

Funding and Spending of U.S. Schools

Shankar Prabhu

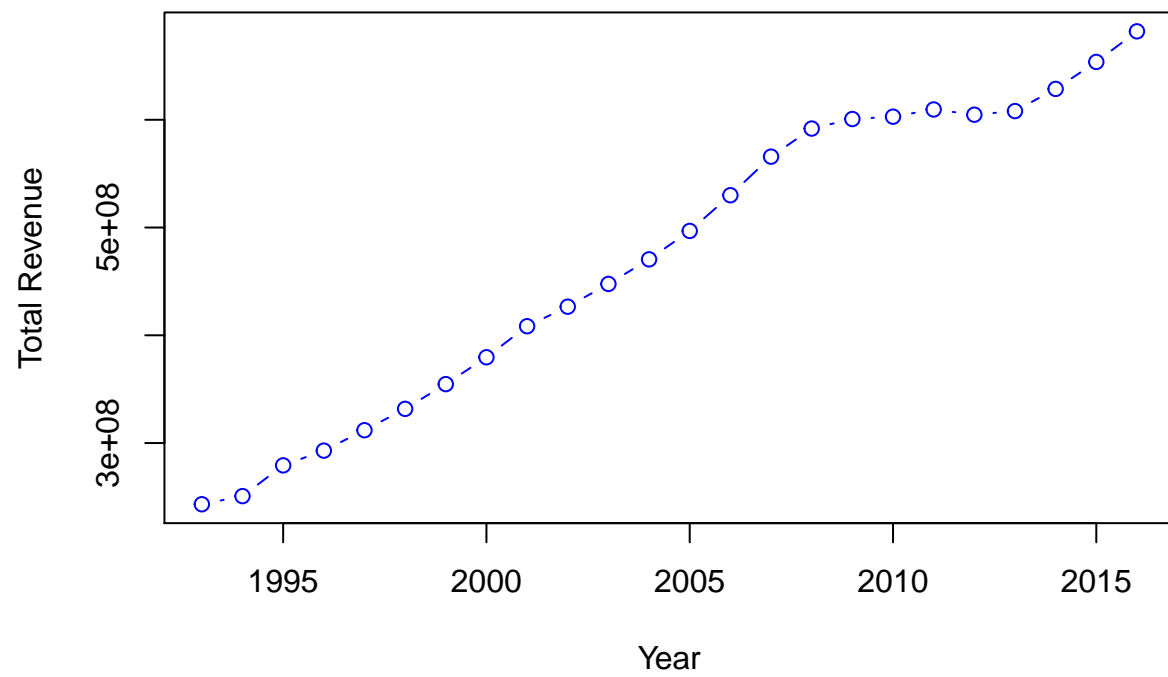
July 2020

In this section, we will take a look at how the financial situations of U.S schools have changed in the past 30 years.

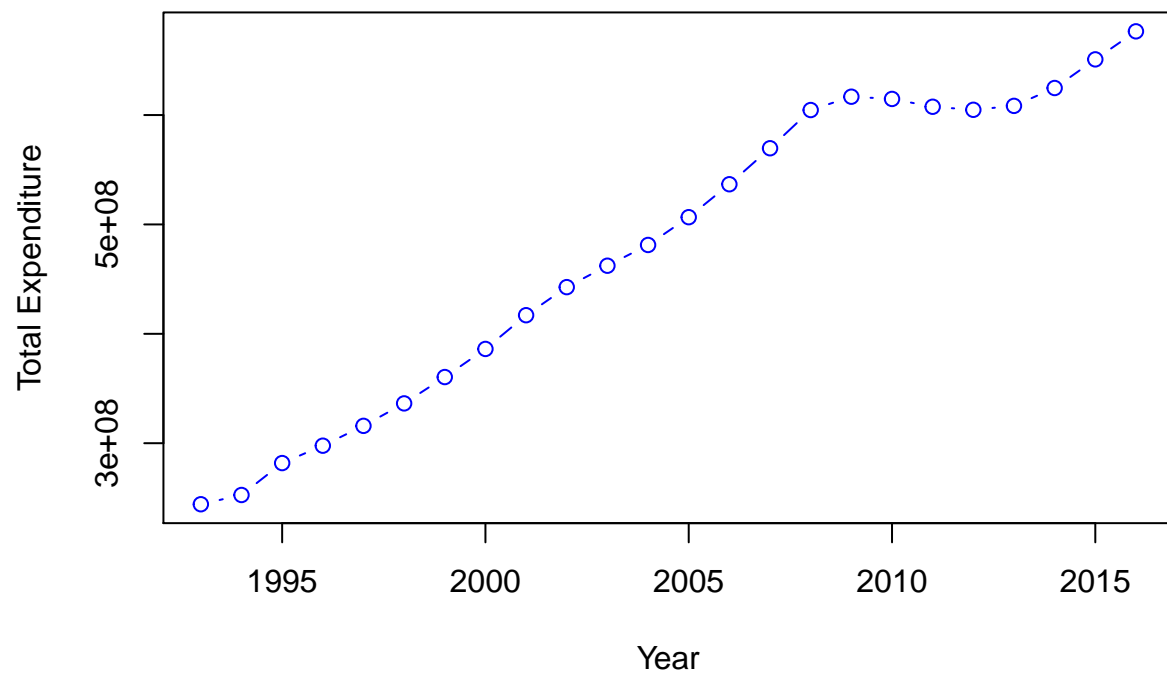
```
data = read.csv(
  "C:/Users/shank/Documents/Data_Science_Projects/education/states_all_extended.csv",
  header = TRUE, sep = ",")
financials_by_year = setNames(aggregate(cbind(data$TOTAL_REVENUE, data$FEDERAL_REVENUE,
  data$STATE_REVENUE, data$LOCAL_REVENUE,
  data$TOTAL_EXPENDITURE,
  data$INSTRUCTION_EXPENDITURE,
  data$SUPPORT_SERVICES_EXPENDITURE,
  data$OTHER_EXPENDITURE,
  data$CAPITAL_OUTLAY_EXPENDITURE, data$ENROLL
), by=list(YEAR=data$YEAR),
  FUN = sum, na.rm = TRUE), c("YEAR",
  "TOTAL_REVENUE", "FEDERAL_REVENUE",
  "STATE_REVENUE", "LOCAL_REVENUE",
  "TOTAL_EXPENDITURE", "INSTRUCTION_EXPENDITURE",
  "SUPPORT_SERVICES_EXPENDITURE", "OTHER_EXPENDITURE",
  "CAPITAL_OUTLAY_EXPENDITURE", "ENROLL"))
```

