The Rising Costs of Higher Education & Its Effects on Enrollment

Aziz Zoher Khilawala, Dandrew Merriweather, Hawra Rabaan, Keyur Kirti Mehta, Lisha Chen, Robert Crouch, Sagar Salvi, and Tariq Khan.

Abstract. Overall costs of tuition for postsecondary education in the United States has been steadily increasing over the past half century. The purpose of the overall study was to analyze the effect increasing tuition costs may be having on enrollment in postsecondary institutions. We hypothesized that the data would show a correlation between enrollment rates dropping for traditional institutions and tuition rates rising. We gathered databases for the study from a multitude of sources, linked to both public and government resources. Due to our varying database sources the data had to be normalized and heavily manipulated to properly visualize the data. The data was then visualized using a variety of methods our team was familiar with. The study proved our hypothesis wrong overall. Although tuition rates rose moderately in recent years, enrollment rates have also continued to rise overall. We did find some correlation between times of economic stress or war and overall percentage of growth, however there still was not a shown decline in enrollment trends. The study successfully visualized trends over the past half century for postsecondary education tuition and enrollment in the United States even though the initial hypothesis was proven wrong.

Keywords: · Education · Tuition · Enrollment · Vocational · Postsecondary · Postgraduate · University · Database · Visualization · mySQL · Python · D3 · Javascript

1 Introduction

Tuition rates nationwide have been on a steady slow climb each year, while federal financial aid has failed to rise at the same rate. This has caused many students to spend more out of pocket for higher education and even question it all together [1]. Purdue University however is in the middle of a six year tuition freeze, yet recently reported their largest incoming class since 2006, before the recession [2]. Is it purely coincidental that Purdue is able to break record enrollment numbers by keeping their tuition costs down, or is there a correlation to be drawn? This question is what the study aims to answer.

1.1 Aim

The study's aim was to evaluate tuition costs in the United States and their correlation with enrollment rates associated with various higher education programs.

1.2 Purpose

The project's purpose was to leverage data in an effort to analyze the trends in tuition and enrollment for higher education around the United States. In leveraging this data the hope was that some trends or insights could be found that correlate cost of education with rate of enrollment. These trends in the data were compared and analyzed using visualizations so as to be more easily deciphered.

1.3 Hypothesis

The increasing cost of tuition for all levels of post secondary school in the United States has had a direct impact on the percentage of eligible individuals' enrollment in those programs.

2 Methodology

The data for the study was gathered from various resources found through research online. There are many institutions, in private and government sectors, that gather data on tuition and enrollment trends. Much of the data that was gathered was organized in vastly different ways when compared to one another. The databases had to be normalized using Microsoft Excel in preparation for use in various data visualization programs. After the data was normalized it was then uploaded to the team mySQL to allow for each data visualization program to access it. As the hypothesis states we were focusing on eligible persons data. The National Center for Education Statistics defines eligible persons in their data as, "Individuals ages 16 to 24 who graduated from high school or completed a GED or other high school equivalency credential [3]." Because of this much of our data was based on eligible persons rather than overall population in the United States.

The process of choosing each visualization tool revolved around the team's various strengths in data visualization. Spreading the data visualizations across multiple software disciplines allowed the team to highlight the versatility of the data and programs used in practice. In the end, visualizations from each respective visualization tool were used to compare against the original hypothesis and determine the study's outcomes.

2.1 Databases

We used a variety of both publicly sourced and government databases in our study. Multiple databases were sourced from the National Center for Educational Statistics whose records dated as far back as 1974. We focused on more modern trends in an effort to see how the modern economy relates to tuition and enrollment rates. Other databases came from public sources such as The Chronicle of Higher Education, and an organization known as The College Board. At times when working with databases from multiple sources, data had to be reworked to properly correlate with data from alternative sources. These methods and specific databases are covered later in the report.

