# python-6.2

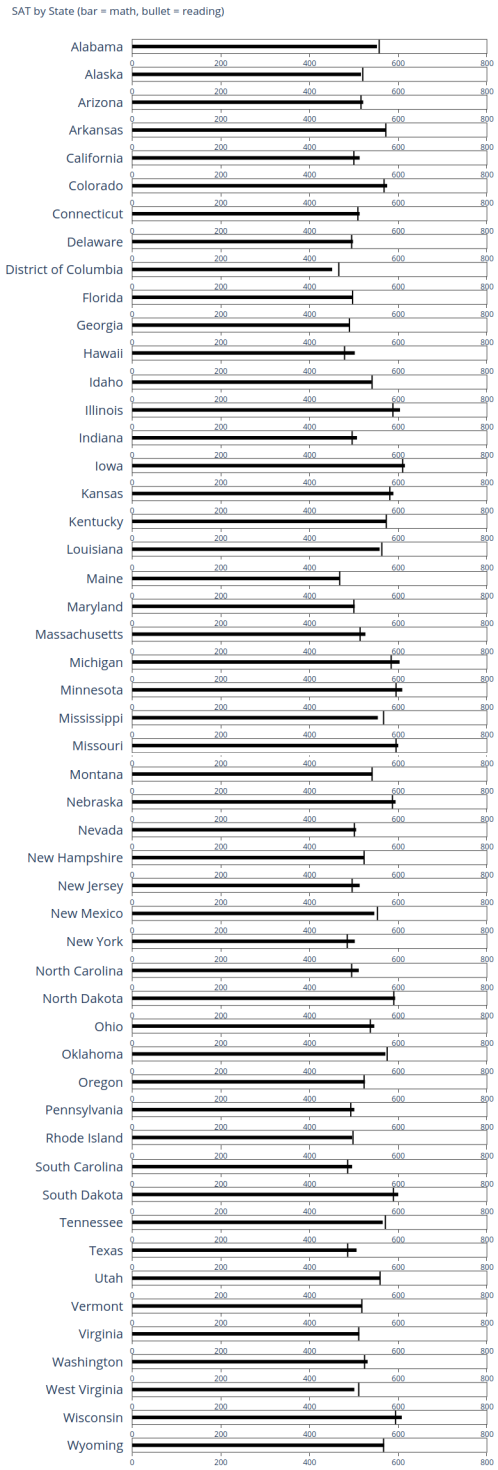
March 2, 2021

[1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import plotly.graph_objects as go
import seaborn as sns
```

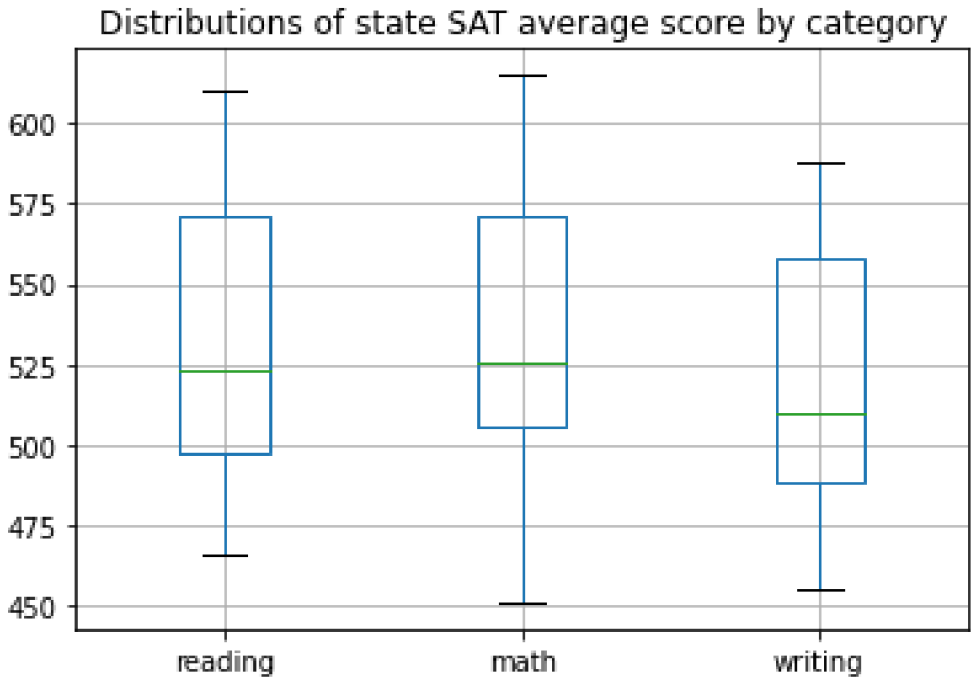
[33]:
```
df = pd.read_csv('education.csv')
```

[35]:
```
# bullet chart

# a function to add a gauge to the plot; I'll call it once per state

def add_bar(state,math,reading,height):
    fig.add_trace(go.Indicator(
    mode = "gauge",
    value = math,
    domain = {'x': [0.2, 0.9], 'y': [height[0], height[1]]}, # controls the␣
 ↪gauge location
    title = {'text' :state}, # bar label
    gauge = {
        'shape': "bullet",
        'axis': {'range': [None, 800]}, # gauge domain
        'threshold': {
            'line': {'color': "black", 'width': 2}, # bullet thickness
            'thickness': 0.95, # bullet height
            'value': reading}, # bullet position
            'bar': {'color': "black"}
    }))

