



Dashboards, Scorecards & Visualization

Team Project ‘Storytelling’

MSIS 2629 – Class of Winter 2017

Project Report submitted to:
Professor: Michael Schermann

Project Report submitted by:
Team#7:

Malya Jain,
Roopa Daga,
Sayari Ghosh,
Vinamrata Dake

Table of Contents

Project outline	4
Project motivation.....	4
Why we chose this project? Problem Statement.....	4
Audience	5
Claim	5
Visualization Analysis	6
Dashboard 1	6
Dashboard 3	7
Dashboard 4	8
Dashboard 5	8
Dashboard 6	8
Use Case 1 - Bob is looking for selecting a university for his graduate studies in Computer Science...	9
Description.....	9
Final Decision.....	14
Use Case 2 - Alice is looking for an undergraduate study in Chemical Engineering	14
Description.....	14
Final Decision.....	19
Calculated Fields and Parameters used.....	20
1. SAT Score consideration for Undergraduate.....	20
2. Undergraduate Total Fees:	21
3. Graduate/Undergraduate Payback period.....	21
4. Undergraduate/Graduate Percent Yield.....	22
5. Graduate/Undergraduate Acceptance Rate.....	22
7. Racial diversity in undergraduate enrollment	23
African/American Enrolled in Undergraduate.....	23
American-Indian Enrolled in Undergraduate	24
Asians Enrolled in Undergraduate.....	24
Hispanic/Latino Enrolled in Undergraduate	24
Native/Hawaiian Enrolled in Undergraduate	24
Nonresident Alien in Undergraduate	25
Two/More Races Enrolled in Undergraduate.....	25

Unknown Race Enrolled in Undergraduate	25
White Enrolled in Undergraduate	25
Project Diary.....	26
Data collection, aggregation and cleaning	26
Analysis and Depiction of SAT 25th and 75th Percentile Scores	26
Analysis of Rankings and Ratings:.....	27
Calculating the Payback Period.....	27
Calculating rise in fees for undergraduate students over four years	28
Location Quotient	28
Data Source and References	29

Project outline

The project consists of analyzing the data from educational institutions in USA for the year 2014-15 and identifying the trends on how prospective students select the universities basis multiple factors. The visualization and story have been prepared using Tableau Desktop software. It is an informative dashboard that helps the students to compare from over 185 universities and almost 20 variables such as location, GRE Score, SAT scores accepted, estimated enrollment, tuition fees, admission yield, living expenses, job opportunities etc. The visualization will help international students to choose schools for graduate and undergraduate studies as per their requirements and criteria.

Project motivation

To lead a happy life and enjoy the good things, we certainly need to get good education. A great job, a good social reputation are few of the many benefits of being an educated person. Education is a must for a promising and secure future and a stable life.

Choosing the right educational institution is of prime importance for a student. Each student has unique values and needs for their education criteria. Certainly, the measure changes for everyone. This visualization will help students choose the best fit for their requirements basis multiple factors such as tuition fee, degrees offered, enrollment and admissions, enrollment by race/ethnicity etc.

Why we chose this project? Problem Statement

As international students, we all struggle to select, shortlist and finalize the universities, which match our requirements. There are no good websites available on the Internet too that comprehensively can help student with all the criteria. Often, student ends up looking at multiple websites and creating an excel sheet of his/her own.

Further, every student has a different set of parameters and criteria to select a university. For instance, for some students fees might be the deciding factor and for others the ranking of the university might be criteria.

Further, a student always wants to apply to limited universities only, as it incurs cost of application fee, which can be overwhelming for a student considering the exchange rates.

Hence, a student prefers to apply to selected universities only.

Further, this was a parameter where each one of our team members could relate. As international students ourselves, we all have gone through this tedious process. Hence, we decided to create dashboards encompassing multiple criteria that help students to shortlist a university and not rely on multiple websites for incomprehensive data. Hopefully this tool helps the future potential students.

Audience

International students who wish to come to US for engineering undergraduate and graduate degrees.

Claim

International Students can shortlist the universities for Graduate or Undergraduate engineering degrees based on their criteria

Visualization Analysis

Dashboard 1

The first dashboard initiates the story by telling the audience the statistics of the inflow of international students in USA for higher education.

Purpose: The purpose of this dashboard is to look into the top home countries of the international students. The countries have been categorized into top 10 or in the range 10 to 20 and so on. The visualization clearly shows that China holds the first spot of international students (328,547) followed by India (165,918), Saudi Arabia, South Korea and Canada. Potential students can filter the category and look which countries are falling in which category and have a better idea of the home countries of the international students.

Top countries of origin for international students in USA

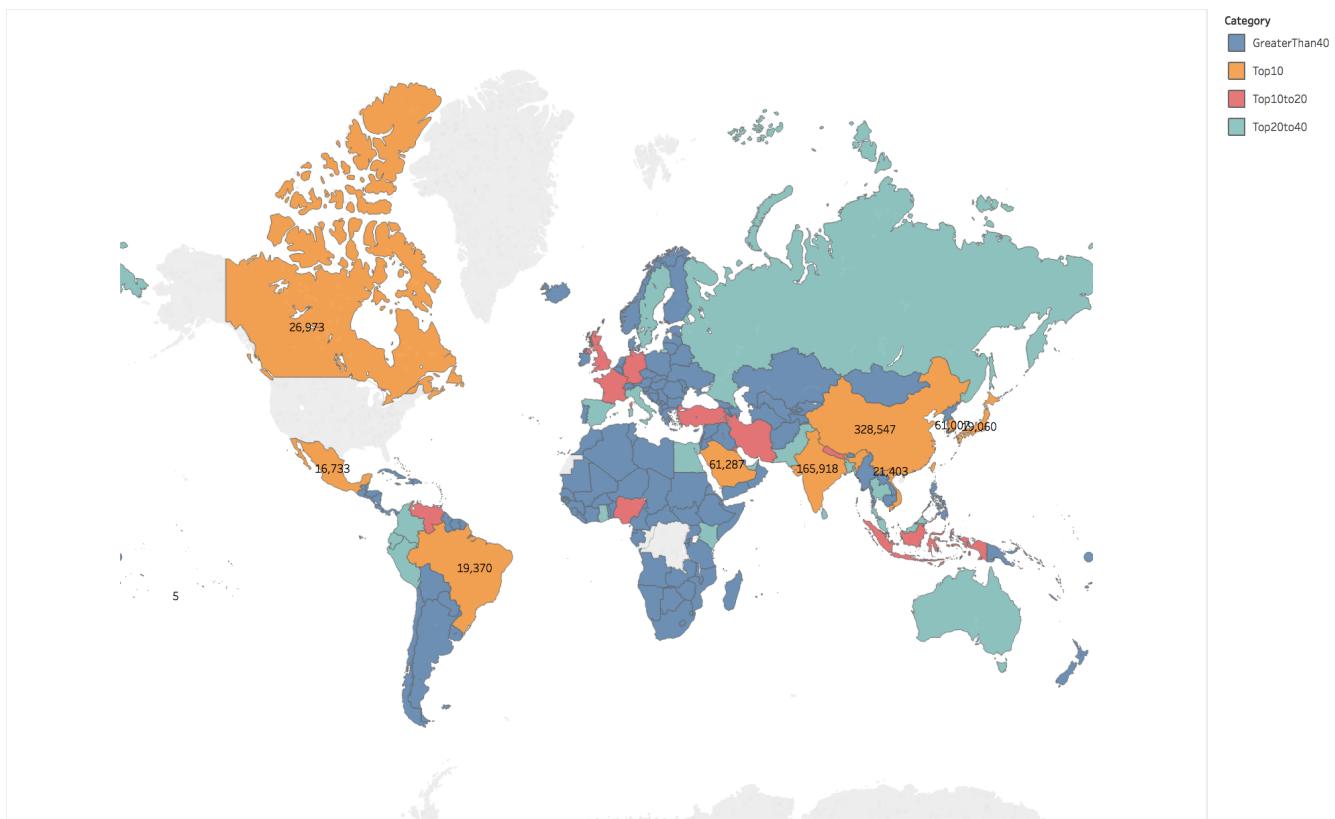


Exhibit 1: Origin of international students coming to USA

Dashboard 2

The next dashboard showcases, which states have the highest inlet of international students.

Purpose: The purpose of this dashboard is to tell the students which states of USA are the hot favorite destinations among the international students. It can be seen that the top 5 destination states are California, New York, Texas, Massachusetts, and Illinois.

Usually, when a student plans to visit US for higher studies, it is important for the student to feel at home. It is extremely important for students to feel secure and know which states have maximum intake of international students.