3 Outcomes and Analysis

3.1 Tuition Cost Trends

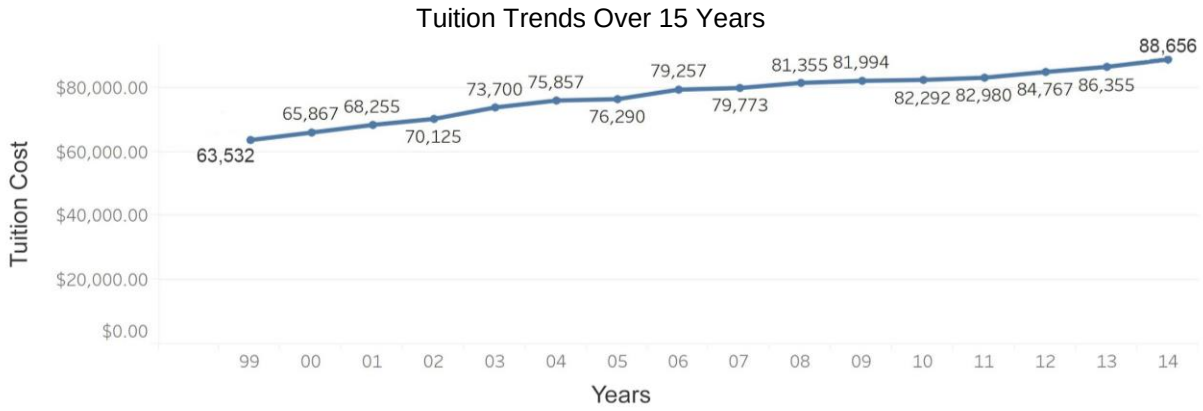


Fig 1. Tuition trends from 1999 to 2014

In the original hypothesis we theorized that tuition has been on the rise, while it may be true that the dollar amount has risen, it could be argued that it hasn't risen at all in the past two decades depending on what it is being compared to. Figure 1 above, shows an increase with rates averaging around $65,000.00 in 1999, while steadily increasing to $88,000.00 in 2014 [4]. Tuition increased on average 1.7% each year which is interestingly less than the average rate of inflation in the United States at 1.9% over those same years [5]. With inflation factored in, one could argue that tuition rates have actually not risen at all over these 15 years, they may have actually dropped in comparison by 0.2%. However with factors such as the modern recession in the United States, it could also be argued that the average tuition rate was lower simply because students were favoring a less expensive education in times of economic stress. In these same years the median household income did not rise, in fact is was on the decline from $58,665 in 1999 to $54,398 in 2014 with an all time low of $53,331 in 2012 [6]. It is no secret that the United States was in a recession and that it certainly affected many aspects of people's lives when it came to spending money. For further explanation we must look at other data.

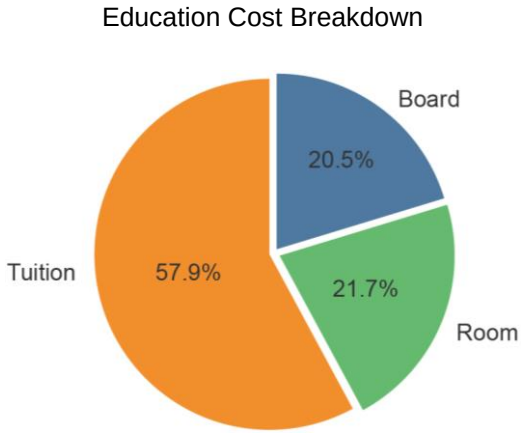


Fig 2. Pie graph of cumulative higher education expenses.

Figure 2 is a breakdown of other expenses that factor into educational costs [7]. While our database analysis focused on tuition costs alone, it is important to know the average students across the United States have other costs directly associated with their higher education decisions. A breakdown of the average costs shows that tuition may account for nearly 58% of all costs, while room and board costs can also weigh heavily on a student's budget, almost evenly splitting the other 42% in education associated expenses.

3.2 University Enrollment Trends

As previously stated, it was assumed that increasing tuition costs in the United States were causing people to pursue alternative routes in higher education and thus enrollment rates in universities would decrease. The data shows that from 2006 to 2016, university enrollment has increased steadily with a few anomalies [3].

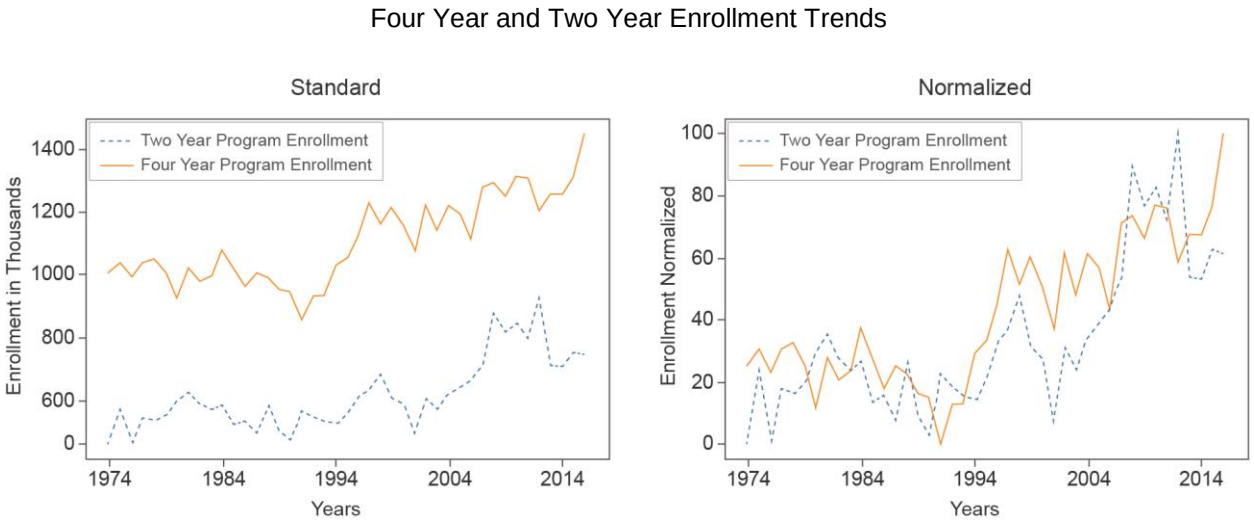


Fig 3. Four year and two year enrollment trends shown normalized on the right from 1974 to 2016