# define the vertical positions of the bars in the plot - called "heights" - a␣
 ↪list of tuples
a=np.linspace(0,1,52)
b=np.linspace(0,1,52)+.01
heights=[(round(a[i],4),round(b[i],4)) for i in range(len(a))]

# create and show the plot
count = 0
```

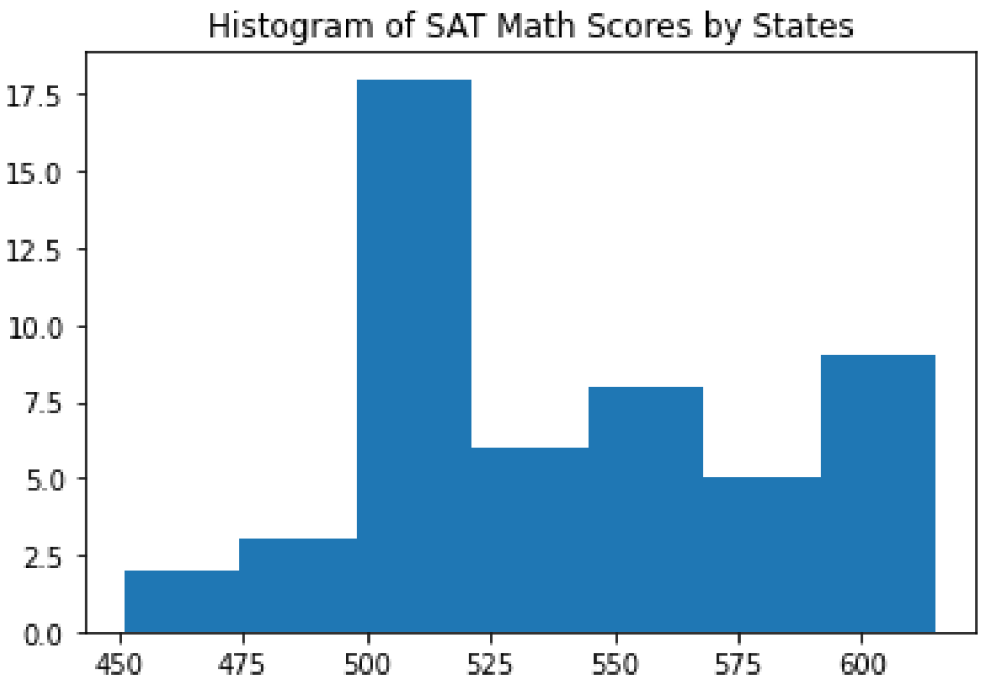
```
fig = go.Figure()
for index, row in df.iterrows():
    add_bar(row[0],row[2],row[1],heights[50-count])
    count+=1
fig.update_layout(height = 2400 ,
                  title="SAT by State (bar = math, bullet = reading)")
fig.show()
```



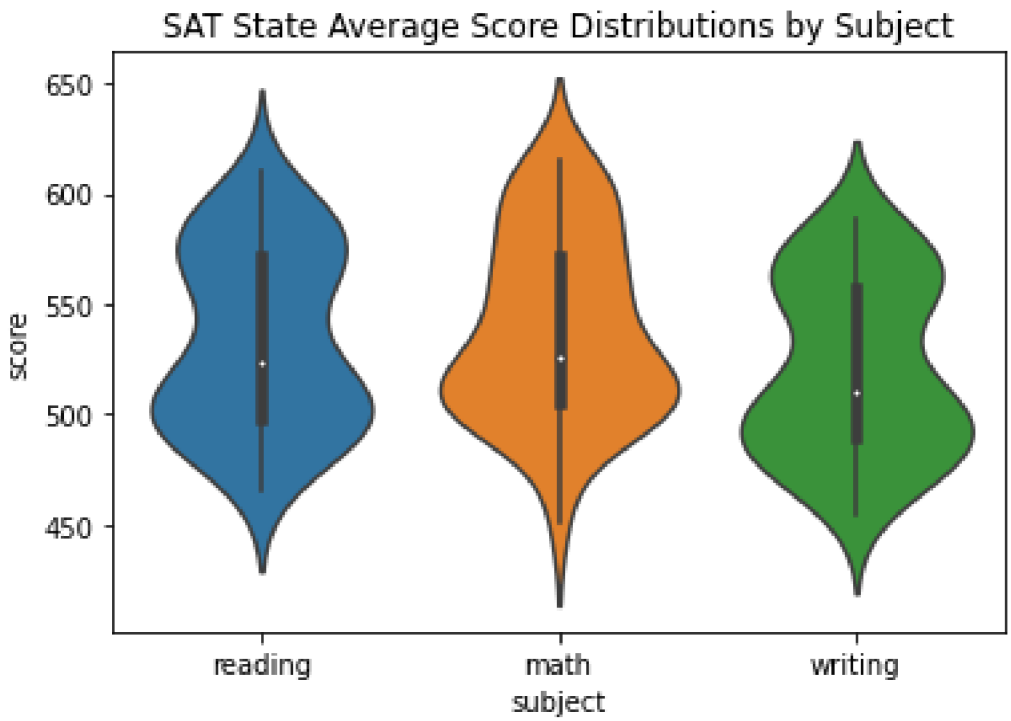
[20]:
```
# box plot

df[['reading','math','writing']].boxplot() # define the boxplot
plt.title('Distributions of state SAT average score by category') # set the␣
 ↪title
plt.show()
```




[24]:
```
# histogram

plt.hist(df.math, bins='auto') # define the histogram
plt.title("Histogram of SAT Math Scores by States") # set the plot title
plt.show()
```




[31]:
```
# violin plot

