9/11/2018

Assignment 1 Submission

Group 8 Tue Batch

INDEX

1. Write Up on Understanding the Data

Well, the assigned College Scorecard data is quite an exhaustive one that can potentially help students and families in analyzing college standings in a holistic way by analyzing the various parameters based on their academic goals. The data is an aggregate of an Integrated Postsecondary Education Data System (IPEDS) by surveys from years 2000-01, 2005-06 and 2009-16 over three collection periods (fall, winter and spring) from all the participating institutes through electronic means and is available for download in US Department of Education Office of Planning, Evaluation, and Policy Development website. It gives insight about the institution, their admission rate based on the ACT/SAT scores, type of courses offered, demographic and other information about the student body, cost of education, federal financial aid, earnings, employment prospects, repayment of loan and completion rate of students.

The College Scorecard data depicts demographic information about their name, minority-serving status and also the SAT and ACT score breakdown for every institution. Moreover, there are attributes which give information about the number of degrees offered by each Department and the total enrollment based on races such as White, Black, Hispanic, Asian, American, native Hawaiian, and non-resident alien students. The data provides more knowledge about the affordability such as the program cost, financial aid such as Pell Grant, Federal loan and school-provided aid which is a boon especially for low-income family. Additionally, it includes valuable information on the performance of the institute such as retention rate of students, earnings of students in terms of median salary after 10 years of entry and earnings over $25000 per year, 6 years after entry. As a comparative analysis of the data we followed a comprehensive approach in terms of arriving at an Index using 5 important parameters that one considers before enrollment into the university. These factors include Cost of attending the college, the completion rate observed in the college, the median earnings of the alumni, the scholarship offerings and the entrance exam scores.

When we compared the set of states allotted to our group 8 with rest of the data with an equal weightage to each of these influencing parameters, we coincidently get to see that our set of states stand at highest index value compared to rest of them (Graph 1 below). It indicates that our set of data has an average index of 45% as against the average no. of 43% for the overall data. We see that the state of Pennsylvania stands highest in terms of the Completion Rate, Earnings recorded and the overall Index (Graph 2 below). The state of Louisiana, in spite of the decent scholarship access scores mediocre in terms of Earnings and completion rate. We also observe that after Pennsylvania, it is the state of West Virginia which stands at impressive levels in terms of the 5 identifies factors. The Graph 3 below also validates our general notion about the public institutions offering higher Return on Investment as compared to Private counterparts. Interestingly, we also get to observe that the completion rate is higher in Private institutions as compared to Public ones, in spite of the higher cost involvement (Graph 4 below). We also explored the probability of observing any possible trend between the Pell Grant, Federal Loans against the Earnings for these two set of institutions i.e., Public and Private. However, apart from the higher earnings aspect for Public institutions we don't get to see any significant relation.

And so, we conclude with the fact that the when we pick up the most influential factors into building the universal index for states, coincidently, the 5 states in our group stand out of

the sample and give the potential readers, who in this case most likely happen to be the aspiring students or/and their parents, who wish to make an informed decision about college selection) a good way of comparing the overall positioning of universities.

2. The Single No. visual for our group of data

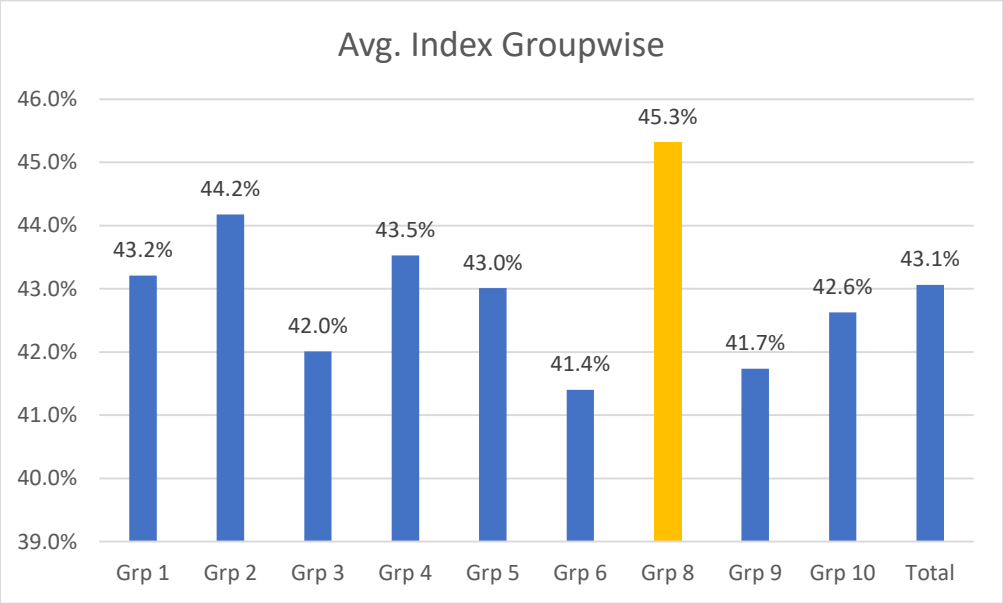
Our group of states stands at the Top among the all 10 groups, in terms of the index-based ranking approach devised here.

