Question 1

Table 1

Data Understanding of Variables

Variables	Data Type	Number of Observations	Summary Statistics
Name	Nominal	231	N.A.
Location	Nominal	231	N.A.
State	Nominal	231	N.A.
Rank	Ordinal	231	N.A.
Tuition and fees	Nominal	228 (Missing Value = 3)	Mean = $33,427 Median = $31,744 Mode = $26,334 Minimum = -$51,265 Maximum = $100,000
Undergrad Enrollment	Nominal	226 (Missing Value = 5)	Mean = 15,009 Median = 12,929 Mode = 6,883 Minimum = 1,001 Maximum = 54,513

Question 2

From **Table 1**, we identified the following data quality issues and the treatment methods.

a) Missing Values for Tuition and fees and Undergrade Enrolment. (**Figure 1**)

- There are 2 methods to input the missing values in Figure 1. We could either use the mean values of the universities' Tuition and fees/Undergrad Enrollment around the same location or search the data from online sources for the input since US universities information is readily available on the internet. The online information is also creditable because they are provided by 2 US Department of Education. (Integrated Postsecondary Education Data System (IPEDS) and College Scorecard data).
- To have more accurate visualisation charts for the studies, we select the 2nd option to use the online information of various universities found on the internet. The main reason for this selection is mainly because Tuition and fees and

Undergrad Enrollment data are published formally on the internet and information backed by the **US Department of Education**.

Figure 1

Missing Values in the dataset (Before)

Name	Location	Sta	Ra	Tuition and fe	Undergrad Enrollmer
University of Kansas	Lawrence	KS	118	$25,932	
Illinois State University	Normal	IL	152	$20,886	
University of California--Merced	Merced	CA	152		6,237.00
Maryville University of St. Louis	St Louis	MO	164		2,795.00
Edgewood College	Madison	WI	171	$27,530	
Lipscomb University	Nashville	TN	176		3,030.00
University of Alaska--Fairbanks	Fairbanks	AK	202	$22,469	
University of Missouri--St. Louis	St. Louis	MO	220	$26,277	

Missing Values in the dataset (After)

Name	Location	Sta	Ra	Tuition and fe	Undergrad Enrollmer
University of Kansas	Lawrence	KS	118	$25,932	18,438
Illinois State University	Normal	IL	152	$20,886	17,955
University of California--Merced	Merced	CA	152	$13,691	6,237
Maryville University of St. Louis	St Louis	MO	164	$27,290	2,795
Edgewood College	Madison	WI	171	$27,530	1,339
Lipscomb University	Nashville	TN	176	$26,000	3,030
University of Alaska--Fairbanks	Fairbanks	AK	202	$22,469	4,259
University of Missouri--St. Louis	St. Louis	MO	220	$26,277	15,192

b) Tuition and fees with negative values. (**Figure 2**)

- As it is generally not possible for negative values in the Tuition and fees variable, affected data of Duke University and Texas A&M university-College Station had to be revised from negative to positive Tuition and fees.

Figure 2

Negative Tuition and fees (Before)

Name	Location	State	Rank	Tuition and fees	Undergrad Enrollment
Duke University	Durham	NC	8	-$51,265	6,639.00
Texas A&M University--College Station	College Station	TX	74	-$28,768	48,960.00

Negative Tuition and fees (After)

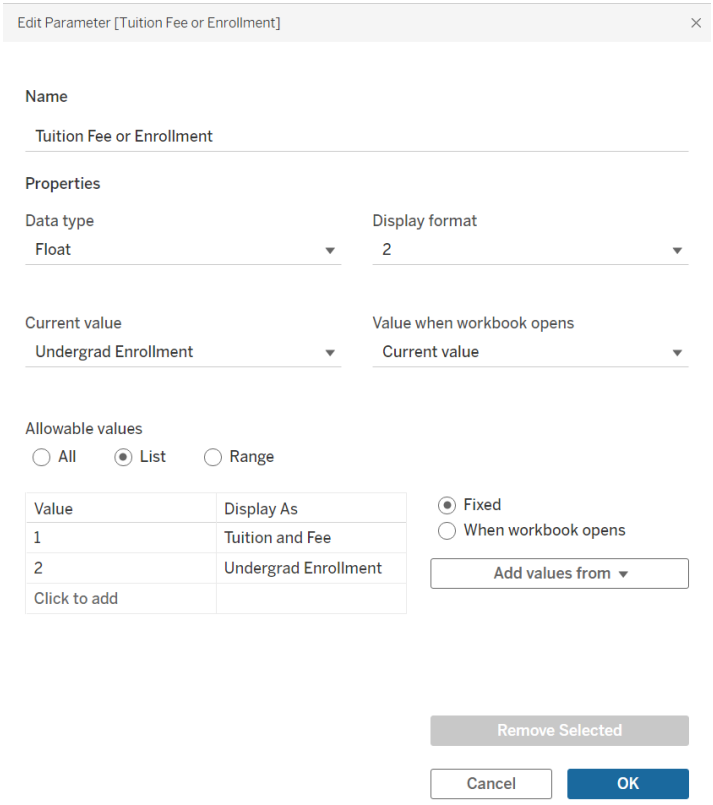
Name	Location	State	Rank	Tuition and fees	Undergrad Enrollment
Duke University	Durham	NC	8	$51,265	6,639
Texas A&M University--College Station	College Station	TX	74	$28,768	48,960

Question 3

In order to be able to find relationship between each of the variables, a parameter & Calculation Field (**Figure 3**) are added for easy viewing of Tuition and fees data or Undergrad Enrollment data.

Figure 3

Parameter & Calculation Field Formula for selection of either Tuition and Fees data or Undergrad Enrollment data on the visualisation charts.