# load the long form of the SAT data
df = pd.read_csv('education_mod.csv')

# define the violin plot
ax = sns.violinplot(x="subject", y="score", data=df)

# set the plot title
ax.title.set_text('SAT State Average Score Distributions by Subject')
```


SAT State Average Score Distributions by Subject
650
600
550
500
450
score
reading
math
writing
subject

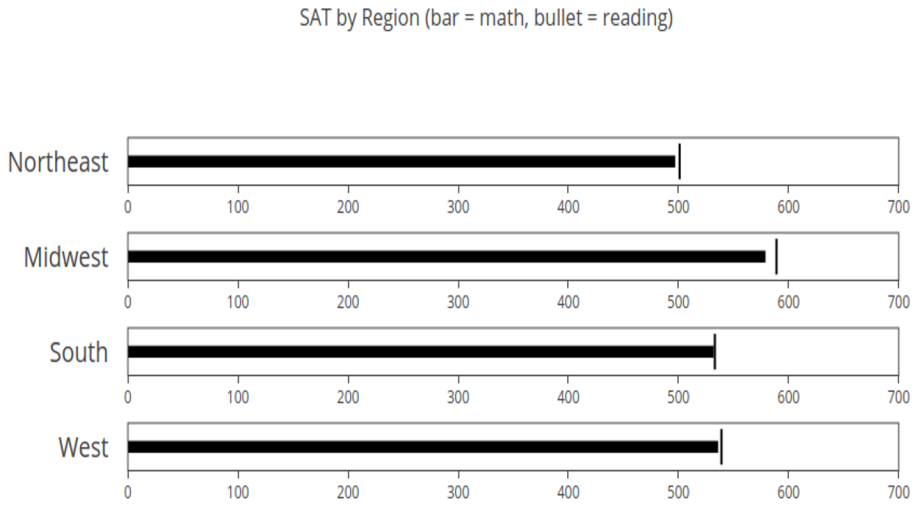
# r-6.2

March 2, 2021

[1]:
```
library(readr)
library(ggplot2)
library(plotly)
```

[50]:
```
# bullet chart

