Math Scores for Different Teaching Styles

Heather Shaw

2022-04-28

Load Necessary Packages and Data

```
knitr::opts_chunk$set(warning = FALSE, message = FALSE)
knitr::opts_chunk$set(echo = TRUE)
```

```
library("tidyr")
library("dplyr")
library("foreign")
library("ggplot2")
#load data
math <- read.spss("1ResearchProjectData.sav", to.data.frame = TRUE)
```

Three teachers at a junior high school have different opinions about which teaching method is most effective for 8th grade math students. Ms. Wesson uses a traditional approach, while Ms. Ruger and Ms. Smith use a standards-based method. The first task is to determine which approach is more effective, based on the math scores of their current students.

In addition, it has been suggested that each teacher has strengths that will make them more effective with students in certain ethnic groups. One teacher believes that students should be divided into classes based on ethnicity. The second task is to determine if there is a difference between performance for each teacher based on student demographics.

The data includes math scores for students in grades 7-8. More information about this data project can be found here.

Cleaning the Data

Tidy Columns

```
head(math)
```

```
##   Student Teacher Gender          Ethnic   Freeredu Score       wesson
## 1       1   Ruger Female           Asian Free lunch    76 Ruger_Smith
## 2       2   Ruger Female        Hispanic Paid lunch    56 Ruger_Smith
## 3       3   Ruger Female African-American Free lunch    34 Ruger_Smith
## 4       4   Ruger Female           Asian Paid lunch    59 Ruger_Smith
## 5       5   Ruger   Male        Hispanic Free lunch    73 Ruger_Smith
## 6       6   Ruger   Male       Caucasian Paid lunch    58 Ruger_Smith
```

```
#rename columns
math <- math %>%
  rename(Method="wesson")
```

```
#change "Freeredu" column to "Lunch"
math <- math %>%
  rename(Lunch="Freeredu")
colnames(math)
```

```
## [1] "Student" "Teacher" "Gender"  "Ethnic"  "Lunch"   "Score"   "Method"
```

```
#drop word "lunch" from lunch status descriptions
math <- math %>%
  mutate(Lunch=gsub(' lunch','',Lunch))

 # change Method factors Ruger_Smith = standards, Wesson = traditional
math$Method <- sub("Ruger_Smith","Standards",math$Method)
math$Method <- sub("Wesson","Traditional",math$Method)

head(math)
```

```
##   Student Teacher Gender           Ethnic Lunch Score    Method
## 1       1   Ruger Female            Asian  Free    76 Standards
## 2       2   Ruger Female         Hispanic  Paid    56 Standards
## 3       3   Ruger Female African-American  Free    34 Standards
## 4       4   Ruger Female            Asian  Paid    59 Standards
## 5       5   Ruger   Male         Hispanic  Free    73 Standards
## 6       6   Ruger   Male         Caucasian Paid    58 Standards
```

Check for Duplicates and Omit Missing Values

```
#check for duplicates
math %>%
  duplicated() %>%
  table()
```

```
## .
## FALSE
##   217
```

```
#none found

#check for NA
math <- na.omit(math)
```

Analysis

Inspect Population

```
#total students
pop_total <- math %>%
  summarize(count=n())
pop_total
```

```
##   count
## 1   216
```

```
#student ethnicity
pop_ethnicity <- math %>%
```

```
  group_by(Ethnic) %>%
  summarize(count=n()) %>%
  mutate(Perc=(count/216)*100)
pop_ethnicity
```

```
## # A tibble: 4 x 3
##   Ethnic            count  Perc
##   <fct>             <int> <dbl>
## 1 Asian                53  24.5
## 2 African-American     52  24.1
## 3 Hispanic             66  30.6
## 4 Caucasian            45  20.8
```

```
#plot
pop_ethnicity_viz <- ggplot(pop_ethnicity,aes(x="",y = Perc, fill = Ethnic,)) +
  geom_col(color="black")+
  scale_fill_manual(values =c("tomato2","khaki3","skyblue4","olivedrab"))+
  geom_text(aes(label = round(Perc,2)),
            position = position_stack(vjust = 0.5)) +
  coord_polar(theta = "y")+
  labs(title="Percentages of Students By Ethnicity")+
  theme_void()

pop_ethnicity_viz
```

Percentages of Students By Ethnicity
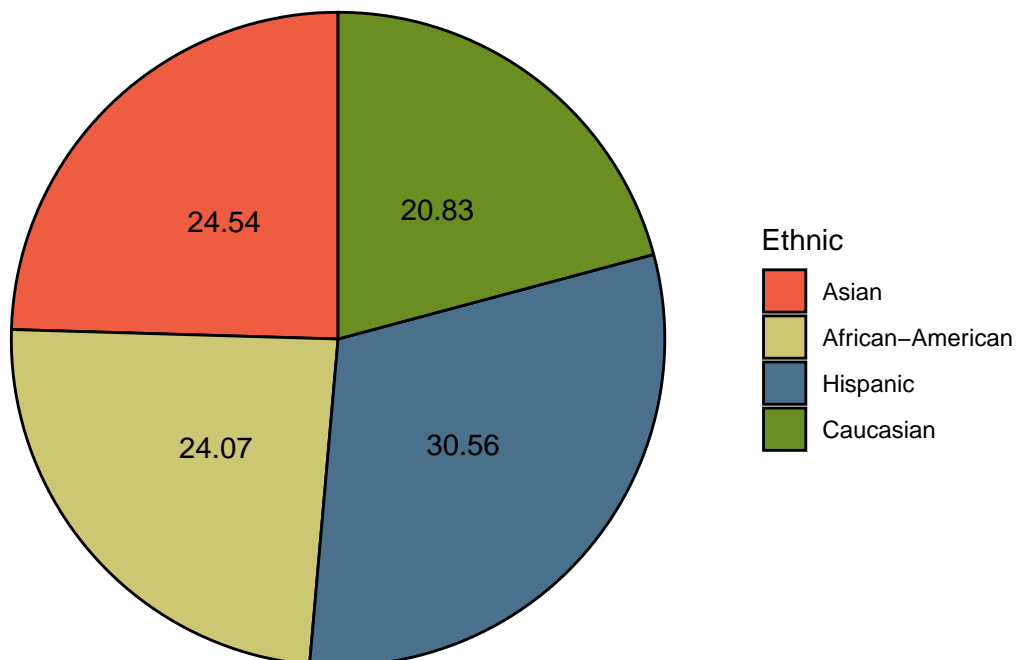

```
#Lunch Status
pop_lunch_status <- math %>%
  group_by(Lunch)%>%
  summarize(count=n())%>%
  mutate(Perc=(count/216)*100)
```