Top destination states for international students coming to USA

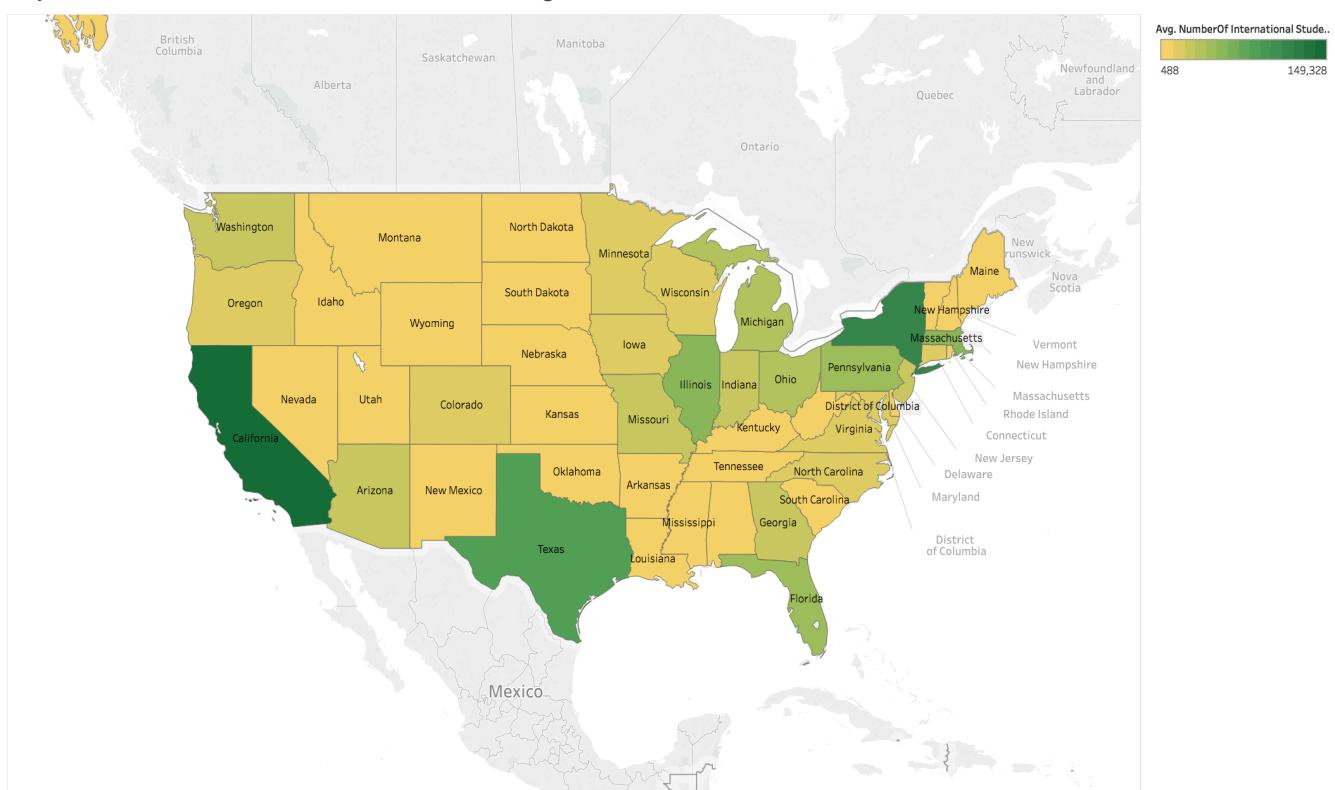


Exhibit 2: Destination States for international students

Dashboard 3

Once the student is convinced that the maximum inflow of international students is from India and China and when the students identify the states where there is the maximum inlet of international students, we move on to the next dashboard.

Purpose: This dashboard enables the student to select the desired majors and see which states have the maximum annual wages. Also, the student can see which states have the maximum

number of employment after the selected engineering degree. This helps the student to select the states that the student can target to go.

Dashboard 4

Once the student selects the target states, the next dashboard helps students to shortlist the universities on basis of multiple criteria such as of ranges of GRE and SAT scores, major of studies, average tuition fees and total cost of living on-campus.

Purpose: The student can input his/her criteria in the interactive filters provided. This will help the students to shortlist 4-6 universities matching their criteria and look at the other details.

Dashboard 5

Next, the student can look at each of these universities and check for more details like ranking, required GRE or SAT scores, tuition fees, enrollment details, acceptance rate, percent yield, number of months it takes to recoup investment made in education and salary details.

Purpose: This dashboard helps the student to shortlist the best two out of the previously obtained list.

Dashboard 6

Finally, the last dashboard enables students to do a comparative analysis and select one final university out of the two taken from the previous dashboard.

Purpose: The student has now shortlisted the final two universities and based on his/her discretion, the student can fix the best university that suits all the requirements.

Use Cases

Let us better understand these by showcasing use cases for graduate and undergraduate potential students.

Use Case 1 - Bob is looking for selecting a university for his graduate studies in Computer Science

Description

Bob completed his undergraduate studies in Computer Science Engineering and after completing his undergraduate studies; he worked in a software development company for two years. He now wants to explore more options and go for higher studies in USA. As he is working while applying for universities, he doesn't have time to look at individual university website and select which university to apply for. Hence, he decides to use our dashboard to get quick results.

Step 1

Bob recently took his GRE exam and obtained a score of 332.

Bob lands on the first dashboard and inputs his Major as Computer Science and gives a GRE range of 329 and 336. The maximum tuition fees in his budget are \$50,000 and he has no restrictions on the living cost. Once he inputs all the criteria, the result he obtains is multiple universities plotted on the map of USA.

He now selects the filter of Public University as compared to Private institution. Since, he has a limited budget, private universities have a higher tuition fees with lesser scholarship and financial aid information. Hence, his criterion is Public university.

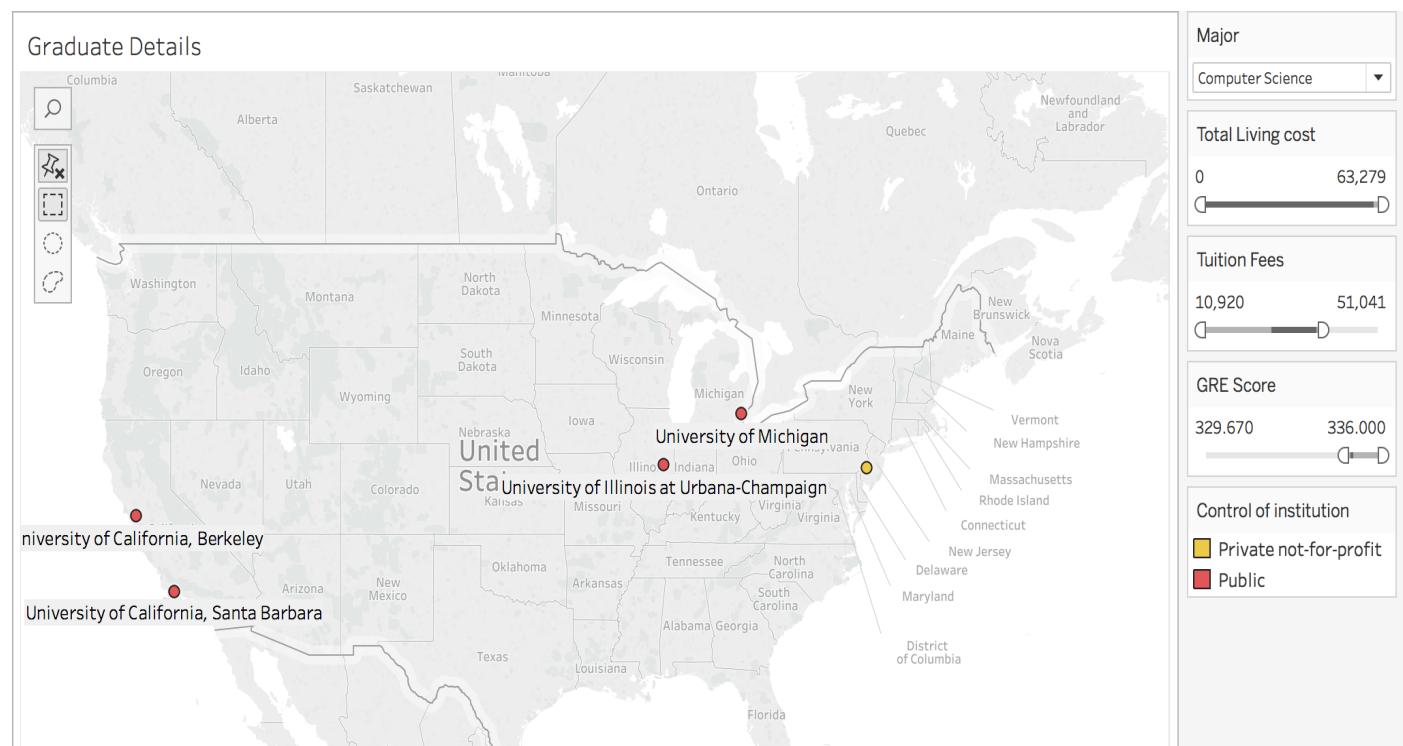


Exhibit 3: Shortlisting of Universities for Graduate studies

Step 2

Next Bob goes to the dashboard where he selects his desired degree i.e. Computer Science and looks at the annual wage and employment distribution in USA.