fig <- plot_ly()  %>% layout(title="SAT by Region (bar = math, bullet =␣
 ↪reading)")
fig <- fig %>%
  add_trace(
    type =  "indicator",
    mode = "gauge",
    value = 497,
    domain = list(x = c(0.1, 0.9), y = c(0.7, 0.8)),
    title = list(text = "Northeast"),
    gauge = list(
      shape = "bullet",
      axis = list(range = list(NULL, 700)),
      threshold = list(
        line = list(color = "black", width = 2),
        thickness = 0.75,
        value = 501),
      bar = list(color = "black")))
fig <- fig %>%
  add_trace(
    type =  "indicator",
    mode = "gauge",
    value = 579,
    domain = list(x = c(0.1, 0.9), y = c(0.5, 0.6)),
    title = list(text = "Midwest"),
    gauge = list(
      shape = "bullet",
      axis = list(range = list(NULL, 700)),
      threshold = list(
        line = list(color = "black", width = 2),
        thickness = 0.75,
        value = 589),
```

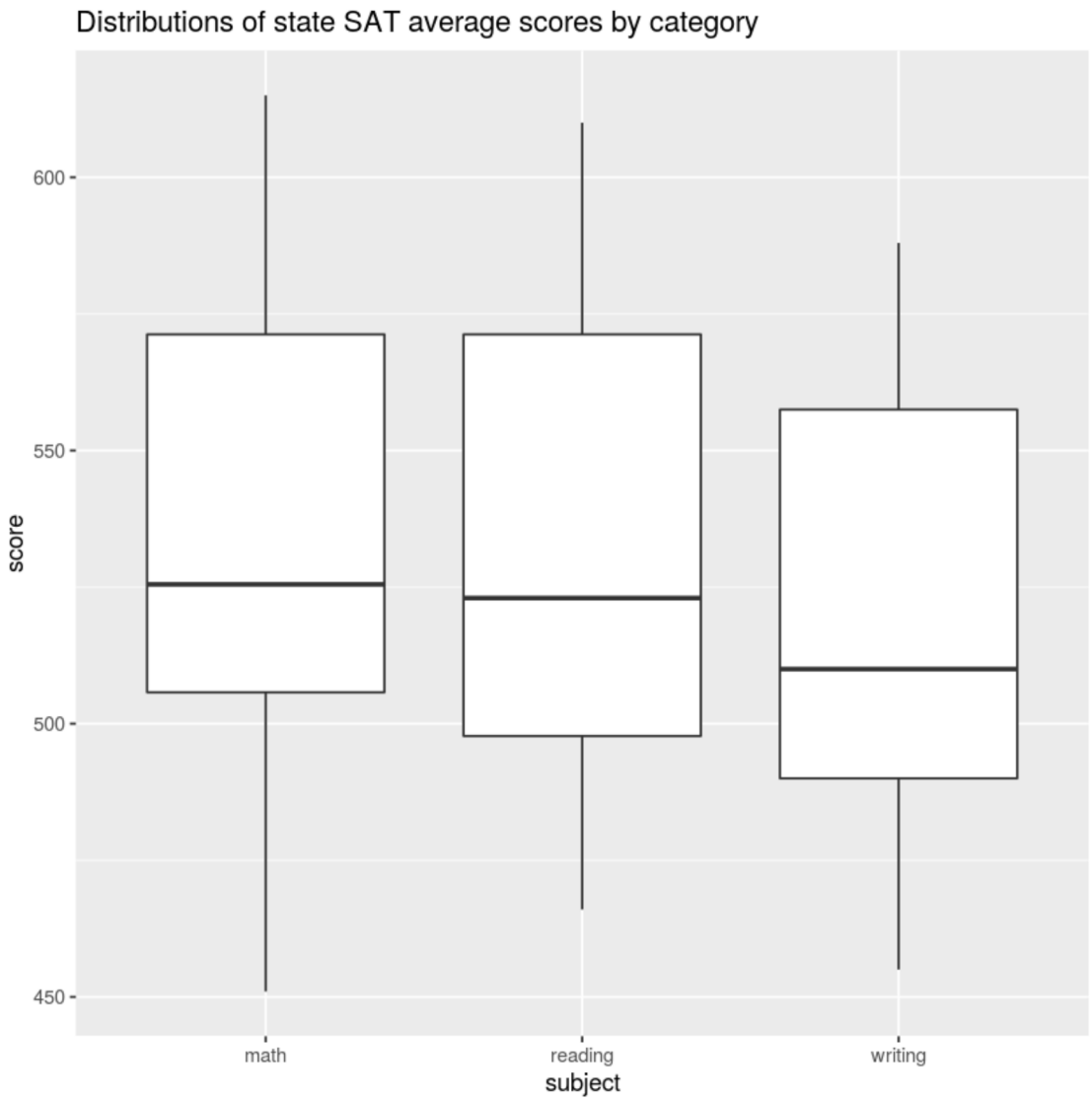
```
      bar = list(color = "black")))
fig <- fig %>%
  add_trace(
    type =  "indicator",
    mode = "gauge",
    value = 532,
    domain = list(x = c(0.1, 0.9), y = c(0.3, 0.4)),
    title = list(text = "South"),
    gauge = list(
      shape = "bullet",
      axis = list(range = list(NULL, 700)),
      threshold = list(
        line = list(color = "black", width = 2),
        thickness = 0.75,
        value = 533),
      bar = list(color = "black")))
fig <- fig %>%
  add_trace(
    type =  "indicator",
    mode = "gauge",
    value = 536,
    domain = list(x = c(0.1, 0.9), y = c(0.1, 0.2)),
    title = list(text = "West"),
    gauge = list(
      shape = "bullet",
      axis = list(range = list(NULL, 700)),
      threshold = list(
        line = list(color = "black", width = 2),
        thickness = 0.75,
        value = 539),
      bar = list(color = "black")))
fig
```




[40]:
```
# box plot

# load the SAT data, long form
df <- read_csv('education_mod.csv')

# generate the box plot
ggplot(df, aes(x=subject, y=score)) +
  geom_boxplot() +
  ggtitle('Distributions of state SAT average scores by category')
```