Calculation Field Formula: CASE [Tuition Fee or Enrollment]
WHEN 1 then SUM([Tuition and fees])
When 2 then SUM ([Undergrad Enrollment])
END

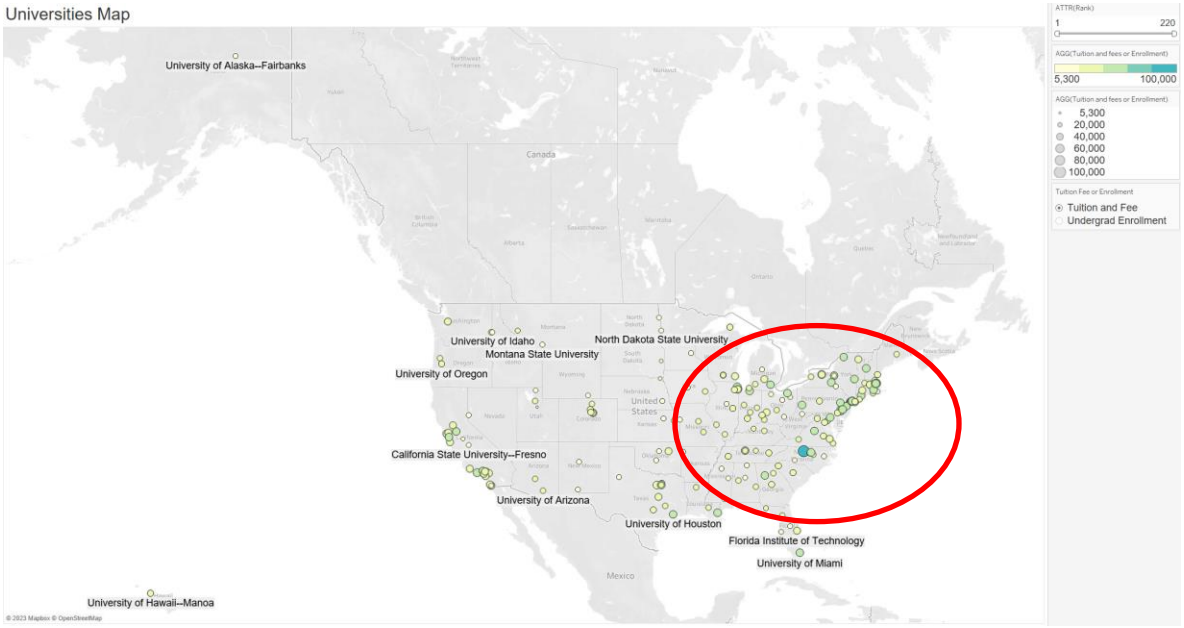
After the above is completed, the following charts can be plotted.

a) A Map is created based on States & City to visualise the universities' location and obtain possible insight (**Figure 4**)
   - Features included is Parameters to allow users to select the map of either Tuition and fees or Undergrad Enrollment (**Figure 4**).

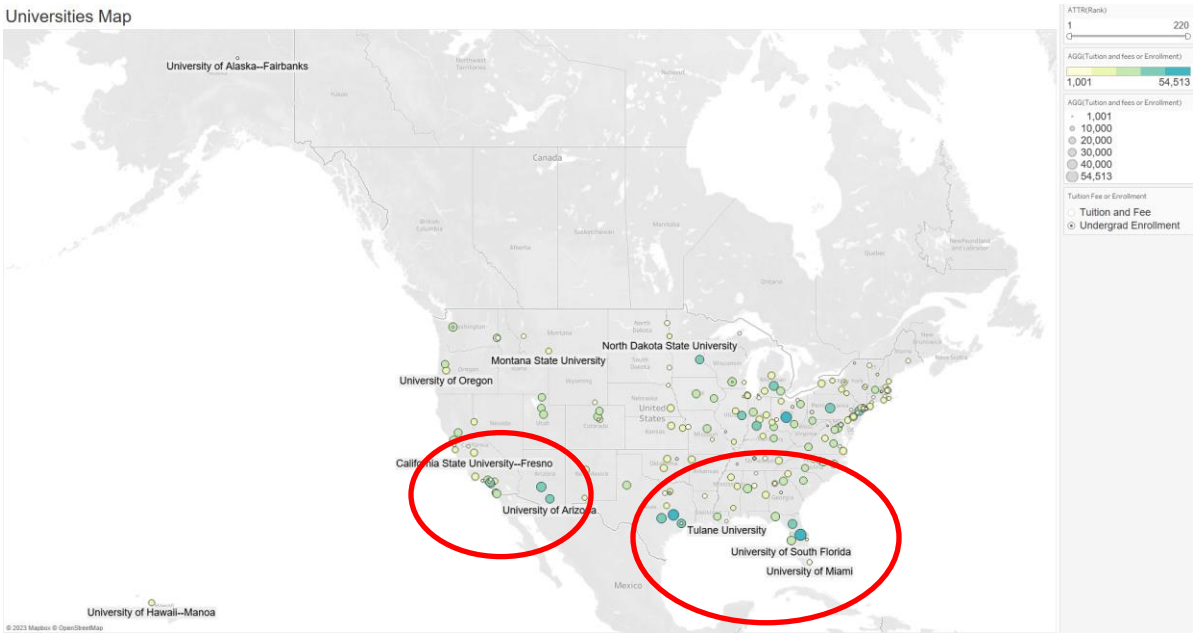
- Key patterns observed are most of the universities are located at the eastern side of USA near to New York. Tuition and fees are generally competitive (Light Green) among the universities. However from the Undergrad Enrollment map, we observed that some of the universities that located far from New York (Dark Green) are more well sought after. Possible reasons might be their popularity in US and school ranking.