In Figure 3 the enrollment graphs show a healthy increase in four year programs at universities between 1990 through 1996 [3]. Although we anticipate there will be natural fluctuation in numbers there is a noticeable drop in enrollment starting in 2000. More on this sudden dip will be covered in a later part of the report. After normalizing the data using the formula covered in Formula 1 below and bringing the two datasets into a common range, it is easier to compare the rates of growth. The normalized datasets are represented in the graph on the right side of figure 3. A noticeable change in trends can be seen in 2006 where two year programs start to gain enrollment at a higher rate than four year programs.
As touched on in outcome section (3.1), we hypothesize this may be a combination of rising tuition rates along with a drop in median household income levels in the United States and economic uncertainty.

$$(X\text{-}min)/(max - min) * (New\_max - New\_min) + New\_min \quad (1)$$

Formula 1 above is what was used to normalize the data. The formula linearly transforms all the data values to aid in visualization representation. The formula converts the original ranges of Minimum and Maximum (Max - Min) in each piece of data to a new range which is user defined. The two datasets are then output into this new normalized format.

Four Year University Enrollment Separated by Gender

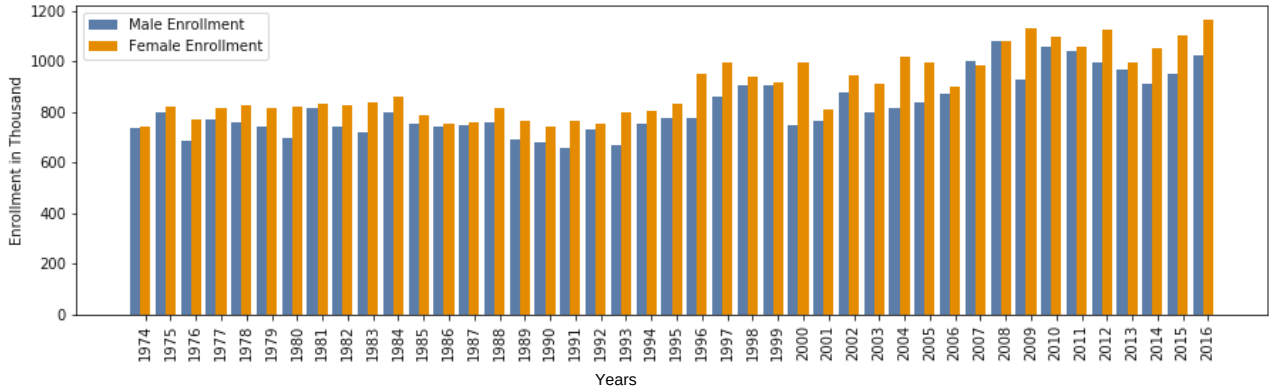


FIG 4. Data covers four year university enrollment from 1974 to 2016

The data used in figures 4 and 5 was derived from a count of highschool graduates between the age of 18-24, and the percentage of these high school graduates enrolled to the different university programs from 1974 to 2016. Enrollment was then categorized on gender basis and then sub categorized as enrolled to 2 year and 4 year university programs. Gender based enrollment shows a steady growth rate and normal fluctuations up until the year 2000, where we see a dramatic drop in male enrollment in Figure 4. This trend holds true several other years off and on through the decade that follows [3].

There is no definitive explanation for this anomaly in the data, however, it could be argued there is a correlation between the 9/11 terrorist attacks on the Twin Towers in New York and male enrollment rates in the united states. Although the drop in enrollment is small, it is correlated with small increases in enlistment in America's varying armed forces after they were experiencing lower numbers than previous years [8]. In 2001 when the attacks occurred we see that the male enrollment seems to remain in an established state of balance with female enrollment. However in 2003 when Operation Iraqi Freedom was officially launched one will notice the enrollment gap starts to increase between the sexes. Furthermore, in 2006 the data shows the averages return to steady balance while enlistment numbers were on the decline for branches of the military [8].

