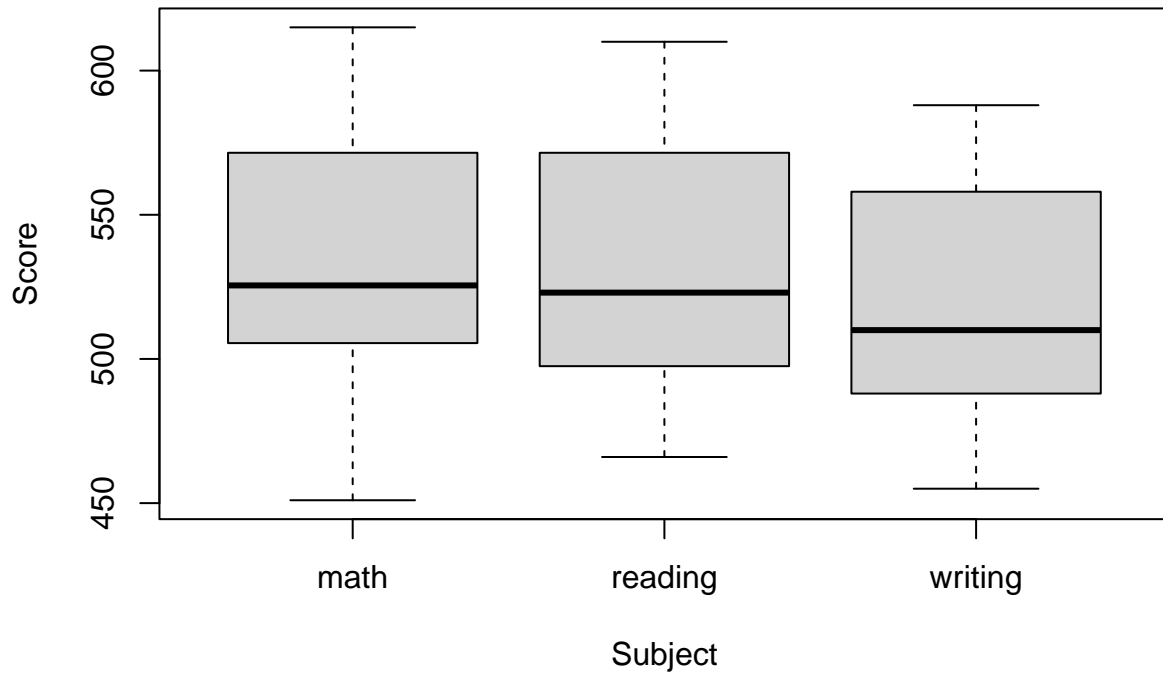
Writing Scores

Box Plot.

Generate a box plot using the education data.

```
boxplot(df2$score~df2$category, ylab = 'Score', xlab = 'Subject',
        main =' US Math, Reading, Writing Scores')
```

US Math, Reading, Writing Scores

Bullet Chart.