Figure 4

Universities Location Maps for Tuition and Fees



Universities Location Maps for Undergrad Enrollment

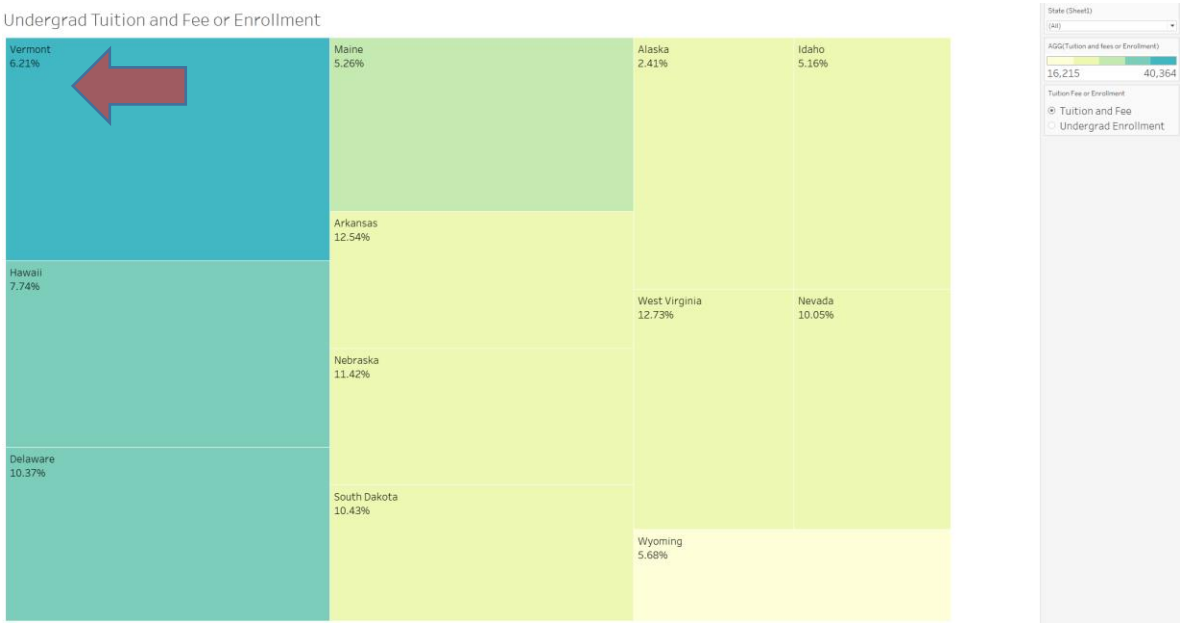


b) Treemap is created to identify the top US states with the higher Tuition and Fees or Undergrad Enrollment.

- Features included is Parameters to allow users to select the map of either Tuition and fees or Undergrad Enrollment (**Figure 5**).
- From the treemap, we observed that Universities with the highest Tuition and Fees will receive lesser Undergrad Enrollment. E.g., Vermont.

Figure 5

US States Treemap for Tuition and Fees



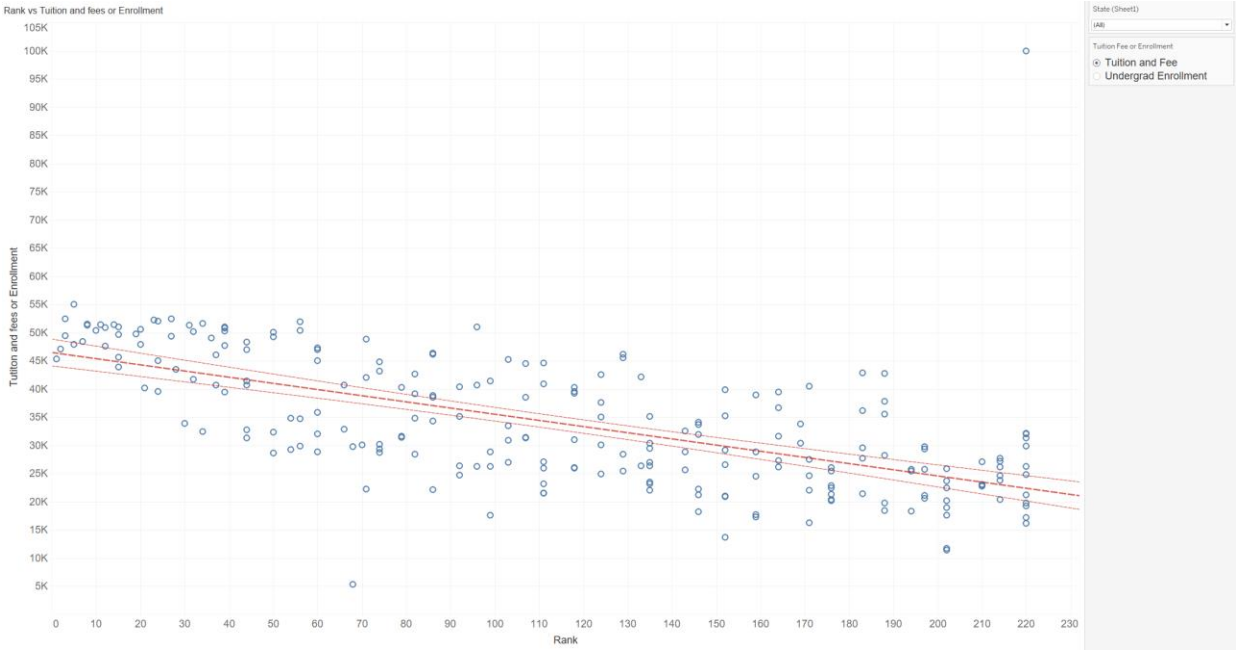
US States Treemap for Undergrad Enrollment

c) Line graph is created to identify the relationship between Universities' Ranking and Tuition and Fees or Undergrad Enrollment.