3

```
pop_lunch_status
```

```
## # A tibble: 2 x 3
##   Lunch count  Perc
##   <chr> <int> <dbl>
## 1 Free    104  48.1
## 2 Paid    112  51.9
```

```
#plot
pop_lunch_viz <- ggplot(pop_lunch_status,aes(x="",y = Perc, fill = Lunch,)) +
  geom_col(color="black")+
  scale_fill_manual(values =c("tomato2","skyblue4"))+
  geom_text(aes(label = round(Perc,2)),
            position = position_stack(vjust = 0.5)) +
  coord_polar(theta = "y")+
  labs(title="Percentages of Students By Lunch Status")+
  theme_void()
pop_lunch_viz
```

## Percentages of Students By Lunch Status
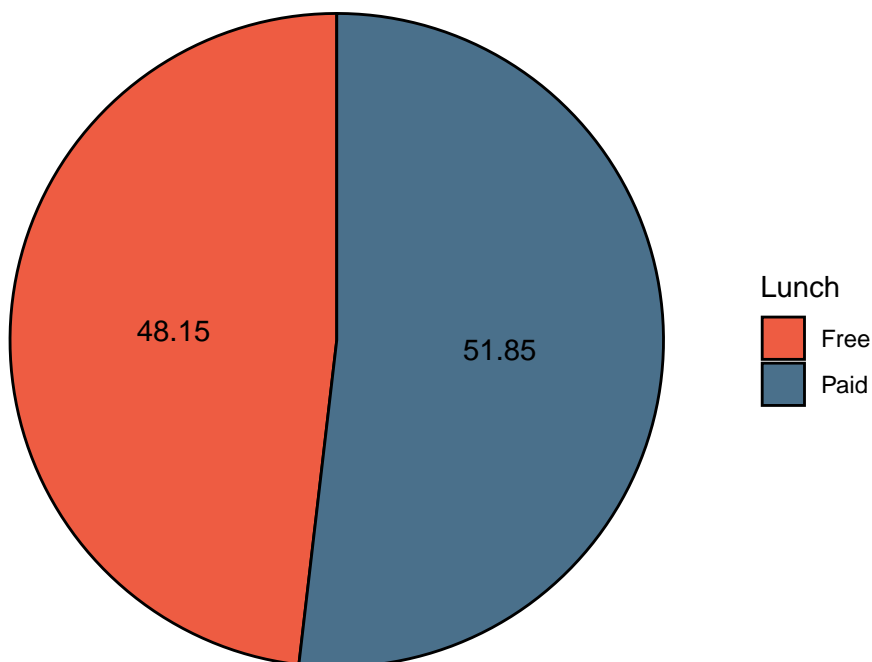
Comparing Methods

Initial comparison of average test scores for each method show higher scores for the traditional method (70.7), compared to the standards-based method (62.3).

```
#avg score by Method
avg_score_method <- math %>%
  group_by(Method) %>%
  summarize(Avg_Score=mean(Score,na.rm=TRUE))
avg_score_method
```

```
## # A tibble: 2 x 2
##   Method     Avg_Score
##   <chr>          <dbl>
## 1 Standards       62.3
## 2 Traditional     70.7
```

```r
#plot
avg_method_viz <- ggplot(data=avg_score_method,aes(x=Method,y=Avg_Score,fill=Method))+
  geom_bar(stat="identity")+
  geom_text(aes(label=round(Avg_Score,2)), vjust=1.25)+
  scale_fill_manual(values=c("tomato2","skyblue4"))+
  labs(title="Average Math Score by Method")
avg_method_viz
```