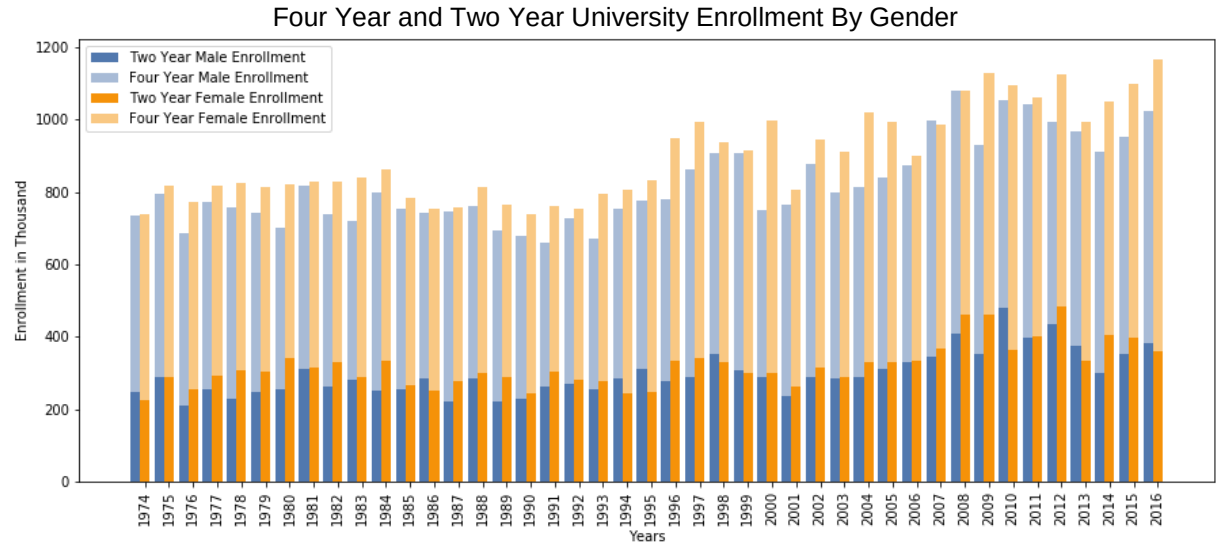


FIG 5. Data covers four year and two year university programs' enrollment from 1974 to 2016

When comparing two year programs with four year programs over the same amount of time, as shown in Figure 5, we see very little differences except the fact that there are many more years where the male population outweighs the female population in enrollment. Stage et al suggests, that the modern education gender gap is caused by

families educational aspirations from an early age and their expectations after high school. Based on their study, families expected less males to go on to college and thus parents would not plan financially for their child to attend, whereas with females, a higher percentage responded that they started planning for college from an early age [9]. This may attribute to the two year versus four year gender gap.

3.3 Vocational School Enrollment Trends

Our original hypothesis assumed that traditional tuition costs continually rising would correlate with an increase in vocational school enrollment. The data shows that this is probable, however, data for trade or vocational based education was somewhat hard to separate from other types of two and four year postsecondary education programs. Our dataset used for this outcome separated it's vocational enrollment into length of enrollment, however, if an institution had both two year and four year programs, that institution was included in the four year column, as seen in Figure 6 [10]. Another issue with this data is that vocational schools go by many names, when started in the 1940s they were most commonly known as Junior College, offering certificates of education and less than baccalaureate level education. Since then, schools have gone by many names such as community college, trade school, or vocational school, which can lead to inaccurate data over time [11]. Other issues include that some of these schools are not recognised by the National Board of Education, and thus were not included in this dataset. Despite these issues, we did still infer about the data and the outcome.

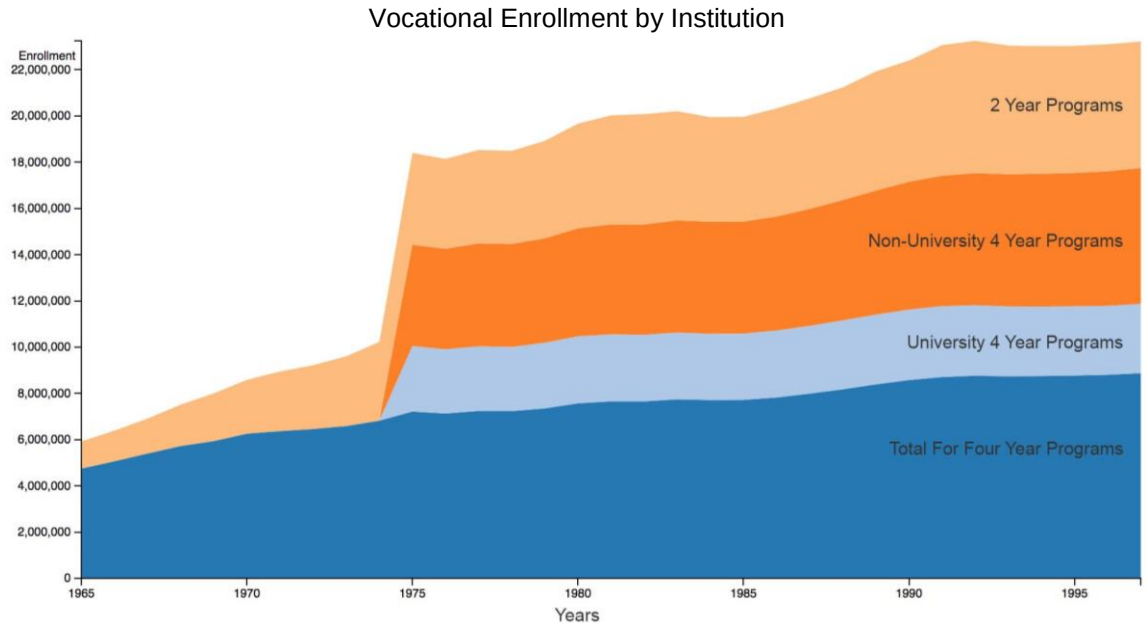


FIG. 6. The above graph shows vocational programs broken up by the ruling bodies enrollment practices.

Enrollment rates in trade schools has increased steadily since 1985. Earlier anomalies such as the 1970s spike in programs is explained by the data's sudden inclusion of new nationally recognized schools. These trends are not dramatically different from the overall traditional college trends. One can theorize that this is due to the fact that there is still an increasing trend towards obtaining an education, trade school or otherwise, despite a drop in financial assistance [12]. Each vocational school category was also plotted as an individual visualisation which can viewed in the appendix, as well as a link to an interactive visualization for further review [13].