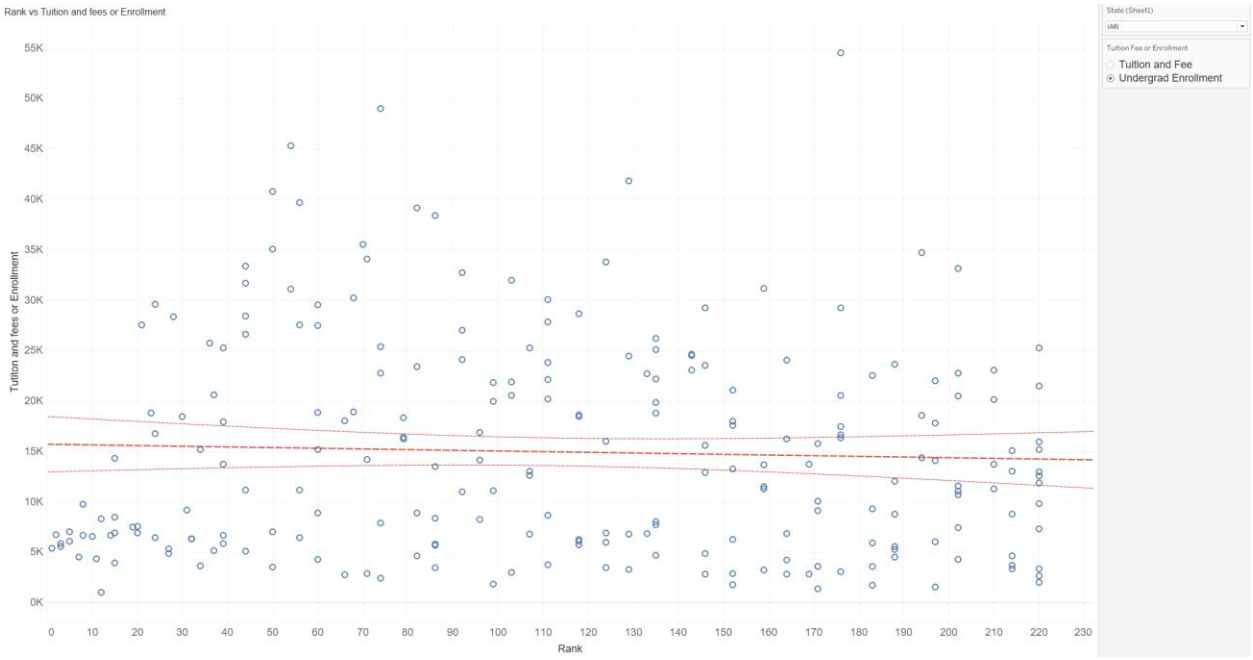
- Features included are Parameters to allow users to select the map of either Tuition and fees or Undergrad Enrollment and trendline to see the behaviour of ranks. (**Figure 6**).
- From the line plot, we observed that Tuition and fees will reduce proportionally when ranking is lower towards 200 and most of the university's fees are near to the trendline. This show that the data is linear. However, the Undergrad Enrollment trend is almost flat showing same number of intakes regardless of ranking. Data points are mostly random and far away from the trendline. This mean that the data is random and not linear. Higher undergrad enrolment can be observed only after Rank 20 and possible reason is due to high Tuition and Fee.

Figure 6

Line plot of Universities' Ranking and Tuition and fees



Line plot of Universities' Ranking and Undergrad Enrollment

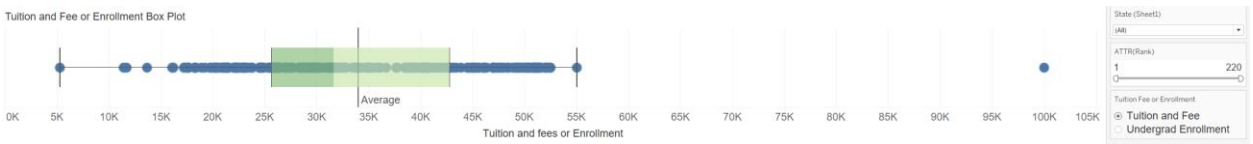


d) Box plot is created to identify the distribution of Tuition and Fees and Undergrad Enrollment.

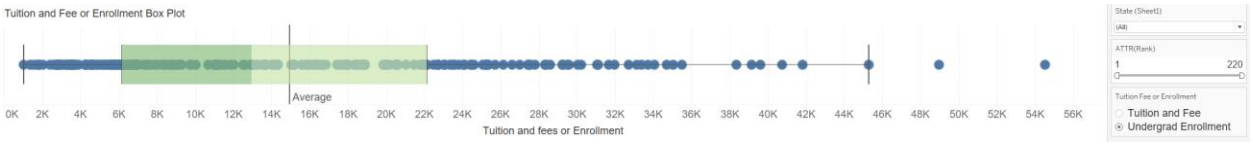
- Features included are Parameters to allow users to select the map of either Tuition and fees or Undergrad Enrollment and Reference line (Average) to see the distribution. (**Figure 7**).
- From the box plot, we observed that most of the universities' Tuition and fees are around $25,693 (Lower Quartile) to $42,821 (Upper Quartile) and one outlier ("University of North Carolina – Greensboro") located out of the box plot. Average Tuition and fees is at $33.976.
- As for Undergrad Enrollment, the lower and upper quartile are wider between 6,170 to 22,146. There are 2 outliers (Texas A&M University-College Station and University of Central Florida). Average Undergrad Enrollment is 14,932.

Figure 7

Box Plot of Tuition and Fees



Box Plot of Undergrad Enrollment

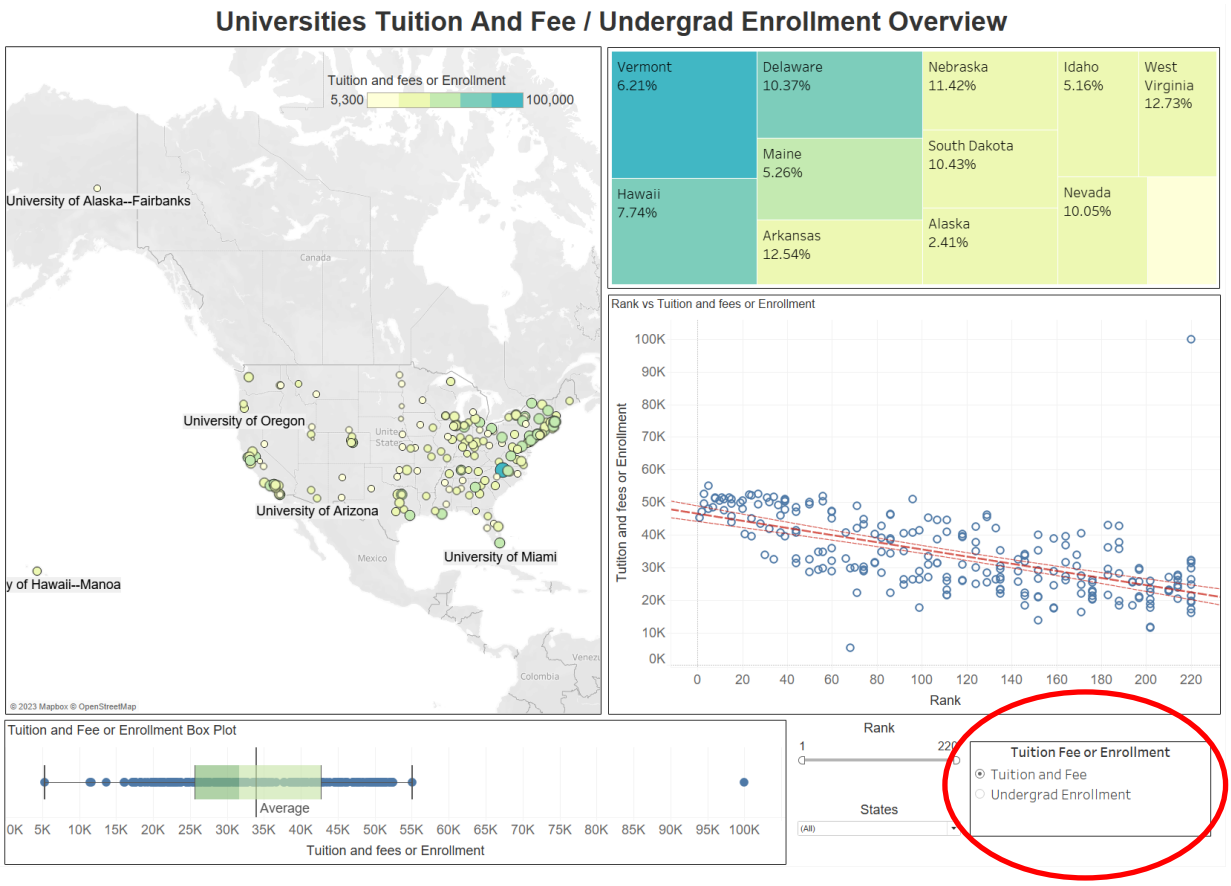


Question 4

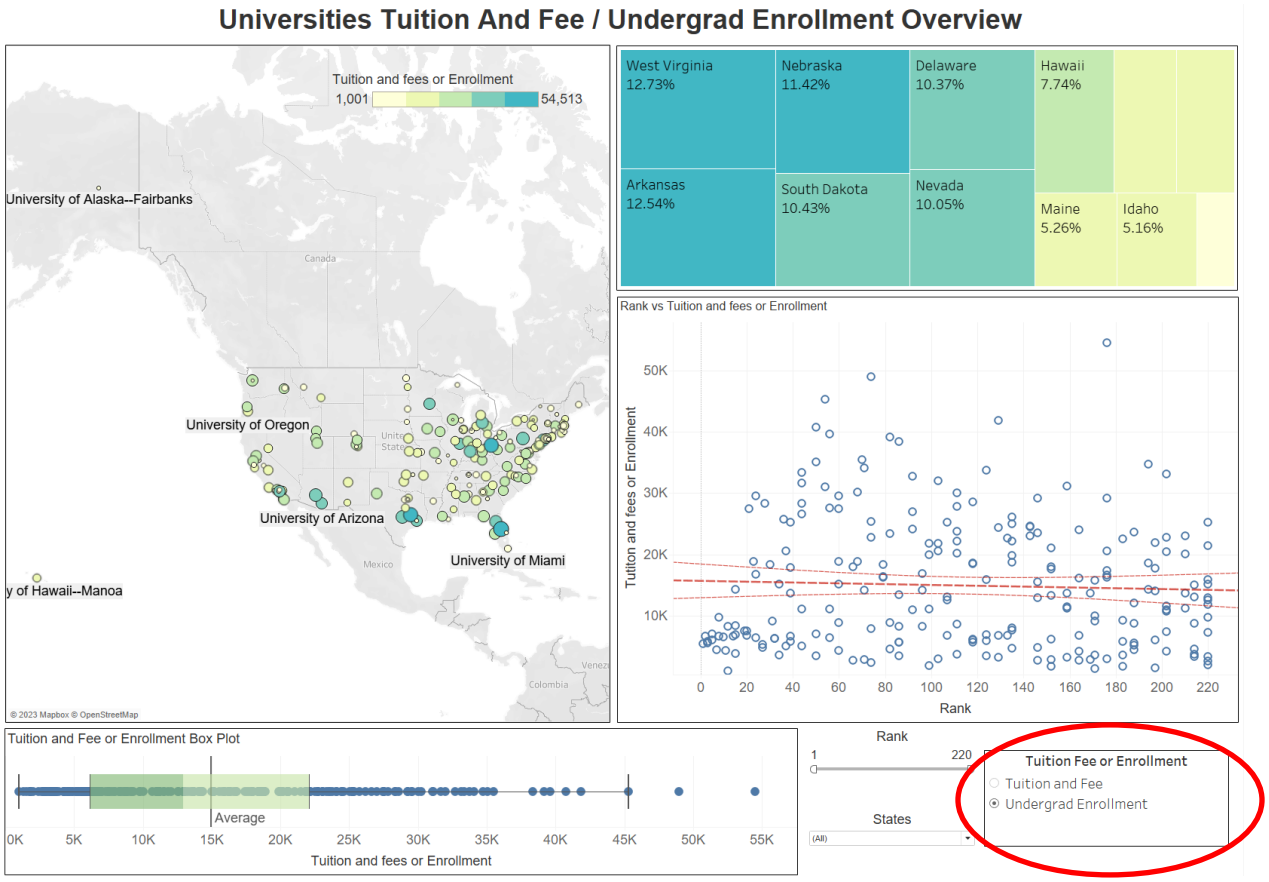
By using Figure 4 to 7 charts, a Tactical Interactive Dashboard can be created as show in **Figure 8**. It is a single page dashboard showcasing an overview of US Universities' Tuition and Fee and Undergrad Enrollment in a glance. A map is placed on the priority zone for quick understanding of the Universities location and its competitiveness against other nearby universities. In order to simplify the usage, a parameter (**Figure 3**) to differentiate Tuition and Fee & Undergrad Enrollment is used and user can easily click on the radio button to switch between both studies.

Figure 8

Business Performance Dashboard – Tuition and Fees



Business Performance Dashboard – Undergrad Enrollment



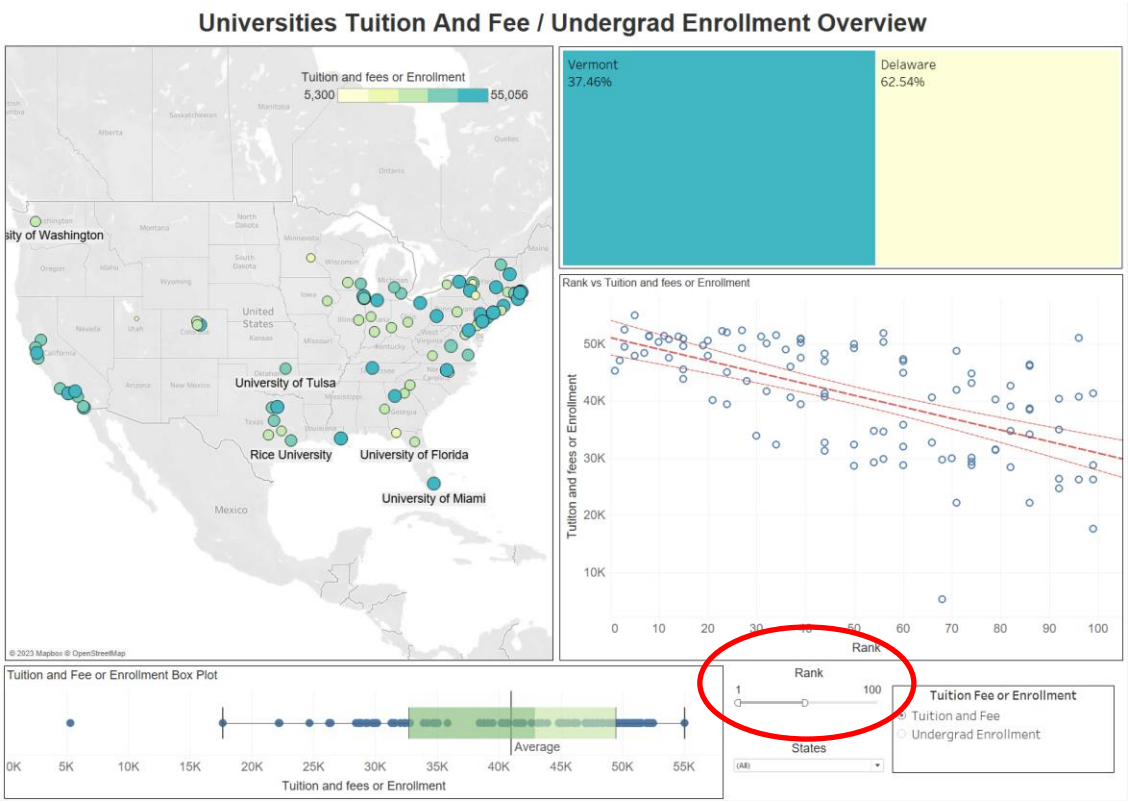
Users can even perform the following actions to obtain insights and understand the position of the universities against other universities in US.

a) By adjusting the Rank Slider bar in **Figure 9**, all the Charts will be filtered to the Slider requirement. For example when we want to know the Top 100 Universities in US, we adjust the Slider from 1 to 100 and the results will be reflected to all the charts as showed in **Figure 9**.

b) By selecting the States in the dropdown menu (**Figure 10**), user can drill down to understand the relationship of the universities in the selected states Tuition and fee and Undergrad Enrollment condition.

Figure 9

Filtering Actions to Universities ranking performance on Tuition and Fees



Filtering Actions to Universities ranking performance on Undergrad Enrollment.