```
  Column specification
```

```
cols(
  state = col_character(),
  subject = col_character(),
  score = col_double()
)
```

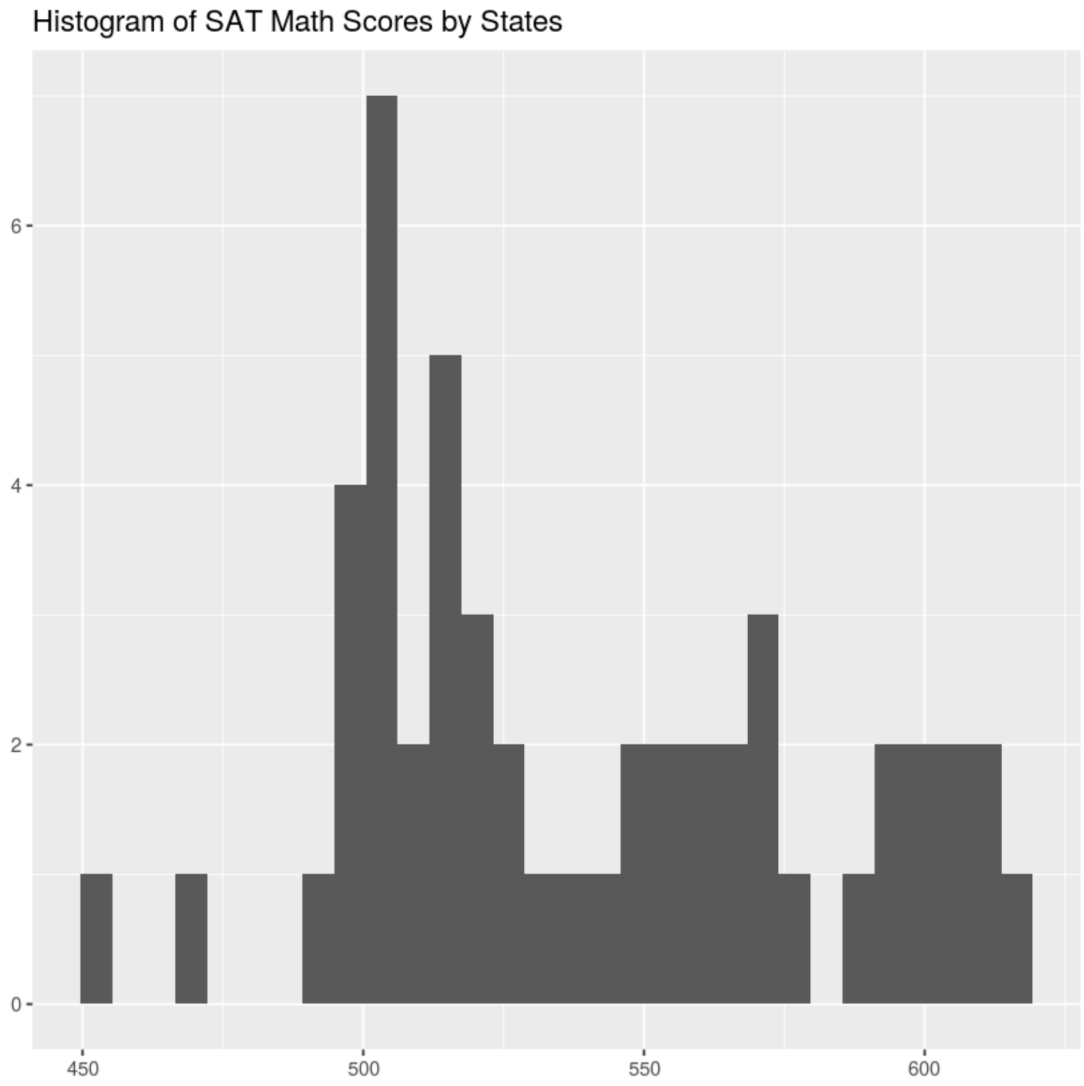


[39]:
```
# histogram

# load the SAT data, wide form
df <- read_csv('education.csv')
ggplot(df, aes(x=math)) +
  labs(title="Histogram of SAT Math Scores by States",x="", y = "") +
  geom_histogram()
```

```
  Column specification
```

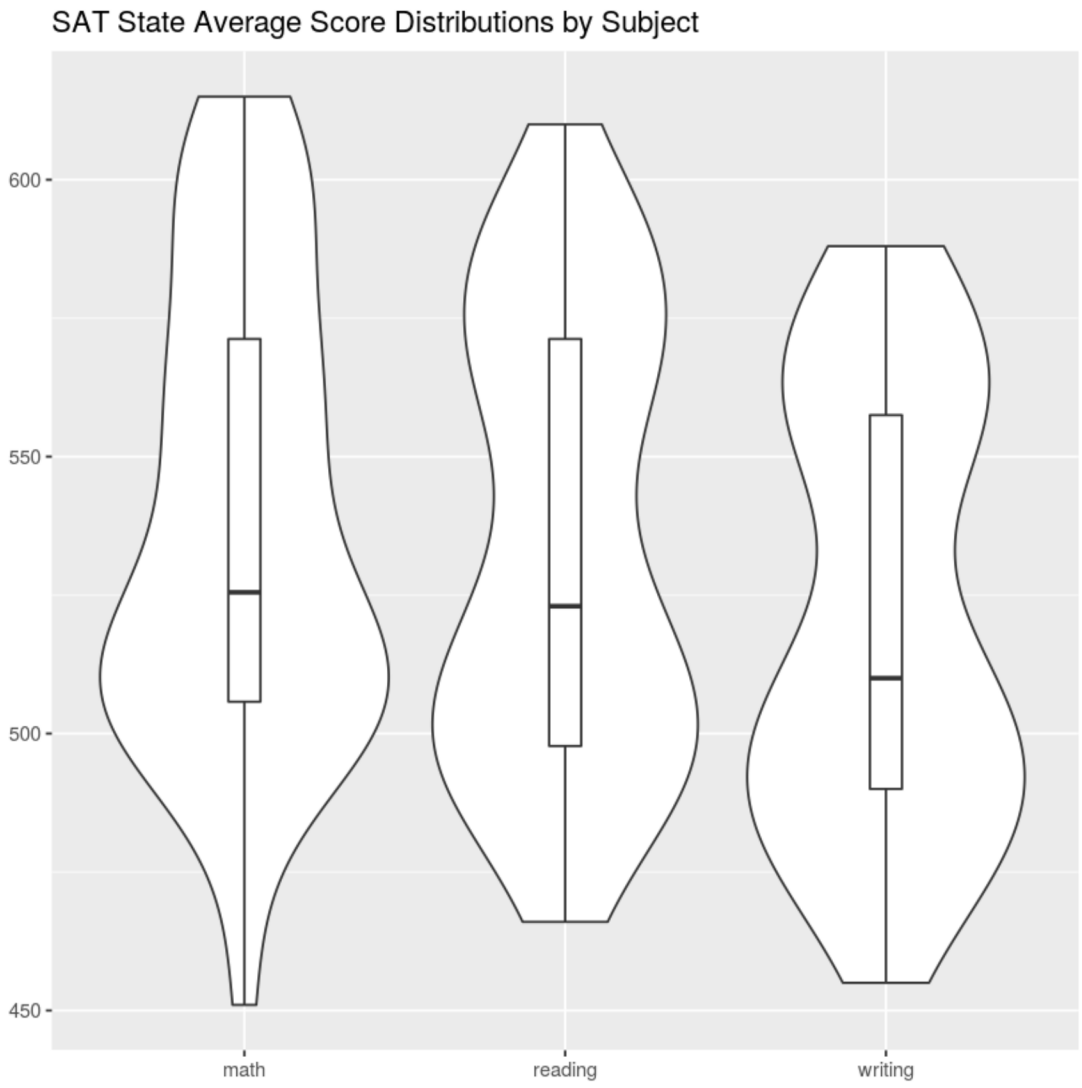
```
cols(
  state = col_character(),
  reading = col_double(),
  math = col_double(),
  writing = col_double(),
  percent_graduates_sat = col_double(),
  pupil_staff_ratio = col_double(),
  dropout_rate = col_double()
)
```