3.4 Postgraduate Enrollment Trends

Despite a decline in postgraduate assistance [12], there is an increase in the current and projected enrollment numbers for postgraduate students [10].

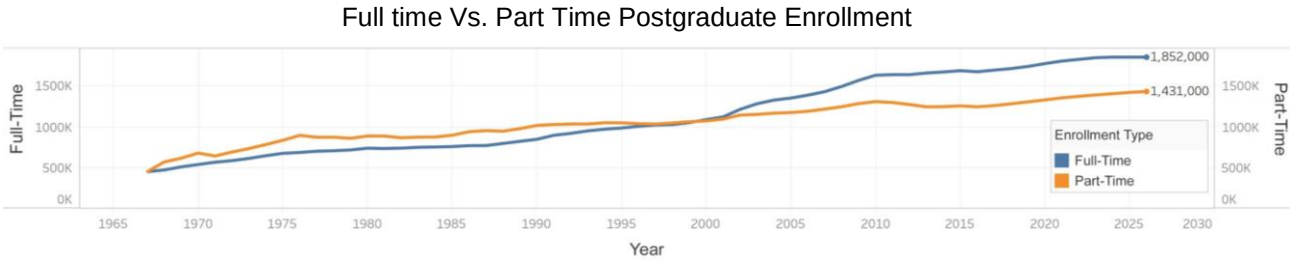


FIG. 7. Full time Vs. Part Time

While the numbers are increasing in enrollment, the data does show that there is an increased interest in attending graduate school full-time. This could be due to the scarcity of jobs, causing undergraduate students to decide on furthering their education immediately after completing their undergraduate degrees rather than entering the workforce. Although the issue may not have anything to do with job availability and everything to do with experience. An article in the Harvard Business Review states that around two-thirds of college graduates have a hard time launching their career on a bachelor's degree, causing many to pursue further schooling at the certificate, masters, or Phd level [14]. It is worth noting that although part-time students may no longer lead in enrollment, they are by no means declining.

3.5 Enrollment vs. Tuition

Finally, we thought it was important to visualize the enrollment rates compared to tuition rates over time. In the previous tuition section the data was focused on a fraction of the data from 1999 to 2014 but below in Figure 8 we see how tuition has trended from 1974 to 2014 with enrollment.

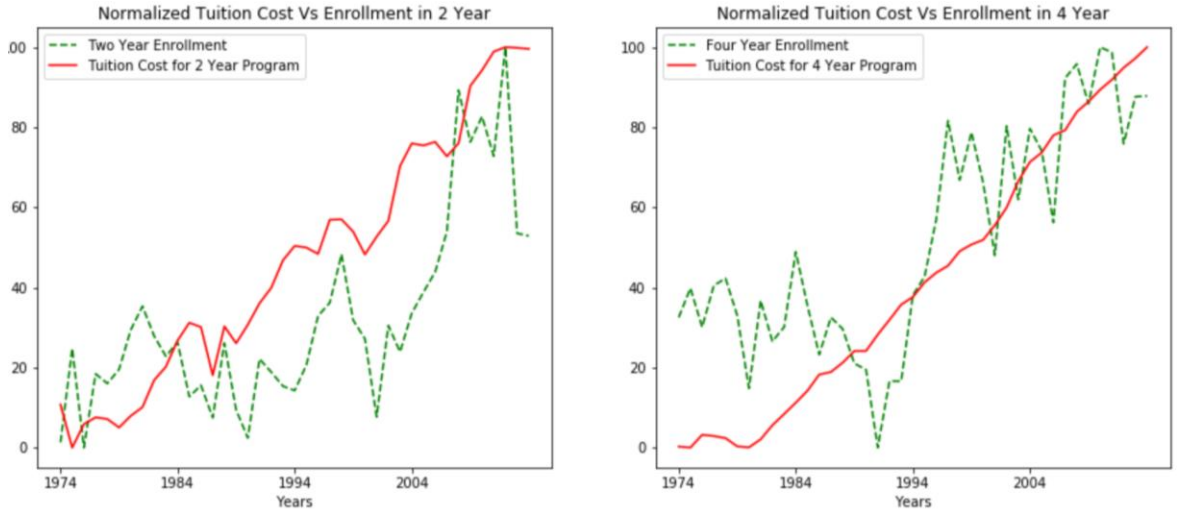


FIG. 8. Tuition Vs. Enrollment Trends from 1974 to 2014 shown with line charts.