He looks at the map and notices that California is the best state in terms of annual wage and maximum employment for Computer Science major. On further research, he found out that Silicon Valley is the hottest spot for software developers and startups. Hence, he will now focus on California universities.

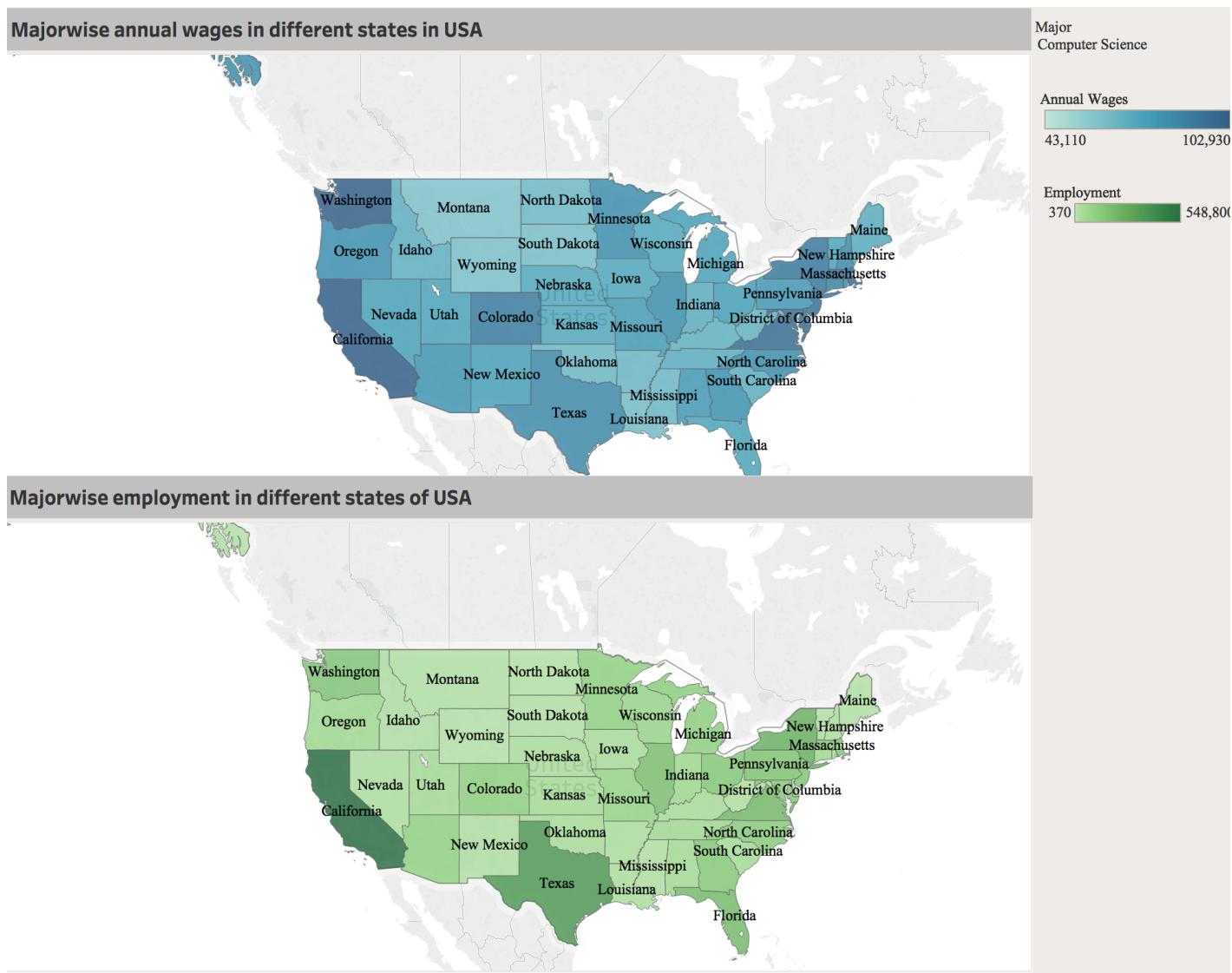


Exhibit 4: Employment and annual wage details for computer science engineering majors in USA

Step 3

Bob now decides to find out the details about individual universities.

He looks at University of California Santa Barbara. He notices that overall rank for this university is 39. However, its acceptance rate is 15. The tuition fees are in range. The payback period is decent of 9 months.

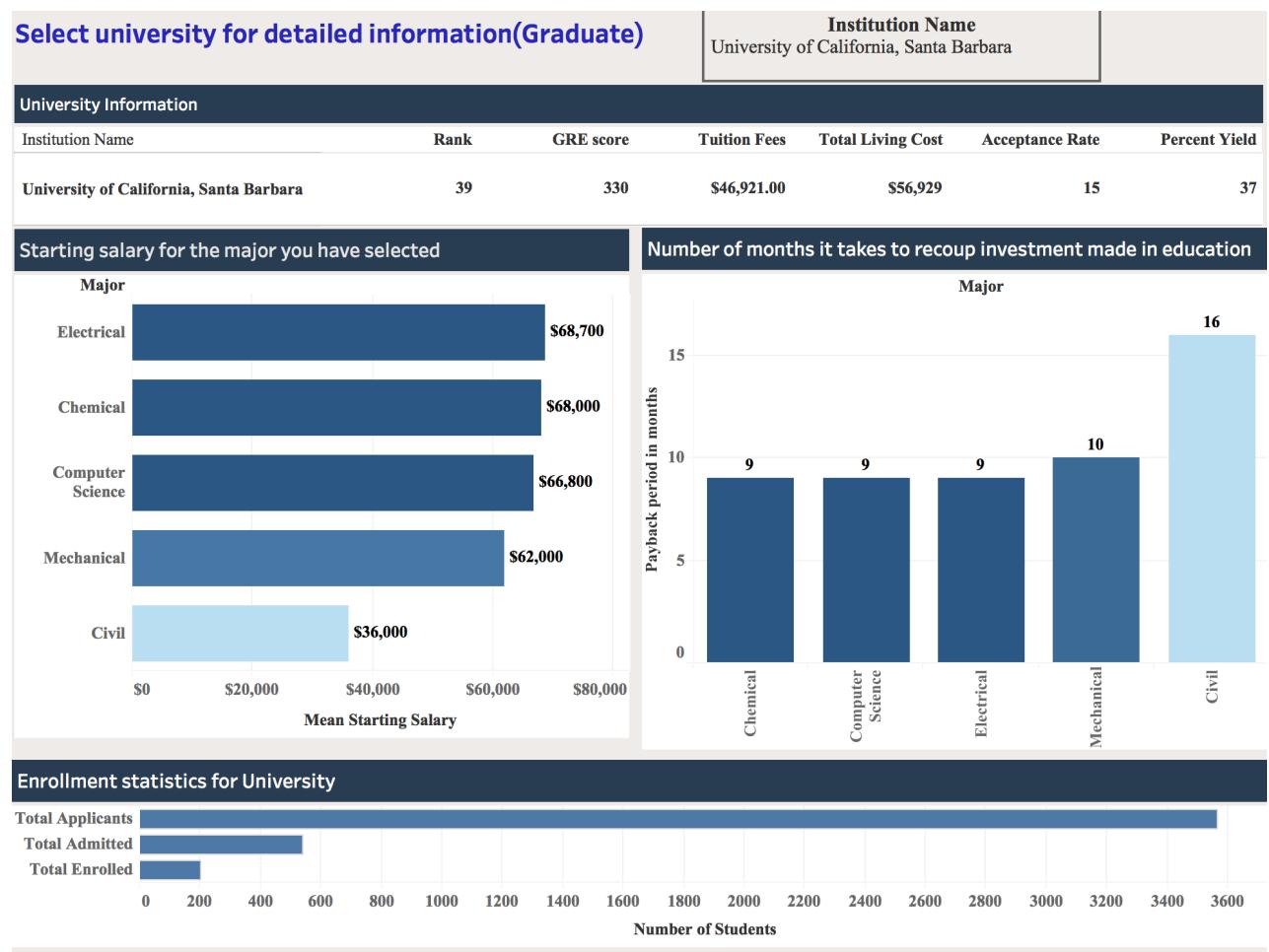


Exhibit 5: University of California, Santa Barbara graduate studies information

Next he looks at details of University of Michigan and tries to compare with the previous university. Similarly, he looks at the individual details of University of Illinois and University of California - Berkeley.