Average Math Score by Method
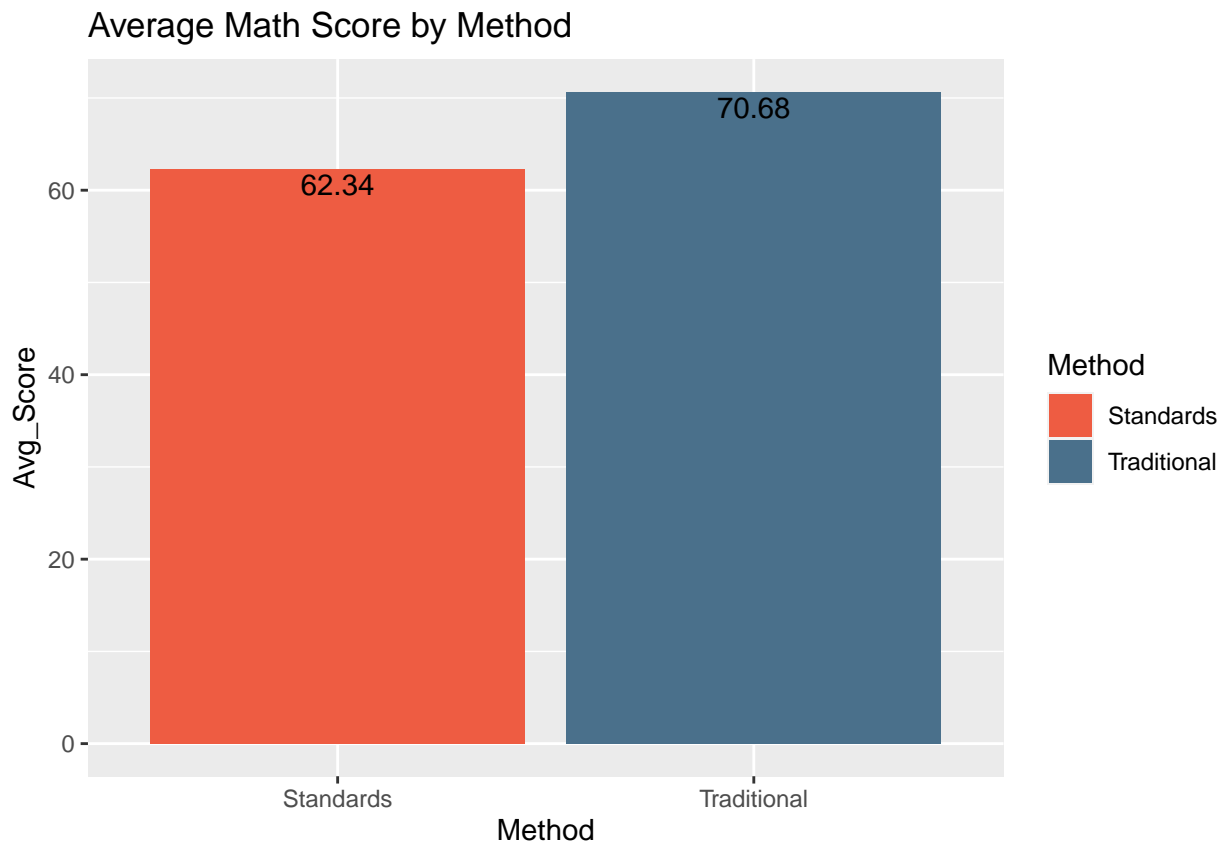
Evaluating the scores by method shows higher scores, with little difference between groups of students based on ethnicity or lunch status.

```r
#avg score by Method and ethnicity
avg_score_method_ethnicity <- math %>%
  group_by(Method,Ethnic) %>%
  summarize(Avg_Score=mean(Score,na.rm=TRUE))
avg_score_method_ethnicity
```

```
## # A tibble: 8 x 3
## # Groups:   Method [2]
##   Method     Ethnic             Avg_Score
##   <chr>      <fct>                  <dbl>
## 1 Standards  Asian                   62.5
## 2 Standards  African-American        62.7
```

```
## 3 Standards   Hispanic            61.6
## 4 Standards   Caucasian           62.9
## 5 Traditional Asian               73.2
## 6 Traditional African-American    72.9
## 7 Traditional Hispanic            70.8
## 8 Traditional Caucasian           66.3
```

```r
#plot
avg_method_ethnicity_viz <- ggplot(data=avg_score_method_ethnicity,aes(x=Method,y=Avg_Score,fill=Method
  geom_bar(stat="identity")+
  geom_text(aes(label=round(Avg_Score,2)), vjust=1.25)+
  facet_wrap(~Ethnic)+
  scale_fill_manual(values=c("tomato2","skyblue4"))+
  labs(title="Average Math Score by Method and Ethnicity")
avg_method_ethnicity_viz
```

Average Math Score by Method and Ethnicity
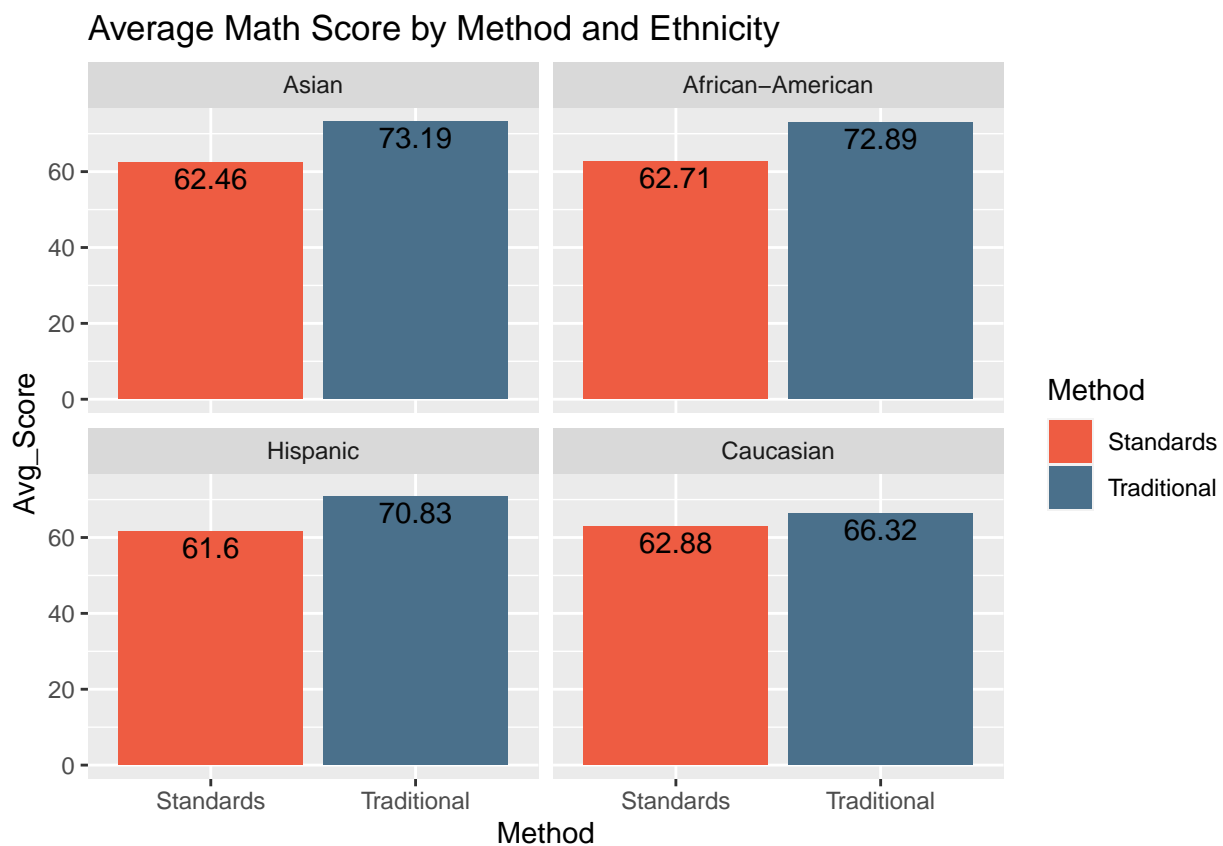

```r
#avg score by Method and lunch status
avg_score_method_lunch <- math %>%
  group_by(Method,Lunch) %>%
  summarize(Avg_Score=mean(Score,na.rm=TRUE))
avg_score_method_lunch
```

```
## # A tibble: 4 x 3
## # Groups:   Method [2]
##   Method      Lunch Avg_Score
##   <chr>       <chr>     <dbl>
## 1 Standards   Free       62.8
## 2 Standards   Paid       62
```

```
## 3 Traditional Free       70.9
## 4 Traditional Paid       70.5
```
```
#plot
avg_method_lunch_viz <- ggplot(data=avg_score_method_lunch,aes(x=Method,y=Avg_Score,fill=Method))+
  geom_bar(stat="identity")+
  geom_text(aes(label=round(Avg_Score,2)), vjust=1.25)+
  facet_wrap(~Lunch)+
  scale_fill_manual(values=c("tomato2","skyblue4"))+
  labs(title="Average Math Score by Method and Lunch Status")
avg_method_lunch_viz
```

