**UTSA AND ASPIRATION DEGREES AWARDED/12 MONTH FTE RATIO**

**DASHBOARD DOCUMENTATION**

**I – Data Source**

IPEDS Data Center

**II – Get The Dataset**

**a. Dataset for 12-month FTE enrollment**

- Click on Compare Institutions

- Select Use final release data

- For Comparison Institution, type in The University of Texas at San Antonio and **Select**

- For Comparison Group, select By Names or UnitIDs, type in the Institution Name, and Select. Repeat the process until the end of list of comparison group.

- Click Continue

- Select **12-Month Enrollment**

**-> 12-month instructional activity and full-time equivalent enrollment: 2003-04 to current year**

**-> Academic year 2003-04 to current year**

**-> Select years for dataset**

**-> Reported 12-month full-time equivalent (FTE) undergraduate enrollment**

- Click **Continuex2**

- Click **Download in comma separated format**

- Rename the downloaded csv file to **fteFallEnroll** and move it to the same directory where the dashboard is

**b. Dataset Degrees Awared**

- Click on Compare Institutions

- Select Use final release data

- For Comparison Institution, type in The University of Texas at San Antonio and **Select**

- For Comparison Group, select By Names or UnitIDs, type in the Institution Name, and Select. Repeat the process until the end of list of comparison group.

- Click Continue

- Select **Frequenly used/Derived variables**

**-> Degrees/awards**

**-> Number of students receiving degress/certificates – 2011-12 to current year**

**-> Select years for dataset**

**-> Number of students receiving a Bachelor’s degree**

- Click **Continuex2**

- Click **Download in comma separated format**

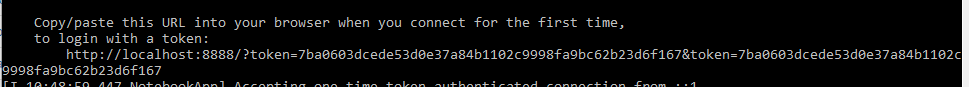
- Rename the downloaded csv file to **Bachelor** and move it to the same directory where the dashboard is

**III – Quick Guide to Customize Dataset**

- Move the dataset into the same folder with DegFTE Ratio.ipynb  
- From Windows symbol, select **Anaconda3 (64-bit) -> Jupyter Notebook** (if you have not had Anaconda installed, here’s the link: <https://www.anaconda.com/download/> (Select the version for Python 3)

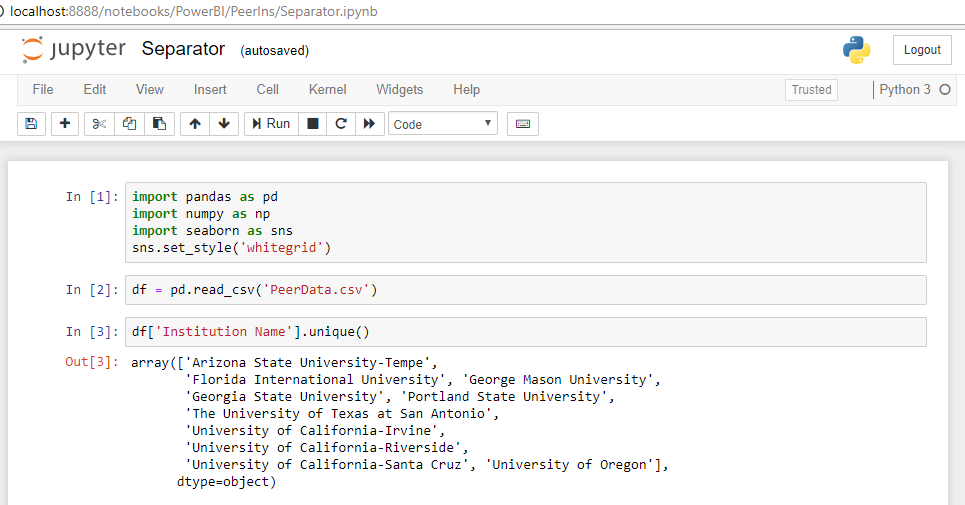
- A picture like this should pop up as an Internet Explorer or Google Chrome window

- If the window does not appear, click on the Jupyter Notebook symbol from Windows tool bar and **select** the similar hyperlink like below, and **open it from a web browser**.



- Navigate to where DegFTE Ratio.ipynb is. If you’re having trouble finding this file from Jupyter Notebook, it should be similar to how you get to it from regular Windows Navigator.