```
`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

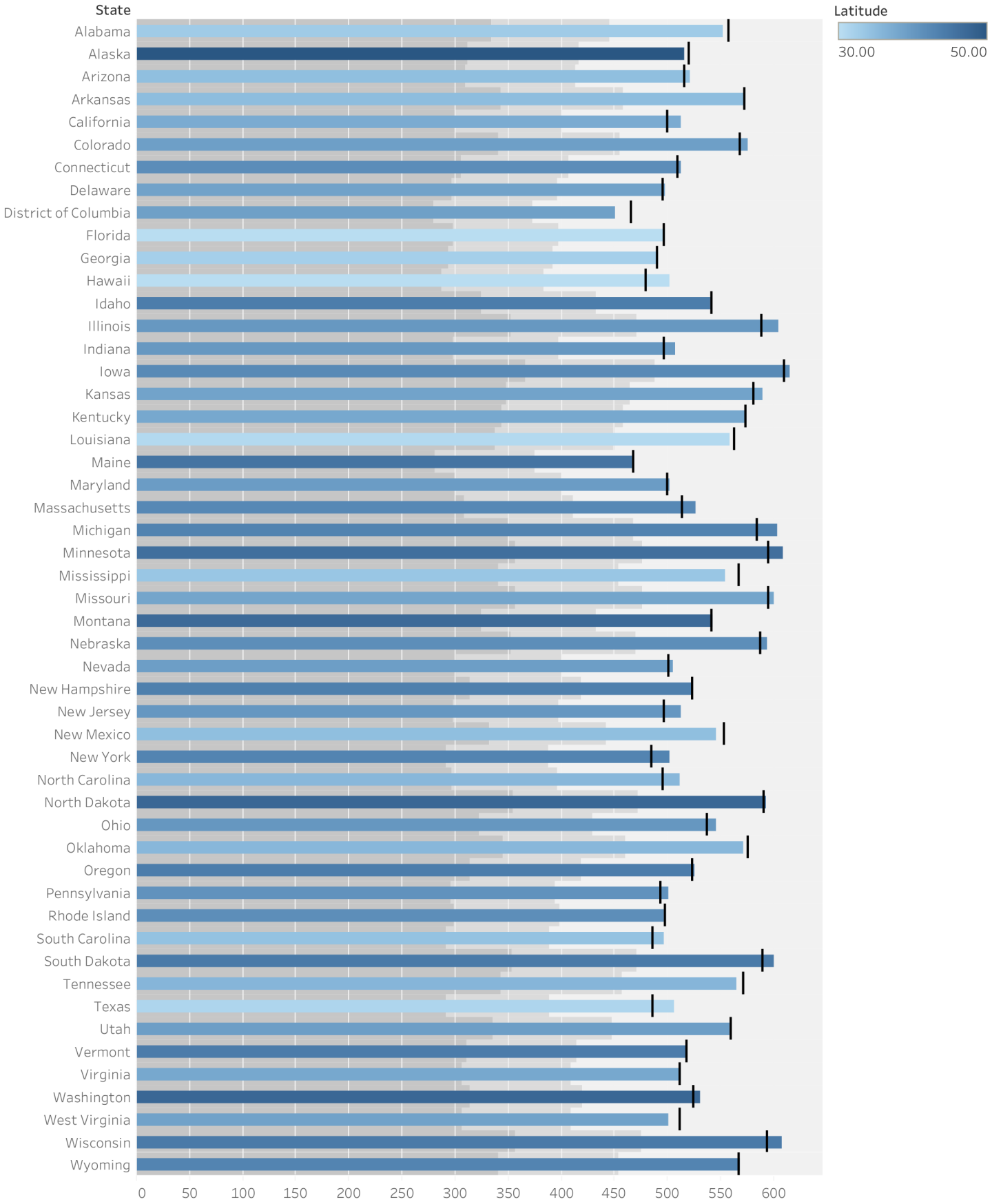



[46]:
```
# violin

# load the SAT data, long form
df <- read_csv('education_mod.csv')

# define the plot
ggplot(df, aes(x=subject, y=score)) + # set aesthetics
  geom_violin() + # set the plot geom
  geom_boxplot(width=0.1) + # add boxplot for additional information
  labs(title='SAT State Average Score Distributions by Subject',x="", y = "") #␣
 ↪set the plot title
```

```
  Column specification


cols(
  state = col_character(),
  subject = col_character(),
  score = col_double()
)
```

