

# Almost Too Easy: A New Way to Wrangle IPEDS Data\*

Kate Aloisio and Kathy Foley, Smith College  
Emma Morgan, Tufts University

AIR Webinar  
August 27<sup>th</sup>, 2019

\*internally referred to as the IPEDS World Domination project



<https://www.smith.edu/about-smith/institutional-research>



**Tufts**  
UNIVERSITY

<https://provost.tufts.edu/institutionalresearch/>

Thank you Smith College and Tufts University for supporting this project!

# Why IPEDS data? Why here? Why now?



# Typical IR Scenario:

Could you compare our institution to 10 specific peer schools on select key metrics from IPEDS?



# Typical IR Scenario:



“Sure thing!”



# Typical IR Scenario:



"To the IPEDS  
Data Center!"

# Typical IR Scenario:

This is fantastic! Now can you show the trend for 10 years? And can you add one more peer school?



# Typical IR Scenario:



I must mentally  
collect myself  
before starting  
over.



# Downloading IPEDS Data Using “Compare Institutions”

## Use the Data

Access IPEDS data submitted to NCES through our data tools or download the data to conduct your research



### IPEDS Survey Components

Learn more about the individual IPEDS survey components: view training videos, infographics, answers to frequently asked questions, survey forms, and more.



### Compare Institutions ?

Download IPEDS data files for more than 7,000 institutions and up to 250 variables. Data files are provided in comma separated value (\*.csv) format.



### Data Trends

Use the Trend Generator to view trends on most frequently asked subject areas including: Enrollment, Completions, Graduation Rates, Employees and Staff, Institutional Revenues, and Financial Aid.



### Survey Data

Data are available starting with the 1980-81 collection year for the *Complete data files* and *Custom data files* functions, which zip the data into comma separated value (\*.csv). Beginning with the 2006-07 collection year, data for each collection year are compiled into an *Access database*.

But this requires you to manually select which variables you want, and which institutions.

# Downloading IPEDS Data as “Complete Surveys”

## Use the Data

Access IPEDS data submitted to NCES through our data tools or download the data to conduct your research



### [IPEDS Survey Components](#)

Learn more about the individual IPEDS survey components: view training videos, infographics, answers to frequently asked questions, survey forms, and more.



### [Data Trends](#)

Use the Trend Generator to view trends on most frequently asked subject areas including: Enrollment, Completions, Graduation Rates, Employees and Staff, Institutional Revenues, and Financial Aid.



### [Look Up an Institution](#) ?

Look up information for one institution at a time. Data can be viewed in two forms: institution profile (similar to College Navigator) and reported data (institution's response to each survey question).



### [Data Feedback Report](#) ?

Download, print, or customize an institution's Data Feedback Report,



### [Compare Institutions](#) ?

Download IPEDS data files for more than 7,000 institutions and up to 250 variables. Data files are provided in comma separated value (\*.csv) format.



### [Survey Data](#)

Data are available starting with the 1980-81 collection year for the *Complete data files* and *Custom data files* functions, which zip the data into comma separated value (\*.csv). Beginning with the 2006-07 collection year, data for each collection year are compiled into an *Access database*.



Select download option
Access database
<b>Complete data files</b> ?
Custom data files ?

And then you realize that each year has to be downloaded as its own data file and an accompanying dictionary.

Complete Data Files

Provisional Release Data (Change)

