Using Ten Years of IPEDS Access Databases to Create National and State Peer Institution Visualizations

Miriam B. Qumsieh, Director of Institutional Research, UH-Clear Lake

Henok Gebrehiwot, Sr. Institutional Analyst, UT-Dallas

Session Outline

During the presentation, we will:

▶ Unveil 10 years of IPEDS data's transformation, shedding light on the methodology behind data formatting, structure, and appending.

▶ Explore the influential role of data visualization in shaping decision-making processes within educational institutions.

▶ Conduct a deep dive into key performance metrics, allowing universities to assess their position in critical areas such as enrollment.

University
of Houston
Clear Lake

Institutional
Research

Session Outcomes

Outcomes:

▶ To demonstrate an enhanced understanding of the value of IPEDS data and data visualization techniques in higher education institutions.

▶ Insight into your university's performance compared to national peer group and state peer institutions, complemented by actionable steps for improvement.

▶ We will provide a sample of case studies and best practices for institutions seeking to optimize their use of IPEDS data.

University
of Houston
Clear Lake

Institutional
Research

IPEDS Projects Objectives

To develop data visualizations that enable peer comparisons for benchmarking and strategic planning.

Data Source: IPEDS

- Integrated Postsecondary Education Data System (IPEDS) is a comprehensive data collection system managed by the National Center for Education Statistics (NCES), part of the U.S. Department of Education.

- IPEDS gathers data from higher education institutions in the United States, including universities, colleges, and technical and vocational institutions..

University of Houston Clear Lake
Institutional Research

Understanding IPEDS Data

IPEDS Data:

- IPEDS releases an Access database each Academic Year.

- Each Access database contains tables of data derived from the various reporting components reported by each institution and includes metadata.

- Tables are labeled with the respective academic year, providing a clear chronological context.

- https://nces.ed.gov/ipeds/use-the-data/download-access-database

Data Downloads:

- We downloaded IPEDS Access databases from 2004 to 2022 from the IPEDS website; the 2023 IPEDS data is not finalized. (Note: Data is available in CSV format for 1980-2003)

Fall Enrollment: Visualization Objectives

The Fall Enrollment dashboard will facilitate a comparative analysis of the data for UHCL and its Peer Universities. The dashboard encompasses the following components:

- Fall Student Enrollment for UHCL and Peer Universities (one fall term at a time)

- Geographical Location Comparison offers a visual comparison of the geographical locations of Peer Universities, aiding in understanding regional enrollment patterns.

- Sex and Race/Ethnicity Comparison data among Peer Universities, to understand the demographic composition across institutions

- 10-Year Trend Analysis of Fall Student Enrollment compares Fall student enrollment for UHCL and Peer Universities, identifying patterns and fluctuations and offering insights into long-term enrollment trends.

University
of Houston
Clear Lake

Institutional
Research

Data Transformation in Power BI

Strategy 1

Working within MS Power BI

▶ Identified tables needed for Fall Enrollment data (EF2012A-EF2022A)

▶ Append Query (DAX) operation in Power BI combined the data from multiple tables in multiple databases, resulting in a MASTER table encapsulating the data spanning all ten years.

▶ Due to the absence of institution names and other institutional characteristics in the Fall Enrollment tables, a lookup table containing UHCL's peer institutions, including UNITID, Institution Name, City, State, etc was created.

▶ Used MetaData in the Access Databases to build additional lookup tables and to label table values (valuesets12-valuesets22)

▶ Imported lookup tables into Power BI, joined tables, and filtered in Power BI.

University
of Houston
Clear Lake

Institutional
Research

Pros and Cons of Strategy 1

Pros: Data Transformation within Power BI

- Simplifies the process by performing data transformation and processing directly within Power BI.

- It's a self-contained workflow, reducing the need for additional tools or platforms.

- It is inexpensive since Power BI has a free desktop version

- In the absence of an on-premises or cloud-based database/data warehouse

  o This approach eliminates the need for additional software, hardware, and person-hours for processing, potentially reducing complexity and resource demands.

University
of Houston
Clear Lake

Institutional
Research

Pros and Cons of Strategy 1

Cons: Challenges in Integrating New Data

- Adding a new year of data may cause data integration issues and inefficiencies.

- Processing data transformations of large data sets causes slow responses or freezing in the Power BI app.

- Unsustainable long-term and labor-intensive

University
of Houston
Clear Lake

Institutional
Research

Data Transformation in Microsoft SQL Server

Strategy 2

University
of Houston
Clear Lake

Institutional
Research

Working within MS SQL Server: Staging Environment

- 18 years of Access DBs data was loaded into the staging on-premises data warehouse: Microsoft SQL Server (2004-2022)

- Identified the tables needed for Fall Enrollment data (EF2012A-EF2022A)
- Identified the required tables for Institutional Characteristics data (HD2012-HD2022)

University
of Houston
Clear Lake

Institutional
Research

Working within MS SQL Server: Staging Environment

- MetaData in the Access Databases is used to build additional lookup tables (dimension tables) to enhance readability used to label table values (valuesets12-valuesets22)

- Create a view of 10 years of Fall Enrollment and a view of Institutional Characteristics data in the staging environment using a UNION query in MS SQL.

- Created a fact table containing UHCL's peer institutions and identifiers not found in the IC tables (UNITID, MSI identifier, MSI designations, etc.)

University
of Houston
Clear Lake

Institutional
Research

Working within MS SQL Server: Production Data Warehouse

- We created materialized queries (fact tables) in the production data warehouse from this view.

- Pushed the necessary lookup tables (dimension tables) and the peer institutions' fact table to the production data warehouse

- Data transformations needed in the production data warehouse are done during the materialized query process

- Power BI reads data directly from the data warehouse; fact and dimension table joins are done in Power BI.

University
of Houston
Clear Lake

Institutional
Research

Fact and Dimension Tables

Fact Tables

- **F_ENROLLMENT** – Fall enrollment materialized query containing 10 years of data

- **F_INSITITUTIONAL_CHARACTERISTICS** – IC materialized query containing 10 years of data

- **F_PEER_INSTITUTIONS** – Table containing UHCL's peer institutions and other characteristics

Dimension Tables

- **L_EFLEVEL** – Student level (i.e. Full-time students, Undergraduate, Degree-seeking)

- **L_HDCARNEGIE** – Carnegie Classification

- **L_HDSECTOR** – Institution Sector Type (i.e. Public, 4-year or above, Public, 2-year)

- **L_HDINSTSIZE** – Institution Size

- **L_MSI** - Minority Serving Institution Types

University
of Houston
Clear Lake

Institutional
Research

Power BI Joins

Table Joins

- F_ENROLLMENT(UNITID, YEAR) <-> F_INSITTUTIONAL_CHARACTERISTICS (UNITID, YEAR)

- F_ ENROLLMENT(EFALEVEL) <-> L_EFALEVEL (EFALEVEL)

- F_ ENROLLMENT(CARNEGIE) <-> L_HDCARNEGIE (CARNEGIE)

- F_ ENROLLMENT(SECTOR) <-> L_HDSECTOR (SECTOR)

- F_ ENROLLMENT(INSTSIZE) <-> L_HDINSTSIZE (INSTSIZE)

- F_ ENROLLMENT (UNITID) <-> F_PEER_INSTITUTIONS (UNITID)

University
of Houston
Clear Lake

Institutional
Research

F_ENROLLMENT

YEAR_UNITID	YEAR	UNITID	EFALEVEL	LINE	SECTION	LSTUDY	EFTOTLT	EFTOTLM	EFTOTLW
2014:225414	2014	225414	1	29	3	4	8665	3280	5385
2014:225414	2014	225414	2	99	3	1	5077	1657	3420
2014:225414	2014	225414	3	99	3	1	5072	1655	3417
2014:225414	2014	225414	4	99	3	1	234	75	159
2014:225414	2014	225414	5	99	3	1	4838	1580	3258
2014:225414	2014	225414	11	99	3	1	5	2	3
2014:225414	2014	225414	12	99	3	3	3588	1623	1965
2014:225414	2014	225414	19	99	3	1	1256	407	849
2014:225414	2014	225414	20	99	3	1	3582	1173	2409
2014:225414	2014	225414	21	14	1	4	4270	1818	2452
2014:225414	2014	225414	22	8	1	1	2374	825	1549
2014:225414	2014	225414	23	6	1	1	2373	825	1548
2014:225414	2014	225414	24	1	1	1	202	66	136
2014:225414	2014	225414	25	99	1	1	2171	759	1412
2014:225414	2014	225414	31	7	1	1	1	0	1
2014:225414	2014	225414	32	11	1	3	1896	993	903
2014:225414	2014	225414	39	2	1	1	620	222	398
2014:225414	2014	225414	40	3	1	1	1551	537	1014
2014:225414	2014	225414	41	28	2	4	4395	1462	2933
2014:225414	2014	225414	42	22	2	1	2703	832	1871

University
of Houston
Clear Lake

Institutional
Research

F_INSTITUTIONAL_CHARACTERISTICS

YEAR_UNITID	YEAR	UNITID	INSTNM	CITY	STABBR	LONGITUD	LATITUDE	CARNEGIE	SECTOR	INSTSIZE
2013:225414	2013	225414	University of Houston-Clear Lake	Houston	TX	-95.0981	29.58252	21	1	3
2014:225414	2014	225414	University of Houston-Clear Lake	Houston	TX	-95.0981	29.58252	21	1	3
2015:225414	2015	225414	University of Houston-Clear Lake	Houston	TX	-95.0981	29.58252	21	1	3
2016:225414	2016	225414	University of Houston-Clear Lake	Houston	TX	-95.0981	29.58252	21	1	3
2017:225414	2017	225414	University of Houston-Clear Lake	Houston	TX	-95.0981	29.58252	21	1	3
2018:225414	2018	225414	University of Houston-Clear Lake	Houston	TX	-95.0981	29.58252	21	1	3
2019:225414	2019	225414	University of Houston-Clear Lake	Houston	TX	-95.0981	29.58252	21	1	3
2020:225414	2020	225414	University of Houston-Clear Lake	Houston	TX	-95.0981	29.58252	21	1	3
2021:225414	2021	225414	University of Houston-Clear Lake	Houston	TX	-95.0981	29.58252	21	1	3
2022:225414	2022	225414	University of Houston-Clear Lake	Houston	TX	-95.0981	29.58252	21	1	3

University
of Houston
Clear Lake

Institutional
Research

F_PEER INSTITUTIONS

UNITID	PEER_UNITID	National_Peer	State_Peer	SACSCOC_Region	MSI	HSI	HBCU	AANAPISI	AANH	NASNTI	TCU	PBI
225414	407009	Y	N	N	N							
225414	165820	Y	N	N	N							
225414	145336	Y	N	N	N							
225414	151342	Y	N	N	N							
225414	224147	Y	N	Y	Y	HSI						
225414	138354	Y	N	Y	N							
225414	161873	Y	N	N	N							
225414	126580	Y	N	N	N							
225414	225414	Y	Y	Y	Y	HSI						
225414	148654	Y	N	N	N							
225414	171137	Y	N	N	N							
225414	222831	N	Y	Y	Y	HSI						
225414	226833	N	Y	Y	N							
225414	228501	N	Y	Y	Y	HSI						
225414	483036	N	Y	Y	Y	HSI						
225414	228714	N	Y	Y	N							
225414	224545	N	Y	Y	N							
225414	228705	N	Y	Y	Y	HSI						
225414	227377	N	Y	Y	Y	HSI						
225414	228802	N	Y	Y	N							
225414	229018	N	Y	Y	Y	HSI						
225414	225432	N	Y	Y	Y	HSI						
225414	225502	N	Y	Y	Y	HSI						
225414	443711	N	Y	Y	Y	HSI						

University
of Houston
Clear Lake

Institutional
Research

L_EFALEVEL

EFALEVEL	EFALEVEL_DESCR
1	All students total
2	All students, Undergraduate total
3	All students, Undergraduate, Degree/certificate-seeking total
4	All students, Undergraduate, Degree/certificate-seeking, First-time
5	All students, Undergraduate, Other degree/certificate-seeking
19	All students, Undergraduate, Other degree/certificate-seeking, Transfer-ins
20	All students, Undergraduate, Other degree/certificate-seeking, Continuing
11	All students, Undergraduate, Non-degree/certificate-seeking
71	All students, Undergraduate, Degree-seeking status unknown - Nondegree-granting institutions
12	All students, Graduate
21	Full-time students total
22	Full-time students, Undergraduate total
23	Full-time students, Undergraduate, Degree/certificate-seeking total
24	Full-time students, Undergraduate, Degree/certificate-seeking, First-time
25	Full-time students, Undergraduate, Degree/certificate-seeking, Other degree/certificate-seeking

University
of Houston
Clear Lake

Institutional
Research

Power BI Relationships
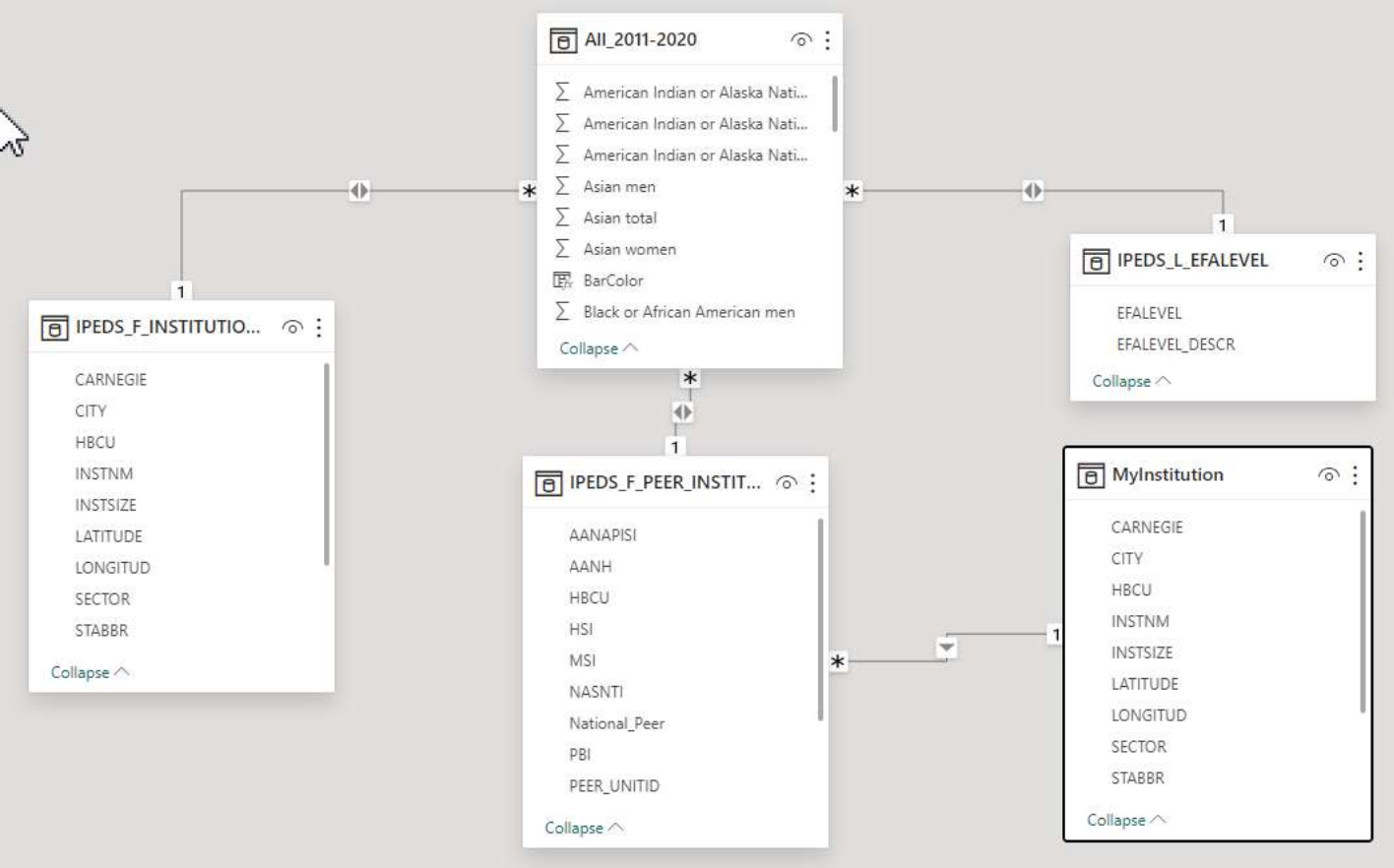
Fall Enrollment Visualization

IPEDS Peer Comparison: Fall Enrollment
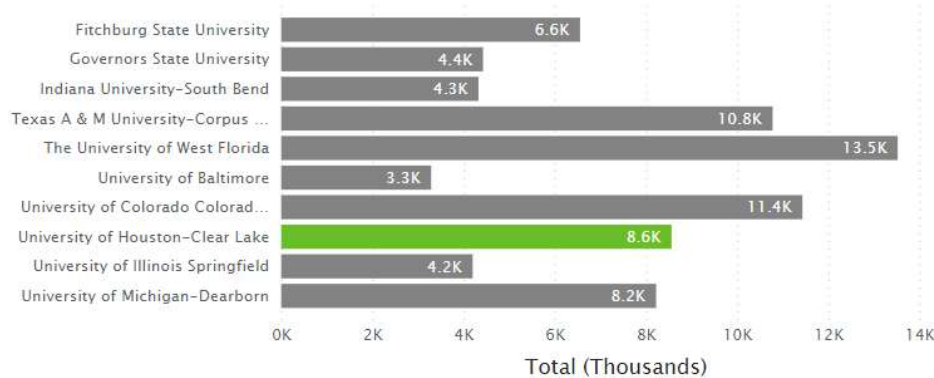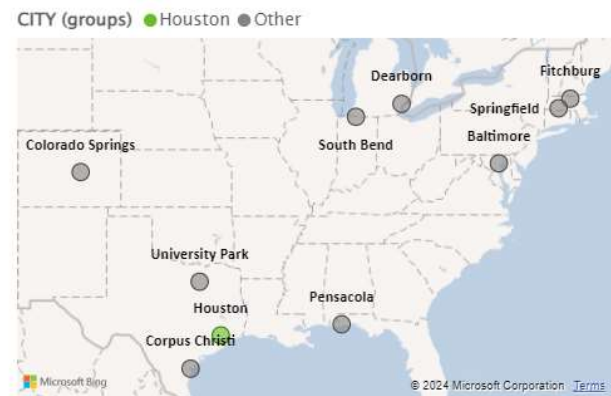This report presents Fall enrollment comparison among national peer groups

Year
2022

Enrollment Category
All students total

Student Enrollment by University

| University	Total (Thousands)
Fitchburg State University	6.6K
Governors State University	4.4K
Indiana University–South Bend	4.3K
Texas A & M University–Corpus ...	10.8K
The University of West Florida	13.5K
University of Baltimore	3.3K
University of Colorado Colorad...	11.4K
University of Houston–Clear Lake	8.6K
University of Illinois Springfield	4.2K
University of Michigan–Dearborn	8.2K

University by Location

CITY (groups) ● Houston ● Other

Peer University by Gender, and Race/Ethnicity

INSTNM	Total Men	Total Women	White	Hispanic	Black	Asian	International	Two and more race	Native Hawaiian	America
Fitchburg State University	2,487	4,067	4,414	689	608	220	265	100	6	
Governors State University	1,600	2,827	1,279	662	1,566	110	553	124	3	
Indiana University–South Bend	1,443	2,883	2,734	785	367	91	80	235	0	
Texas A & M University–Corpus Christi	4,237	6,541	3,833	5,043	518	439	473	262	11	
The University of West Florida	5,135	8,386	8,154	1,617	1,481	501	398	730	30	
University of Baltimore	1,212	2,076	1,223	281	1,254	170	91	161	2	
University of Colorado Colorado Springs	5,163	6,268	6,554	2,085	491	425	164	781	31	
University of Houston–Clear Lake	3,146	5,416	2,658	3,609	665	643	617	263	5	

Fall Enrollment 10-Year Trend Visualization
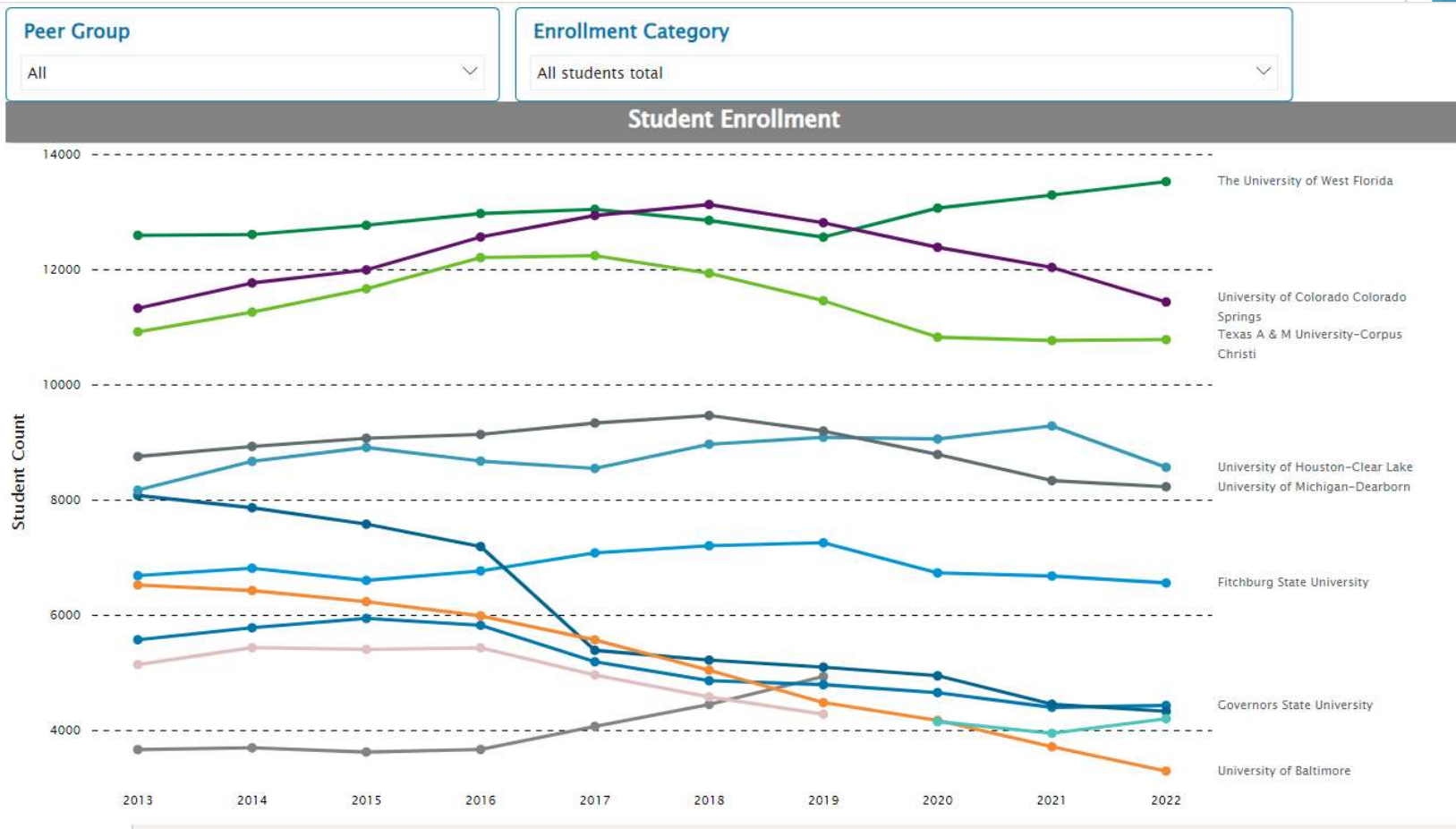
Project Challenges

- **Data Formatting:** The raw data in the IPEDS Access databases is not easily digestible. It takes time to understand the file structure and the metadata.

- **Data Structure:** The IPEDS data is structured as aggregate row data. Hence, the data is not normalized. This makes it challenging to aggregate and apply slicer rules effectively.

- **Data Appending:** Each academic year's data is stored in a separate database. Creating a dashboard that spans 10 years requires importing and integrating data from 10 different databases.

- **Long-term Data Integration:** There are many factors to consider and audit when adding a new year of data; the more automated the process, the more efficient and sustainable. Note: There were changes in the Fall Enrollment table from 2020 to 2021.

University
of Houston
Clear Lake

Institutional
Research

Future Improvements and Project Developments

- Enhance the Fall Enrollment dashboard to meet UHCL's data visualization standards.

- Promote the Scholarship and Allowance dashboard into production

- Develop a 10-year Completions dashboard using IPEDS Completions data

- Develop a Full-time/Part-time faculty and staff by occupational category dashboard using the IPEDS Human Resources data

- Develop a Graduation Rates dashboard using the IPEDS Graduation Rate and Graduation Rate 200 data

- Investigate the development of Finance and Student Financial Aid dashboards

- Investigate if other institutions are interested in having access to our work
  - UHCL has IPEDS Data for **All Colleges and Universities** who report to IPEDS accessible in our data warehouse

University
of Houston
Clear Lake

Institutional
Research

Thank you

University of Houston-Clear Lake

Office of Institutional Research

- ir@uhcl.edu
- www.uhcl.edu/ir

Special thanks to:

- Kurt Lund, DMO
- Steven Fernandez, Client Server DB Analyst

Special recognition to:

- Henok Gebrehiwot, Sr. IR Analyst at UT-Dallas
- Shalini Dhamodharan, Intern: Data Science MS

University of Houston Clear Lake

Institutional Research