Exhibit 6: University of Michigan graduate studies information

Step 4

Since Bob wants to pursue higher education in California, he wants to do a comparative analysis of UC Santa Barbara and UC Berkeley. The final step is now to compare these two universities and conclude. This is a tough match for Bob. He wants to quickly payback the loan he intends to take and at the same time he wants to spend the least on living cost.

On looking at the informative sheets, he notes that the GRE Score is the same but the tuition fees vary by a considerable amount. Further, it would take 7 months to payback the fees if he chooses UC Berkeley and the median starting salary is also better.

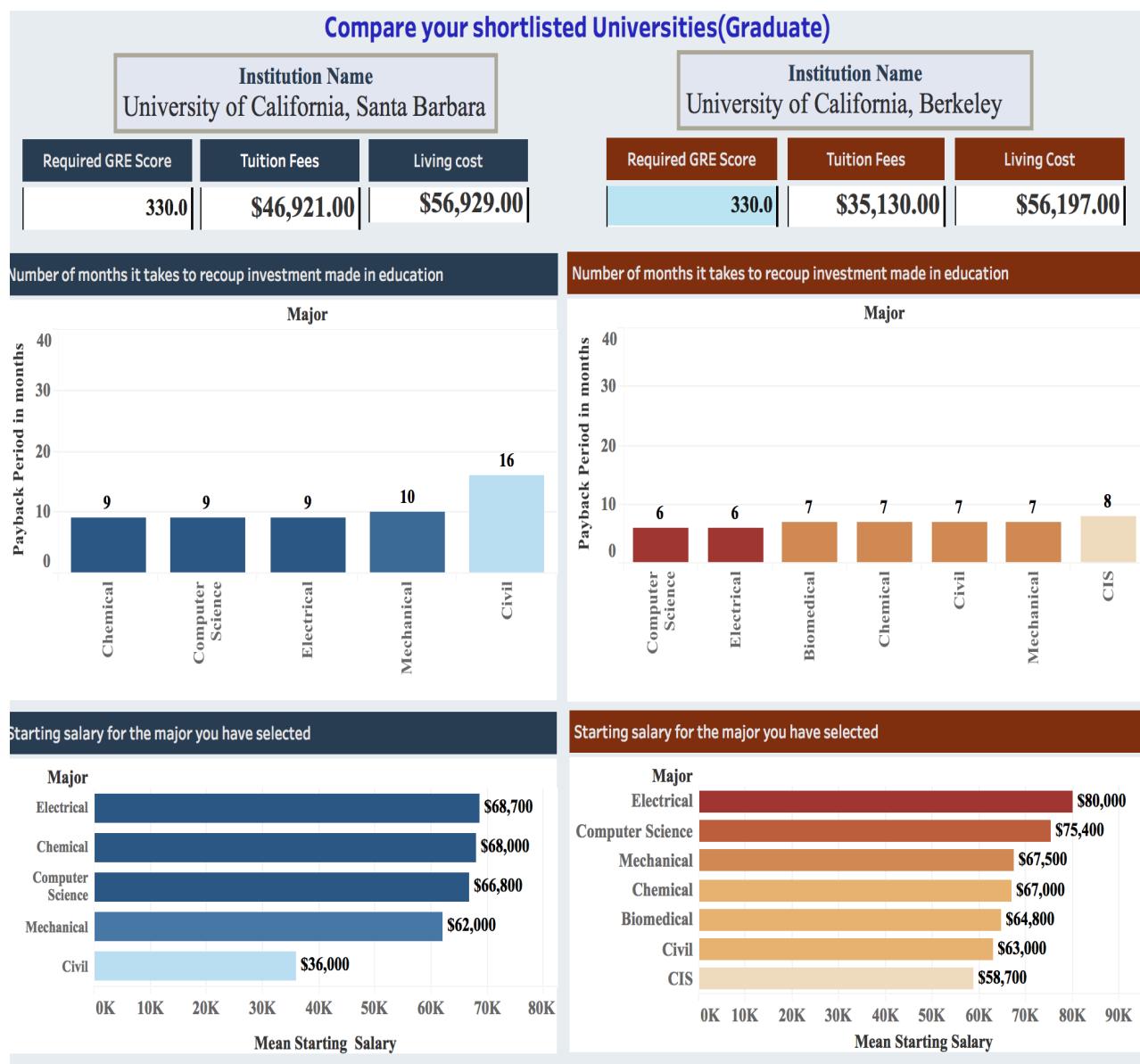


Exhibit 7: Comparison between Universities for graduate studies

Final Decision

Upon some critical thinking and judgments, Bob selects UC Berkeley for pursuing his graduate studies in computer science.

Use Case 2 - Alice is looking for an undergraduate study in Chemical Engineering

Description

Alice is an international student, just graduated from High School. She recently appeared for her SAT exam and now wants to shortlist the universities for her undergraduate studies in Chemical Engineering.

Step 1

Alice appeared for her SAT exam and she got a good score. On the first dashboard, she picks up the following criteria - Reading Score 500, Writing Score 500, Math Score 550. She wants her major to be Chemical Engineering. Further, she wants the range of tuition fee to be an upper limit of \$32,000 and average living cost to be maximum \$38,000. Once she inputs these parameters in the dashboard, she gets a list of 15 universities matching her criteria.

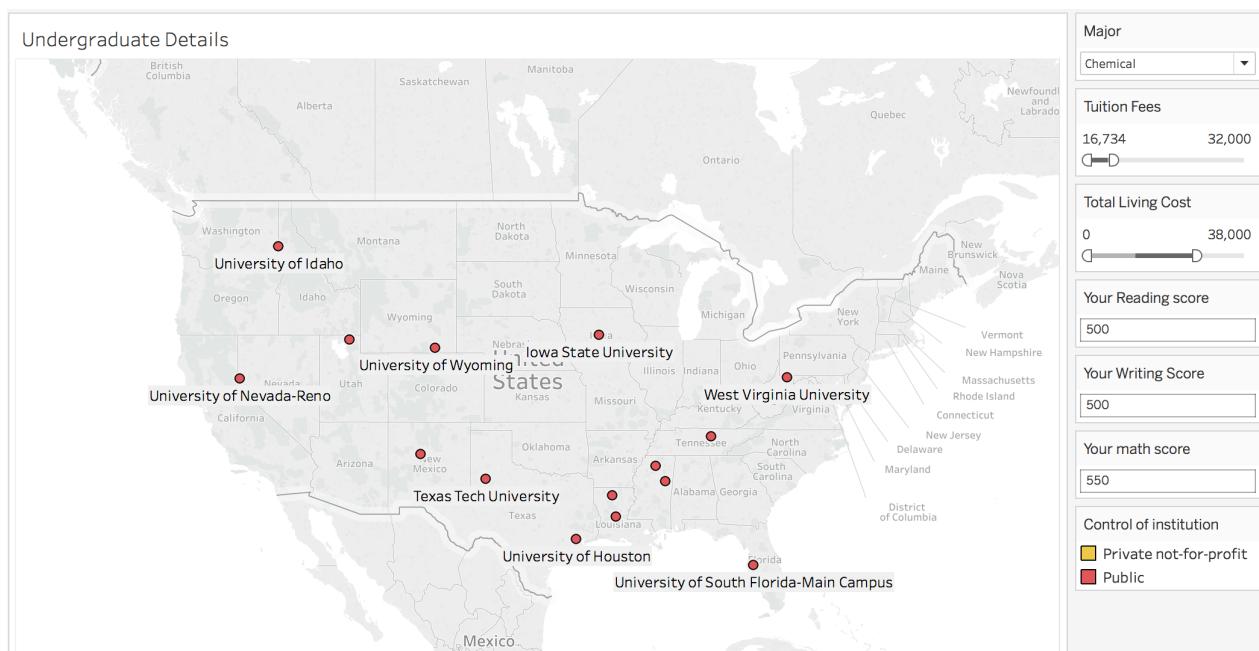


Exhibit 8: Shortlisting of Universities for undergraduate studies

Step 2

Next she looks at the dashboard providing the annual wages and the employment data. She puts the filter as chemical engineering and sees that the best state salary wise and employment wise is Texas for chemical engineers.