Years & Surveys

All years

Fall Enrollment

Continue

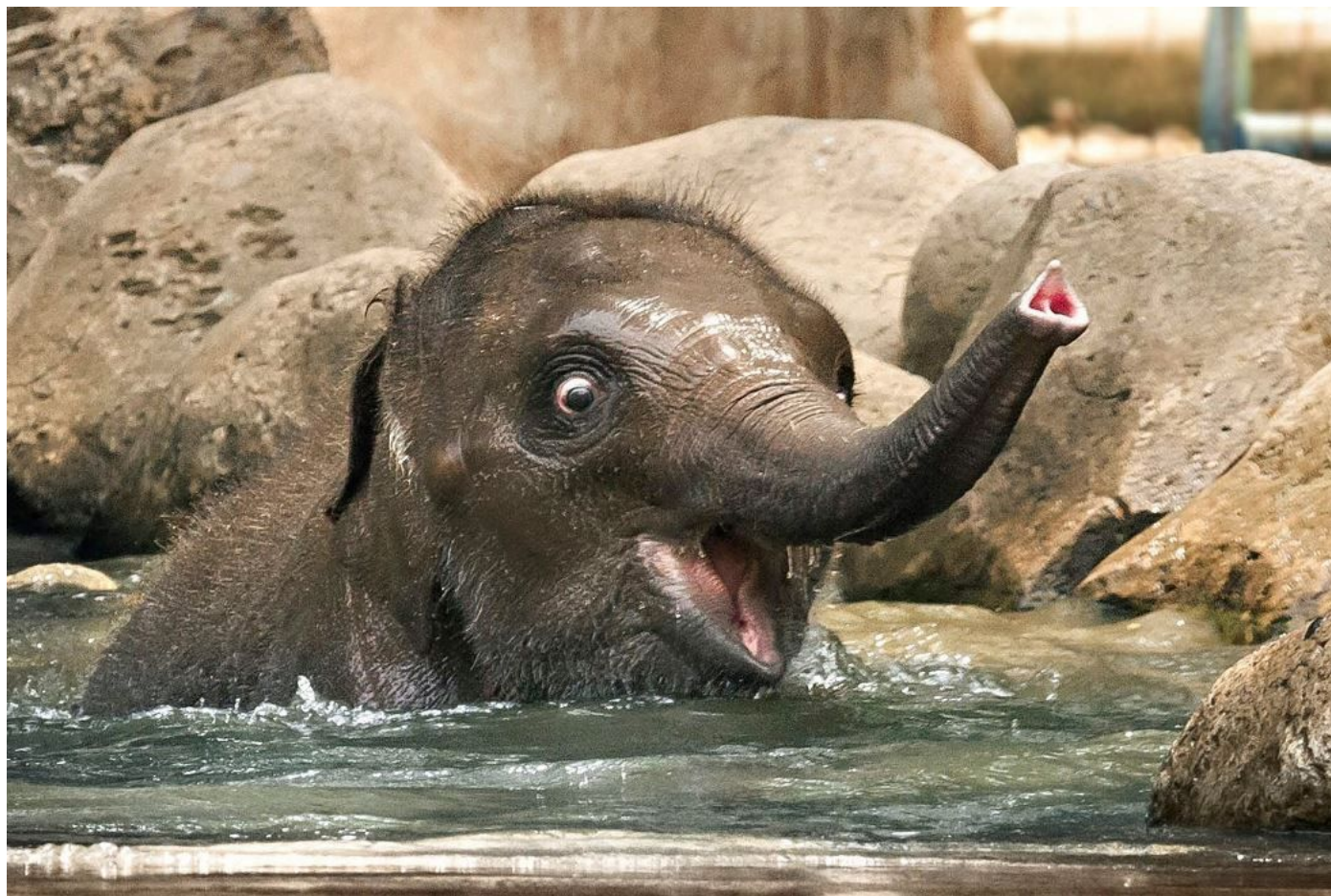
Data files are available in ZIP format.

Year	Survey	Title	Data File	Stata Data File	Programs	Dictionary
2017	Fall Enrollment	Race/ethnicity, gender, attendance status, and level of student: Fall 2017	<a href="#">EF2017A</a>	<a href="#">EF2017A STATA</a>	<a href="#">SPSS</a> , <a href="#">SAS</a> , <a href="#">STATA</a>	<a href="#">Dictionary</a>
2017	Fall Enrollment	Age category, gender, attendance status, and level of student: Fall 2017	<a href="#">EF2017B</a>	<a href="#">EF2017B STATA</a>	<a href="#">SPSS</a> , <a href="#">SAS</a> , <a href="#">STATA</a>	<a href="#">Dictionary</a>

And that the data come down like this:

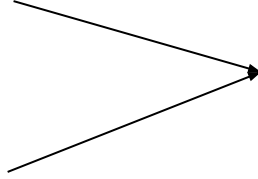
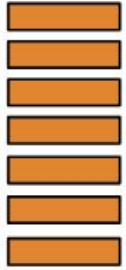
UNITID	XAPPLCN	APPLCN	XAPPLCNM	APPLCNM	XAPPLCNW	APPLCNW	XADMSSN	ADMSSN	XADMSSNM	ADMSSNM	XADMSSNW	ADMSSNW	ADMCON1
100654	R	8610	R	3066	R	5544	R	7772	R	2704	R	5068	1
100663	R	7555	R	2729	R	4826	R	6936	R	2487	R	4449	1
100706	R	4454	R	2466	R	1988	R	3618	R	2082	R	1536	1
100724	R	6842	R	2136	R	4515	R	6696	R	2071	R	4438	1
100751	R	38129	R	14705	R	23424	R	20321	R	7992	R	12329	1
100830	R	2474	R	841	R	1633	R	2042	R	715	R	1327	1
100858	R	18072	R	7841	R	10231	R	15168	R	6576	R	8592	1
100937	R	2559	R	1098	R	1461	R	1583	R	726	R	857	1
101189	R	2335	R	878	R	1457	R	1191	R	484	R	707	1
101365	R	107	R	36	R	71	R	81	R	34	R	47	3
101435	R	2074	R	1119	R	955	R	1161	R	585	R	576	1
101453	R	6	R	6	Z	0	R	6	R	6	A		3





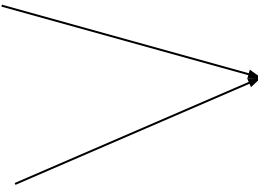
# There is a better way

Data files for  
each year

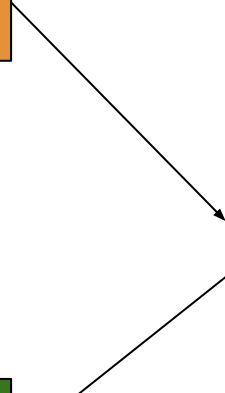


**Compiled  
raw data**

Dictionary files  
for each year



**Compiled  
dictionary**



**Beautiful csv**

## So what starts looking like this:

UNITID	ADMCON7	XAPPLCN	APPLCN	XAPPLCNM	APPLCNM	XAPPLCNW	APPLCNW	XADMSSN	ADMSSN	XADMSSNM	ADMSSNM	XADMSSNW	ADMSSNW
100654	1	R	8610	R	3066	R	5544	R	7772	R	2704	R	5068
100663	1	R	7555	R	2729	R	4826	R	6936	R	2487	R	4449
100706	1	R	4454	R	2466	R	1988	R	3618	R	2082	R	1536
100724	1	R	6842	R	2136	R	4515	R	6696	R	2071	R	4438
100751	1	R	38129	R	14705	R	23424	R	20321	R	7992	R	12329
100830	1	R	2474	R	841	R	1633	R	2042	R	715	R	1327
100858	1	R	18072	R	7841	R	10231	R	15168	R	6576	R	8592
100937	1	R	2559	R	1098	R	1461	R	1583	R	726	R	857
101189	1	R	2335	R	878	R	1457	R	1191	R	484	R	707
101365	2	R	107	R	36	R	71	R	81	R	34	R	47
101435	1	R	2074	R	1119	R	955	R	1161	R	585	R	576



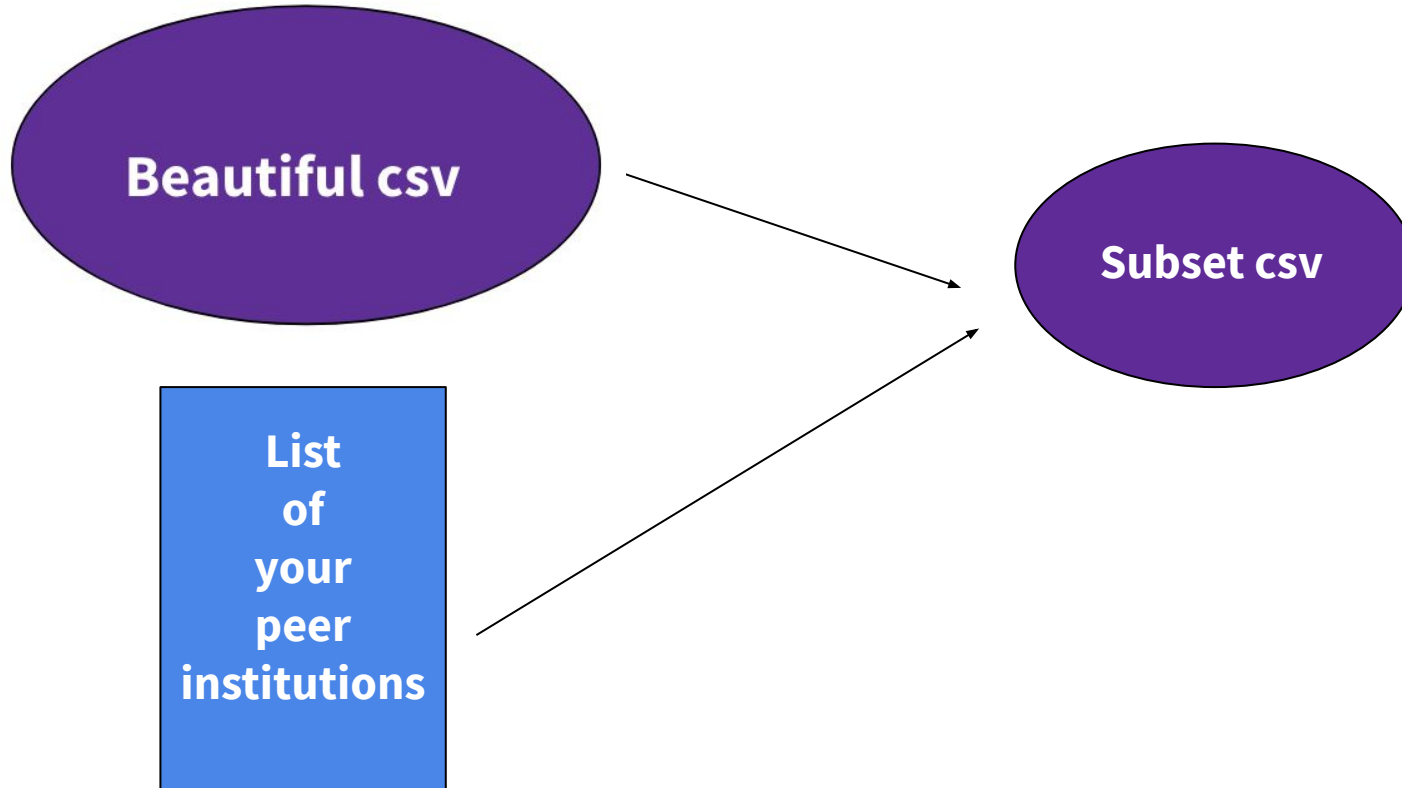
## So what starts looking like this:

UNITID	ADMCON7	XAPPLCN	APPLCN	XAPPLCNM	APPLCNM	XAPPLCNW	APPLCNW	XADMSSN	ADMSSN	XADMSSNM	ADMSSNM	XADMSSNW	ADMSSNW
100654	1	R	8610	R	3066	R	5544	R	7772	R	2704	R	5068
100663	1	R	7555	R	2729	R	4826	R	6936	R	2487	R	4449
100706	1	R	4454	R	2466	R	1988	R	3618	R	2082	R	1536
100724	1	R	6842	R	2136	R	4515	R	6696	R	2071	R	4438
100751	1	R	38129	R	14705	R	23424	R	20321	R	7992	R	12329
100830	1	R	2474	R	841	R	1633	R	2042	R	715	R	1327
100858	1	R	18072	R	7841	R	10231	R	15168	R	6576	R	8592
100937	1	R	2559	R	1098	R	1461	R	1583	R	726	R	857
101189	1	R	2335	R	878	R	1457	R	1191	R	484	R	707
101365	2	R	107	R	36	R	71	R	81	R	34	R	47
101435	1	R	2074	R	1119	R	955	R	1161	R	585	R	576

## Ends up looking like this:

UNITID	ACAD_YEAR	FILE_NAME	Admission test scores	Applicants total	Applicants men	Applicants women	Admissions total	Admissions men	Admissions women
100654	2007	ic2006.csv	Required	9663	3774	5889	3979	1656	2323
100654	2008	ic2007.csv	Required	6470	2507	3963	2059	799	1260
100654	2009	ic2008_rv.csv	Required	3961	1664	2297	1941	837	1104
100654	2010	ic2009_rv.csv	Required	5705	2273	3432	2696	1059	1637
100654	2011	ic2010_rv.csv	Required	5986	2557	3429	3070	1264	1806
100654	2012	ic2011_rv.csv	Required	7156	2847	4309	3585	1442	2143
100654	2013	ic2012_rv.csv	Required	10877	4219	6658	5915	2246	3669
100654	2014	ic2013_rv.csv	Required	6142	2401	3741	5521	2100	3421
100654	2015	adm2014_rv.csv	Required	9901	3852	6049	5204	1977	3227
100654	2016	adm2015.csv	Required	7901	3030	4871	5166	1999	3167
100654	2017	adm2016.csv	Required	7043	2725	4318	6154	2276	3878
100654	2018	adm2017.csv	Required	8610	3066	5544	7772	2704	5068

# But you likely don't want all 7,000 institutions.....



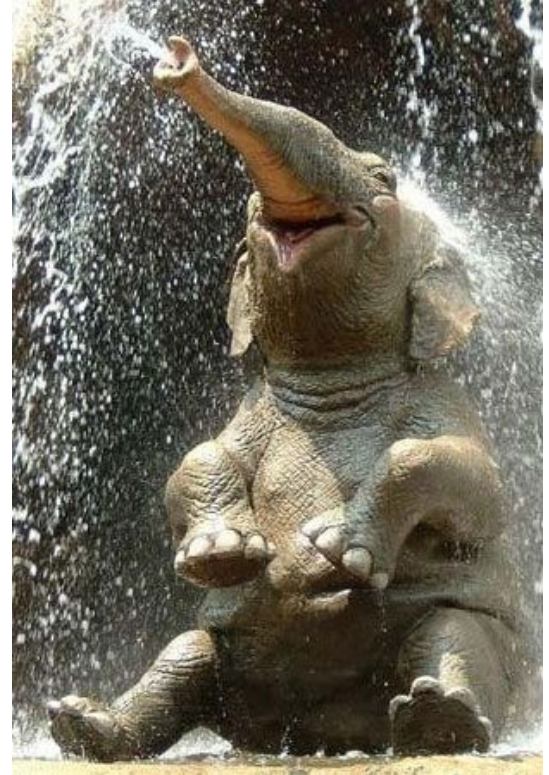
# Do I need to know R?



# Do I need to know R?



# Nope!



# Demo

Let us Show you How it Works!



# IPEDS Data Compiler

Compiler

[FAQ](#)

## Welcome to the IPEDS Data Compiler!

Follow the steps below to generate and download a csv file of longitudinal IPEDS data for a custom group of institutions.

**Step 1:** Upload your list of peers as a csv. Please make sure it contains NCES ID in a column called UNITID. Please note, any columns that you have included in your uploaded file will be joined onto the resulting csv that you download.

See the FAQ for an example peer list.

Choose csv file:

Browse...

No file selected



Choose csv file:

Browse...

small\_bacc\_MA.csv

Upload complete

Show 5 entries

Search:

	UNITID	INSTITUTION NAME	CARNEGIE CLASSIFICATION 2015: BASIC (HD2017)	SECTOR OF INSTITUTION (HD2017)	INSTITUTION SIZE CATEGORY (HD2017)	STATE ABBREVIATION (HD2017)
1	164465	Amherst College	21	2	2	MA
2	166124	College of the Holy Cross	21	2	2	MA
3	165671	Emmanuel College	21	2	2	MA
4	165936	Gordon College	21	2	2	MA
5	166018	Hampshire College	21	2	2	MA

Showing 1 to 5 of 10 entries

Previous

1

2

Next

Your peer list contains 10 institutions.

**Your peer list contains 10 institutions.**

**Step 2:** Select an IPEDS survey from the dropdown list.

Select survey:

Applications, admissions, enrollees and te: ⚡

**Step 3:** Press the button below to see a preview of the first 10 columns of the dataset.

It may take a couple of minutes for your file to load.

🚀 Dominate the World!

	UNITID	ACAD_YEAR	FILE_NAME	TABLE_TRIM	Academic	Academic_value	Academic/career counseling service
1	164465	2007	ic2006.csv	IC	Yes	1	Yes
2	165671	2007	ic2006.csv	IC	Yes	1	Yes
3	165936	2007	ic2006.csv	IC	Yes	1	Yes
4	166018	2007	ic2006.csv	IC	Yes	1	Yes
5	166124	2007	ic2006.csv	IC	Yes	1	Yes
6	166939	2007	ic2006.csv	IC	Yes	1	Yes
7	167996	2007	ic2006.csv	IC	Yes	1	Yes
8	168218	2007	ic2006.csv	IC	Yes	1	Yes
9	168281	2007	ic2006.csv	IC	Yes	1	Yes
10	168342	2007	ic2006.csv	IC	Yes	1	Yes

Showing 1 to 10 of 120 entries

Previous 1 2 3 4 5 ... 12

**This survey contains academic (fiscal) years from 2007 to 2018. Please see the FAQ for more information about the years contained in this file.**

**Step 4:** Once the table renders, press the button below to download your compiled csv file.

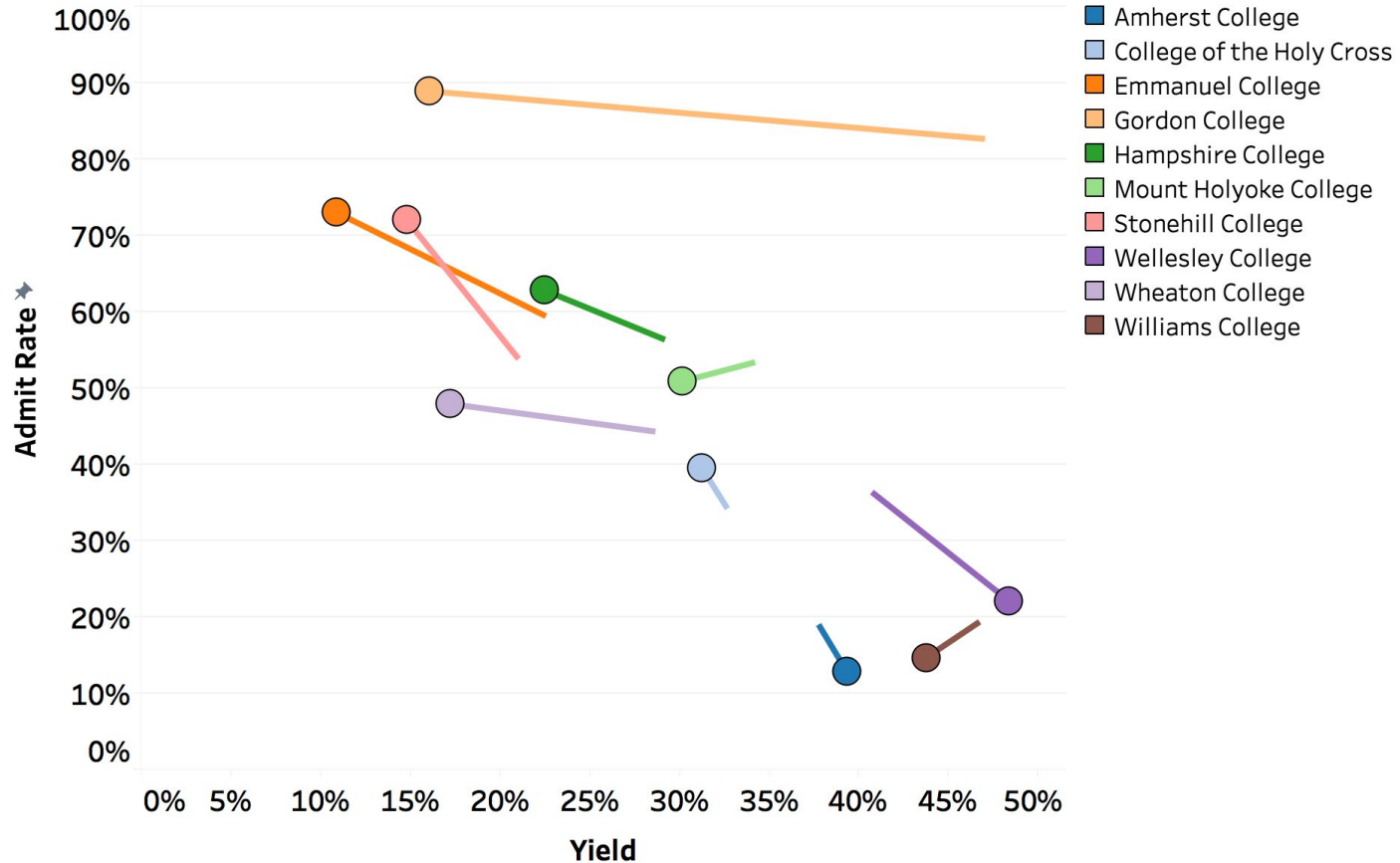
 Download CSV

The screenshot displays the Tableau Desktop interface. At the top, the title bar reads "Tableau - Demo Viz". Below the title bar is a toolbar with various icons for navigation and analysis. The main interface is divided into several panes:

- Data Pane:** Located on the left, it shows a table named "adm\_compiled".
- Dimensions Pane:** Also on the left, it lists various fields under the "Dimensions" section, including "Acad Year", "Academic", "Admission test scores", "Adult basic remedial or high school equivalent", "Advanced placement (AP) credits", "All programs offered completely via distance education", "Any alternative tuition plans offered by institution", "Application fee is required", "Associate's degree", "Bachelor's degree", and "c\_fall\_spring".
- Columns and Rows Shelves:** Located at the top of the main workspace, they are currently empty.
- Filters Shelf:** Located below the Columns and Rows shelves, it is also empty.
- Marks Shelf:** Located below the Filters shelf, it shows a dropdown menu set to "Automatic" and buttons for "Color", "Size", "Text", "Detail", and "Tooltip".
- Main Workspace:** The central area where the visualization is created. It contains a large text area with the prompt "Drop field here".

# Admit Rate and Yield

## Change from 2007 to 2018



A	B	C	D	E	F
UnitID	Institution Name	Carnegie Clas	Sector of inst	Institution siz	State abbrevi
164465	Amherst College	21	2	2	MA
166124	College of the Holy Cross	21	2	2	MA
165671	Emmanuel College	21	2	2	MA
165936	Gordon College	21	2	2	MA
166018	Hampshire College	21	2	2	MA
166939	Mount Holyoke College	21	2	2	MA
167996	Stonehill College	21	2	2	MA
168218	Wellesley College	21	2	2	MA
168281	Wheaton College	21	2	2	MA
168342	Williams College	21	2	2	MA
167835	Smith College				



Choose csv file:

Browse...

small\_bacc\_MA.csv

Upload complete

Show 5 entries

Search:

	UNITID	INSTITUTION NAME	CARNEGIE CLASSIFICATION 2015: BASIC (HD2017)	SECTOR OF INSTITUTION (HD2017)	INSTITUTION SIZE CATEGORY (HD2017)	STATE ABBREVIATION (HD2017)
1	164465	Amherst College	21	2	2	MA
2	166124	College of the Holy Cross	21	2	2	MA
3	165671	Emmanuel College	21	2	2	MA
4	165936	Gordon College	21	2	2	MA
5	166018	Hampshire College	21	2	2	MA

Showing 1 to 5 of 11 entries

Previous

1

2

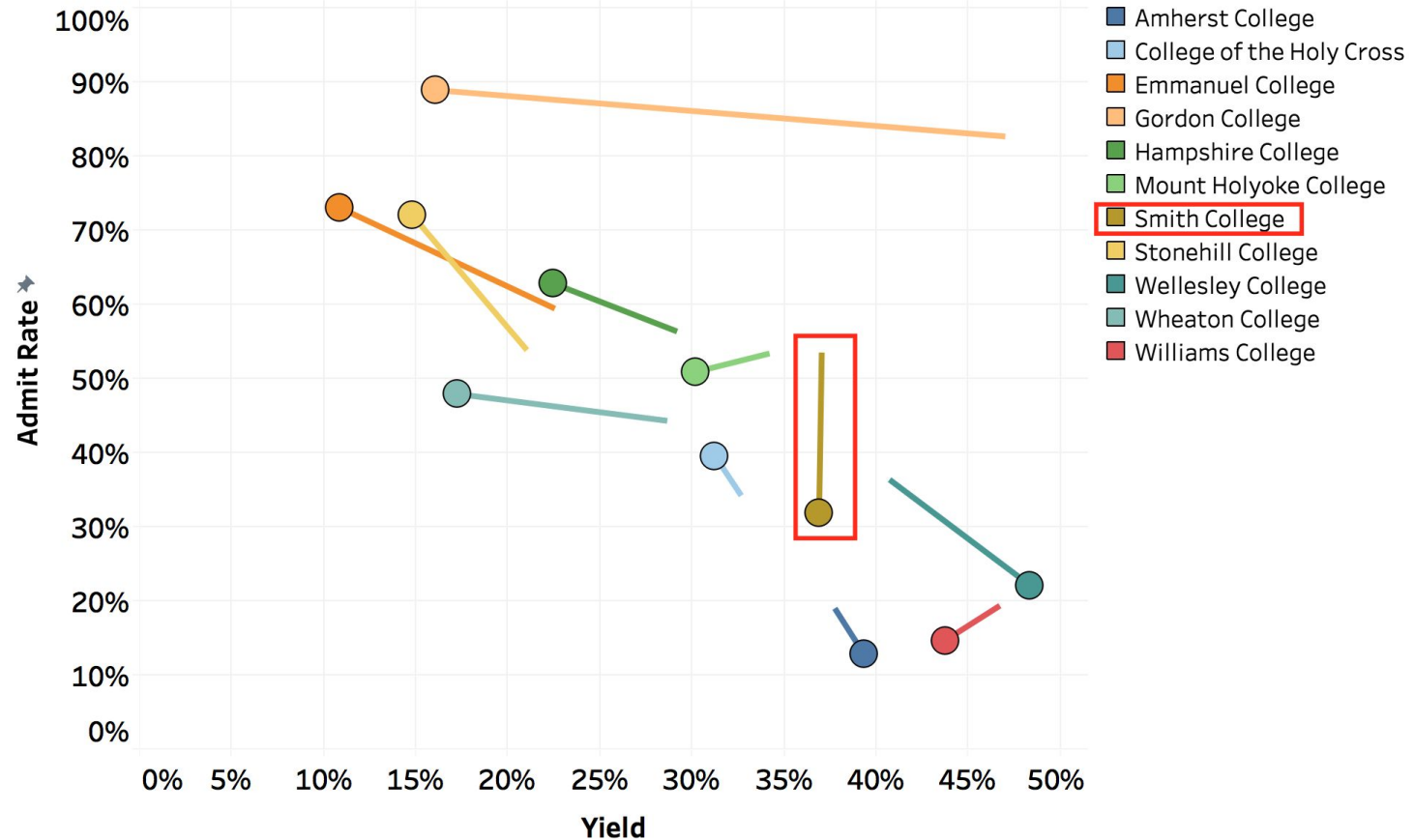
3

Next

**Your peer list contains 11 institutions.**

# Admit Rate and Yield

## Change from 2007 to 2018



# Future Design Enhancements

- Option for user to select specific years



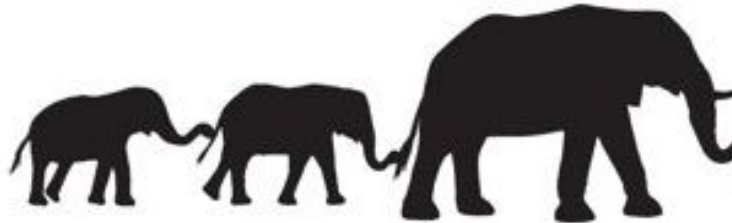
# Future Design Enhancements

- Option for user to select specific years
- Inclusion of additional years



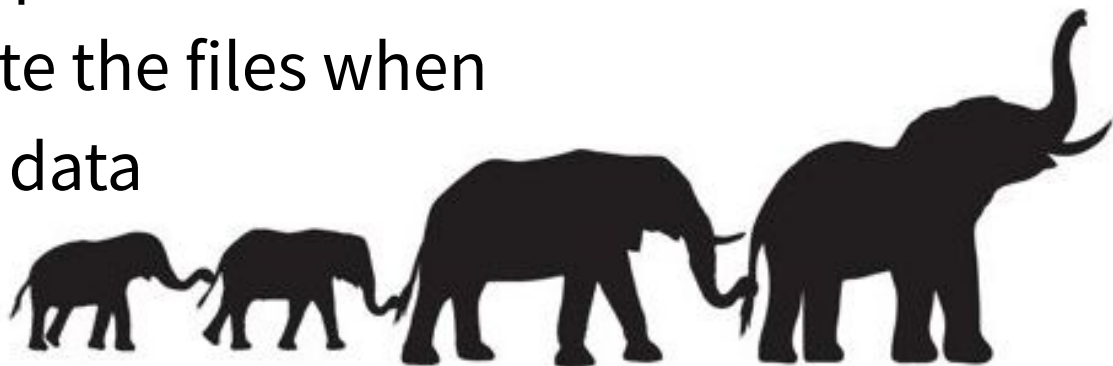
# Future Design Enhancements

- Option for user to select specific years
- Inclusion of additional years
- Expanded peer selection options
  - Based on IPEDS variables
  - User manually inputs list



# Future Design Enhancements

- Option for user to select specific years
- Inclusion of additional years
- Expanded peer selection options
  - Based on IPEDS variables
  - User manually inputs list
- Automatically update the files when IPEDS releases new data





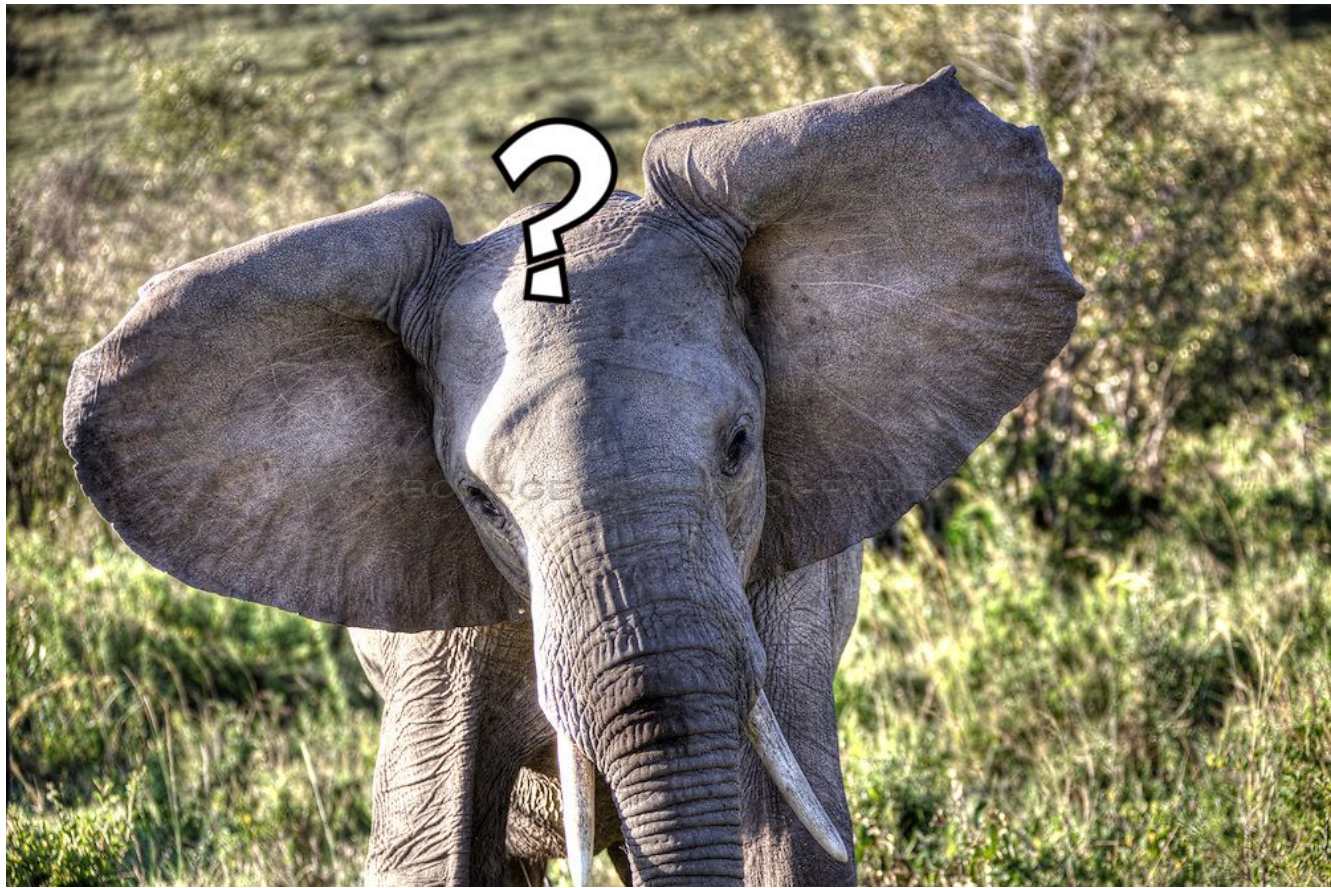
# Wanna use it?

You only need two things --

- (1) Your peer list, saved as a csv
- (2) The link to the app!

**<https://tinyurl.com/ipeds-compiler>**

# Questions?



<https://tinyurl.com/ipeds-compiler>

# Thank You!

Kate Aloisio  
Smith College  
[kaloisio@smith.edu](mailto:kaloisio@smith.edu)

Kathy Foley  
Smith College  
[kfoley@smith.edu](mailto:kfoley@smith.edu)

Emma Morgan  
Tufts University  
[emma.morgan@tufts.edu](mailto:emma.morgan@tufts.edu)



<https://tinyurl.com/ipeds-compiler>