SDS 192 Mini-Project 1

SDS 192

Tabitha Githinji Ayla Hauser Author Three

2024-10-28

```
scorecard <- sc_init() |>
  sc_year(2018) |>                #Note how we are looking at only 2021 data here!
  sc_filter(stabbr == "MA") |>     #Note how we are looking at only Massachusetts data here!
  #The line below shows variable selection (there are lots of variables)
  sc_select(unitid, instnm, city, highdeg, control, ugds, adm_rate, costt4_a, costt4_p, pcip2
  sc_get()
```

```
Large request will require: 1 additional pulls.

Request additional chunk 1

Request complete!
```

```
#  Here's an example of how to recode the control variable

# We are renaming the column control_text from control
scorecard$control_text <-
  #The recode function does the work. It calls 1 "Public", 2 " Private nonprofit",etc.
  recode(
    scorecard$control,
    "1" = "Public",
    "2" = "Private nonprofit",
    "3" = "Private for-profit",
    .default = NA_character_
  )
```

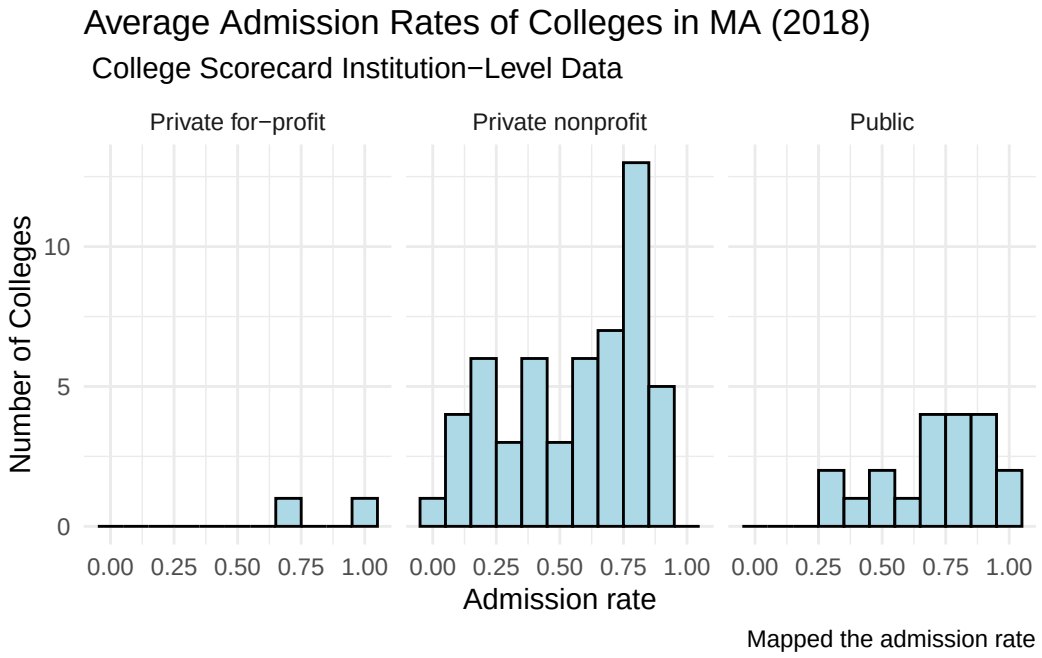
Plot 1

How do the cost of attendance, and admission rates differ between public and private colleges in Massachusetts for the year 2018?

```
# Create plot 1 here
#This histogram displays the distribution of admission rates across colleges in Massachusetts
ggplot(data = scorecard,
  aes(x = adm_rate)) +
  geom_histogram(binwidth = 0.1, fill = "lightblue", color = "black") +
  labs(title = "Average Admission Rates of Colleges in MA (2018)",
       x = "Admission rate",
       y = "Number of Colleges",
       subtitle = " College Scorecard Institution-Level Data ",
       caption = "Mapped the admission rate",
       ) +

   facet_wrap(~ control_text)+
  theme_minimal()
```

```
Warning: Removed 74 rows containing non-finite outside the scale range
(`stat_bin()`).
```