Let's take a look at overall school finances

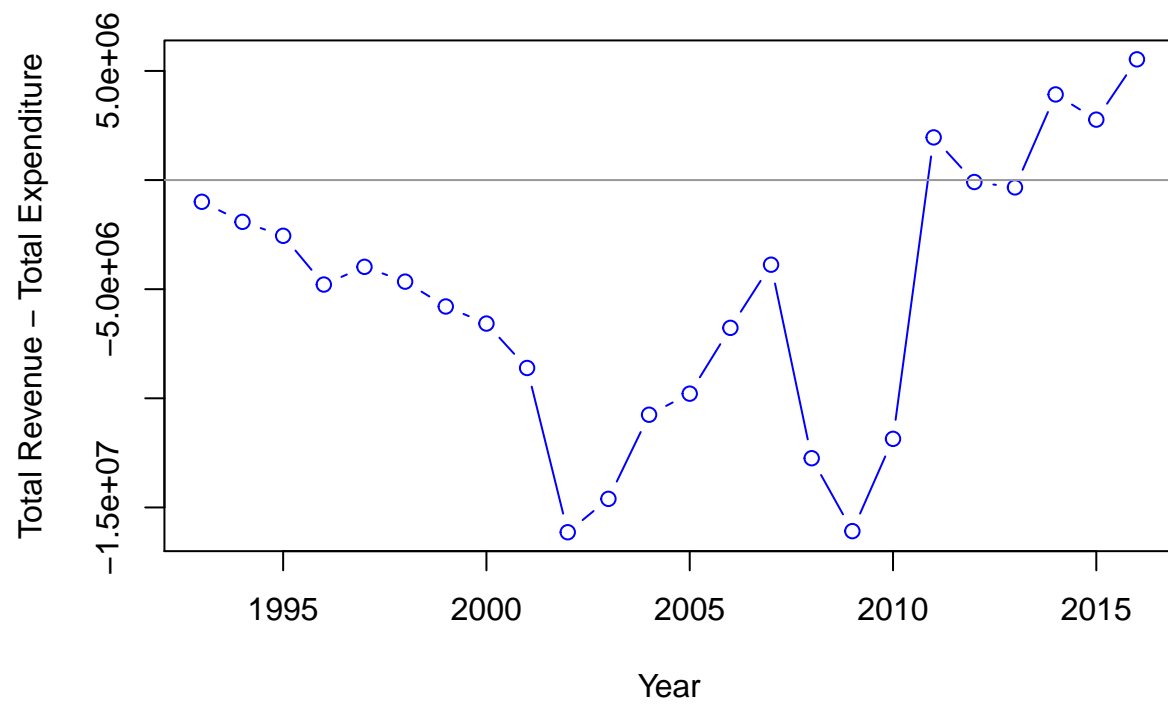
```
financials_by_year = financials_by_year[-c(1, 2, 3, 4, 5, 6, 7, 32, 33),]
plot(financials_by_year$TOTAL_REVENUE ~ financials_by_year$YEAR, type = 'b',
  col = "blue", xlab = "Year", ylab = "Total Revenue")
```



```
plot(financials_by_year$TOTAL_EXPENDITURE ~ financials_by_year$YEAR, type = 'b',  
     col = "blue", xlab = "Year", ylab = "Total Expenditure")
```

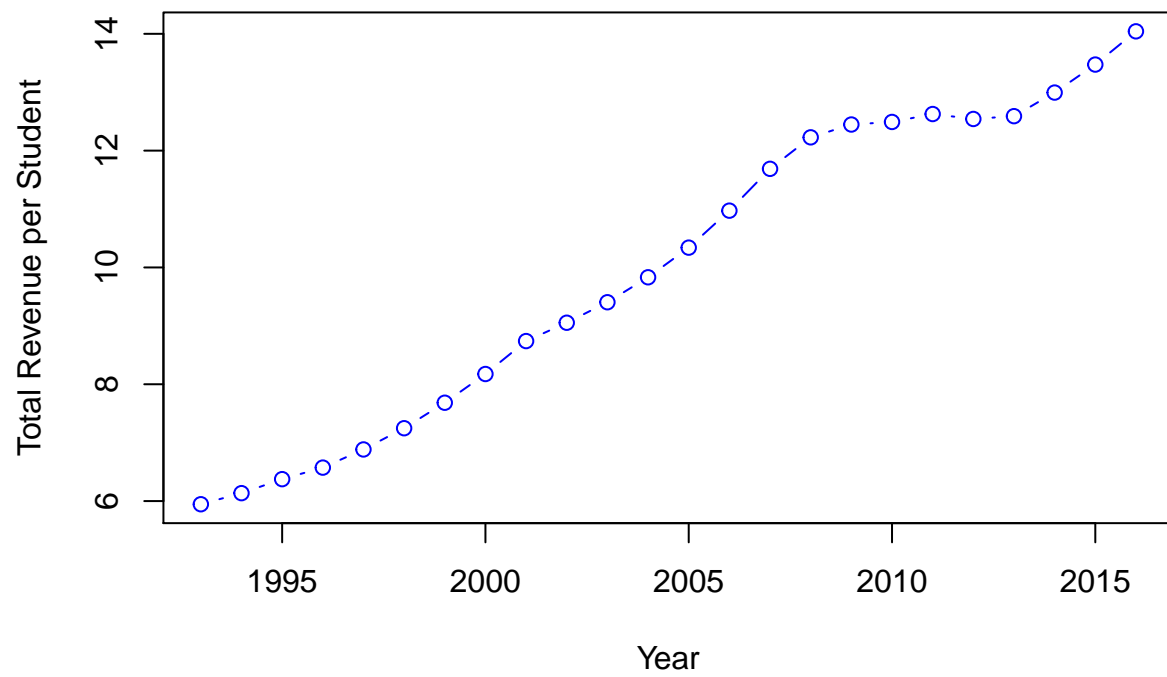


```
plot((financials_by_year$TOTAL_REVENUE - financials_by_year$TOTAL_EXPENDITURE) ~ financials_by_year$YEAR,
     type = 'b', col = "blue", xlab = "Year", ylab = "Total Revenue - Total Expenditure")
abline(a = 0, b = 0, col = "gray60")
```