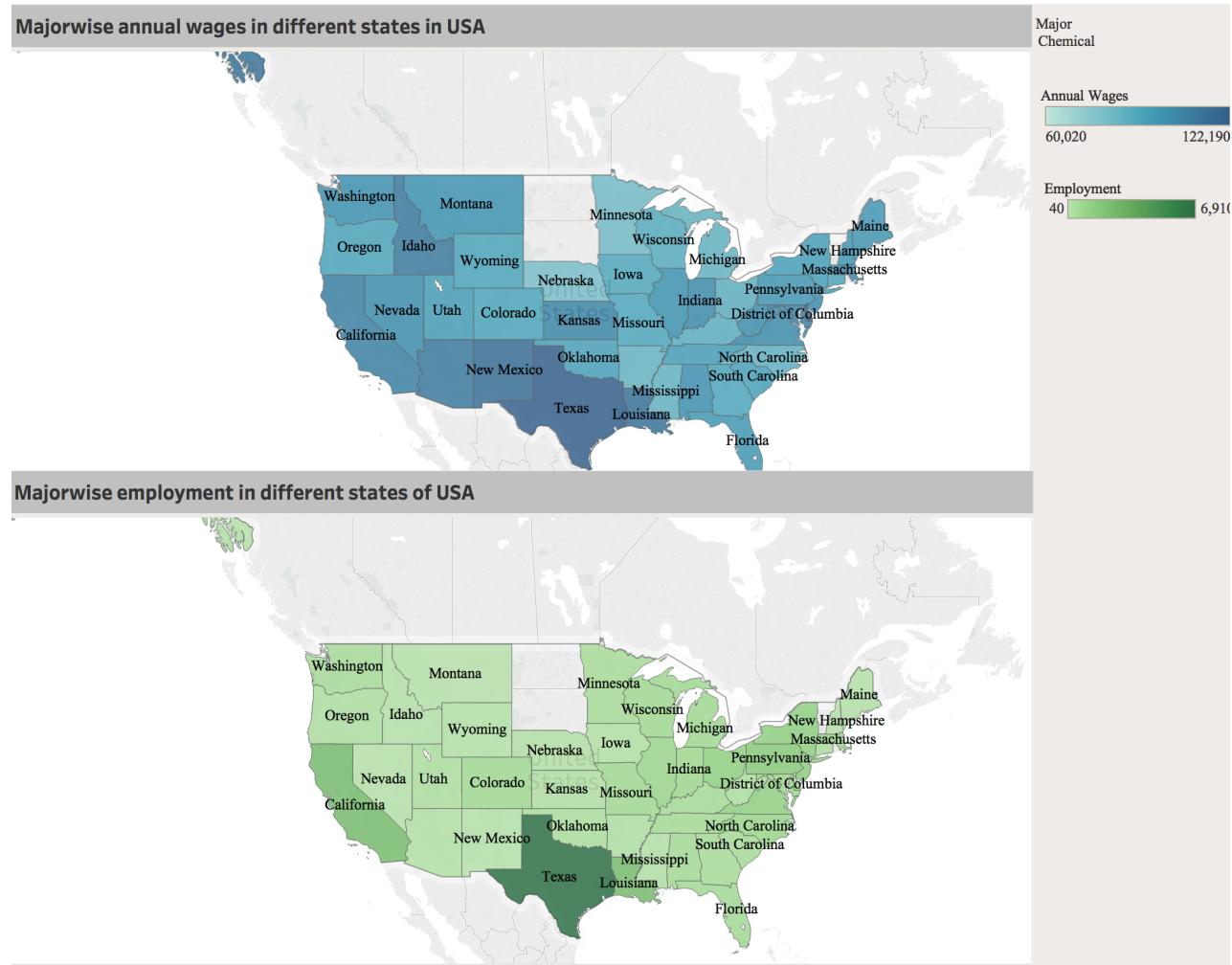


Exhibit 9: Employment and annual wage details for chemical engineering majors in USA

Step 3

Now, she selects 4 universities from the results and checks the information of each and every university. Now she gets a better understanding of details like what are the salaries per major and most importantly how many months it will take her to recoup the investment made in her education.

University of Houston

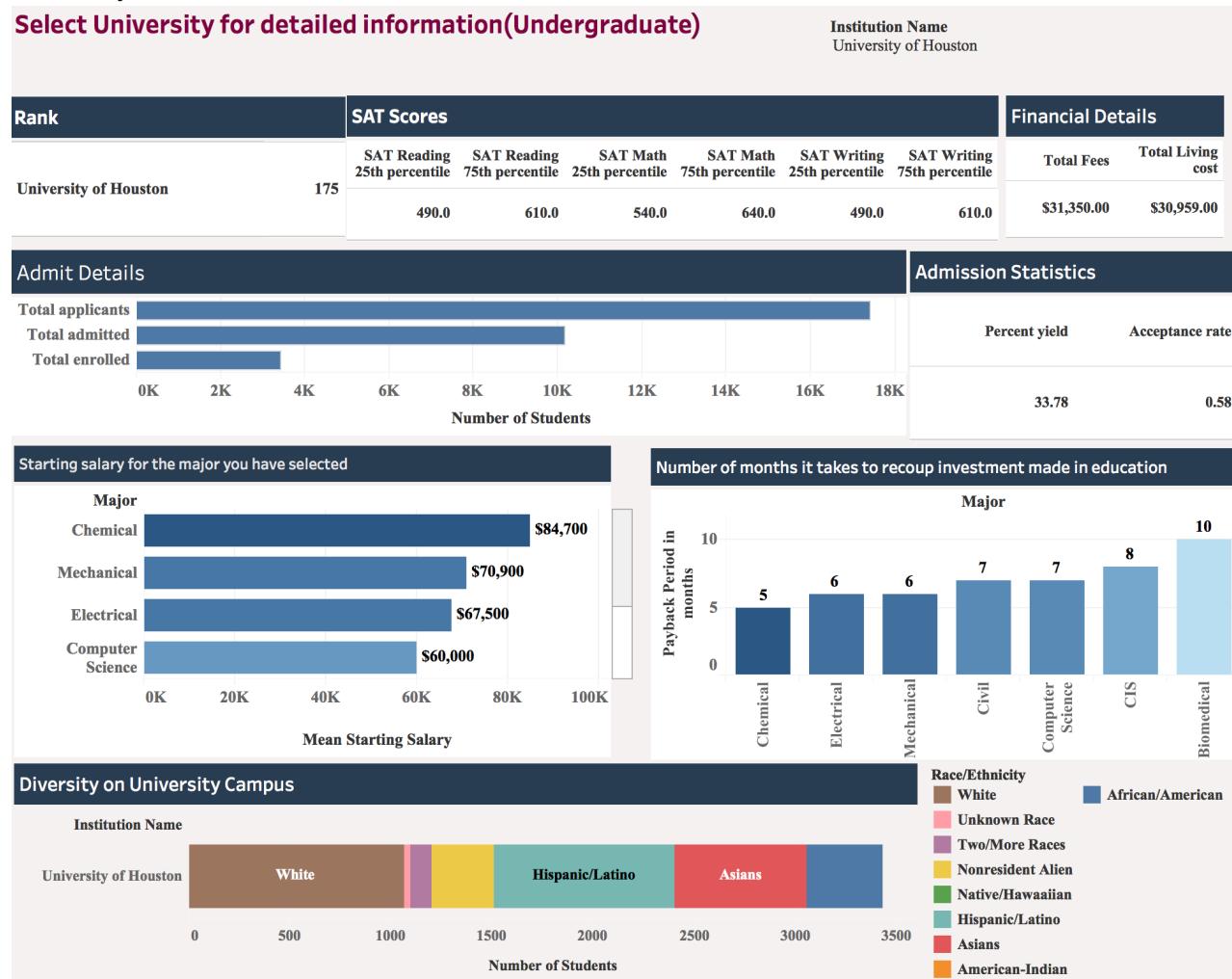


Exhibit 10: University of Houston undergraduate studies information

Louisiana Tech University

Select University for detailed information(Undergraduate)