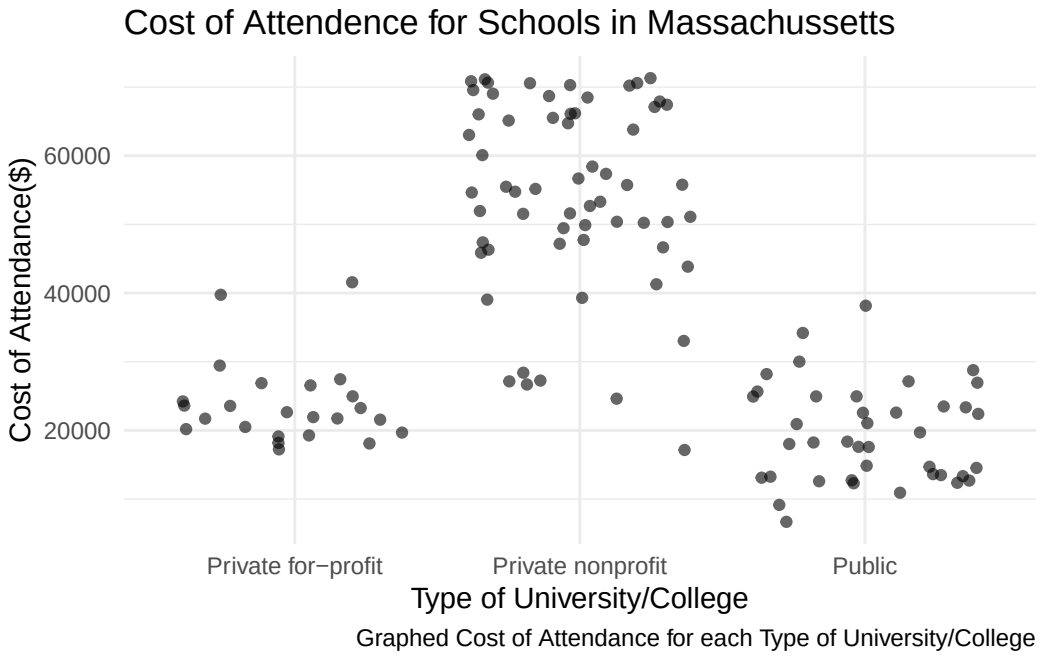


Plot 2

```
#Mutation creating a cost of attendance that combines both costs in the data set and filters
#Jitter plot with the x axis equal to the categorical variable of the type of university and

scorecard <- scorecard |>
  mutate(cost_attendance = (costt4_a + costt4_p)) |>
  filter(cost_attendance > 0)

ggplot(data = scorecard,
       aes(x = control_text,
       y = cost_attendance)) +
  geom_jitter(alpha = 0.6,
              ) +
  theme_minimal() +
  labs(x = "Type of University/College",
       y = "Cost of Attendance($)",
       caption = "Graphed Cost of Attendance for each Type of University/College",
       title = "Cost of Attendence for Schools in Massachussetts")
```



Blog post

In this blog post, we delve into the admission rates and costs of attendance for colleges in Massachusetts. We used data from the College Scorecard, a comprehensive dataset produced by the U.S. Department of Education. This dataset provides information on various metrics related to higher education institutions, including tuition costs, graduation rates, and admission statistics. For our analysis, we focused specifically on the year 2018, examining how admission rates and costs of attendance differ between public universities, private nonprofit universities, and private profit universities. Our findings are illustrated through two visualizations: a histogram depicting admission rates and a jitter plot showcasing costs of attendance.

The first visualization, a histogram, reveals important insights into the distribution of admission rates among Massachusetts colleges. From the histogram, we observe that public universities generally exhibit higher admission rates, clustering between 70% and 90%, while private nonprofit universities display a wider range of admission rates, with many institutions falling below 60%. This indicates that public universities tend to be more accessible to a broader demographic of applicants compared to their private counterparts. The faceting of the histogram by university type allows us to easily compare these patterns, highlighting the distinct roles that public and private institutions play in the higher education landscape.