Average Math Score by Method and Lunch Status
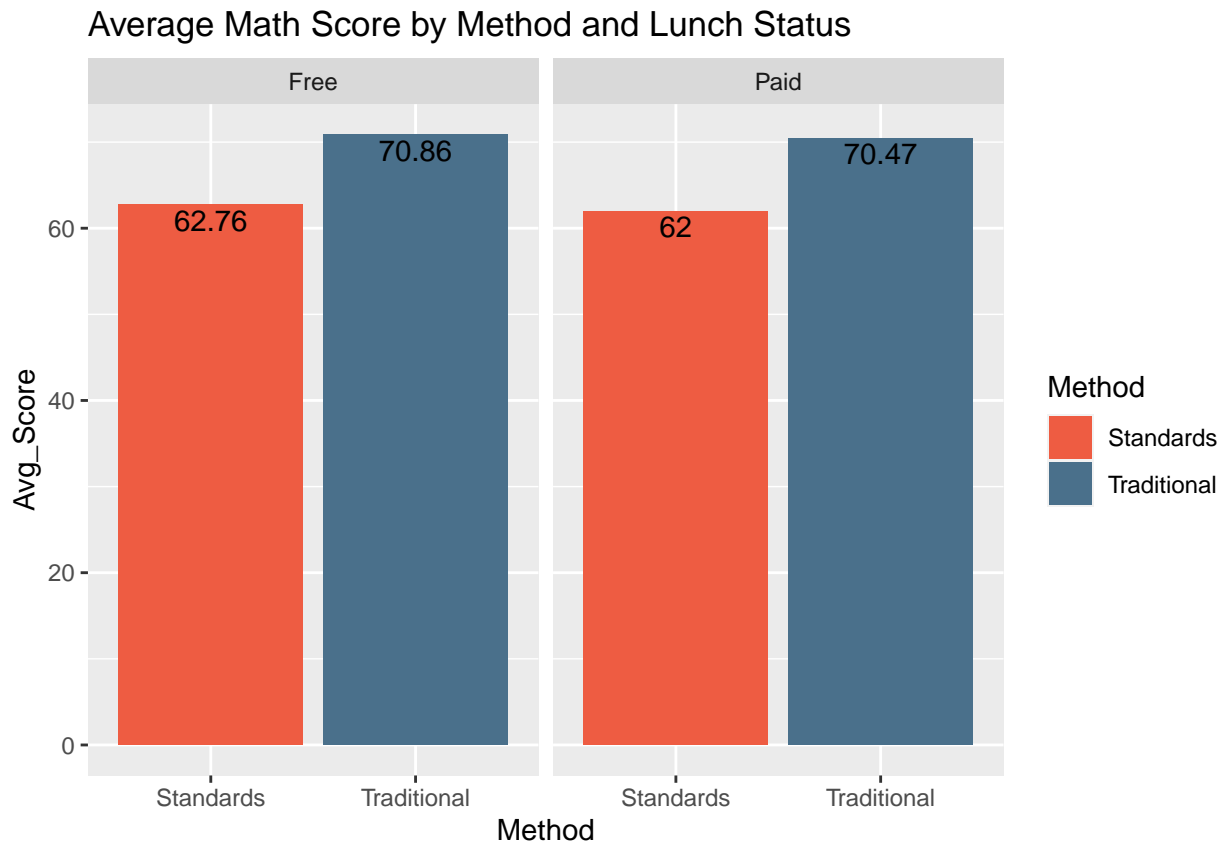
Comparing Teachers

We should also compare student scores based on their math teacher.