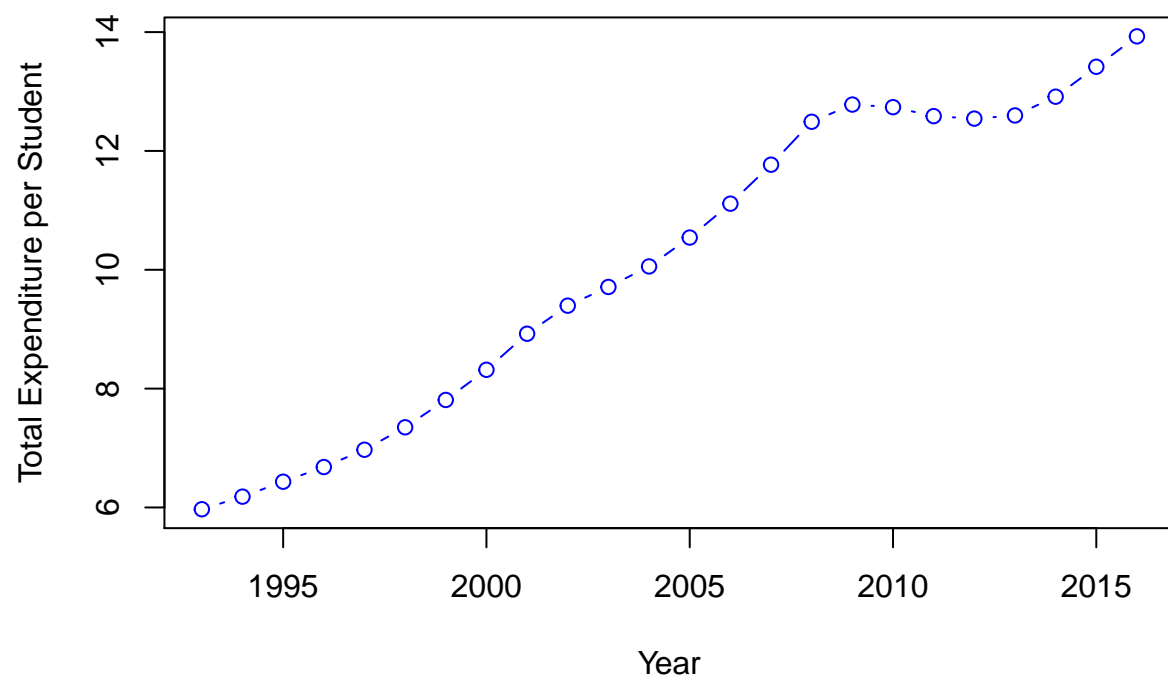


Let's try and understand this on a per-student basis

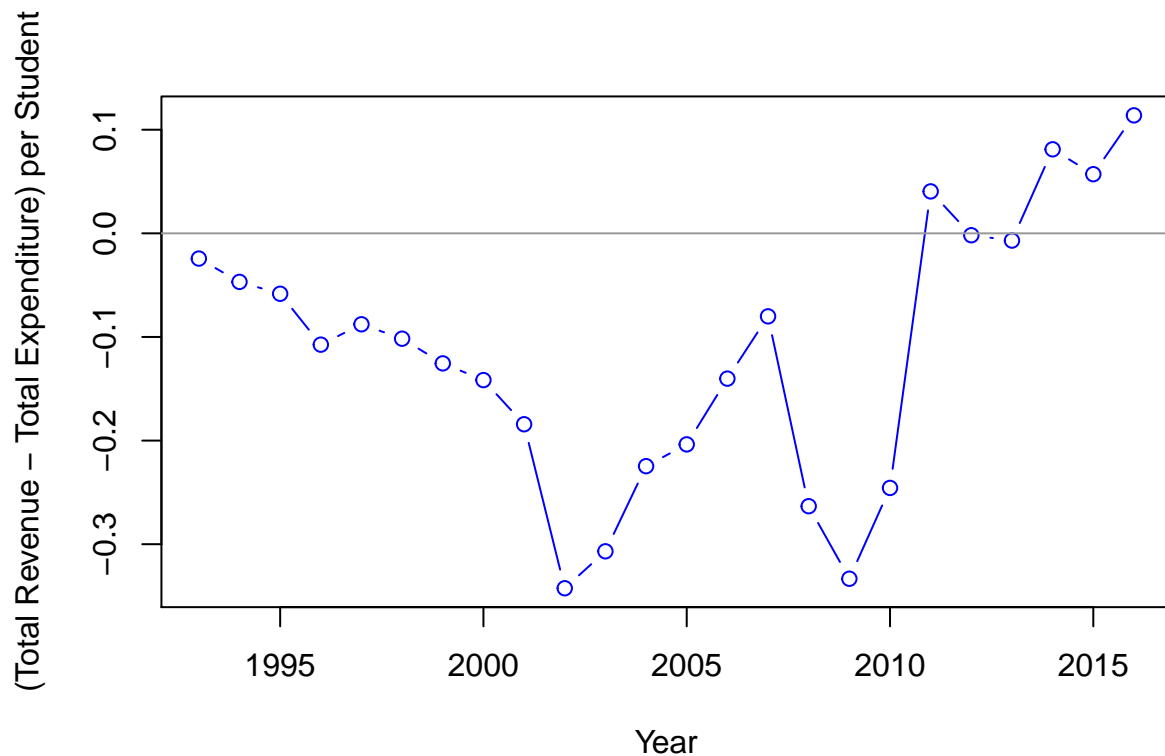
```
plot(financials_by_year$TOTAL_REVENUE/financials_by_year$ENROLL ~ financials_by_year$YEAR, type = 'b',
     col = "blue", xlab = "Year", ylab = "Total Revenue per Student")
```



```
plot(financials_by_year$TOTAL_EXPENDITURE/financials_by_year$ENROLL ~ financials_by_year$YEAR, type = 'l',  
     col = "blue", xlab = "Year", ylab = "Total Expenditure per Student")
```



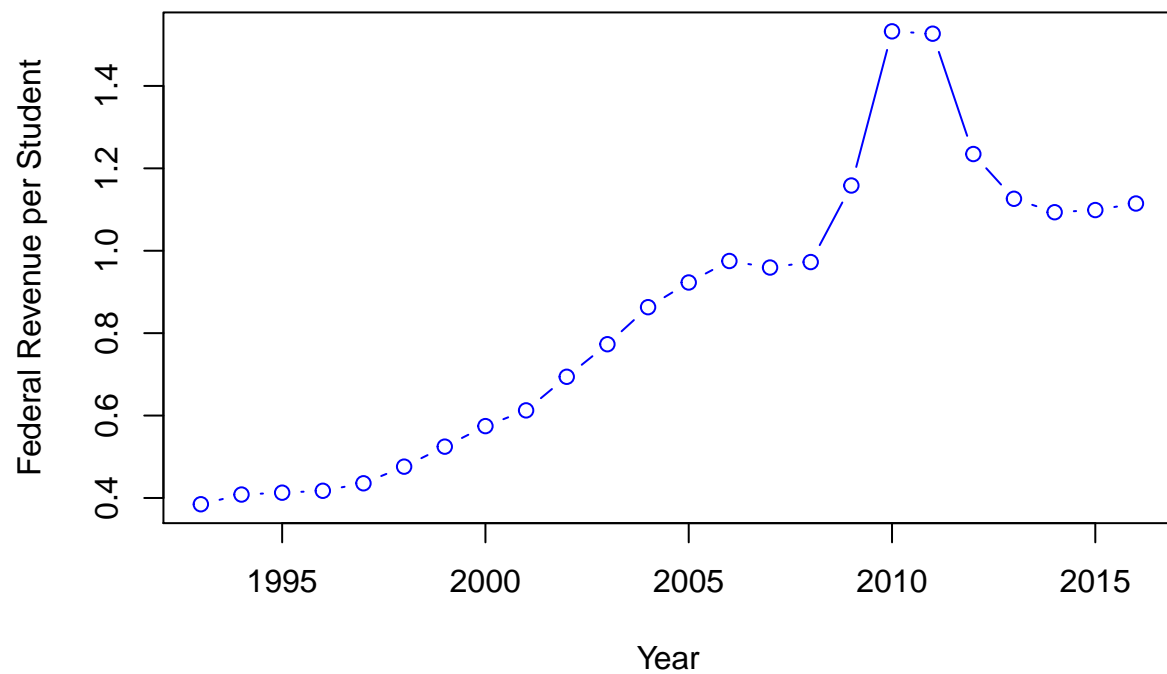
```
plot((financials_by_year$TOTAL_REVENUE - financials_by_year$TOTAL_EXPENDITURE)/financials_by_year$ENROLLED_STUDENTS,
     col = "blue", xlab = "Year", ylab = "(Total Revenue - Total Expenditure) per Student")
abline(a = 0, b = 0, col = "gray60")
```



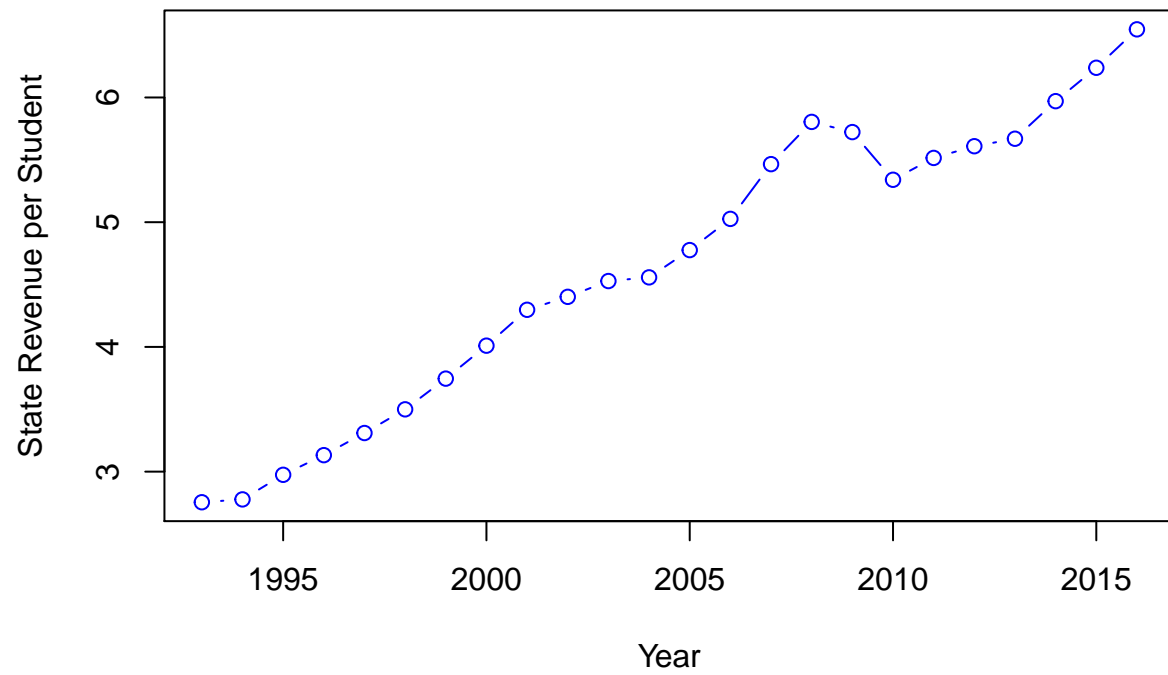
Interestingly, the total graphs and per-student graphs look nearly identical! This means that the rising funding and rising expenditures are not all attributable to enrollment growth. So, we have to ask the question, what is driving these increases? Is it only federal revenue or state revenue, are teachers earning more or are schools spending on real estate & equipment (capital outlay)? We can dig into the total revenue/expenditure figures to get more granular and identify drivers.