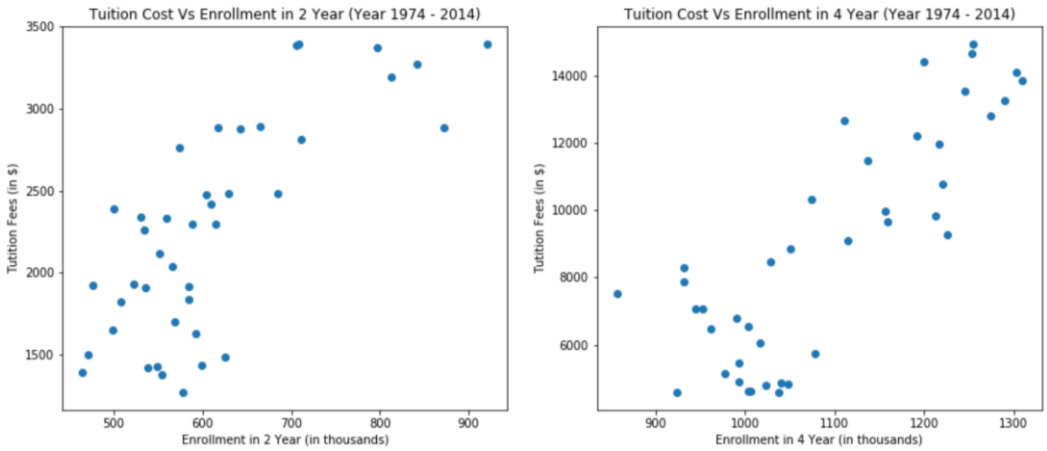


FIG. 9. Tuition vs. Enrollment Trends from 1974 to 2014 shown in scatter plots.

One may notice immediately that four year tuition rates rise at a much steadier predictable rate as opposed to two year institutions. The normalized line chart in figure 8 is quite revealing in regards to how tuition may influence enrollment rates. Four year institutions enrollment seem to be completely independent from tuition rates whereas two year enrollment rates highly correlate. One could theorize that persons looking to enroll in two year programs are more likely to make that decision based on cost of tuition. It would seem that after each sharp drop in two year enrollment the institutions lower tuition as an incentive to raise enrollment once again. The same cannot be said for four year enrollment, each time we see a drop in enrolment tuition rates continue to increase over time.

Our above observations hold true when reviewing the scatter plots in Figure 9. It is clear that those enrolling in two year institutions favor lower tuition rates, when referencing the normalized line charts we can infer that the dots indicating higher cost enrolment are the more recent years where we have observed spikes in enrollment. Although we see some fluctuation in the four year plot overall both tuition and enrollment are considered to be steadily rising. Which in the end, disproves our hypothesis for four year enrollment, although it could be argued that the hypothesis is proven true when applied to two year enrollment and tuition.

4 Conclusion

Our original hypothesis assumed that the rise in tuition rates in recent years, would lead to a de-incentivised population of students in the United States to pursue higher education, resulting in pursuing alternative routes of education such as vocational schools, or directly entering the workforce. Our data however, shows that despite the hike in tuition, enrollment is steadily rising denoting that more people are pursuing higher education specifically in four year programs. Interestingly though we found that with two year programs, enrollment was directly correlated with cost. The data also showed that there are times during economic hardship where tuition rates increased at a lower rate, but nonetheless did rise. In times of war we observed a slight decrease in enrollment overall, however numbers seem to quickly recover and continue their increasing trend. It was also noted that there is increased pressure on current students to pursue a higher education past their undergraduate degree because the current job market demands a certain level of specialization for entry level jobs. Thus, in order to have a more successful and stable future, they must pursue a postgraduate education despite tuition hikes. All of this is to say that, in the end, we found that tuition costs had very little impact on the majority of one's decision to pursue a postsecondary education. Further research into four year and two year programs separately could potentially reveal more factors that directly impact enrollment rates in The United States.

5 Appendix

A.1 Total Tuition Filter by Different Institutions (Using D3.js):

```
d3.csv("tuition_1973.csv", function(error, data) {
  data.forEach(function(d) {
    d.year = parseDate(d.year);
    d.cost = +d.cost;
    d.value = +d.value;
  });

  var dataNested = d3.nest()
    .key(function (d) { return d.type })
    .entries(data)

    // var arr = ["1: All Institute","2:Public Institute","3:Private
Institute","4:Private Non profit","5:Private for profit"];
  // div.append('select')
  d3.select('select')
      .attr('id','variableSelect')
      .on('change',variableChange)
    .selectAll('option')
      .data(dataNested).enter()
      //.data(arr)
    .append('option')
      .attr('value',function (d) { return d.key })
      .text(function (d) { return d.key })

var dataFiltered = dataNested.filter(function (d) { return
d.key===d3.select('#variableSelect').property('value') })
   console.log(dataFiltered);
```

A.2 Vocational School Enrollment Comparison Area Chart (Using D3.js):

```
 var browser = svg.selectAll(".browser")
      .data(browsers)
    .enter().append("g")
      .attr("class", "browser");

  browser.append("path")
      .attr("class", "area")
      .attr("d", function(d) { return area(d.values); })
      .style("fill", function(d) { return color(d.name); });

  browser.append("text")
      .datum(function(d) { return {name: d.name, value:
d.values[d.values.length - 1]}; })
      .attr("transform", function(d) { return "translate(" +
x(d.value.year) + "," + y(d.value.y0 + d.value.y / 2) + ")"; })
```