Ms. Wesson and Ms. Smith have students with average scores that are very similar (70.7 and 69.6, respectively). Ms. Ruger's students average score is 55.26.

```
#avg score by teacher
avg_teacher_score <- math %>%
    group_by(Teacher) %>%
    summarize(Avg_Score=mean(Score,na.rm=TRUE))
avg_teacher_score
```
```
## # A tibble: 3 x 2
##   Teacher Avg_Score
##   <fct>       <dbl>
## 1 Ruger        55.3
```

```
## 2 Smith          69.6
## 3 Wesson         70.7
```

```r
#plot
avg_teacher_viz <- ggplot(data=avg_teacher_score,aes(x=Teacher,y=Avg_Score,fill=Teacher))+
  geom_bar(stat="identity")+
  geom_text(aes(label=round(Avg_Score,2)), vjust=1.25)+
  scale_fill_manual(values=c("tomato2","khaki3","skyblue4"))+
  labs(title="Average Math Score by Teacher")
avg_teacher_viz
```
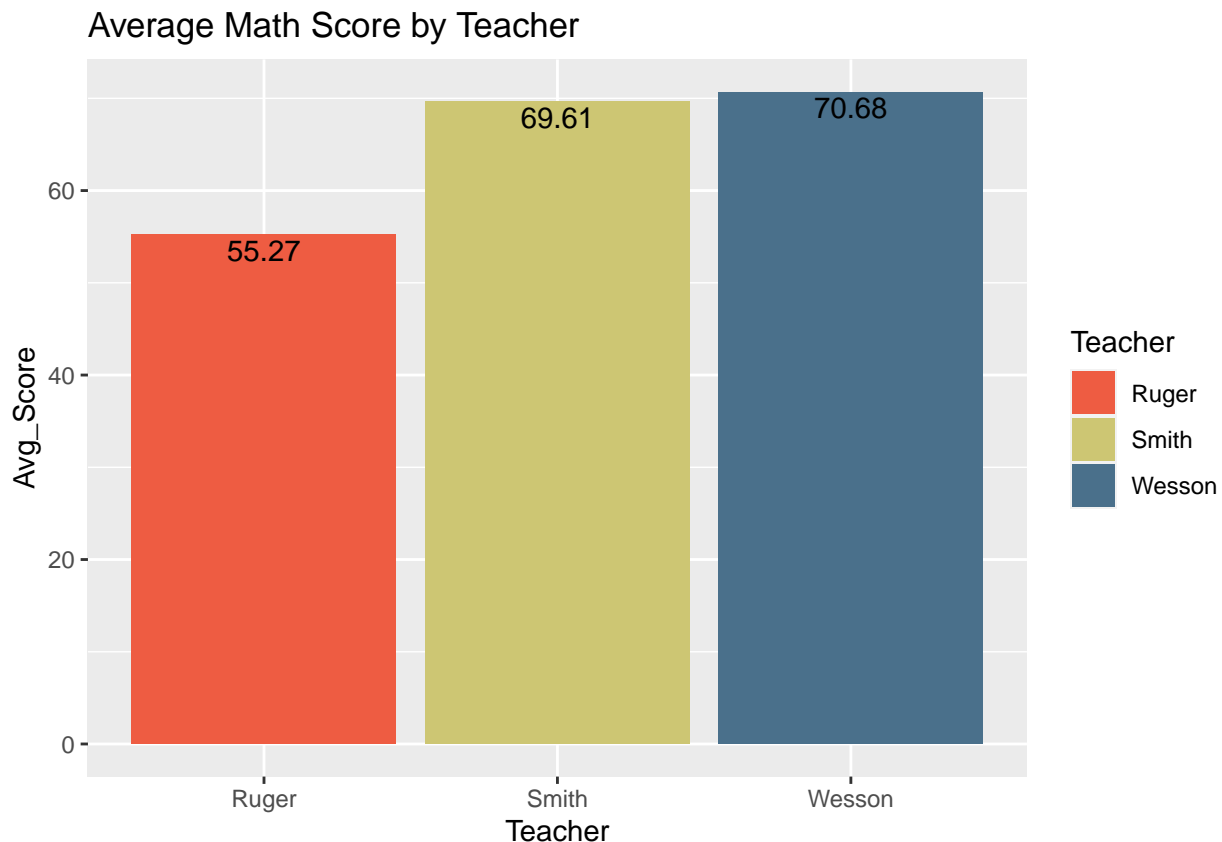

Next, we will examine how each teacher's students perform when grouped by ethnicity and lunch status.