Institution Name
Louisiana Tech University

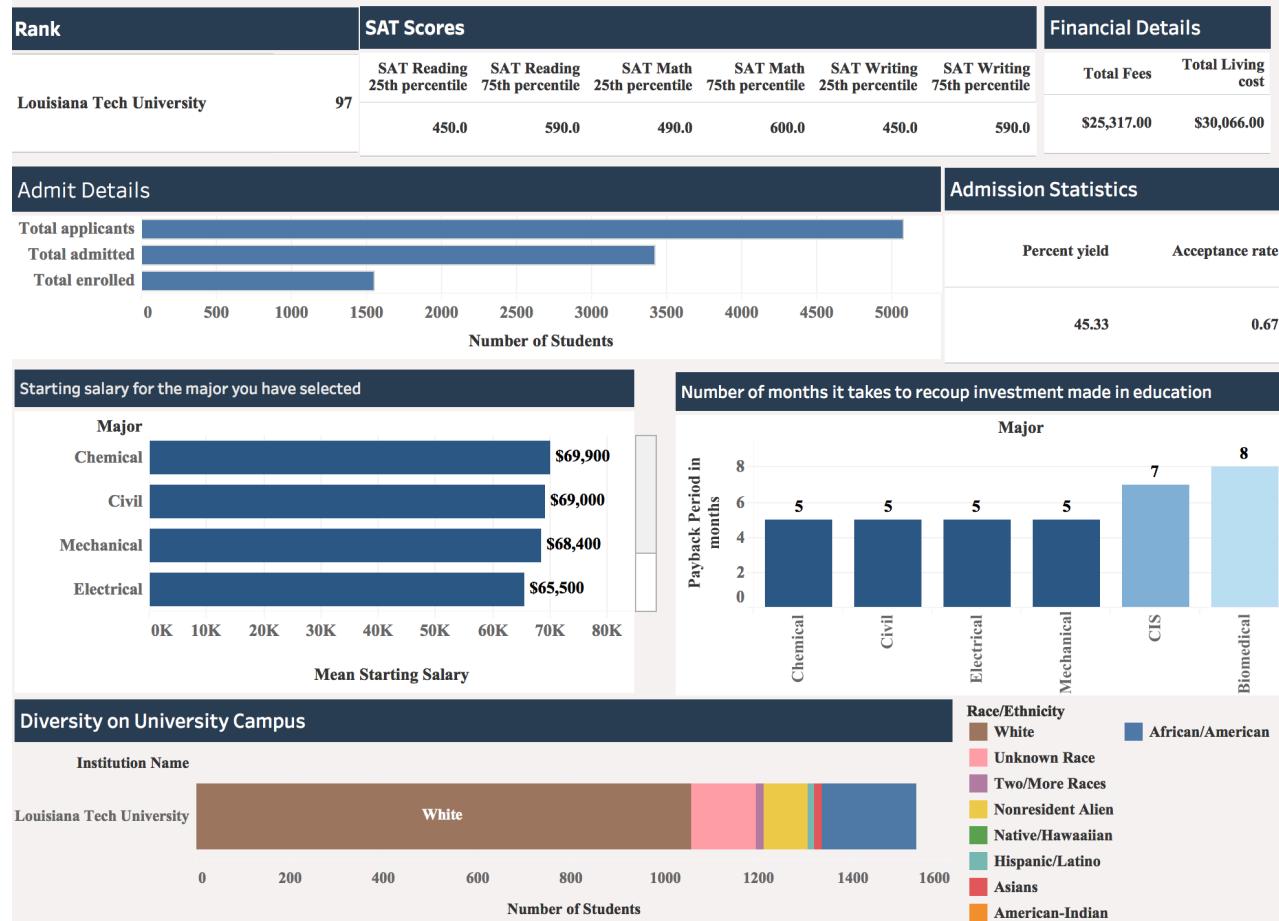


Exhibit 11: Louisiana Tech university undergraduate studies information

Step 4

Looking at the details of individual universities, she is still confused about two Universities University of South Florida and University of Houston and can't pick one. To help her in this dilemma, she can look at the final interactive dashboard where she inputs those two universities.

The parameters such as total fees, ranking, payback period is in favor of University of South Florida. However, the parameters such as median starting salary and SAT scores are in favor of University of Houston.

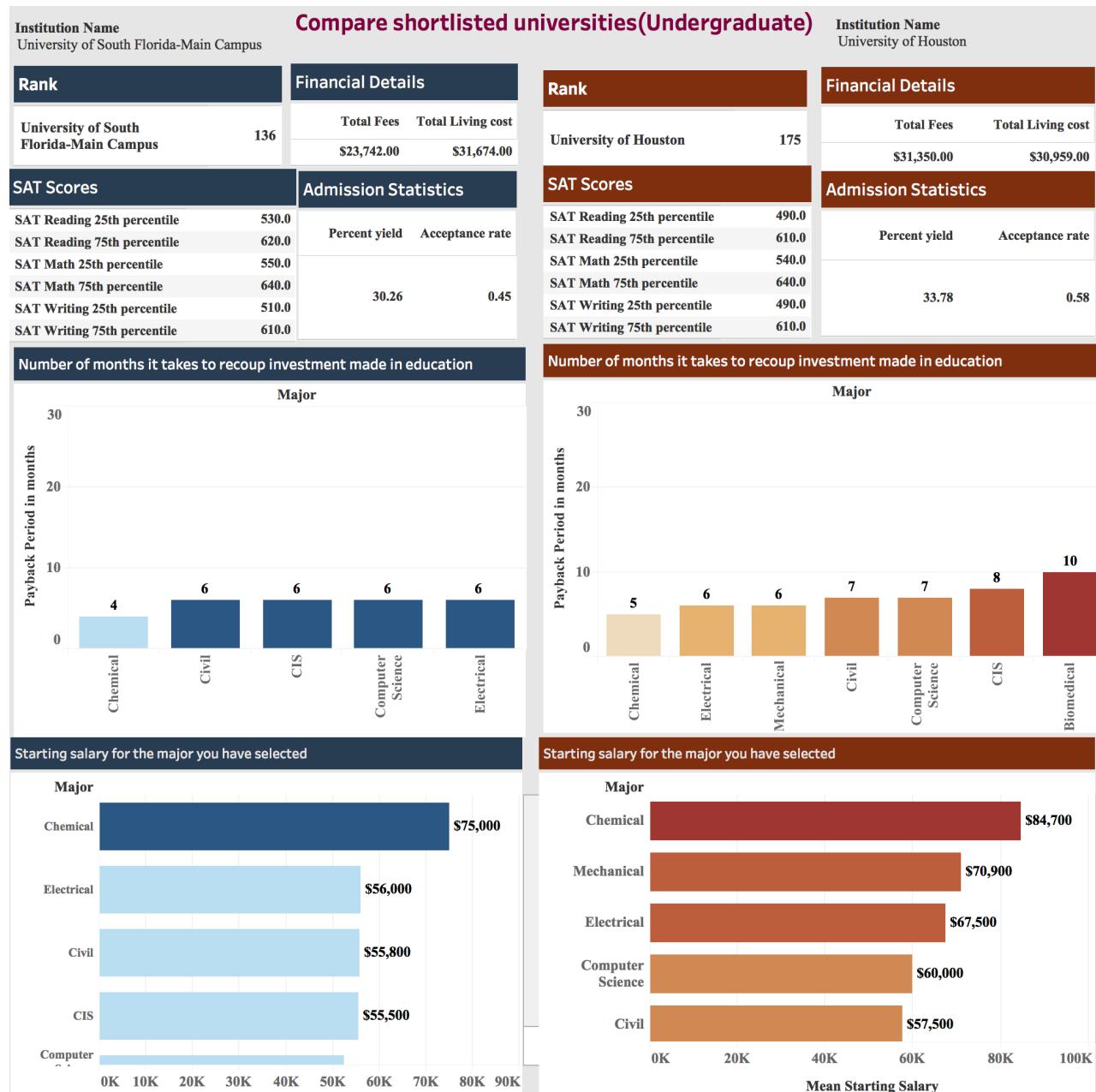


Exhibit 12: Comparison between Universities for undergraduate studies

Step 5

Alice has another dashboard that can help her decide which university to select from which is the comparing the growth rate of tuition fees over 4 years. She notices that percent growth in tuition fees is lesser for University of South Florida as compared to University of Houston.