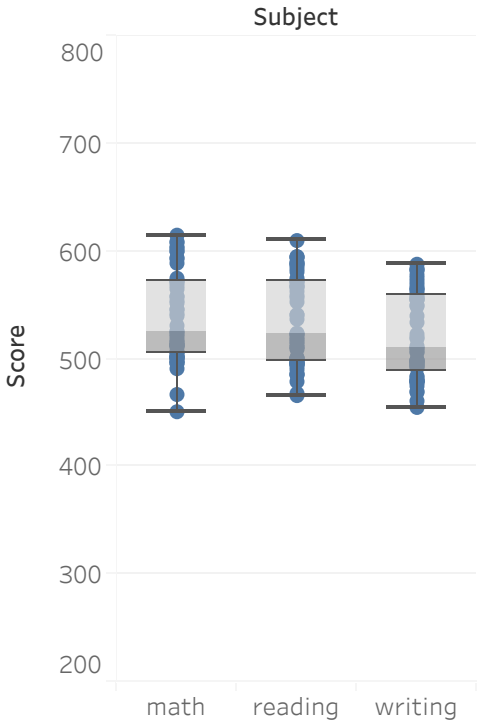


# SAT Reading and Math by State (and latitude)



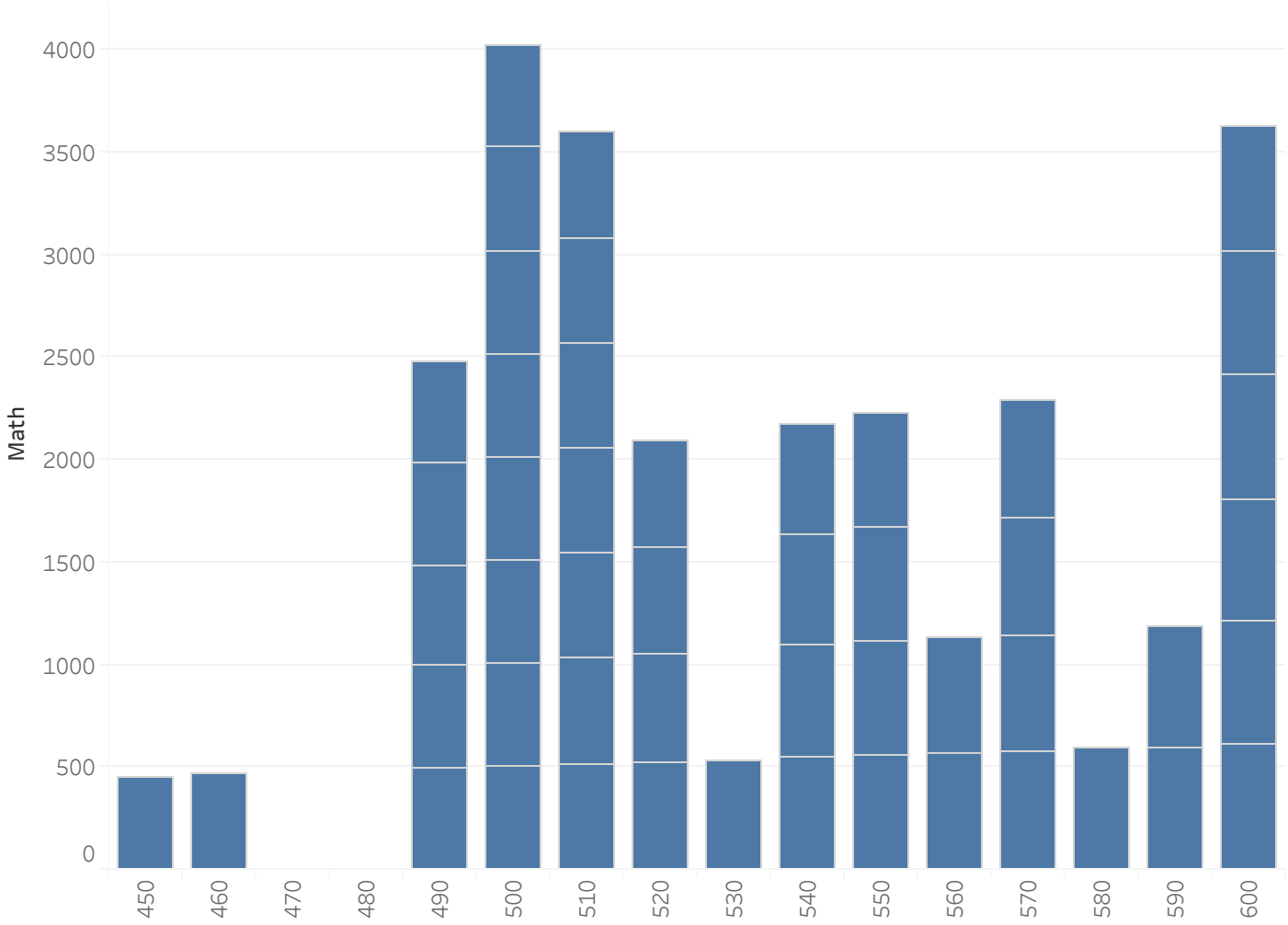

Sum of Math for each State. Color shows Latitude (generated). The view is filtered on State, which has multiple members selected.