There is little differences in overall performance between ethnic groups. While there are minor differences between average student scores of Ms. Smith and Ms. Wesson when grouped according to ethnicity, average student scores in Ms. Ruger's class were over 10-16 points lower than the other two student groups for all ethnic groups.
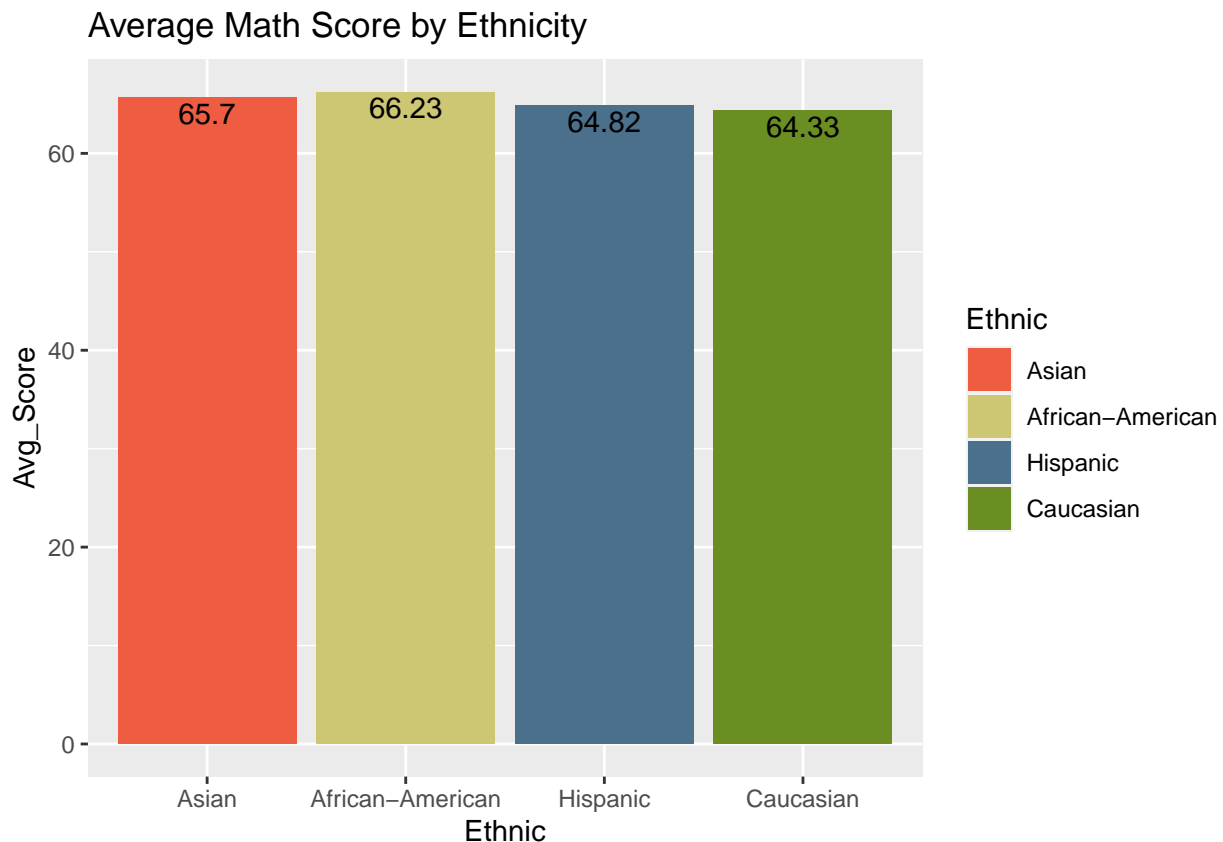
```r
#avg score by ethnicity
avg_score_ethnicity <- math %>%
    group_by(Ethnic) %>%
    summarize(Avg_Score=mean(Score,na.rm=TRUE))
avg_score_ethnicity
```

```
## # A tibble: 4 x 2
##   Ethnic            Avg_Score
##   <fct>                 <dbl>
## 1 Asian                  65.7
## 2 African-American       66.2
```

```
## 3 Hispanic              64.8
## 4 Caucasian             64.3
```

```r
#plot
avg_ethnicity_viz <- ggplot(data=avg_score_ethnicity,aes(x=Ethnic,y=Avg_Score,fill=Ethnic))+
  geom_bar(stat="identity")+
  geom_text(aes(label=round(Avg_Score,2)), vjust=1.25)+
  scale_fill_manual(values=c("tomato2","khaki3","skyblue4","olivedrab"))+
  labs(title="Average Math Score by Ethnicity")
avg_ethnicity_viz
```


Average Math Score by Ethnicity

```r
#avg score by teacher for each ethnicity
avg_teacher_score_ethnicity <- math %>%
  group_by(Ethnic,Teacher) %>%
  summarize(Avg_Score=mean(Score,na.rm=TRUE))
avg_teacher_score_ethnicity
```

```
## # A tibble: 12 x 3
## # Groups:   Ethnic [4]
##    Ethnic           Teacher Avg_Score
##    <fct>            <fct>       <dbl>
##  1 Asian            Ruger        53.7
##  2 Asian            Smith        69.9
##  3 Asian            Wesson       73.2
##  4 African-American Ruger        55.2
##  5 African-American Smith        71.1
##  6 African-American Wesson       72.9
##  7 Hispanic         Ruger        55.6
```

```
##  8 Hispanic         Smith         69.9
##  9 Hispanic         Wesson        70.8
## 10 Caucasian        Ruger         56.9
## 11 Caucasian        Smith         67.3
## 12 Caucasian        Wesson        66.3
```

```r
#plot
avg_teacher_ethnic_viz <- ggplot(data=avg_teacher_score_ethnicity,aes(x=Teacher,y=Avg_Score,fill=Teacher
  geom_bar(stat="identity")+
   geom_text(aes(label=round(Avg_Score,2)), vjust=1.25)+
  facet_wrap(~Ethnic)+
  scale_fill_manual(values=c("tomato2","khaki3","skyblue4"))+
  labs(title="Average Math Score by Teacher and Ethnicity")
avg_teacher_ethnic_viz
```