Let's start with revenue.

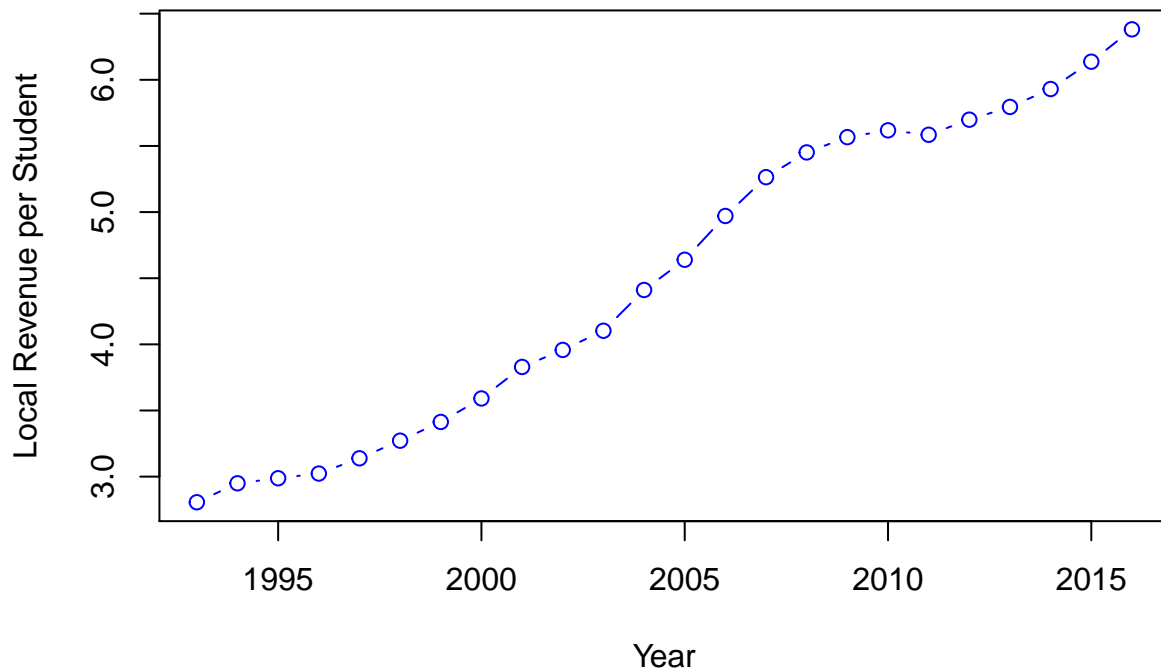
```
plot(financials_by_year$FEDERAL_REVENUE/financials_by_year$ENROLL ~ financials_by_year$YEAR, type = 'b',
     col = "blue", xlab = "Year", ylab = "Federal Revenue per Student")
```



```
plot(financials_by_year$STATE_REVENUE/financials_by_year$ENROLL ~ financials_by_year$YEAR, type = 'b',  
     col = "blue", xlab = "Year", ylab = "State Revenue per Student")
```

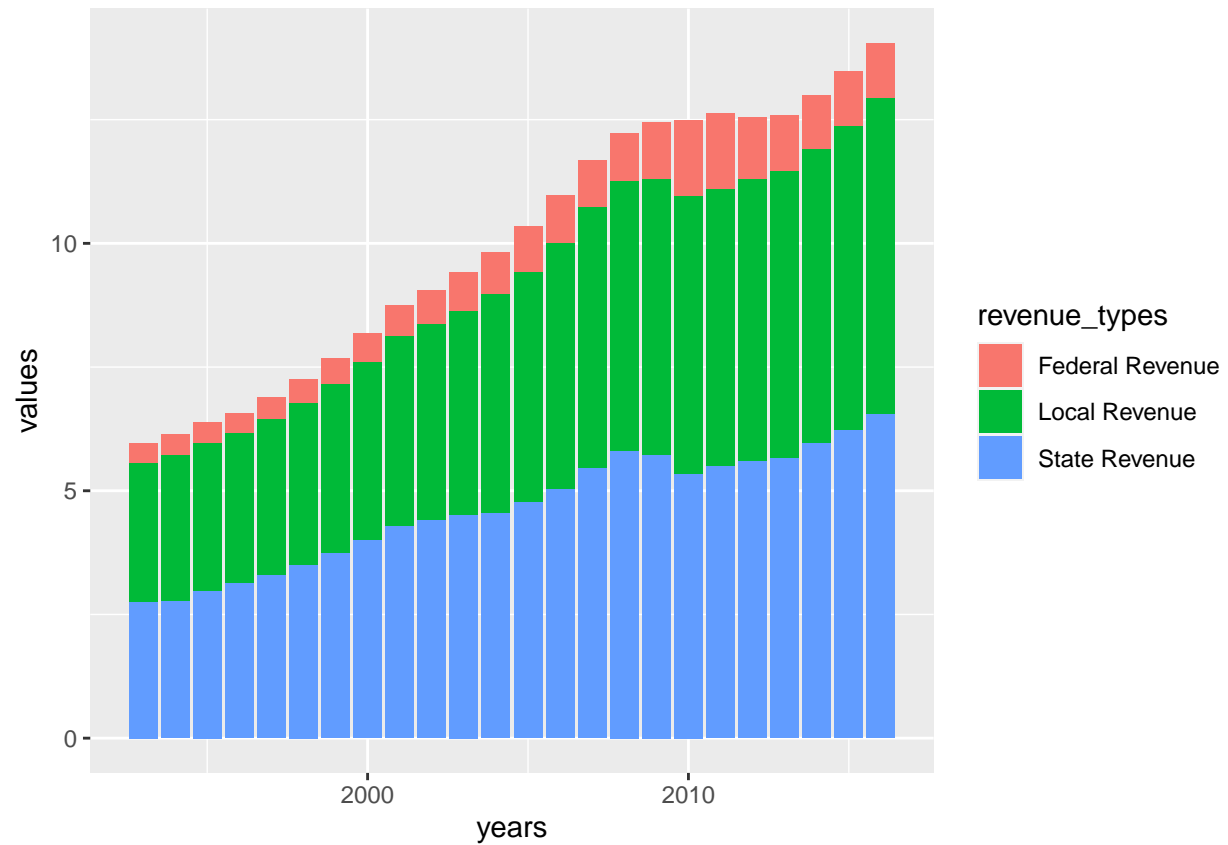
```
plot(financials_by_year$LOCAL_REVENUE/financials_by_year$ENROLL ~ financials_by_year$YEAR, type = 'b',  
     col = "blue", xlab = "Year", ylab = "Local Revenue per Student")
```