# SAT Area Scores by State



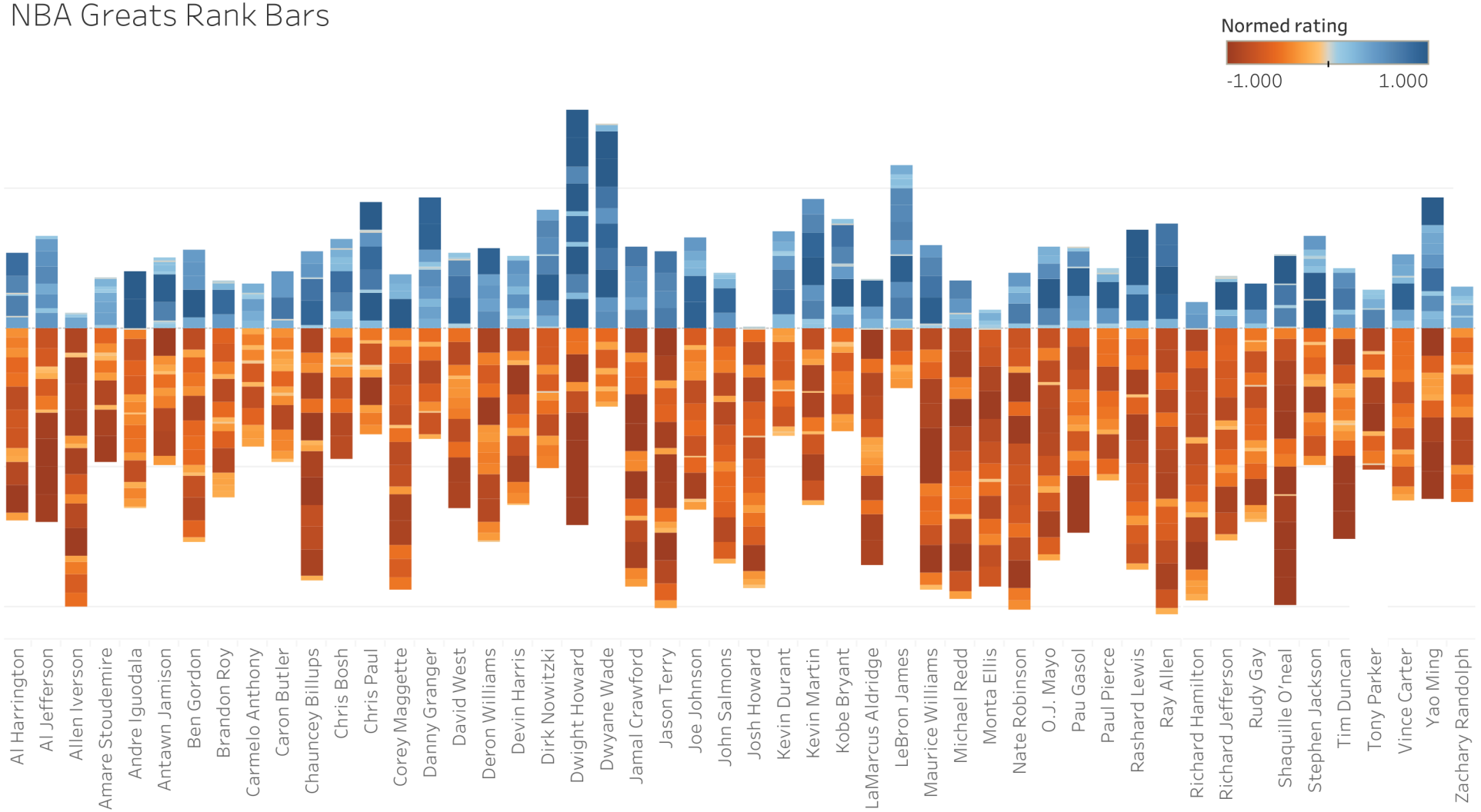

Score for each Subject.

# SAT Math Scores by State




Sum of Math for each Math (bin). Details are shown for State. The view is filtered on State, which has multiple members selected.

# NBA Greats Rank Bars




3PA, 3PM, 3PP, AST, BLK, DRB, FGA, FGM, FGP, FTA, FTM, FTP, G, MIN, ORB, PF, PTS, STL, TO and TRB for each Name. Color shows 3PA, 3PM, 3PP, AST, BLK, DRB, FGA, FGM, FGP, FTA, FTM, FTP, G, MIN, ORB, PF, PTS, STL, TO and TRB. Details are shown for 3PA, 3PM, 3PP, AST, BLK, DRB, FGA, FGM, FGP, FTA, FTM, FTP, G, MIN, ORB, PF, PTS, STL, TO and TRB. The view is filtered on Name, which has multiple members selected.