Average Math Score by Teacher and Ethnicity
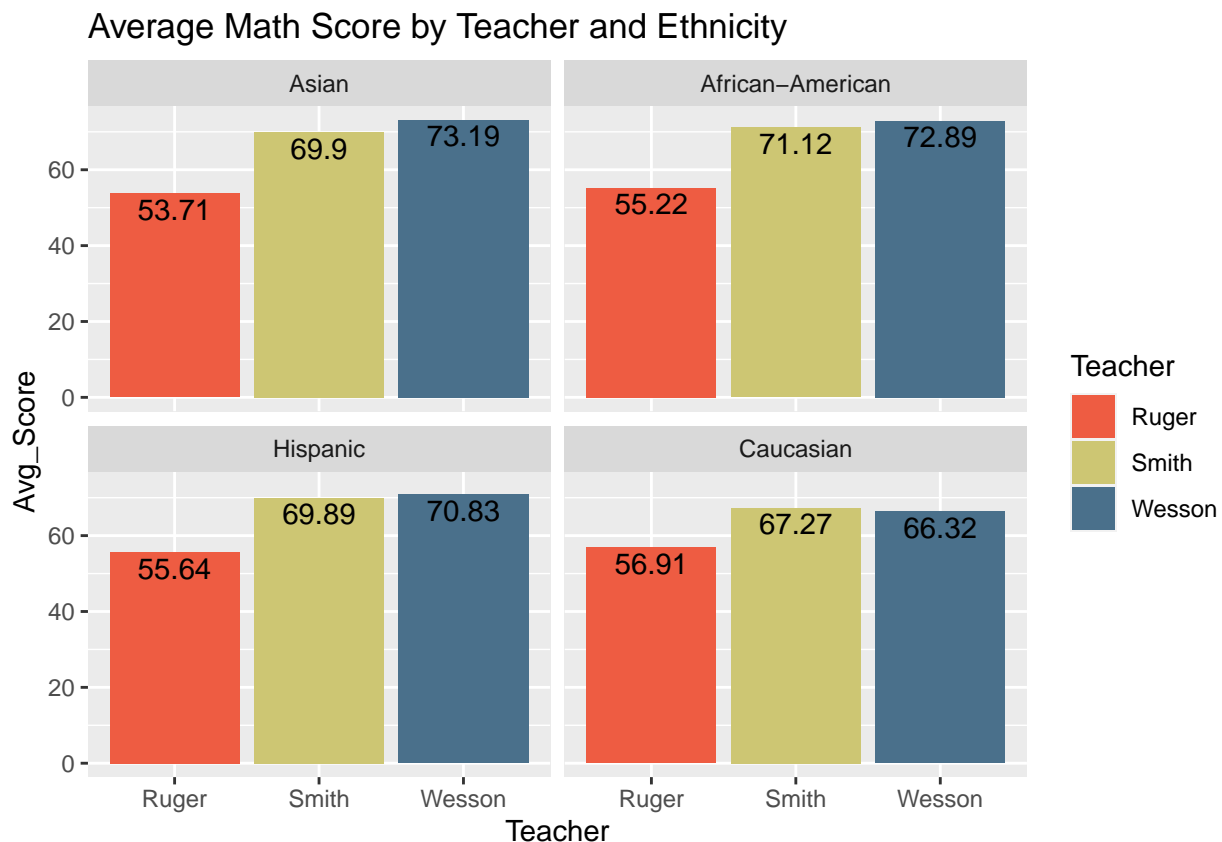
Again we find there is no large difference in performance between all students grouped by lunch status, but there remains a visible gap between the performance of students in both groups when comparing teachers.

```r
#avg score by lunch status
avg_score_lunch <- math %>%
  group_by(Lunch)%>%
  summarize(Avg_Score=mean(Score,na.rm=TRUE))
avg_score_lunch
```

```
## # A tibble: 2 x 2
##   Lunch Avg_Score
##   <chr>     <dbl>
## 1 Free       66.0
## 2 Paid       64.6
```

```
#plot
avg_lunch_viz <- ggplot(data=avg_score_lunch, aes(x=Lunch,y=Avg_Score,fill=Lunch))+
    geom_bar(stat="identity")+
    geom_text(aes(label=round(Avg_Score,2)), vjust=1.25)+
    scale_fill_manual(values=c("tomato2","skyblue4"))+
    labs(title="Average Math Score by Lunch Status")
avg_lunch_viz
```

## Average Math Score by Lunch Status
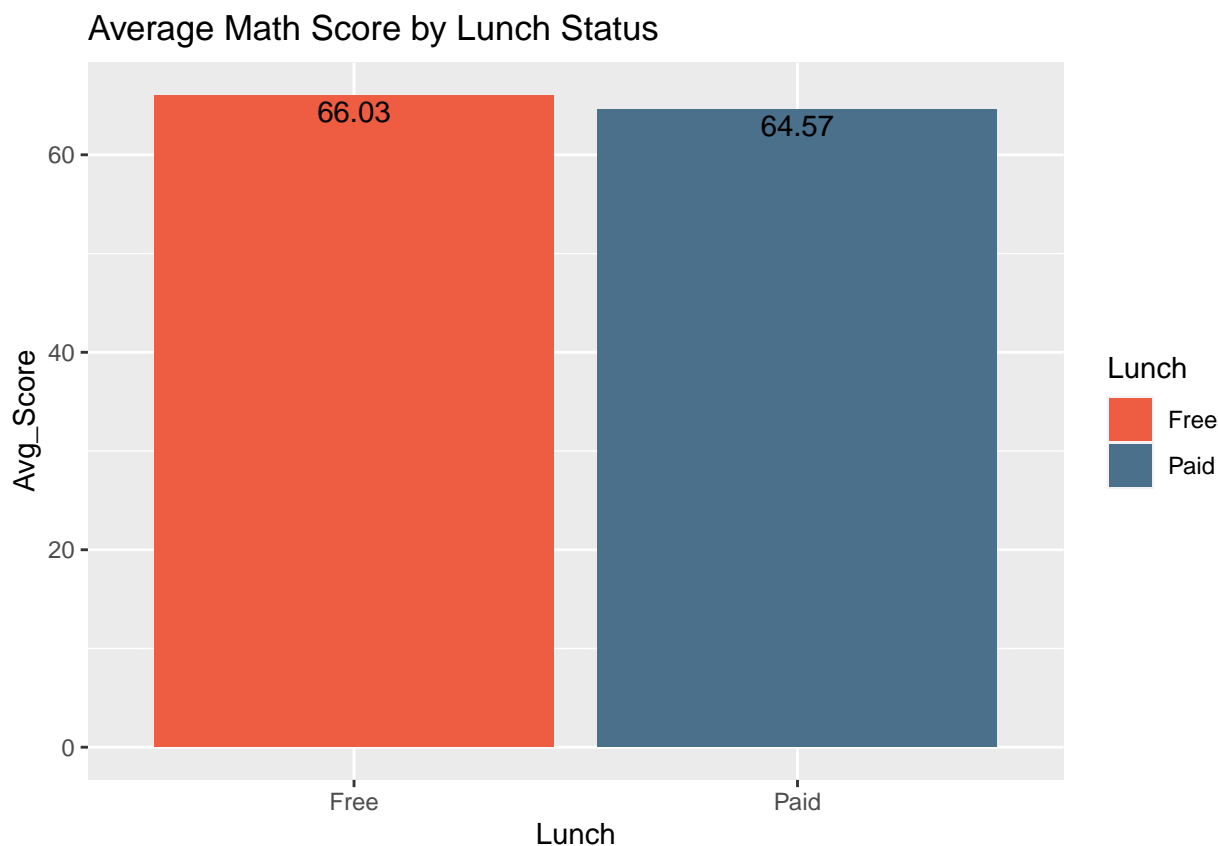

```
#avg score by teacher for lunch status
avg_teacher_score_lunch <- math %>%
  group_by(Teacher,Lunch) %>%
  summarize(Avg_Score=mean(Score,na.rm=TRUE))
avg_teacher_score_lunch
```

```
## # A tibble: 6 x 3
## # Groups:   Teacher [3]
##   Teacher Lunch Avg_Score
##   <fct>   <chr>     <dbl>
## 1 Ruger   Free       56.2
## 2 Ruger   Paid       54.4
## 3 Smith   Free       71.3
## 4 Smith   Paid       68.5
## 5 Wesson  Free       70.9
## 6 Wesson  Paid       70.5
```

```
#plot
avg_teacher_lunch_viz <- ggplot(data=avg_teacher_score_lunch,aes(x=Teacher,y=Avg_Score,fill=Teacher))+
  geom_bar(stat="identity")+
```

```
  geom_text(aes(label=round(Avg_Score,2)), vjust=1.25)+
  facet_wrap(~Lunch)+
  scale_fill_manual(values=c("tomato2","khaki3","skyblue4"))+
  labs(title="Average Math Score by Teacher and Lunch Status")
avg_teacher_lunch_viz
```