- Open DegFTE Ratio.ipynb. A window like this should appear:



- Select **Kernel -> Restart & Run All**

- Scroll down to the last 2 cells, select the 40th or 41st if you want to get the data in excel or csv format

- Select all of the content within the selected cell and Ctrl + /

- Shift + Enter

- The new excel file should now be in the same directory with DegFTE Ratio.ipynb

- Remember to delete the first columns (index columns) on the two sheets for Power Bi to pick up the first row as parameters row

**IV. Step-by-Step Jupyter Notebook Tutorial (Cell by cell)**

**A. Introduction**

- About Jupyter Notebook: <http://jupyter.org/about>

- About Python libraries for data science and tutorials:

a. Pandas

Tutorial: <https://www.analyticsvidhya.com/blog/2016/01/complete-tutorial-learn-data-science-python-scratch-2/>

Documentation: <https://pandas.pydata.org/pandas-docs/stable/>

b. Numpy

Documentation: <https://docs.scipy.org/doc/>

c. Seaborn

Tutorial: <https://www.youtube.com/playlist?list=PLE50-dh6JzC7X8VFX40yoIXnhctF2bR8F>

Documentation: <https://seaborn.pydata.org/>

**B. Tutorial (Cell by cell)**

1. Importing necessary libraries, I used pd, np and sns because they are universal aliases for these libraries.

3. Using method ‘read\_csv’ from library pandas to read a file call ‘Bachelor.csv’ and ‘fteFallEnroll.csv’ as Pandas Dataframe objects called ‘bac’ and ‘fte’.

Your csv should be in the same directory with your Jupyter Notebook because it would be easier to read the csv file.

4,5. Create a new list called ‘datasets’ that has its elements as ‘bac’ and ‘fte’ – Drop unnecessary collumns: ‘Unnamed: 7’ and ‘UnitID’

6. Create a new dataframe object from ‘bac’ that does not contain UTSA and drop ‘Institution Name’ column for data manipulation. This object is called ‘peers\_bac’.

7. Create a new dataframe object from ‘fte’ that does not contain UTSA and drop ‘Institution Name’ column for data manipulation. This object is called ‘peers\_fte’.

8. Create two Python dictionary objects

9. Assign the columns of ‘peers\_bac’ as values for key ‘Year’ in median\_bac

Extract the medians from each column of ‘peers\_bac’ and append as values for ‘Median’ key

10. Display median\_bac

11. Repeat step 9 to fill in data for ‘median\_fte’

12. Display median\_fte

13. Transform ‘median\_bac’ into a Python dataframe and call it ‘bac\_df’

14. Transform ‘medit\_fte’ into a Python dataframe and call it ‘fte\_df’

15. Create a new dataframe object from ‘fte’ that does contains UTSA and drop ‘Institution Name’ column for data manipulation. This object is called ‘utsa\_fte’.

16. Create a new dataframe object from ‘bac that does contains UTSA and drop ‘Institution Name’ column for data manipulation. This object is called ‘utsa\_bac’.

17. Get the row of data from ‘utsa\_fte’, assign the numbers to a list and call it ‘utsa\_fte\_list’

18. Get the row of data from ‘utsa\_bac’, assign the numbers to a list and call it ‘utsa\_bac\_list’

19. Create a new column ‘UTSA’ in ‘bac\_df’ object and assign values from ‘utsa\_bac\_list’

20. Create a new column ‘UTSA’ in ‘fte\_df’ object and assign values from ‘utsa\_fte\_list’

21. Rename the rows in column ‘Year’ of ‘bac\_df’ to the number of that year only

22. Rename the rows in column ‘Year’ of ‘fte\_df’ to the number of that year only

23. Create new column names for bac\_df and fte\_df

24. Merge ‘bac\_df’ and ‘fte\_df’ to create a new dataframe and call it ‘found’

25. Display ‘found’

26. Extract columns 'Median Degrees', 'Median FTE', 'Year' and assign as a dataframe

Extract columns 'UTSA Degrees', 'Median FTE', 'Year' and assign as a dataframe

27. Calculate ratio for each year in each dataframe.

28. Sort values of ‘median\_ratio’ by column ‘Year’

29. Sort values of ‘UTSA\_ratio’ by column ‘Year’

30. Plot and compare ratios of UTSA and Aspirational Peers

31. To uncomment this cell, select all of the content of this cell, select Ctrl + / and select Shift + Enter to run the cell. This cell will create an Excel file called ‘UTSA and Peers.xlsx’

If the cell can’t be run, delete ‘UTSA and Peers.xlsx’ in the directory and run the cell again