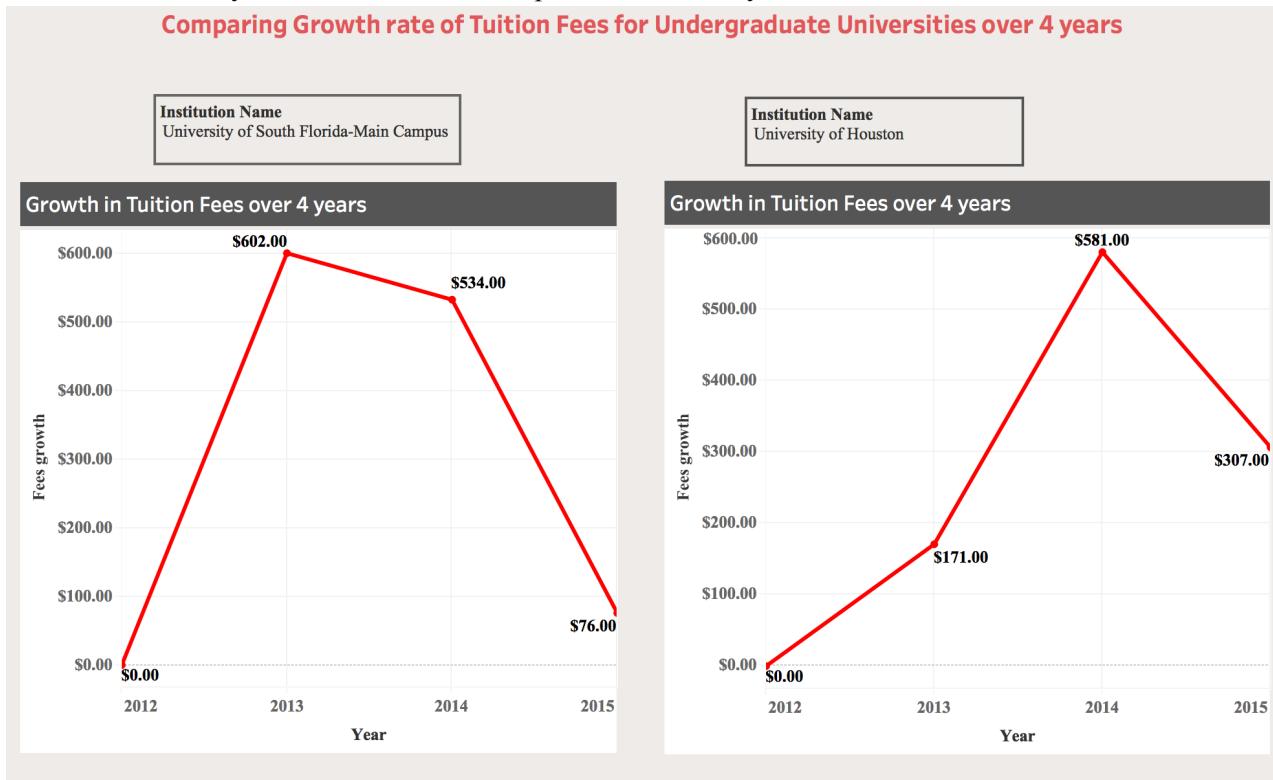


Exhibit 13: Comparison between growth rate in fees between two universities for undergraduate studies

Final Decision

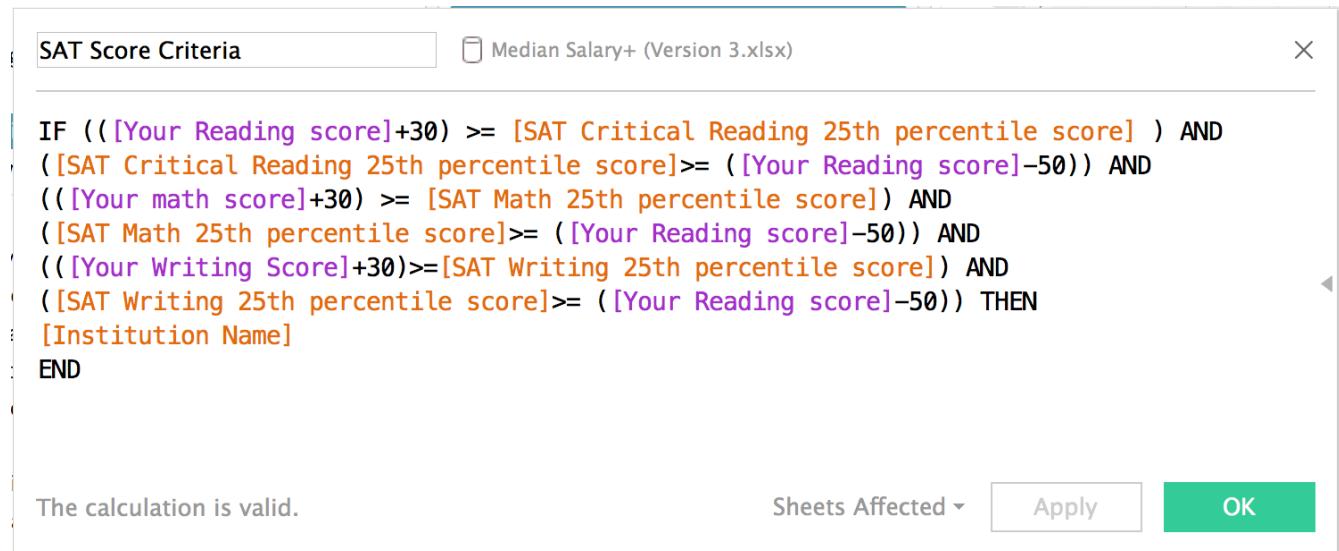
Further, coupled by the decision that Texas is better in employment and annual wages, she selects University of Houston as her final selection of the university.

Calculated Fields and Parameters used

1. SAT Score consideration for Undergraduate

SAT scores are divided into three categories: Reading, Math and Writing. The data shows the score in which 25th percentile of students of that University falls and scores, in which 75 percentile of students of that University fall for each category.

The chances of getting an admit from a good university increases, if the students SAT reading, math and writing scores are in the range of university's accepted scores. Based on the scores entered by student in parameter fields, this calculated field is created.



Below are the parameter fields, which take input as SAT scores from any student:

The screenshot shows a list of parameter fields on the left and their corresponding names on the right:

Parameter	Value
Your Reading score	500
Your Writing Score	500
Your math score	550

On the right, under the heading "Parameters", are the names of the parameters used in the calculated field:

- # Your math score
- # Your Reading score
- # Your Writing Score

2. Undergraduate Total Fees:

This calculated field is created for the summation of undergraduate fees for the 4 years as we had data for each year starting from 2012 to 2015. This field Undergraduate_Total Fees is then used to calculate the Undergraduate Payback period (another calculated field defined below).

The screenshot shows a dialog box for creating a calculated field named "UG_Total Fees". The formula entered is $[UG\ Tuition\ and\ fees,\ 2012]+[ug\ Tuition\ and\ fees,\ 2013]+[UG\ Tuition\ and\ fees,\ 2014]+[UG\ Tuition\ and\ fees,\ 2015]$. Below the formula, a message says "The calculation is valid." There are "Apply" and "OK" buttons at the bottom.

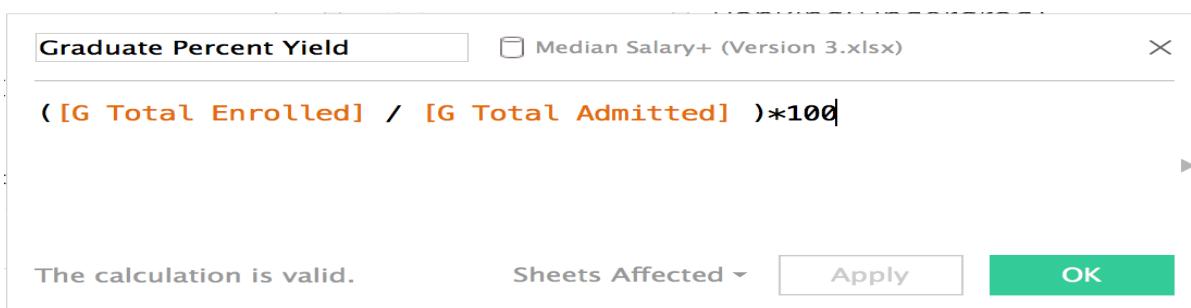
3. Graduate/Undergraduate Payback period

The below calculated field gives us the payback period in months of fees invested in the education. We are not taking into consideration the living cost of the student because this is very subjective and varies from person to person.

The screenshot shows two dialog boxes for calculated fields. The top one is for "Graduate Payback period" with the formula $CEILING(([Graduate\ Tuition\ Fees])/[Start\ Salary])*12$. The bottom one is for "UG_Payback Period" with the formula $CEILING(([UG_Total\ Fees])/[Start\ Salary])*12$. Both dialogs show a message "The calculation is valid." and have "Apply" and "OK" buttons at the bottom.

4. Undergraduate/Graduate Percent Yield

This calculated field provides the yield percentage of that particular university to the international students, which is actually the percent of students who choose to enroll in a particular college or university after having been offered admission. This factor is also important as if it is high that means if given choice students will prefer other universities.



5. Graduate/Undergraduate Acceptance Rate

The below calculated fields provide the international students with the acceptance rate of that university which means what percent of the students who applied to the university is actually admitted. If the value is too low it means that the possibility of getting an admission in that university is less. So the student will decide accordingly.



The calculation is valid.

Sheets Affected ▾ Apply OK

6. Undergraduate Fees growth:

The below calculated field will give the growth in fees over four years for undergraduate students. This will give an idea of growth in fees so that students can decide which schools to opt for in monetary terms.

Results are computed along Table (across).

`ZN(SUM([Fees]) - LOOKUP(ZN(SUM([Fees])), -1))`

The calculation is valid.

Default Table Calculation Sheets Affected ▾ Apply OK

7. Racial diversity in undergraduate enrollment

African/American Enrolled in Undergraduate

The below calculated fields provide us with the percentage of students by their ethnicity studying in that particular university. It will provide the international students with idea about it and they will feel more comfortable going into an university, which has a higher percentage for their own ethnicity.

The different ethnicities we have are: **African, American Indian, Asian, Hispanic/Latino, Native/Hawaiian, Nonresident Alien, Two/More Races, Unknown Race and White.**

African/American Enrolled in UG X

$\text{ABS}([\text{Percent of total enrollment that are Black or African American}]*[\text{UG Total Enrolled}]/100)$

The calculation is valid. Sheets Affected

American-Indian Enrolled in Undergraduate

American-Indian Enrolled in UG X

$\text{ABS}([\text{Percent of total enrollment that are American Indian or Alaska Native}]*[\text{UG Total Enrolled}]/100)$

The calculation is valid. Sheets Affected

Asians Enrolled in Undergraduate

Asians Enrolled in UG X

$\text{ABS}([\text{Percent of total enrollment that are Asian}]*[\text{UG Total Enrolled}]/100)$

The calculation is valid. Sheets Affected

Hispanic/Latino Enrolled in Undergraduate

Hispanic/Latino Enrolled in UG X

$\text{ABS}([\text{Percent of total enrollment that are Hispanic/Latino}]*[\text{UG Total Enrolled}]/100)$

The calculation is valid. Sheets Affected

Native/Hawaiian Enrolled in Undergraduate

Native/Hawaiian Enrolled in UG X

$\text{ABS}([\text{Percent of total enrollment that are Native Hawaiian or Other Pacific Islander}]*[\text{UG Total Enrolled}]/100)$

The calculation is valid. Sheets Affected

Nonresident Alien in Undergraduate

Nonresident Alien in UG X

ABS([Percent of total enrollment that are Nonresident Alien]*[UG Total Enrolled]/100)

The calculation is valid.

Two/More Races Enrolled in Undergraduate

Two/More Races Enrolled in UG X

ABS([Percent of total enrollment that are two or more races]* [UG Total Enrolled]/100)

The calculation is valid.

Unknown Race Enrolled in Undergraduate

Unknown Race Enrolled in UG X

ABS([Percent of total enrollment that are Race/ethnicity unknown]* [UG Total Enrolled]/100)

The calculation is valid.

White Enrolled in Undergraduate

White Enrolled in UG X

ABS([Percent of total enrollment that are White]* [UG Total Enrolled]/ 100)

The calculation is valid.

Project Diary

Data collection, aggregation and cleaning

The most difficult part of this project was data collection. Since the inception of the idea where we wanted to create an informative dashboard on the most important criteria an international student. Being a team of international students ourselves, we could relate to the topic. As a team we brainstormed on what are the most important parameters that an international student takes into consideration before shortlisting the university. We listed the criteria on the decreasing order of priority: exam score (GRE and SAT), Fees, Acceptance rate, racial ethnicity, public and private, area, median salaries per degree etc. This gave us an idea about the top major selection criteria of international students. Secondly, it was required that we collect the relevant data. But first, the list of universities was to be shortlisted. We looked up the references online and shortlisted a total of 185 universities, which had both graduate and undergraduate engineering degrees. The next step was the most difficult one, which was finding the data from the individual universities. We started off by dividing the universities and going on individual websites of the universities and started on data collaboration.

The main problem faced here was that comprehensive data was not available even on the university websites. Hence, we had to look for additional websites to collect data. Those links are mentioned in the references.

Further, we kept on adding scope during the course of the project. The further data added was for inflow of international students per state.

Analysis and Depiction of SAT 25th and 75th Percentile Scores

The main challenge in depicting the scores for Undergraduate was how to show the values of the 25th and 75th percentile scores of the universities in our Dashboard#1.

The student's get one comprehensive score and one score each in Reading, Writing and Critical Reasoning sections. As per the data we gathered, there is data for 25th and 75th percentile for each of the university for each of the three section. We took the following approaches:

1. First we calculated the average of the SAT score of the 25th and 75th percentile scores for each of the sections, and matched it to the SAT Score of the student and gave the output of the list of universities above that range. On further discussions with the team we noted that this data will not accurately tell the students chances of getting into that university as most universities have a cut off of the individual section scores.
2. Second approach we took was to let the student input all the three scores obtained in the sections and compare them with the 25th and 75th percentile scores that is present in the data. We created three categories of High, Medium; Low so that the student can get to know the chances of getting into the university basis the three section scores. The problem with this approach was that it will give the categorizations for each university thrice, which is again not a correct parameter to select a university.
3. Third approach for this was that we take a slider filter for the student to put their ranges of scores in the 25th and 75th percentile, but this made us make 6 filters only for the SAT scores and this is an indication of an ineffective and redundant dashboards.
4. Final approach we finalized was that the student enters the score in all the three sections and we compared all three with the 25th percentile of the university. We created a parameter to give output of only those universities where the SAT score of students falls in the range of the Universities who accepts that score range. Hence, this approach gives the list of all universities, which are in the category of "Safe" for the student.

Analysis of Rankings and Ratings:

The main issue the team was facing was how to show the rankings and ratings of universities. First issue with this attribute was that from the websites where we picked data, the ratings are mentioned basis the contributory parameters like financial aid, faculty ratings etc. Further, it consisted of a smart rating which is a cumulative of the graduate and undergraduate rating. Also, there are few universities, which have a higher ranking for undergraduate and a lower ranking for graduate degrees and vice versa.

Hence, we came to the conclusion that we should take into consideration the ranking instead. Hence, we researched about the top universities for graduate and undergraduate and ranked the universities differently for graduate and undergraduate degrees. As our audience is different, it will give us a better picture.

Calculating the Payback Period

The idea we wanted to represent on our dashboard was to calculate the return on investment. Upon discussing with the team, we noted that the more relatable factor for international students is the payback period. When international students come to the USA for higher studies, because of the high conversion rates to USD, the students usually are more concerned that in how many years they can payback the cost they incur for the studies.

For this we thought of creating a calculated fee for Payback Period. The most important question is how to create the formula. First consideration is that what is the start salary of the students who graduate from a particular degree. Depending on the field of degree, we have the data for the median starting salary. So we calculated the field as what are the fees of the students and then the Median Starting Salary.

Further, we are not taking into consideration the living cost of the student because this is very subjective and varies from person to person. For instance, there is a student who has close relatives who live in the same city, for them, the cost of living is negligible. Also, if a student is used to a lavish lifestyle, the cost of living could be more than double of a student who wants to live in a tight budget. Hence, we have not taken this criteria into consideration and we took only the median starting salary.

For undergraduate, we created a calculated field on the sum of fees from the year 2011 to 2015, because it is a four-year course.

For graduate, we have the data for the full tuition fees for the term of the course

Hence the formula for the Payback Period is:

Fees/Median Starting Salary per degree

Upon plotting this data, we noted that the payback period is coming out to be in years and were very close like 1.02 and 1.2. This is not a correct representation of the payback period. Also, the student will have to calculate that 0.2 years means not 2 months but 2.5 months. Hence, we altered the formula so that we get the payback period in Months.

Hence the formula for the Payback Period is:

(Fees/Median Starting Salary per degree)*12 Months

Calculating rise in fees for undergraduate students over four years

This dashboard requires the analysis of the rise in fees of undergraduate students over the four years. This will give an idea to the students about the average tuition fees of the four years and also incorporate the YOY rise in the tuition fees. This will give students an idea about the estimated increase in the fees. For this we have used the following parameter.

$\text{ZN}(\text{SUM}([\text{Fees}]) - \text{LOOKUP}(\text{ZN}(\text{SUM}([\text{Fees}])), -1))$

The result is a line graph, which can show the trend. Further, the student can select the universities from a dropdown filter provided.

Location Quotient

The location quotient is an idea, which tells the students which states give the best annual wages and best employment for select degrees.

Data Source and References

- <http://nces.ed.gov/ipeds/datacenter/Default.aspx>
- http://www.nafsa.org/_File/_econvalue2016_natl.pdf
- <http://www.iie.org/Research-and-Publications/Open-Doors/Data/International-Students/All-Places-of-Origin/2014-16>
- <https://www.bls.gov/data/home.htm>
- <https://data.bls.gov/oes/#/home>
- <http://graphics.wsj.com/international-students/>
- <https://www.usnews.com/news/articles/2016-05-17/more-stem-degrees-going-to-foreign-students>
- <https://www.statista.com/statistics/233880/international-students-in-the-us-by-country-of-origin/>
- <https://www.usnews.com/education/blogs/college-rankings-blog/2013/06/18/top-ranked-universities-that-grant-the-most-stem-degrees>
- <https://www.ed.gov/stem>
- <http://www.worldwidelearn.com/education-rankings/25-stem-majors-with-the-best-value-2015.html>
- <https://github.com/fivethirtyeight/data/tree/master/college-majors>
- <https://www.niche.com/colleges/rankings/best-colleges-for-engineering/>
- <https://www.brookings.edu/interactives/the-geography-of-foreign-students-in-u-s-higher-education-origins-and-destinations/>
- <http://www.pewresearch.org/fact-tank/2015/06/18/growth-from-asia-drives-surge-in-u-s-foreign-students/>
- <https://www.usnews.com/best-colleges/princeton-university-2627>