Generate a bullet chart using the education data.

```r
ex_df <- bind_rows(
  tibble(
  name = rep("Reading", 2),
  group = c("Qualitative", "Measure"),
  color = c("grey", "black"),
  value = c(600, 536),
  width = c(0.9, 0.5),
  target = rep(580, 2),
  ymin = rep(0.7, 2),
  ymax = rep(1.3, 2)
  ),
  tibble(
  name = rep("Writing", 2),
  group = c("Qualitative", "Measure"),
  color = c("grey", "black"),
  value = c(600, 523),
  width = c(0.9, 0.5),
  target = rep(580, 2),
  ymin = rep(1.7, 2),
  ymax = rep(2.3, 2)
  )
)
```

```
ex_df %>%
  ggplot(aes(x = value, y = name, fill = color)) +
  geom_col(width = c(0.9, 0.5, 0.9, 0.5)) +
  geom_linerange(
    aes(x = target, ymin = ymin, ymax = ymax),
    size = 2, color = "red"
    ) +
  coord_cartesian(ylim = c(0.3, 2.7)) +
  scale_fill_identity() +
  theme_minimal() +
  theme(panel.grid.major.y = element_blank()) +
  labs(x = "Score", y = "Subject", title = "US Scores for Reading and Writing")
```

```
## Warning: position_stack requires non-overlapping x intervals
```
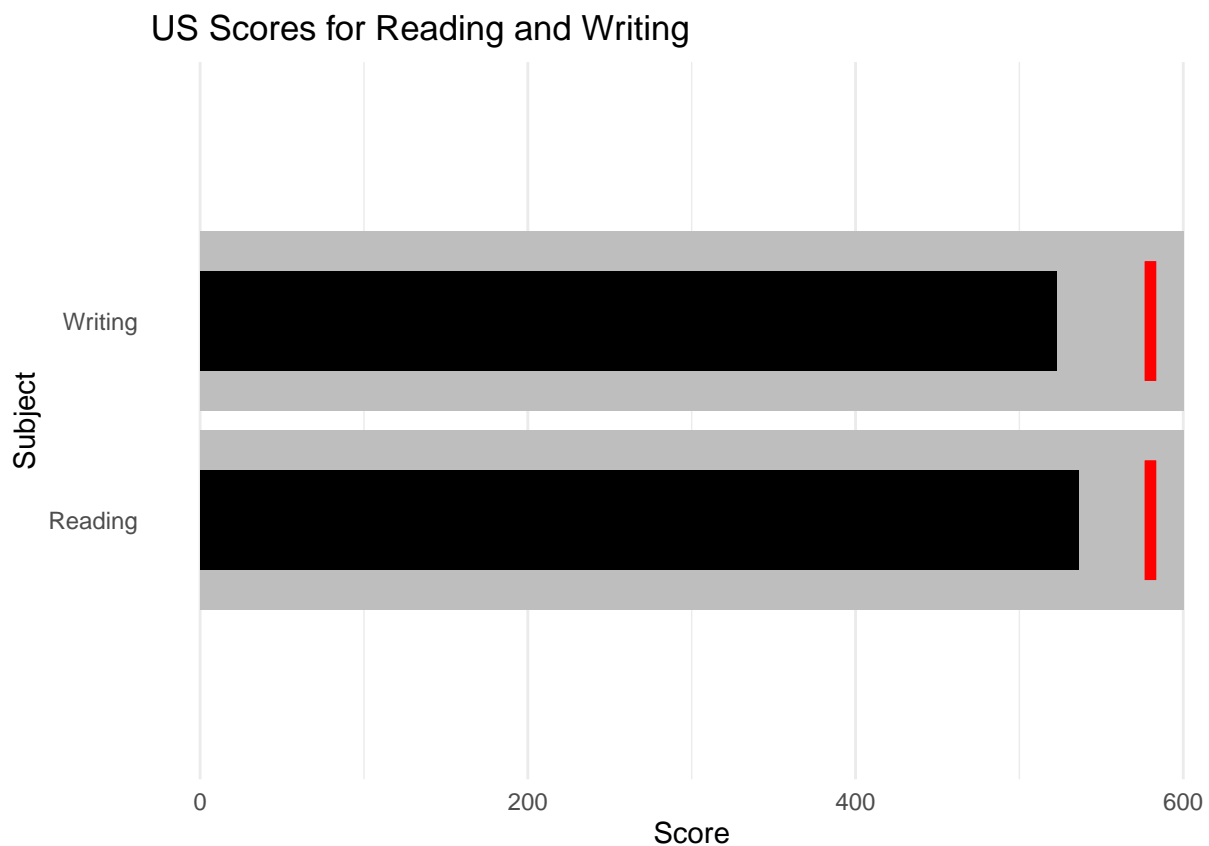

Lollipop Chart.

Generate a lollipop chart with ggplot2.

```
# Create the data for the chart.
data <- data.frame(
  x=LETTERS[1:26], y = abs(rnorm(26))
)

# Generate the plot.
ggplot(data, aes(x=x, y=y)) +
```

```
geom_segment( aes(x=x, xend=x, y=0, yend=y), color="grey") +
geom_point( color="green", size=5) +
theme_light() +
theme(
  panel.grid.major.x = element_blank(),
  panel.border = element_blank(),
  axis.ticks.x = element_blank()
) +
xlab("") +
ylab("Value of Y") +
ggtitle('26 Random Values - Lollipop Chart')
```


26 Random Values – Lollipop Chart