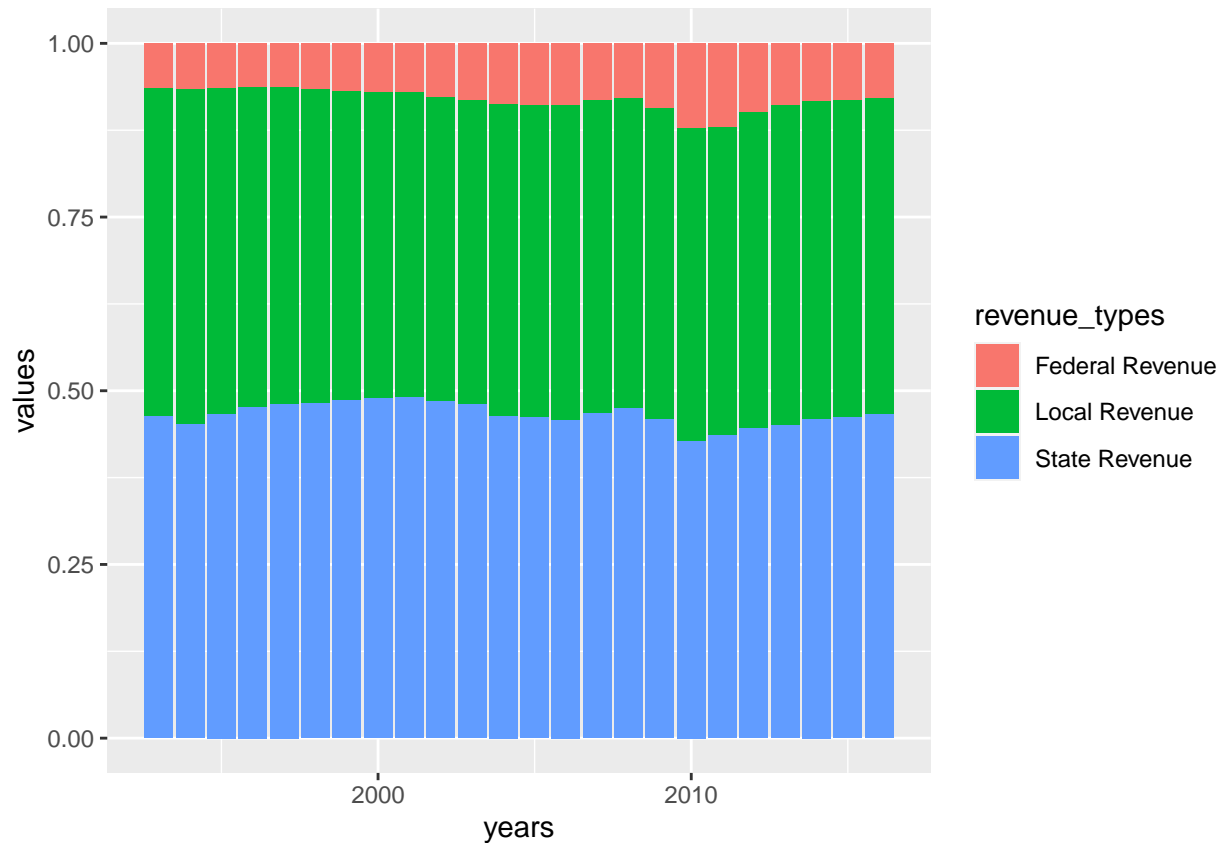
Looks like all three are increasing, with few plateaus and dips mixed in. Let's take a look at the three types of revenue on a per student basis in a stacked bar chart to get a sense of their relative size.

```
years = rep(financials_by_year$YEAR, 3)
revenue_types = c(rep("Federal Revenue", 24), rep("State Revenue", 24), rep("Local Revenue", 24))
values = c(financials_by_year$FEDERAL_REVENUE/financials_by_year$ENROLL,
           financials_by_year$STATE_REVENUE/financials_by_year$ENROLL,
           financials_by_year$LOCAL_REVENUE/financials_by_year$ENROLL)
revenues = data.frame(years, revenue_types, values)

library(ggplot2)
ggplot(revenues, aes(fill=revenue_types, y=values, x=years)) +
  geom_bar(position="stack", stat="identity")
```



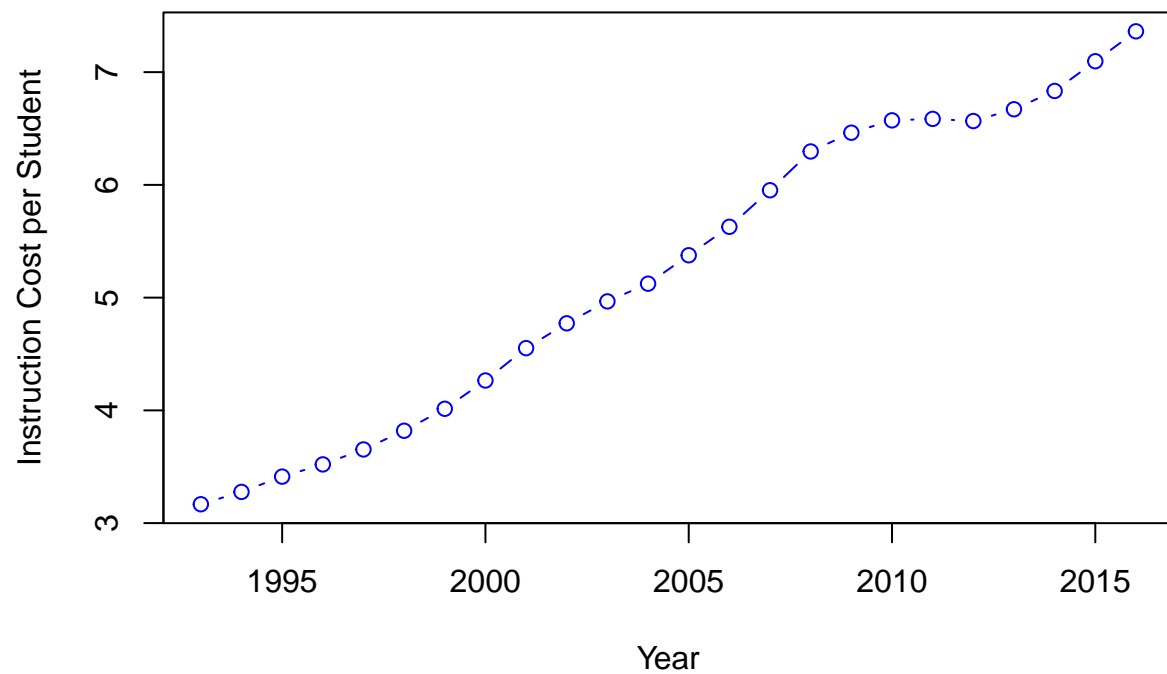
```
ggplot(revenues, aes(fill=revenue_types, y=values, x=years)) +  
  geom_bar(position="fill", stat="identity")
```



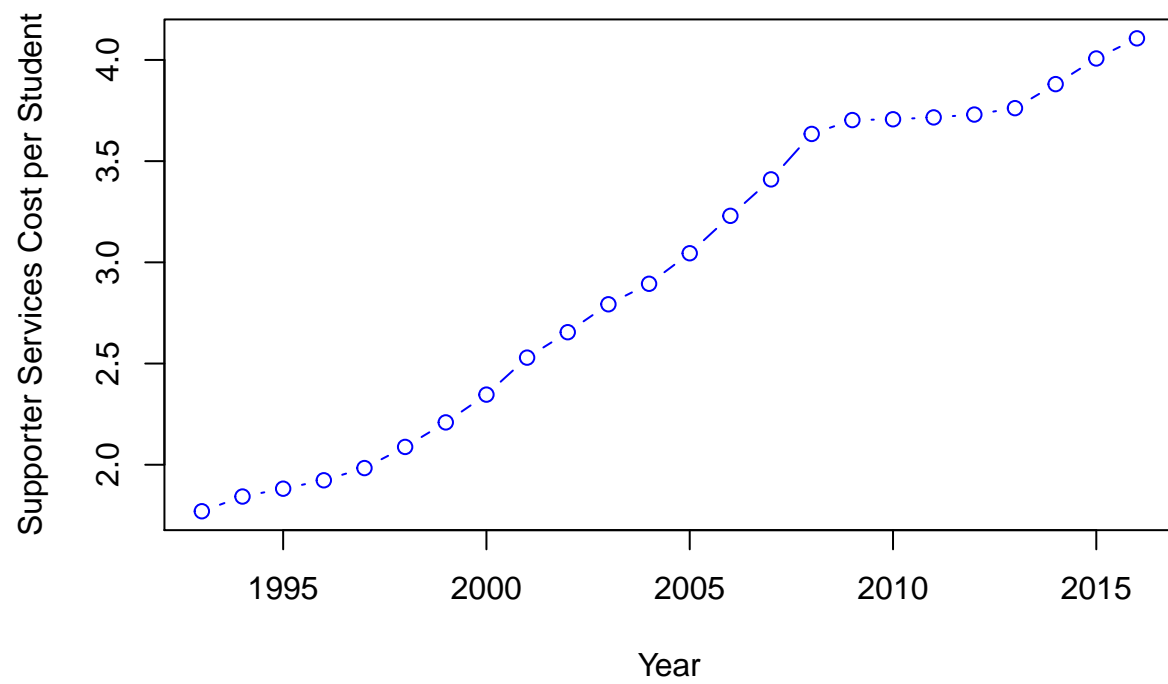
We see that state and local revenue make up the large bulk of revenue per student. Also, the increase in revenue per student is primarily driven by state and local revenue, with federal revenue remaining mostly stagnant in later years.

Now, we can dig in to the expenditure side of things and see if there is anything else to uncover.

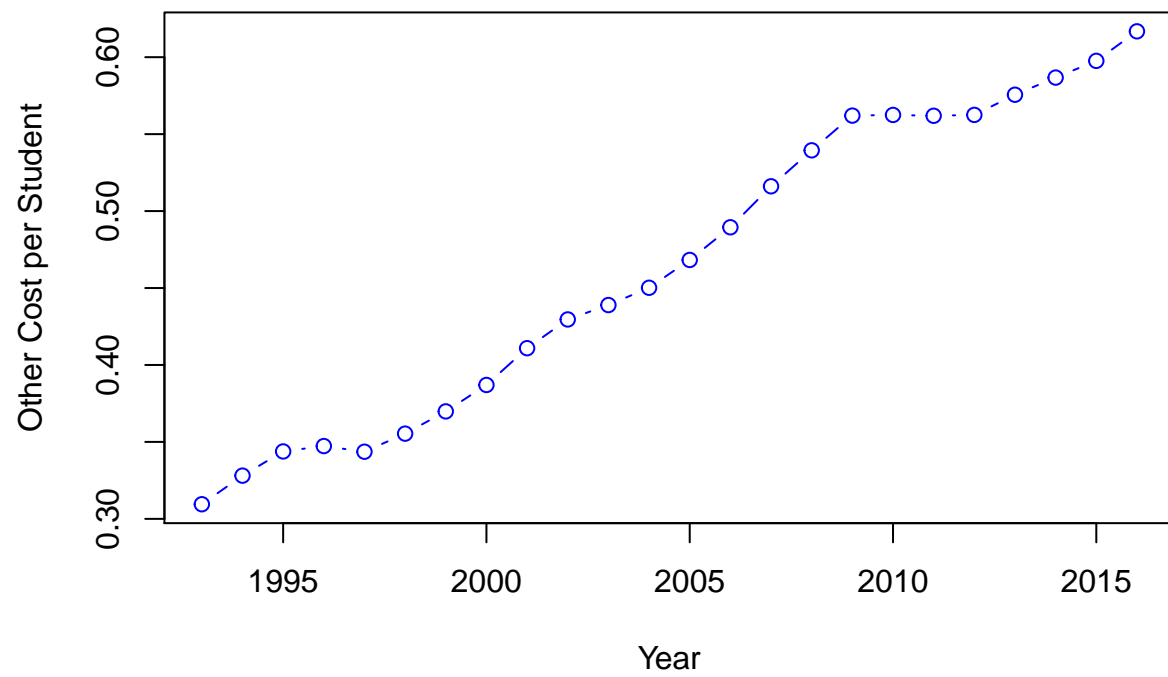
```
plot(financials_by_year$INSTRUCTION_EXPENDITURE/financials_by_year$ENROLL ~ financials_by_year$YEAR, type = "n",
     col = "blue", xlab = "Year", ylab = "Instruction Cost per Student")
```



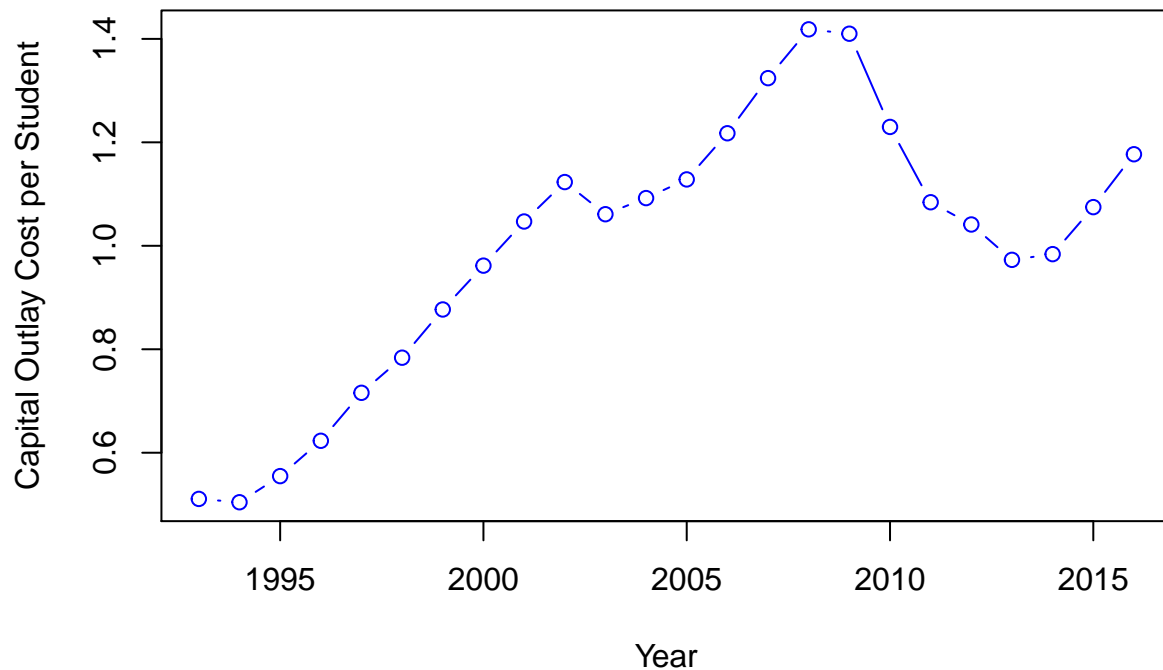
```
plot(financials_by_year$SUPPORT_SERVICES_EXPENDITURE/financials_by_year$ENROLL ~ financials_by_year$YEAR,
     col = "blue", xlab = "Year", ylab = "Supporter Services Cost per Student")
```



```
plot(financials_by_year$OTHER_EXPENDITURE/financials_by_year$ENROLL ~ financials_by_year$YEAR, type = 'l',  
     col = "blue", xlab = "Year", ylab = "Other Cost per Student")
```