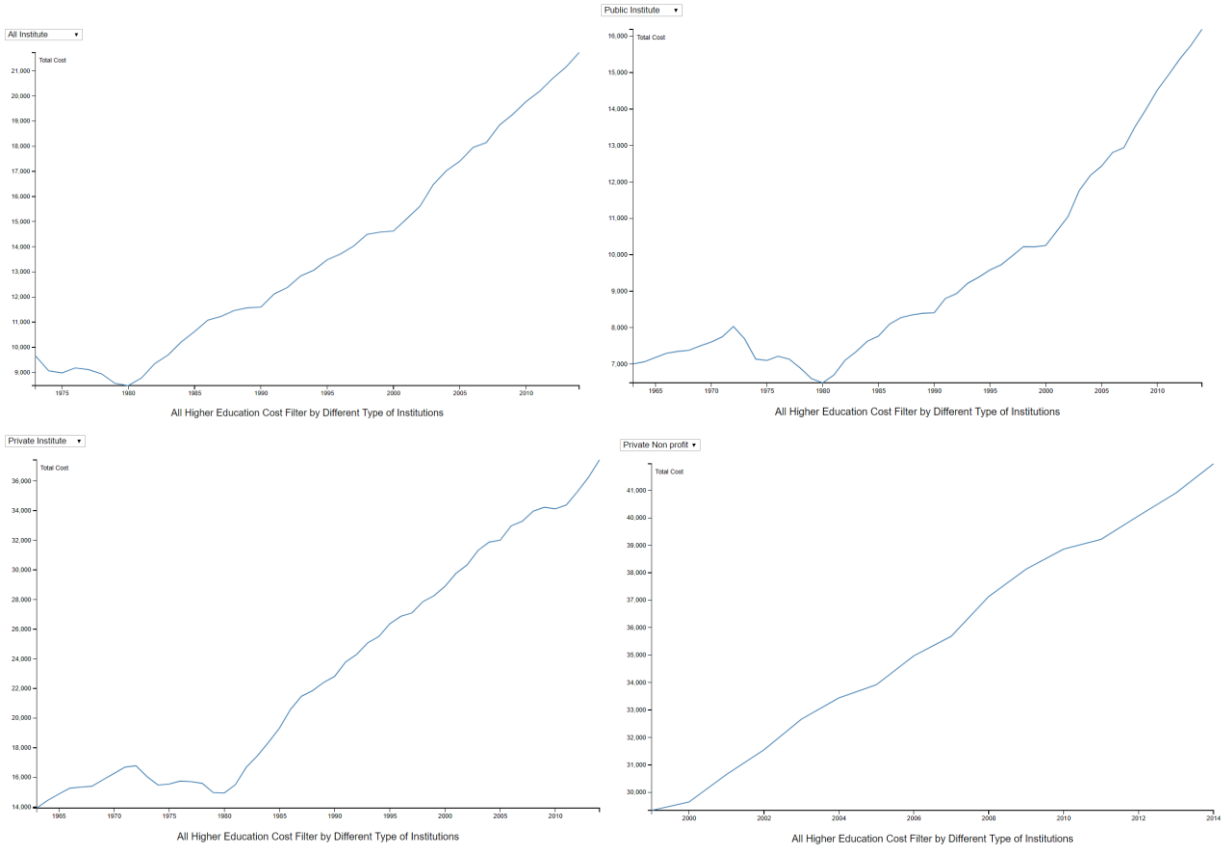
```
        .attr("x", -6)
        .attr("dy", ".35em")
        .text(function(d) { return d.name; });

    svg.append("g")
        .attr("class", "x axis")
        .attr("transform", "translate(0," + height + ")")
        .call(xAxis);

    svg.append("g")
        .attr("class", "y axis")
        .call(yAxis)
        .append("text")
        .attr("transform", "rotate(0)")
        .attr("y", 6)
        .attr("x", -5)
        .attr("dy", "1em")
        .style("text-anchor", "end")
        .text("Enrollment");

    svg.append("text")
    .attr("x", (width / 2))
    .attr("y", (height + margin.bottom + 20))
    .attr("text-anchor", "middle")
    .style("font-size", "16px")
    .style("padding-top","20px")
    .text("Trade School Enrollment Comparison");
});
```

A.3 Additional Vocational School Charts



A.Fig. 1. Filtered results for vocational enrollment, for interactive filtered chart see source fifteen in references [15].

A.4 Code Snippet for Student Enrollment

Loading Data From mySql to Python

```
      conn      =      MySQLdb.connect(host="localhost",      user="mehtake",
passwd="info501", db="sp18i501_5")
      cursor = conn.cursor()
      cursor.execute('SELECT                                                year,
      High_School_Completer,Per_enrolled_to_college_Total,           Two_Year,
      Four_Year,\
      ROUND(Per_enrolled_to_college_Total  *  High_School_Completer/  100)  AS
Total_enrolled_to_college,\
      ROUND(Two_Year       *       High_School_Completer/       100)       AS
TwoY_enrolled_to_college,\
      ROUND(Four_Year       *       High_School_Completer/       100)       AS
      FourY_enrolled_to_college  FROM  university_enrollment_2  WHERE  year  >
      1973');
      data = cursor.fetchall()        #gives the resultset
      cursor.close()

      year = []
      Total_enrolled_to_uni = []
      TwoY_enrolled_to_uni = []
```

```
    FourY_enrolled_to_uni = []

    for row in data:
        year.append(row[0])
        Total_enrolled_to_uni.append(int(row[5]))
        TwoY_enrolled_to_uni.append(int(row[6]))
        FourY_enrolled_to_uni.append(int(row[7]))
```

Normalizing the Student's Enrollment Count

```
    nor_TwoY_enrolled_to_uni = []
    min_Twoyear = min(TwoY_enrolled_to_uni)
    max_Twoyear = max(TwoY_enrolled_to_uni)

    for i in TwoY_enrolled_to_uni:
        nor_TwoY_enrolled_to_uni.append((i-min_Twoyear)/(max_Twoyear-
min_Twoyear)*(100-0) + 0)
```

Plotting Line Chart

```
    plt.plot(year,    nor_TwoY_enrolled_to_uni,    linestyle='--',    color
='green', label="Two Year Enrollment")
    plt.plot(year, nor_FourY_enrolled_to_uni, linestyle='-', color = 'red',
label='Four Year Enrollment')
    plt.xlabel('Years')
    plt.ylabel('Enrollment normalized')
    plt.title("University Enrollment in diff streams in normalized way")
    plt.xticks(np.arange(min(year),max(year), 10))
    plt.legend()
```

Plotting Multi-Stack Bar Chart

```
    fig = plt.figure(figsize=(15,6))
    ax = fig.add_subplot(111)
    width = 0.4
    years = np.arange(len(year))

    rects1  =  ax.bar(years,  Twoy_males_count,width,  color='b',label="Two
Year Male Enrollment")
    rects2  =  ax.bar(years,  Foury_males_count,width,  color='b',alpha  =
0.5,label="Four Year Male Enrollment",
              bottom=Twoy_males_count)

    rects3      =      ax.bar(years+width,      Twoy_females_count,width,
color='r',label="Two Year Female Enrollment")
    rects4 = ax.bar(years+width, Foury_females_count,width, color='r',alpha
    = 0.5,label="Four Year Female Enrollment", bottom=Twoy_females_count)

    ax.set_xlabel('Years')
```

```
ax.set_ylabel('Enrollment in Thousand')
ax.set_title("University Enrollment")
ax.set_xticks(years+width)
ax.set_xticklabels( [i for i in year],rotation =90)
```

Plotting Pie Chart

```
import matplotlib.pyplot as plt
import numpy as np
import csv

tuition = []
dorm = []
board = []
labels = ['Tuition', 'Dormatory', 'Board']
file = open('Tution_fees.csv','r')
reader = csv.reader(file, delimiter=',')
file.readline()

for row in reader:
    tuition.append(int(row[5]))
    dorm.append(int(row[8]))
    board.append(int(row[11]))

sum_tuition = sum(tuition)
sum_dorm = sum(dorm)
sum_board = sum(board)
pie_chart = [sum_tuition,sum_dorm, sum_board]

plt.close()
plt.pie(pie_chart, labels = labels, shadow = True, autopct='%1.1f%%')
plt.title("Breakdown of average cost of postsecondary institutions in
the US\n\n")
plt.axis('equal')
plt.xticks(())
plt.yticks()
plt.show()
```

6 References

1. Patterson, A.: Purdue sets record for student enrollment, graduation rate, https://www.purdue.edu/newsroom/releases/2017/Q3/purdue-sets-record-for-student-enrollment,-graduation-rate.html.

2. Seltzer, R.: Net Price Keeps Creeping Up, https://www.insidehighered.com/news/2017/10/25/tuition-and-fees-still-rising-faster-aid-college-board-report-shows.

3. National Center for Education Statistics: Table 302.10. Recent high school completers and their enrollment in college, by sex and level of institution: 1960 through 2016, https://nces.ed.gov/programs/digest/d17/tables/dt17_302.10.asp.

4. The Chronicle of higher Education: Tuition and Fees, 1998-99 Through 2017-18, https://www.chronicle.com/interactives/tuition-and-fees.

5. Inflation Calculator: Inflation Rate between 1999-2014, http://www.in2013dollars.com/1999-dollars-in-2014?amount=65000.

6. Statista: Average household income 1990-2016, https://www.statista.com/statistics/200838/median-household-income-in-the-united-states/.

7. The College Board: Published Prices – National, https://trends.collegeboard.org/college-pricing/figures-tables/published-prices-national.

8. History in Pieces: U.S. Military Personnel 1954-2014, https://historyinpieces.com/research/us-military-personnel-1954-2014.

9. Stage, F.K., Hossler, D.: Differences in family influences on college attendance plans for male and female ninth graders. Research in Higher Education. 30, 301–315 (1989).

10. National Center for Education Statistics: Table 303.80. Total postbaccalaureate fall enrollment in degree-granting postsecondary institutions, by attendance status, sex of student, and control of institution: 1967 through 2026, https://nces.ed.gov/programs/digest/d16/tables/dt16_303.80.asp.

11. Cohen, A.M., Brawer, F.B.: The American community college. Jossey-Bass, San Francisco (2003).

12. Thompson, D.: This Is the Way the College 'Bubble' Ends, https://www.theatlantic.com/business/archive/2017/07/college-bubble-ends/534915/.

13. Chen, L.: Part Time Interactive Chart, http://lishaux.com/d3/part_time.html.

14. Selingo, J.J., Helvey, K., Heifetz, J., Gielan, M.: Two-Thirds of College Grads Struggle to Launch Their Careers, https://hbr.org/2016/05/two-thirds-of-college-grads-struggle-to-launch-their-careers.

15. Chen, L.: Vocational Enrollment Filtered Chart, http://lishaux.com/d3/fliter.html.