Our second visualization is a jitter plot that provides a clearer picture on the costs of attendance for these colleges. The plot illustrates that public universities typically have lower costs compared to private nonprofit and private profit institutions. By representing the cost of attendance on the y-axis and the type of university on the x-axis, the plot reveals that many public colleges offer a more affordable option for students. The use of jittering helps to visualize the concentration of costs within each category, showcasing the financial accessibility of each type of institution.

The key takeaway from our analysis is the difference in both admission rates and costs of attendance between public and private colleges in Massachusetts. Public universities generally provide higher admission rates and lower costs, making them more accessible to a wider range of students. In contrast, private nonprofit and private profit institutions often have more selective admission processes and higher tuition fees, which may limit access for some students. However, it is crucial to approach this data with ethical considerations in mind. One primary concern is the representation of marginalized student groups in the dataset. While the College Scorecard provides valuable insights, it may not fully capture the unique barriers faced by students from low-income backgrounds or underrepresented communities. Furthermore, the dataset relies heavily on quantitative metrics, which can overlook qualitative aspects of students' experiences and the broader context of their educational journeys.

In summary, our visualizations of admission rates and costs of attendance among Massachusetts colleges underscore important trends in accessibility and affordability. As we analyze such datasets, it is essential to remain aware of the ethical implications and potential biases inherent

in the data collection process. This ensures that our findings contribute to a more equitable understanding of higher education access and opportunities for all students.

In this blog post, we delve into the admission rates and costs of attendance for colleges in Massachusetts, using data from the College Scorecard, a comprehensive dataset produced by the U.S. Department of Education. This dataset provides a wealth of information on various metrics related to higher education institutions, including tuition costs, graduation rates, and admission statistics. For our analysis, we focus specifically on the year 2018, examining how admission rates and costs of attendance differ between public universities, private nonprofit universities, and private profit universities. Our findings are illustrated through two visualizations: a histogram depicting admission rates and a jitter plot showcasing costs of attendance.

The first visualization, a histogram, reveals important insights into the distribution of admission rates among Massachusetts colleges. From the histogram, we observe that public universities generally exhibit higher admission rates, clustering between 70% and 90%, while private nonprofit universities display a wider range of admission rates, with many institutions falling below 60%. This indicates that public universities tend to be more accessible to a broader demographic of applicants compared to their private counterparts. The faceting of the histogram by university type allows us to easily compare these patterns, highlighting the distinct roles that public and private institutions play in the higher education landscape.

In addition to admission rates, our second visualization—a jitter plot—provides a clearer picture of the costs of attendance for these colleges. The jitter plot illustrates that public universities typically have lower costs compared to private nonprofit and private profit institutions. By representing the cost of attendance on the y-axis and the type of university on the x-axis, the plot reveals that many public colleges offer a more affordable option for students. The use of jittering helps to visualize the concentration of costs within each category, showcasing the financial challenges that may be faced by students attending private colleges.

The key takeaway from our analysis is the stark difference in both admission rates and costs of attendance between public and private colleges in Massachusetts. Public universities generally provide higher admission rates and lower costs, making them more accessible to a wider range of students. In contrast, private nonprofit and private profit institutions often have more selective admission processes and higher tuition fees, which may limit access for some students. However, it is crucial to approach this data with ethical considerations in mind. One primary concern is the representation of marginalized student groups in the dataset. While the College Scorecard provides valuable insights, it may not fully capture the unique barriers faced by students from low-income backgrounds or underrepresented communities. Furthermore, the dataset relies heavily on quantitative metrics, which can overlook qualitative aspects of students' experiences and the broader context of their educational journeys.

In summary, our visualizations of admission rates and costs of attendance among Massachusetts colleges underscore important trends in accessibility and affordability. As we analyze such datasets, it is essential to remain aware of the ethical implications and potential

biases inherent in the data collection process, ensuring that our findings contribute to a more equitable understanding of higher education access and opportunities for all students.