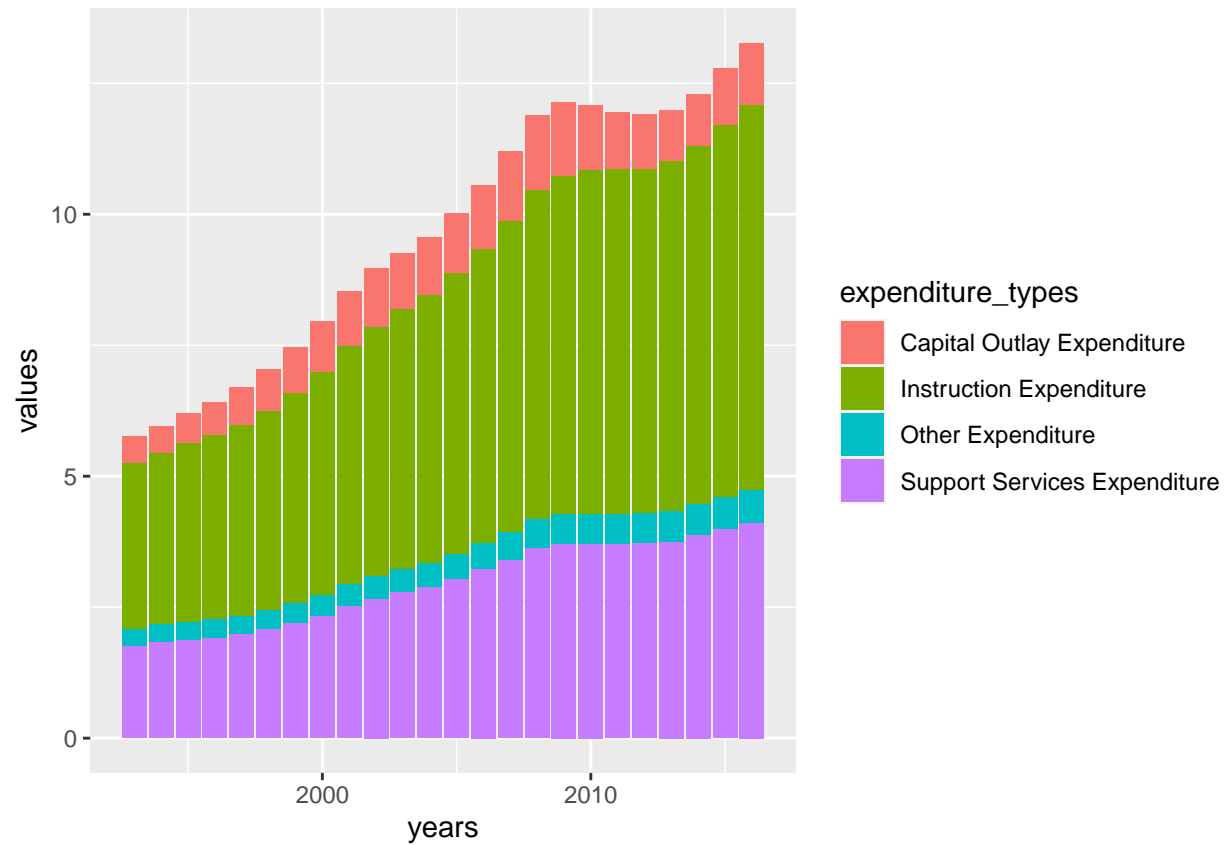
```
plot(financials_by_year$CAPITAL_OUTLAY_EXPENDITURE/financials_by_year$ENROLL ~ financials_by_year$YEAR,  
     col = "blue", xlab = "Year", ylab = "Capital Outlay Cost per Student")
```



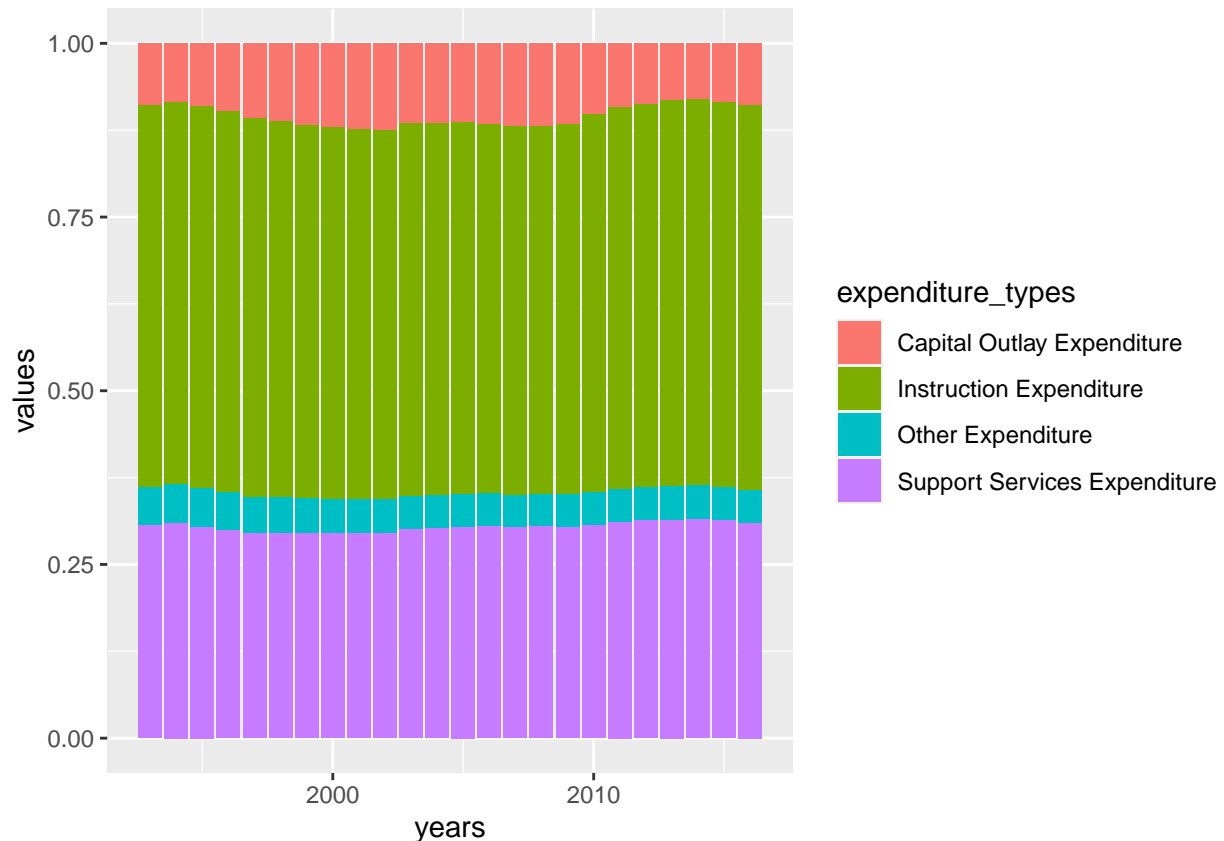
Looks like all four are increasing, with few plateaus and dips mixed in. Let's take a look at the four types of expenditure on a per student basis in a stacked bar chart to get a sense of their relative size.

```
years = rep(financials_by_year$YEAR, 4)
expenditure_types = c(rep("Instruction Expenditure", 24), rep("Support Services Expenditure", 24),
                      rep("Other Expenditure", 24), rep("Capital Outlay Expenditure", 24))
values = c(financials_by_year$INSTRUCTION_EXPENDITURE/financials_by_year$ENROLL,
            financials_by_year$SUPPORT_SERVICES_EXPENDITURE/financials_by_year$ENROLL,
            financials_by_year$OTHER_EXPENDITURE/financials_by_year$ENROLL,
            financials_by_year$CAPITAL_OUTLAY_EXPENDITURE/financials_by_year$ENROLL)
revenues = data.frame(years, expenditure_types, values)

library(ggplot2)
ggplot(revenues, aes(fill=expenditure_types, y=values, x=years)) +
  geom_bar(position="stack", stat="identity")
```

```
ggplot(revenues, aes(fill=expenditure_types, y=values, x=years)) +  
  geom_bar(position="fill", stat="identity")
```