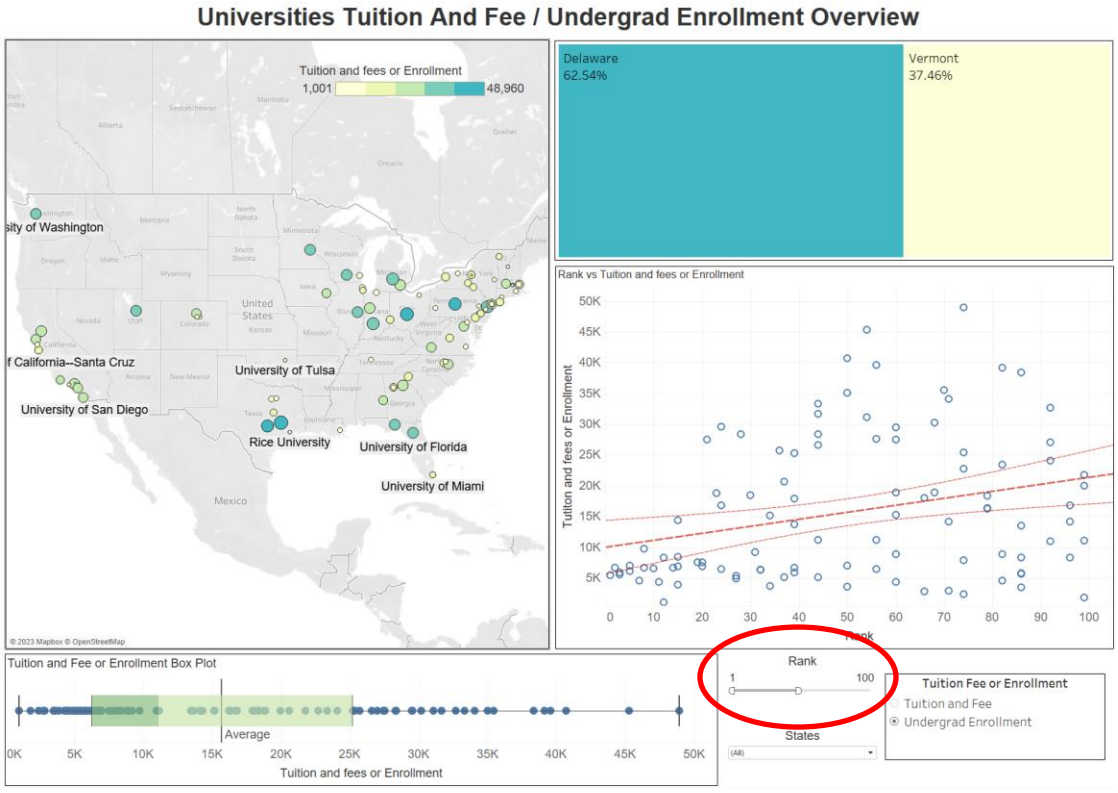
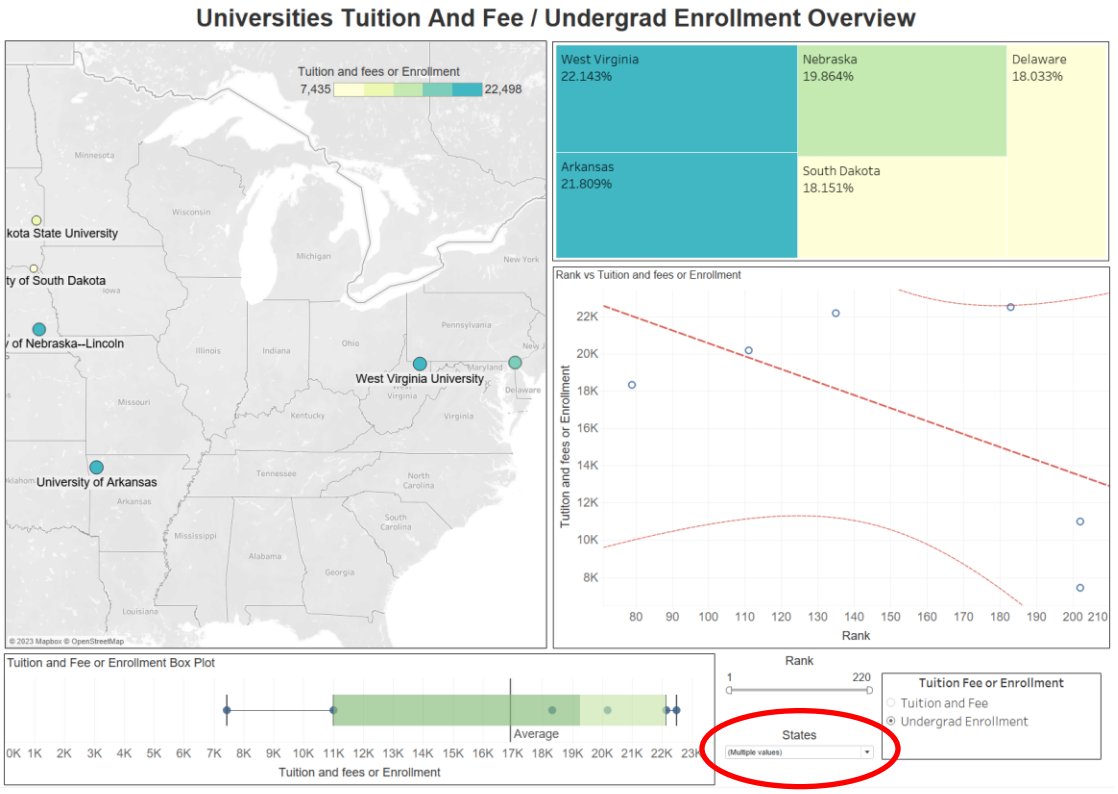
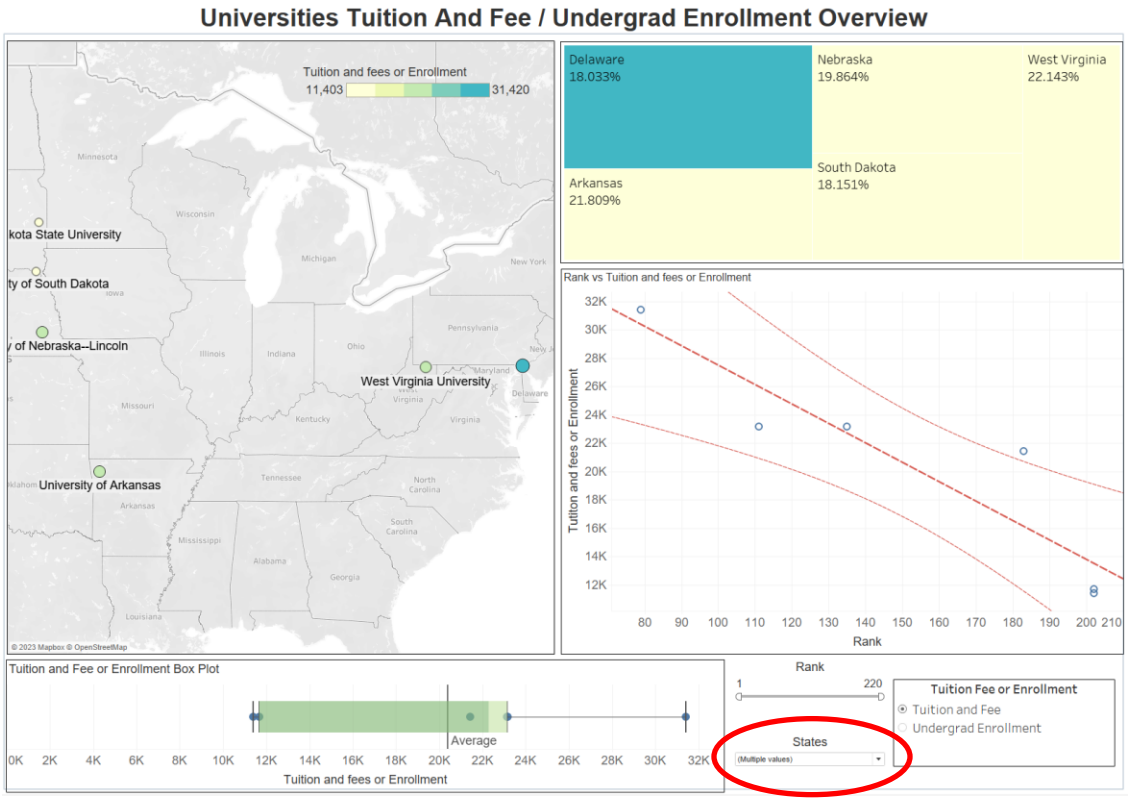