Average Math Score by Teacher and Lunch Status
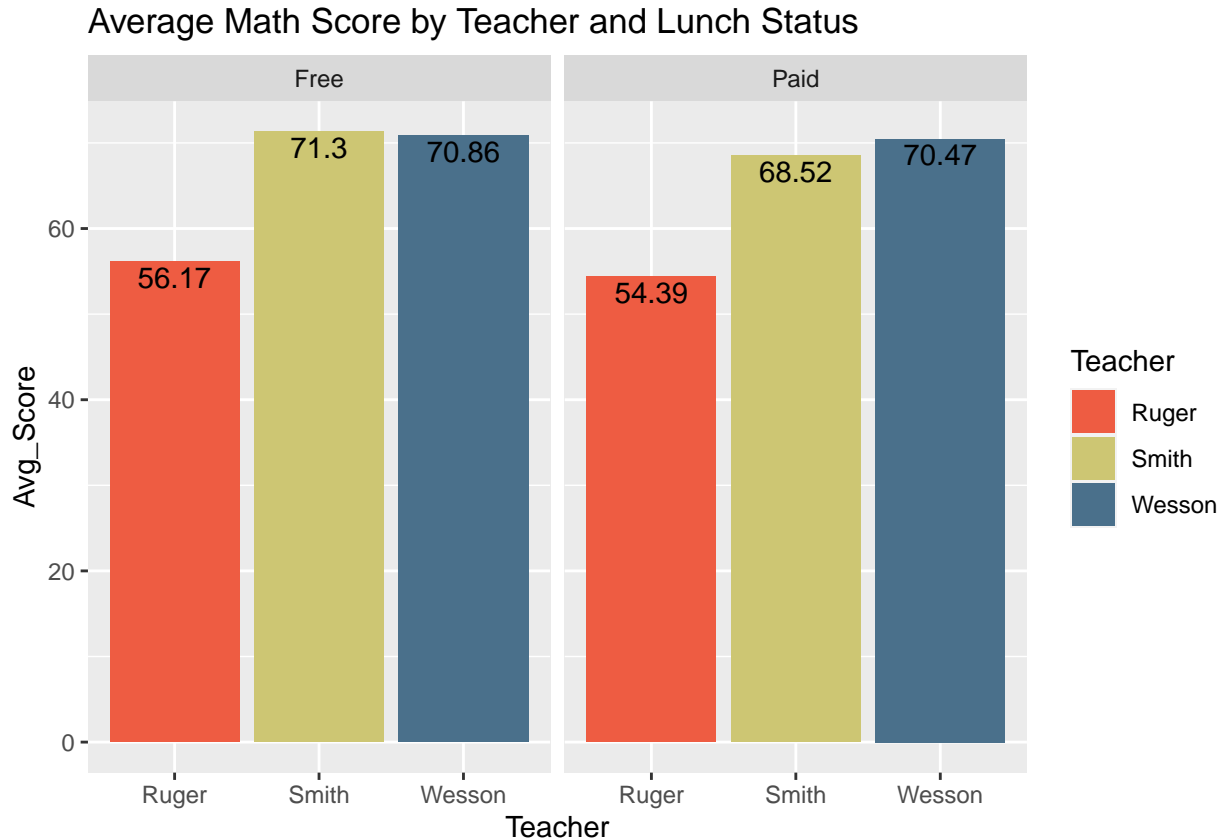
Conclusions

When comparing traditional and standards-based methods among all students, the traditional method seems to result in higher scores. However, the difference between the scores of students taught by Ms. Wesson and Ms.Smith is less than the difference between those two classes and Ms. Ruger's students.

Ms. Wesson and Ms. Ruger both use the standards-based method, and Ms. Smith uses the traditional method. When ranking the performance of students in each class, the standards-based method ranks first and last, while the traditional method is a close second place. Considering the discrepancy between the two standards-based classrooms, we cannot definitively conclude that the teaching method is the determining factor affecting student performance.

The data does not show any benefit for students being assigned to a specific teacher according to their ethnicity or lunch status that would outweigh the ethical issues that such a grouping would create. It does show that Ms. Ruger's students do not perform as well as those taught by Ms. Smith or Ms. Wesson, even when taking ethnicity and lunch status into account.

Based on this analysis, the data does not support grouping students by ethnicity or mandating one teaching method over another. The school administration should consider evaluating and offering professional development support for Ms. Ruger. The differences between student scores in each class can be seen on a graph, but may not be statistically significant. There could be other factors causing an appearance of low

performance. Hypothesis testing needs to be done to determine the level of significance for these results before any other steps are taken.