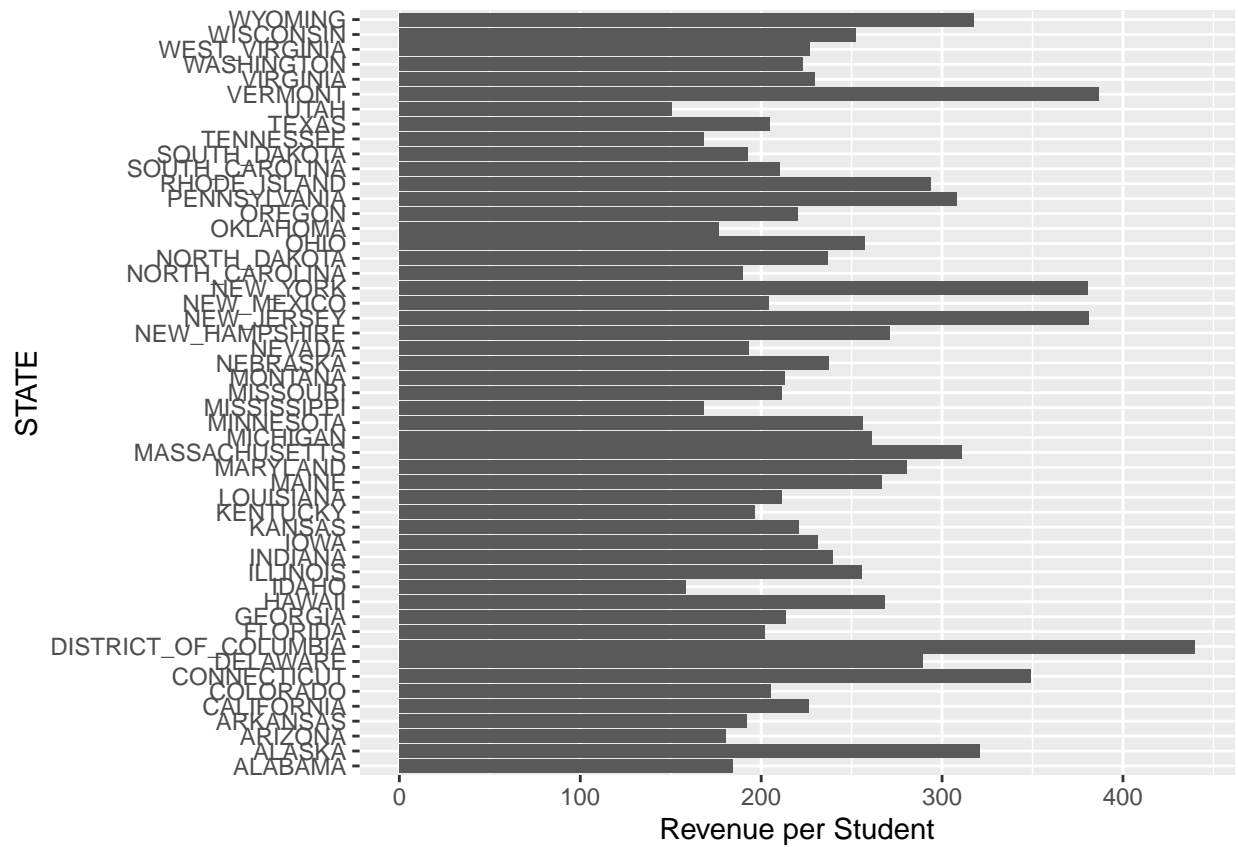


Looks like the large portion of expenditure comes from instruction and support services, which seems reasonable. The percentage mix of the different types of expenses remains pretty constant throughout all years, so we can conclude that all 4 categories are increasing in dollar value at roughly the same rate; they are all somewhat equal drivers of rising expenditures for schools.

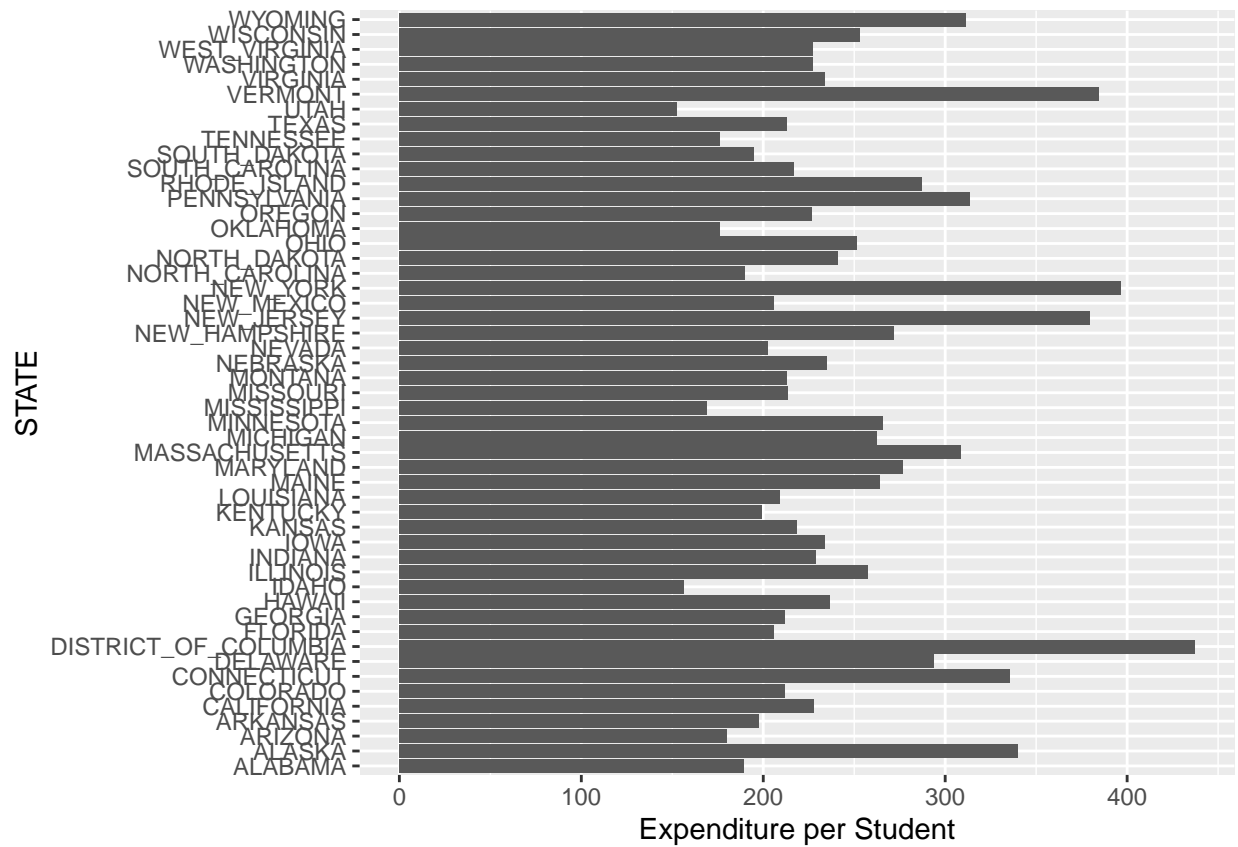
Lastly, let's try and understand how total revenue and expenditure per student varies between different states.

```
state_data = setNames(aggregate(cbind(data$TOTAL_REVENUE/data$ENROLL,
                                      data$TOTAL_EXPENDITURE/data$ENROLL,
                                      data$ENROLL
                                ), by=list(STATE=data$STATE),
                                FUN = sum, na.rm = TRUE), c("STATE",
                                                            "TOTAL_REVENUE_per_Student",
                                                            "TOTAL_EXPENDITURE_per_Student", "ENROLL"))
state_data = state_data[-c(10,29),]
state_data = state_data[order(state_data$TOTAL_REVENUE_per_Student),]

library(ggplot2)
ggplot(state_data, aes(x=STATE, y=TOTAL_REVENUE_per_Student )) +
  geom_bar(stat="identity") +
  ylab("Revenue per Student") +
  coord_flip()
```



```
ggplot(state_data, aes(x=STATE, y=TOTAL_EXPENDITURE_per_Student )) +
  geom_bar(stat = "identity") +
  theme(legend.position="none") +
  ylab("Expenditure per Student") +
  coord_flip()
```



Overall, we can see significant variation between states on both the revenue and expenditure sides. It is interesting that the funding of education differs so significantly between different states, even ones that are pretty close geographically.