Figure 10

Drilling Down Actions to States Universities performance on Tuition and Fees

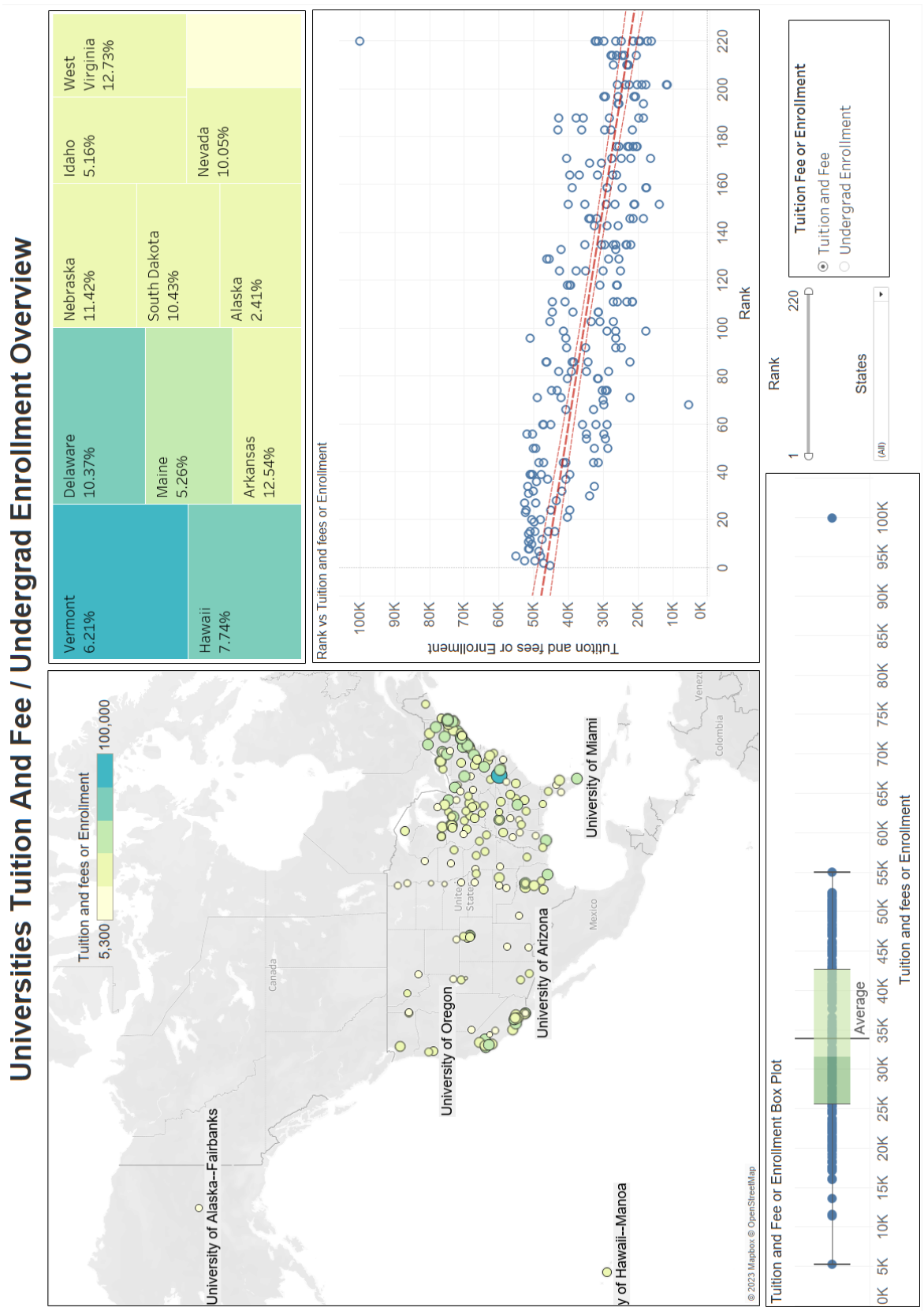


Drilling Down Actions to States Universities performance on Undergrad Enrollment



Appendix 1

Business Performance Dashboard – Tuition and Fee View



Appendix 2

Business Performance Dashboard – Undergrad Enrollment