3. The One Visualization through IBM-Watson

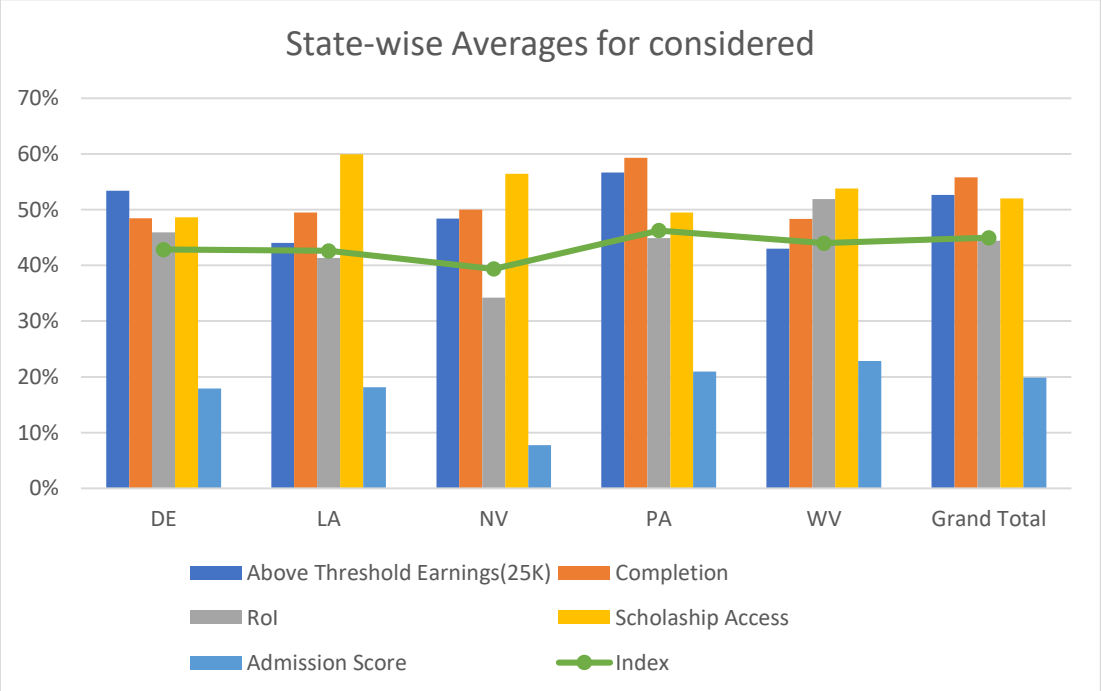
The picture above helps us infer that the **Index value (Purple)** value for **Group 8** states comes out to be the highest with an comprehensive collection of all influential factors. And as evident from the graph above, this is primarily because it scores almost the highest on at least 3 of the 5 parameters considered.

4. Visuals for Analytical Findings using Excel (Additonal)

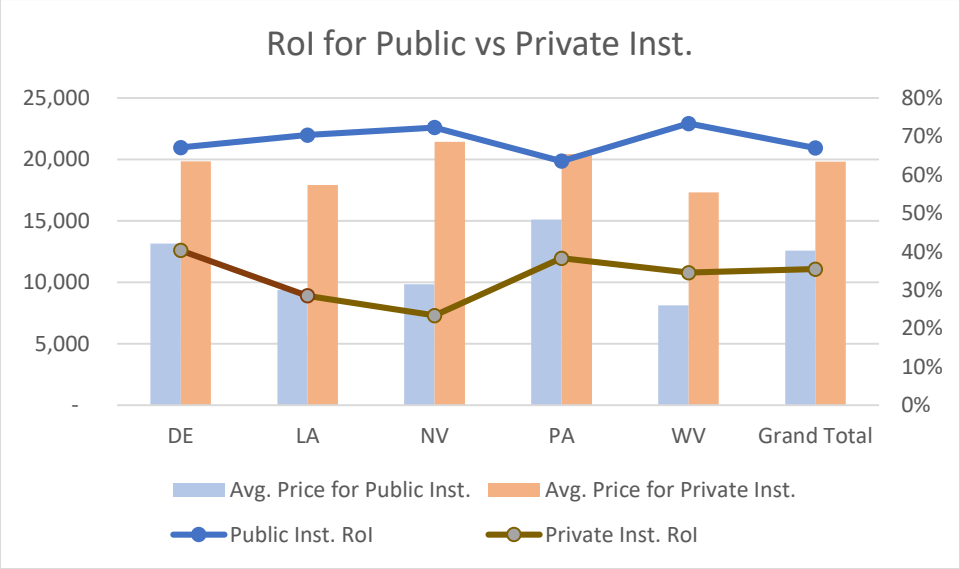
Graph 1:



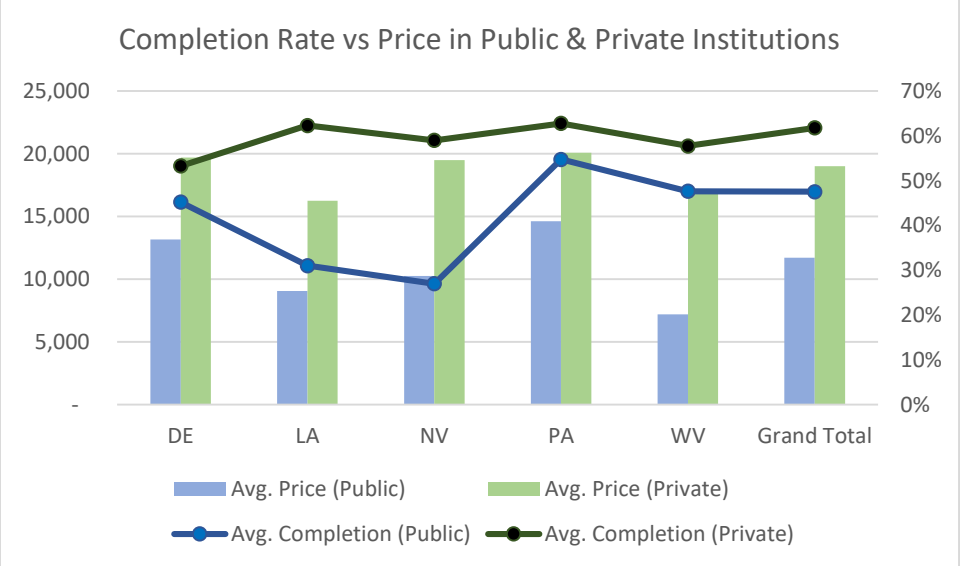
Graph 2:



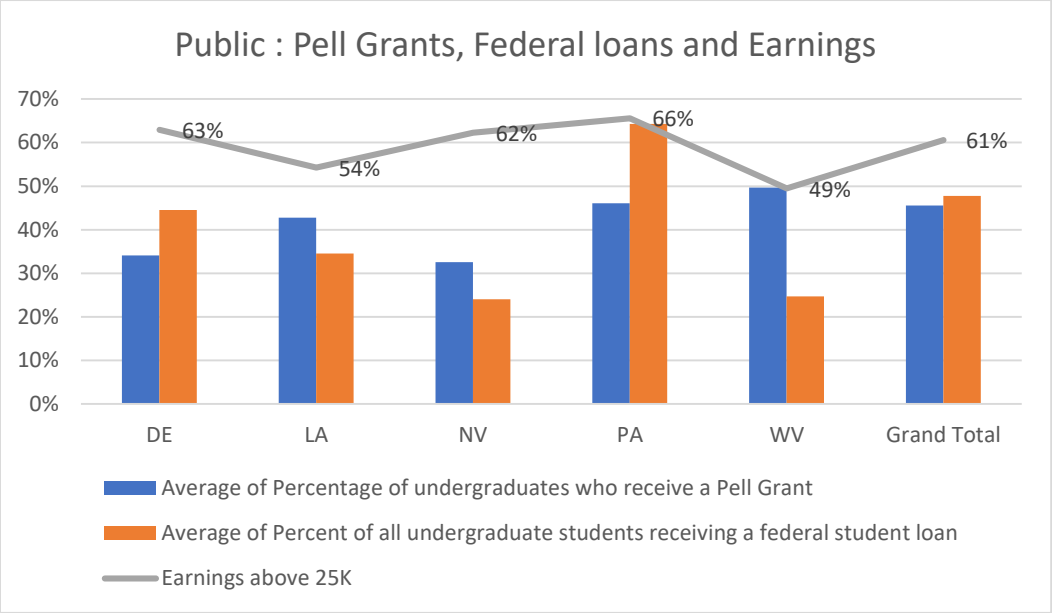
Graph 3:



Graph 4:



Graph 5:



Graph 6: