Understanding, Creating and using OLAP cubes for multidimensional reporting: *The* coolest *visualization you're probably not using*

John Tong and Andrea Halow

Enterprise Information Technology Services
UNIVERSITY OF GEORGIA

eVisions conference

About the University of Georgia

Chartered in 1785 we are America's first state chartered and state supported land grant university

Fall 2017 enrollment totaled 37,606

Undergraduates 28,848

Graduate and Professional 8,758

Work force of 10,370 faculty, staff and administration

UGA went live with Banner and Argos in Fall of 2014

John Tong

Application Analyst Principal

Student Reporting Team Lead (URR/ODS/EDW)

Responsible for reporting and improving data quality and service with functional areas

20+ years of Banner/Higher Education experience

Part-time Door Guard

John Tong

Andrea Halow

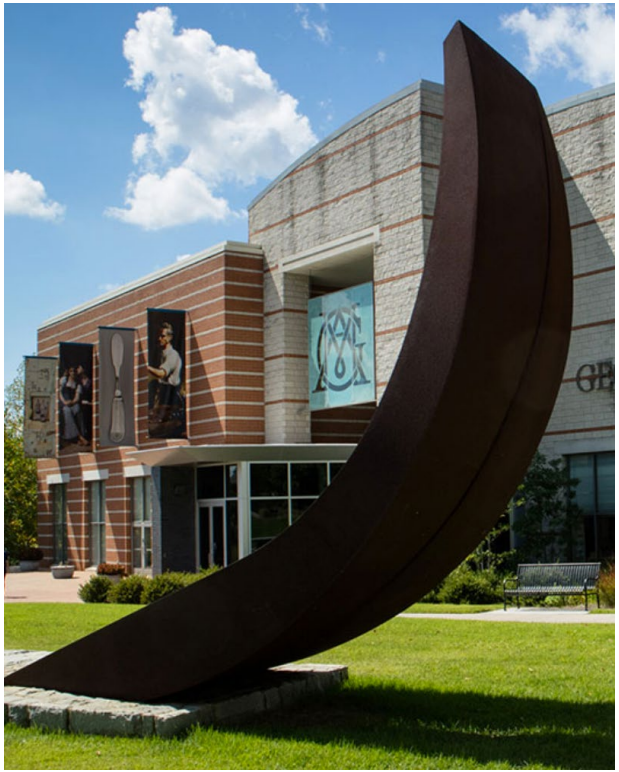
Application Analyst Specialist

Administrator for Argos reporting and development

Student Reporting and Administration (ODS/EDW)

15 years of higher education experience 12 of those years working in the Registrar's office

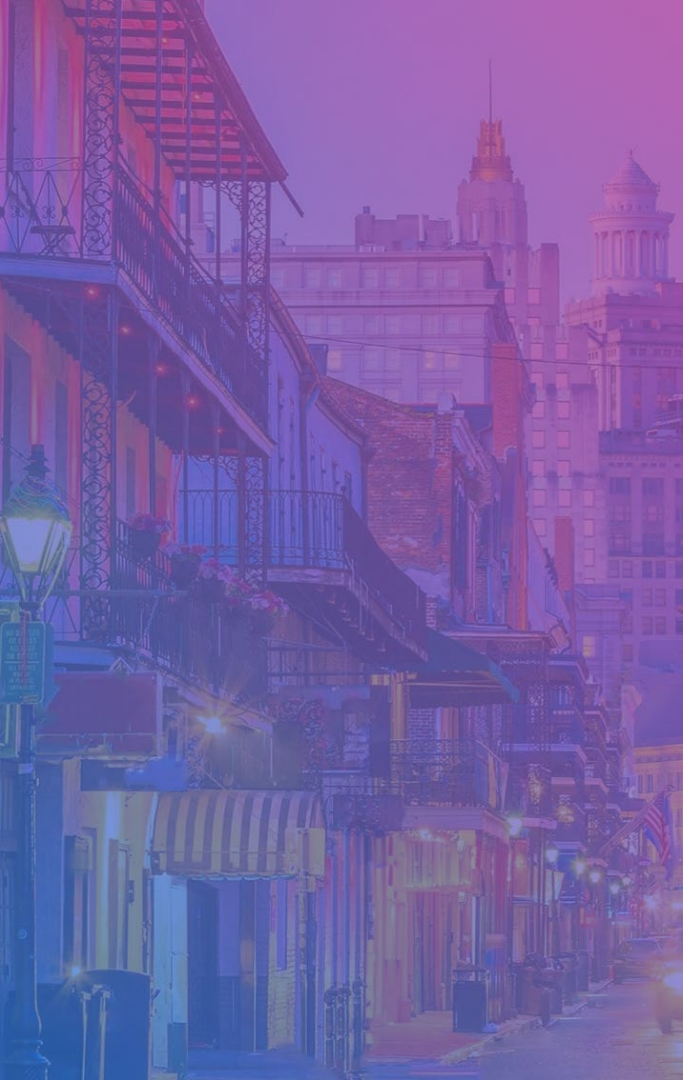
No part time job



Quick poll

How knowledgeable are you about OLAP cubes?

1. Know almost nothing
2. Heard of them, they might be useful
3. I've seen and used them, but don’t know how to make one
4. I know a pretty good bit, just here to see if the presenters do



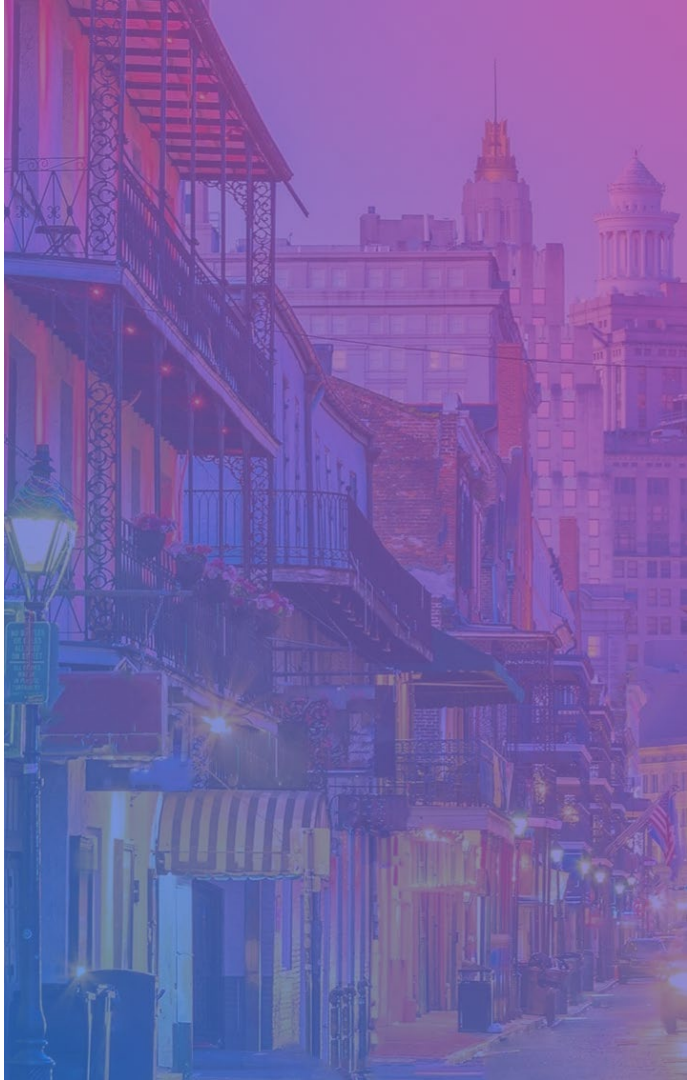
AGENDA

evisions conference

#Evisions19

1. OLAP Overview
2. Argos Demonstration
3. Creating an OLAP
4. Common Pitfalls
5. Tips and Tricks
6. Summary
7. Q & A

OLAP Cube Overview



WHAT IS AN OLAP CUBE?

OLAP is an acronym for **O**n**l**ine **A**nalytical **P**rocessing.

OLAP cubes are used to analyze a multidimensional dataset.

OLAPs are used for analysis and to look for insights or relations.

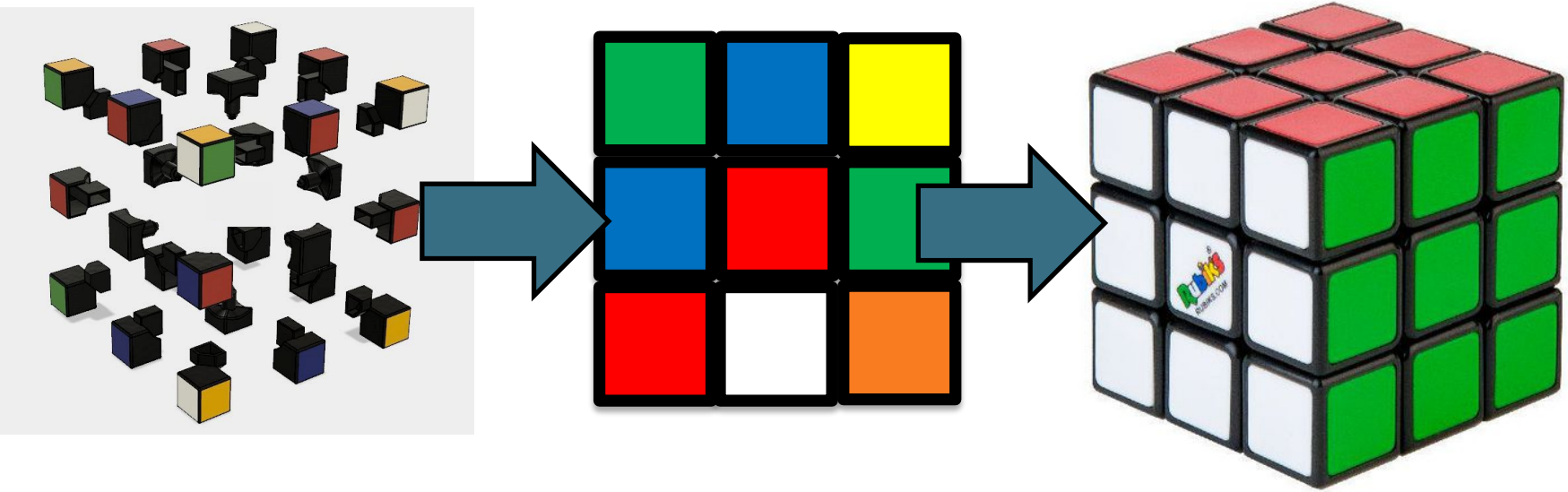
- A visualization tool that allows users to analyze data real time
- AKA 1000 reports in a box AKA Pivot Tables on Steroids
- Originally created for browsing EDW data, but can be adapted to ODS or Transactional data.

evisions conference

Your data

Your report

Your data in the cube

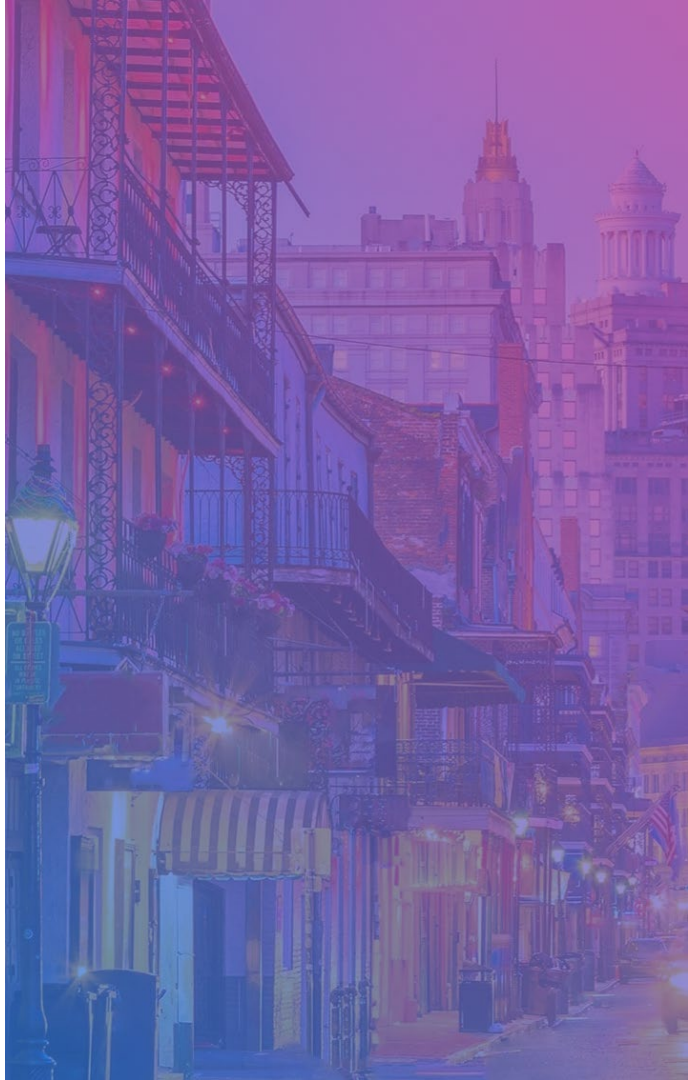


BEST USES FOR AN OLAP?

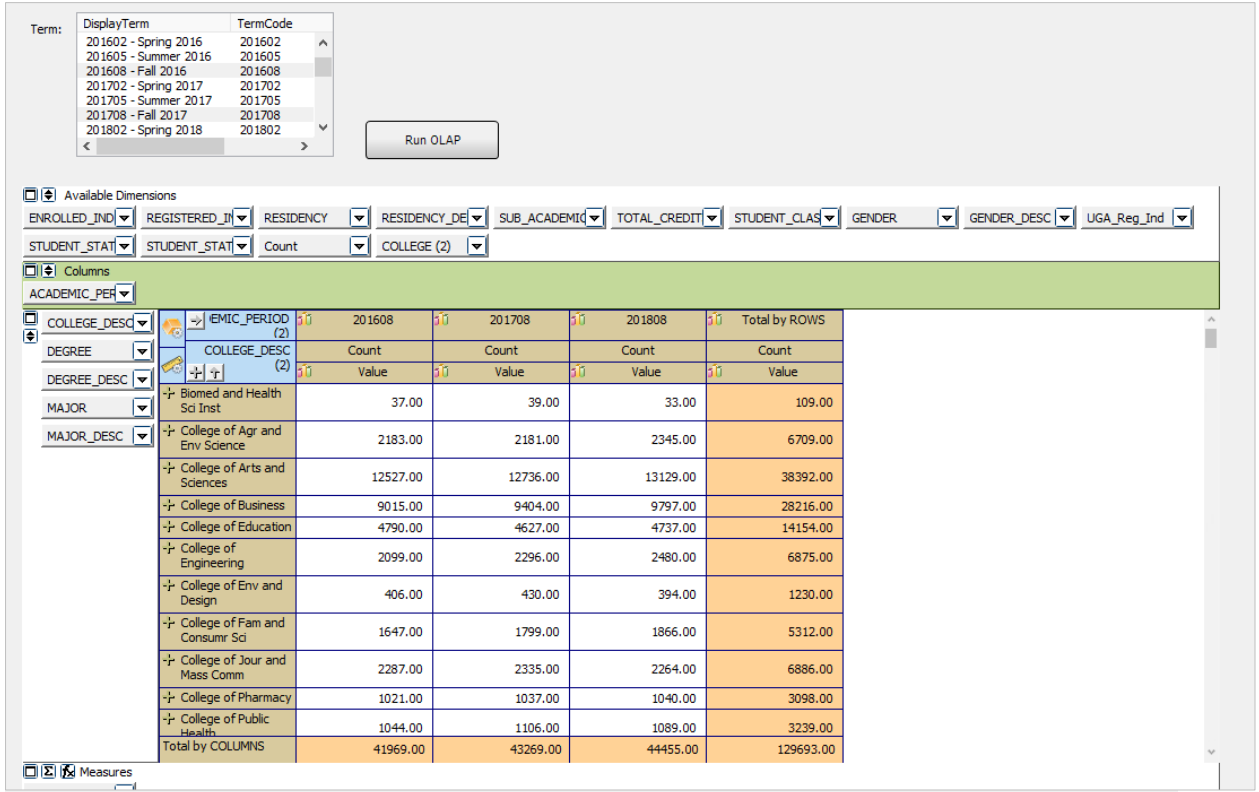
- Focused scope of data, but large in volume
- When dynamic analysis and trends are more important than straight lists
- For users who want more control/flexibility than a static report
- Where the data needs (like filters) may change frequently

evisions conference

Demo of OLAP Cubes

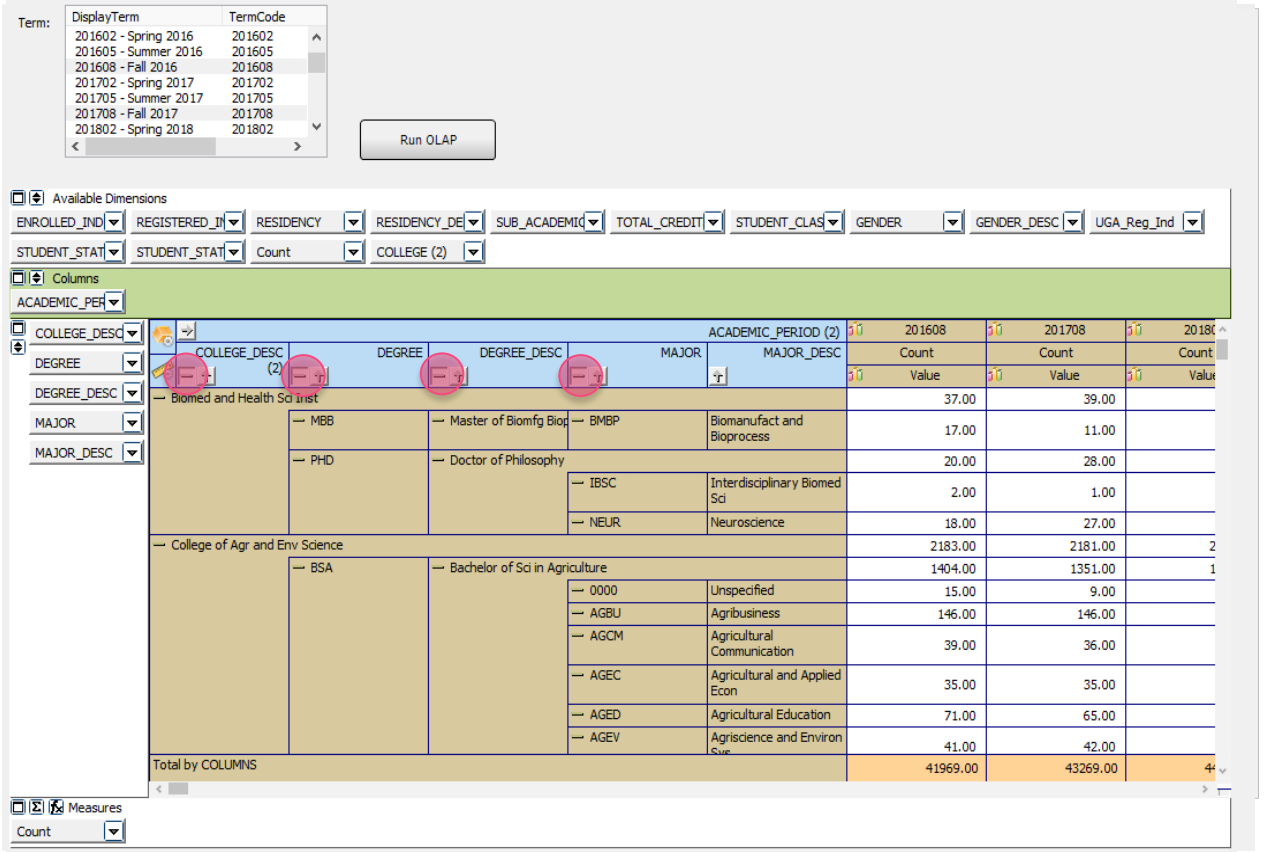


Demo


Term:
DisplayTerm
TermCode
201602 - Spring 2016
201605 - Summer 2016
201608 - Fall 2016
201702 - Spring 2017
201705 - Summer 2017
201708 - Fall 2017
201802 - Spring 2018
Run OLAP
Available Dimensions
ENROLLED_IND
REGISTERED_IN
RESIDENCY
RESIDENCY_DE
SUB_ACADEMIC
TOTAL_CREDIT
STUDENT_CLAS
GENDER
GENDER_DESC
UGA_Reg_Ind
STUDENT_STAT
STUDENT_STAT
Count
COLLEGE (2)
Columns
ACADEMIC_PER
COLLEGE_DESC
DEGREE
DEGREE_DESC
MAJOR
MAJOR_DESC
EMIC_PERIOD (2)
COLLEGE_DESC (2)
201608
201708
201808
Total by ROWS
Count
Value
Biomed and Health Sci Inst 37.00 39.00 33.00 109.00
College of Agr and Env Science 2183.00 2181.00 2345.00 6709.00
College of Arts and Sciences 12527.00 12736.00 13129.00 38392.00
College of Business 9015.00 9404.00 9797.00 28216.00
College of Education 4790.00 4627.00 4737.00 14154.00
College of Engineering 2099.00 2296.00 2480.00 6875.00
College of Env and Design 406.00 430.00 394.00 1230.00
College of Fam and Consumr Sci 1647.00 1799.00 1866.00 5312.00
College of Jour and Mass Comm 2287.00 2335.00 2264.00 6886.00
College of Pharmacy 1021.00 1037.00 1040.00 3098.00
College of Public Health 1044.00 1106.00 1089.00 3239.00
Total by COLUMNS 41969.00 43269.00 44455.00 129693.00
Measures

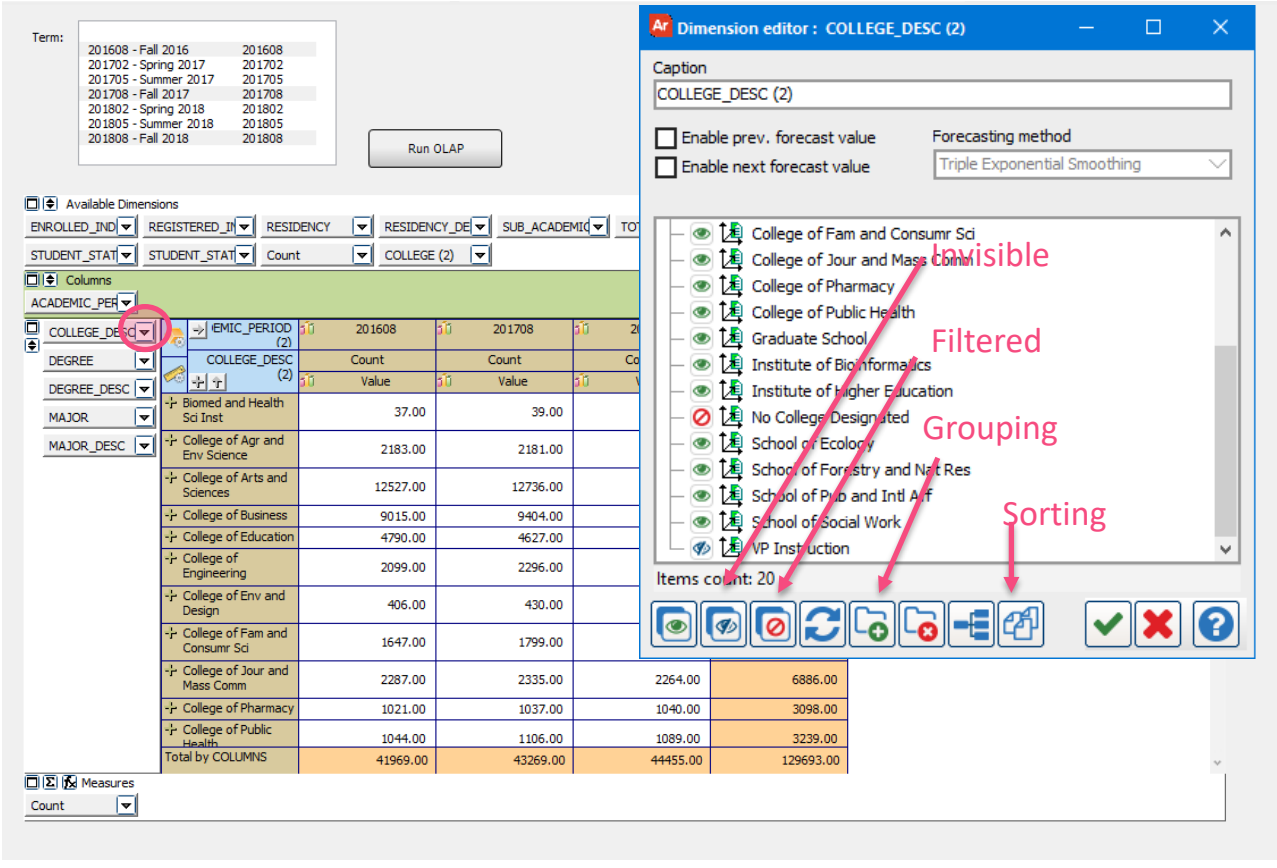
Demo

Drill Down



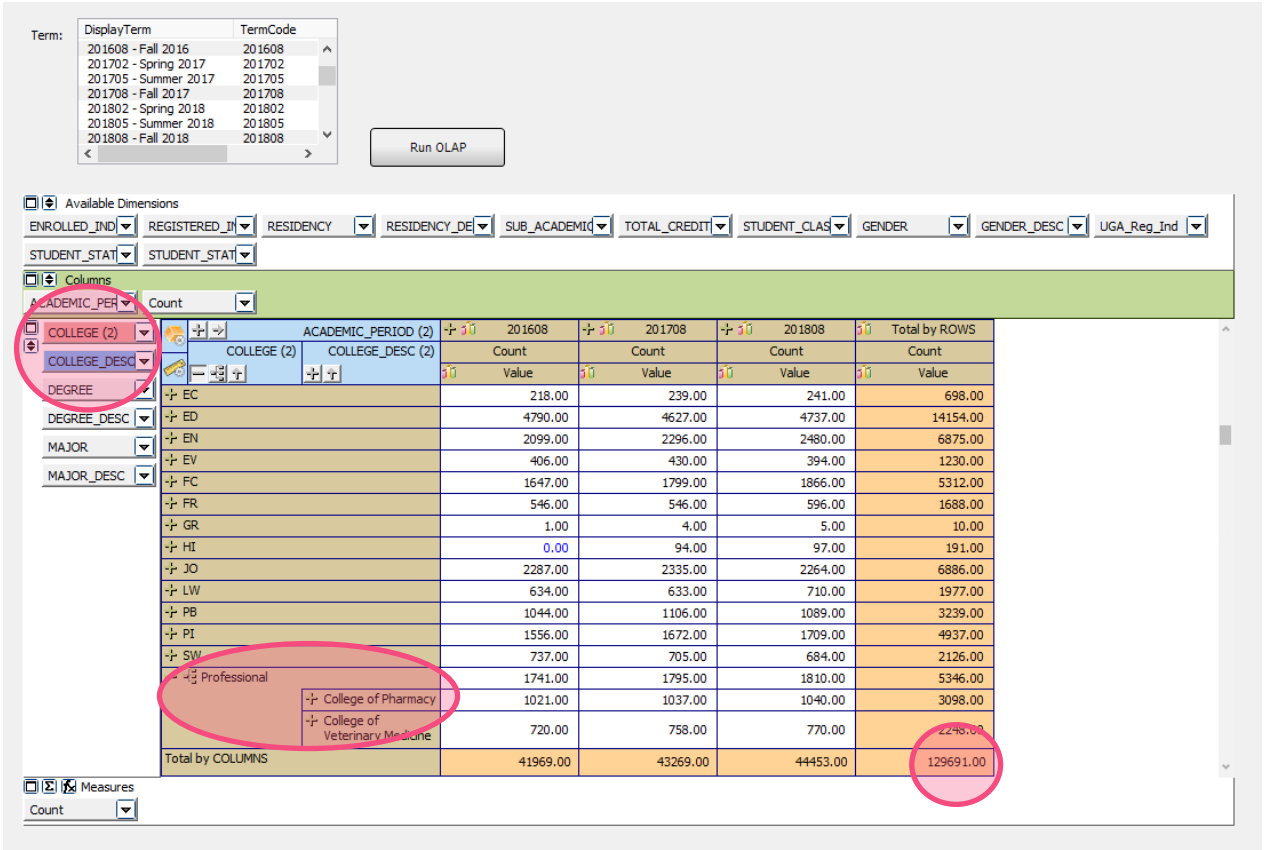
Demo

Filtering
Sorting
and
Grouping



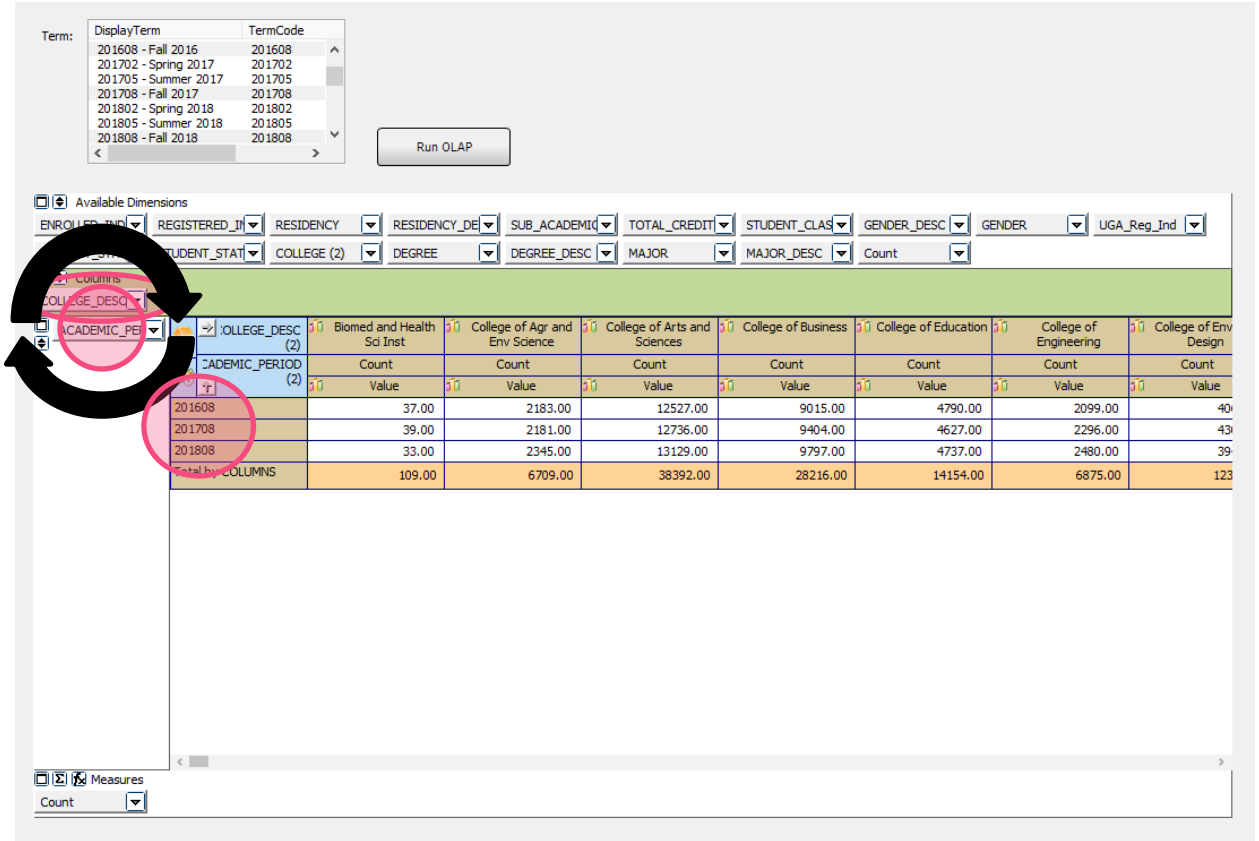
Demo

Filtering
Sorting
and
Grouping



Demo

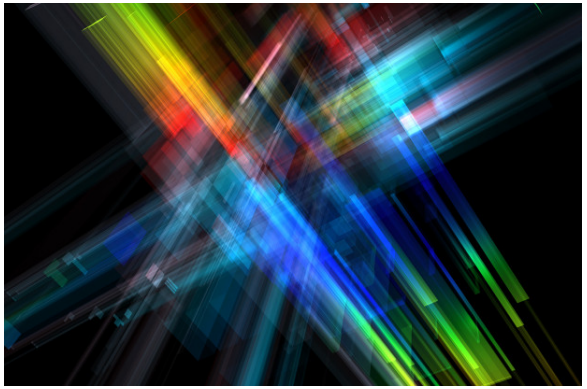
Rotating the Cube

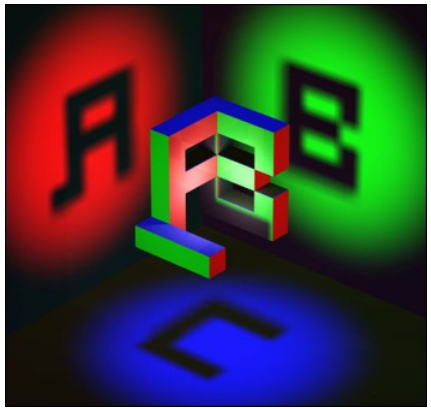


OLAP Vocabulary Summary

- **Dimension** – ‘Bucket’ or category you want to be able to slice/dice by. For example, Demographics, Major, Account Type. Usually comes from a Dimension table in Star Schemas.
- **Measures** – Numerics that you want to track. Usually a count or an amount like money.
- **Hierarchy** – Breakdown or ‘Roll up’ categories for Dimensions.
- **Granularity/Grain** – The lowest level of detail you can get to on a Cube. For example, for student accounts it might be Per Account Detail Transaction, or a summary Per Student. Too high level and you can’t get at the details, too low and you pay a performance cost.

DIMENSIONS

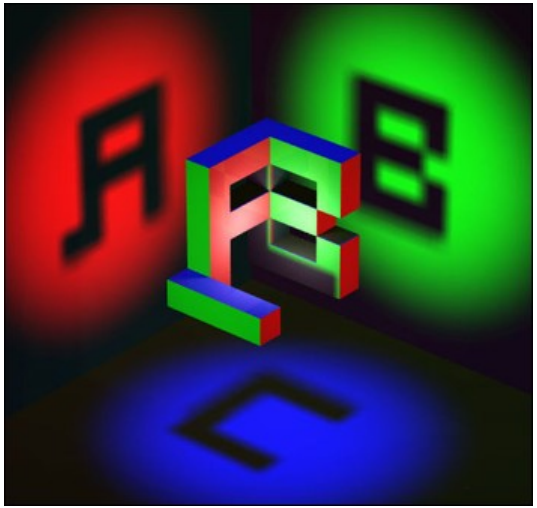




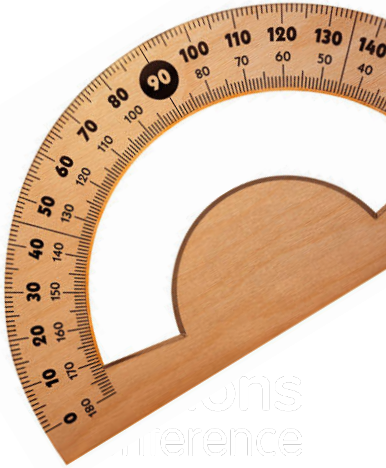
evisions conference

Dimension Techniques

- Dragging and Dropping
- Drilling down
- Rotating the cube
- Filtering
- Sorting
- Grouping



FACTS/ MEASURES



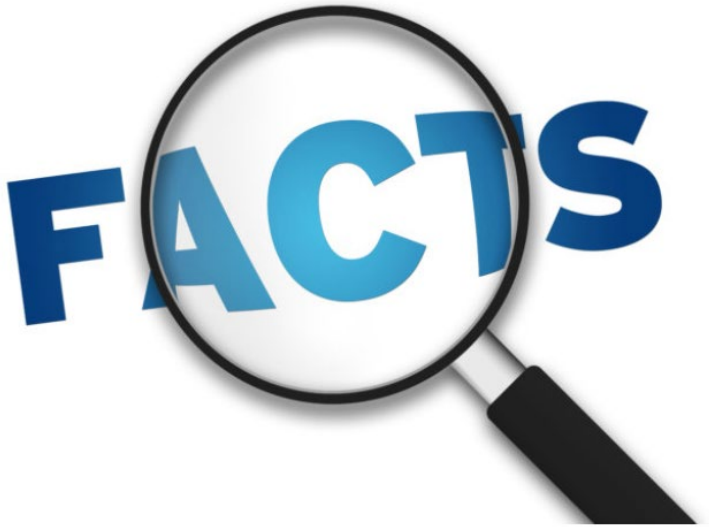
Measure Techniques

Can be calculated or built in:

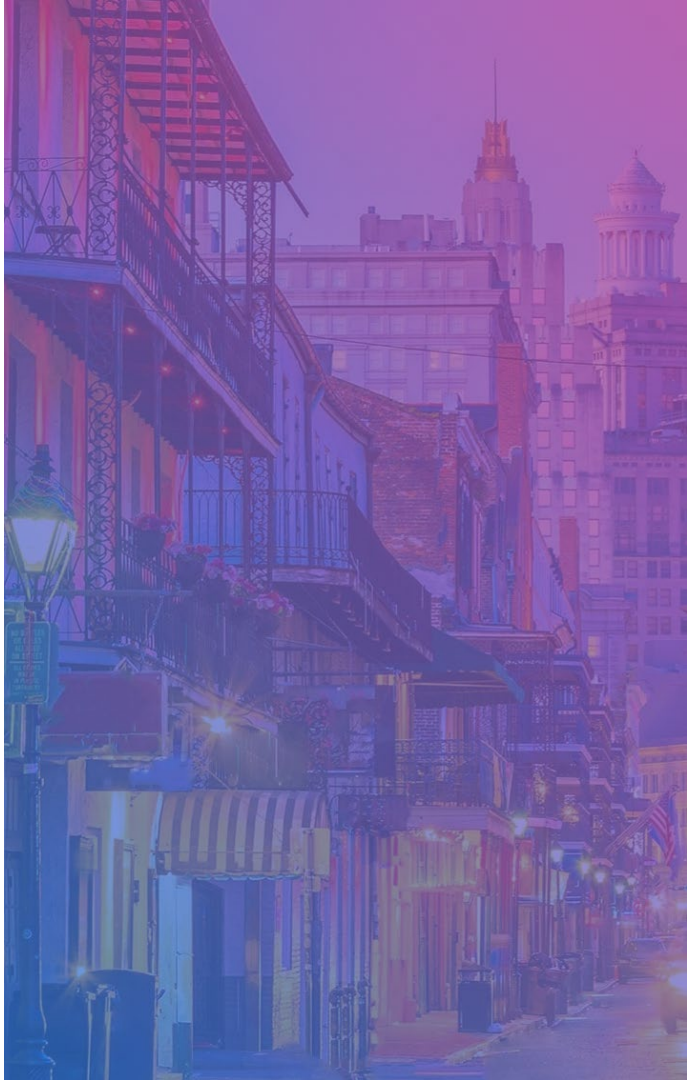
- Count
- Average
- Sum
- Percentages
- Create formulas

Types of Facts

- Factless facts
- Additive
- Semi-additive
- Derived



Designing an OLAP



OLAP Design Process

1. What questions are we trying to answer?

Involve functional from the start

OLAP Design Process

2. What kind of data do we need to answer the questions?

What measures and dimensions?
What data sources?

OLAP Design Process

3. What is the lowest level of detail needed?

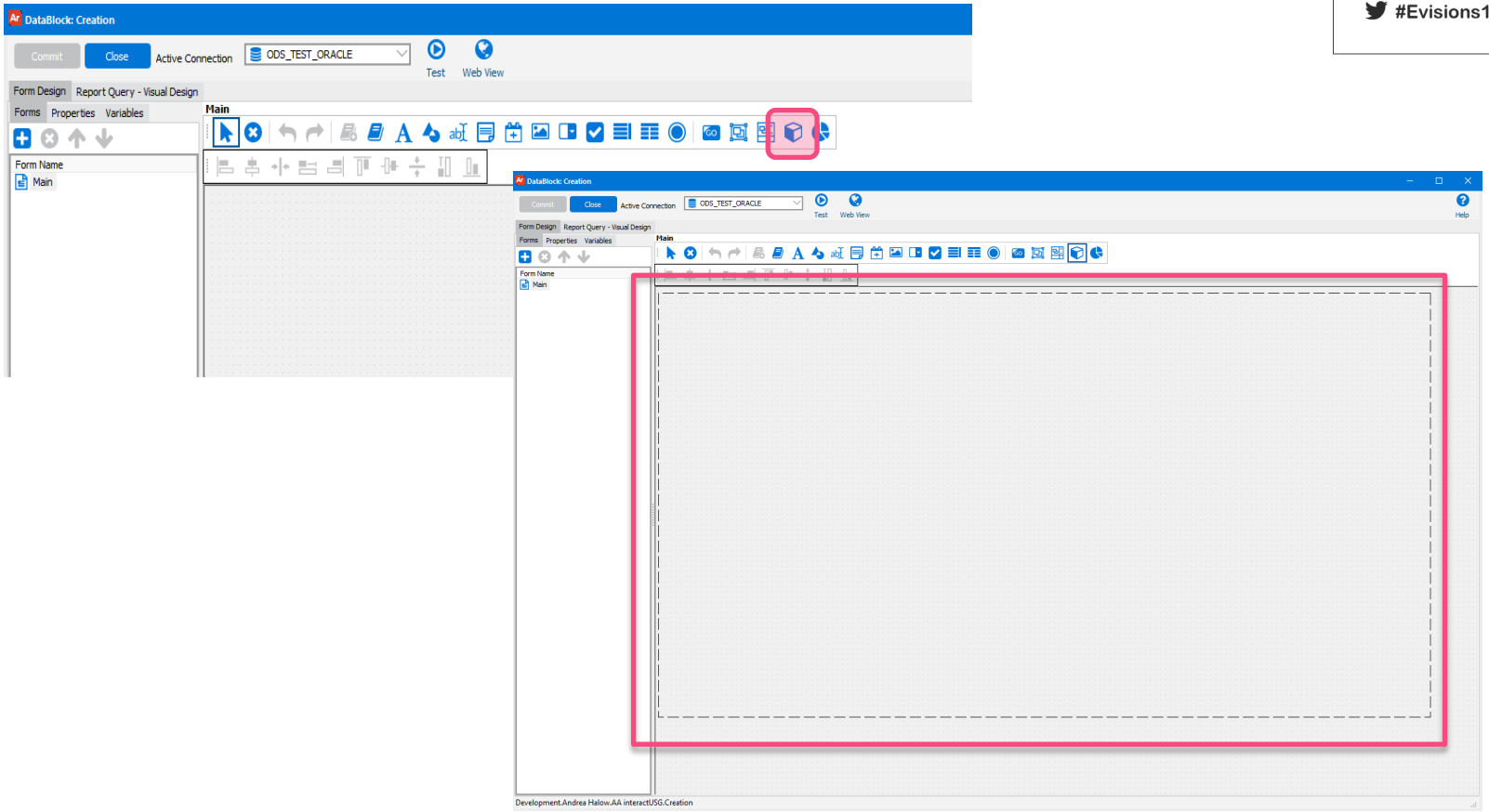
Granularity: Be flexible!

BUILD EXAMPLE IN ARGOS

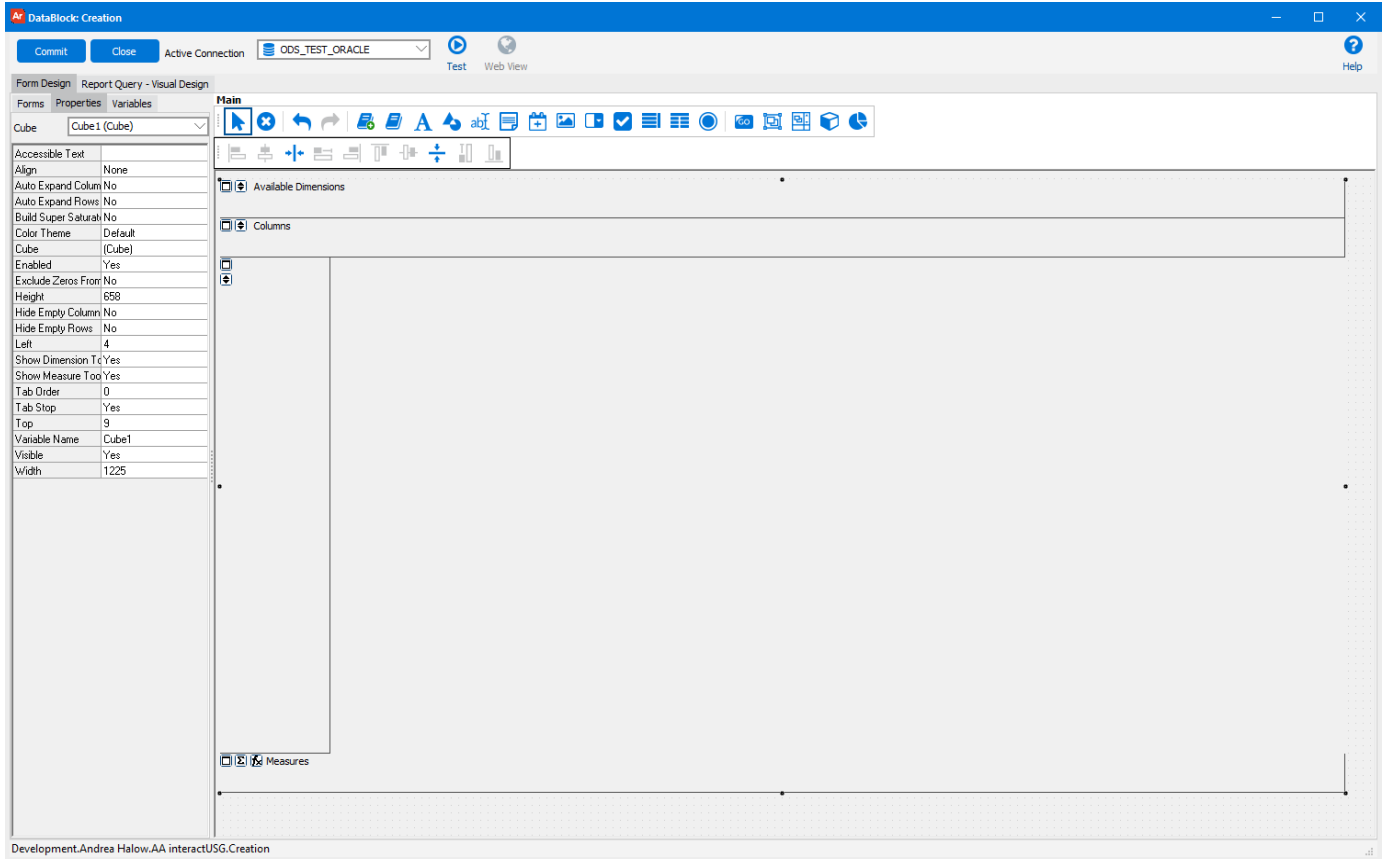
evisions
conference

Building an Argos OLAP example

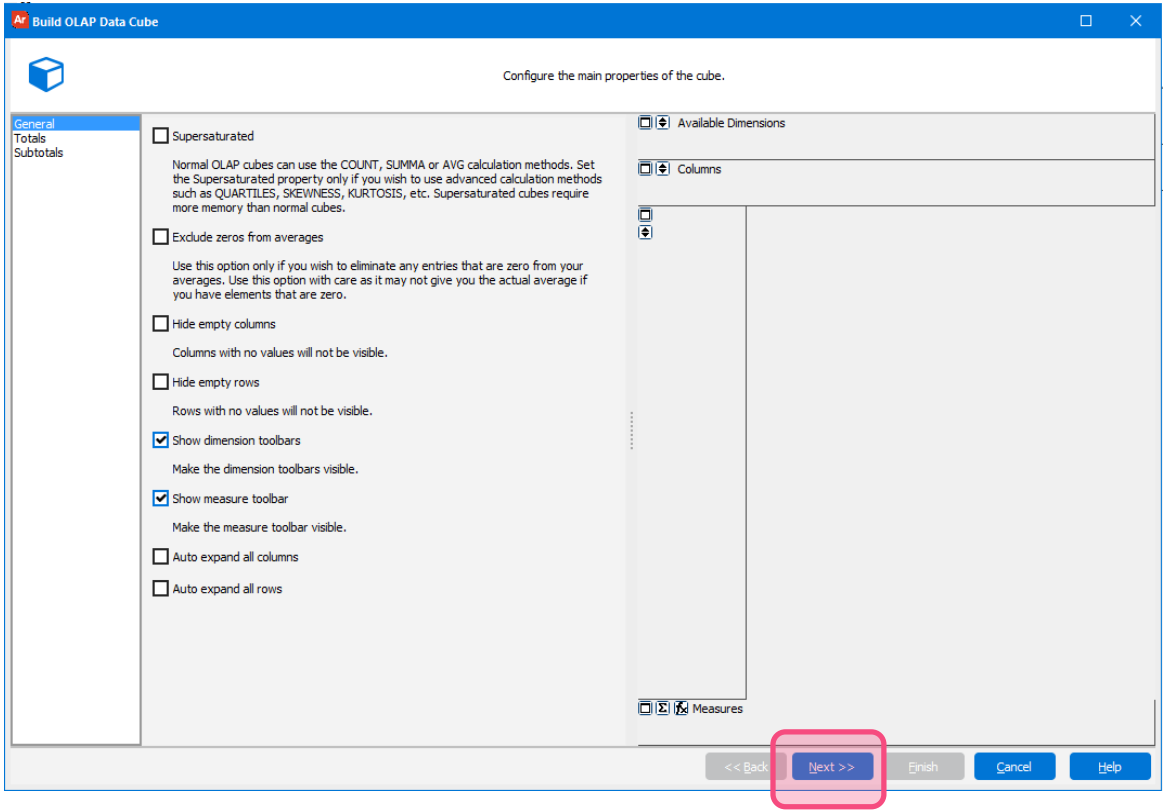
evisions conference
#Evisions19



Building an Argos OLAP example

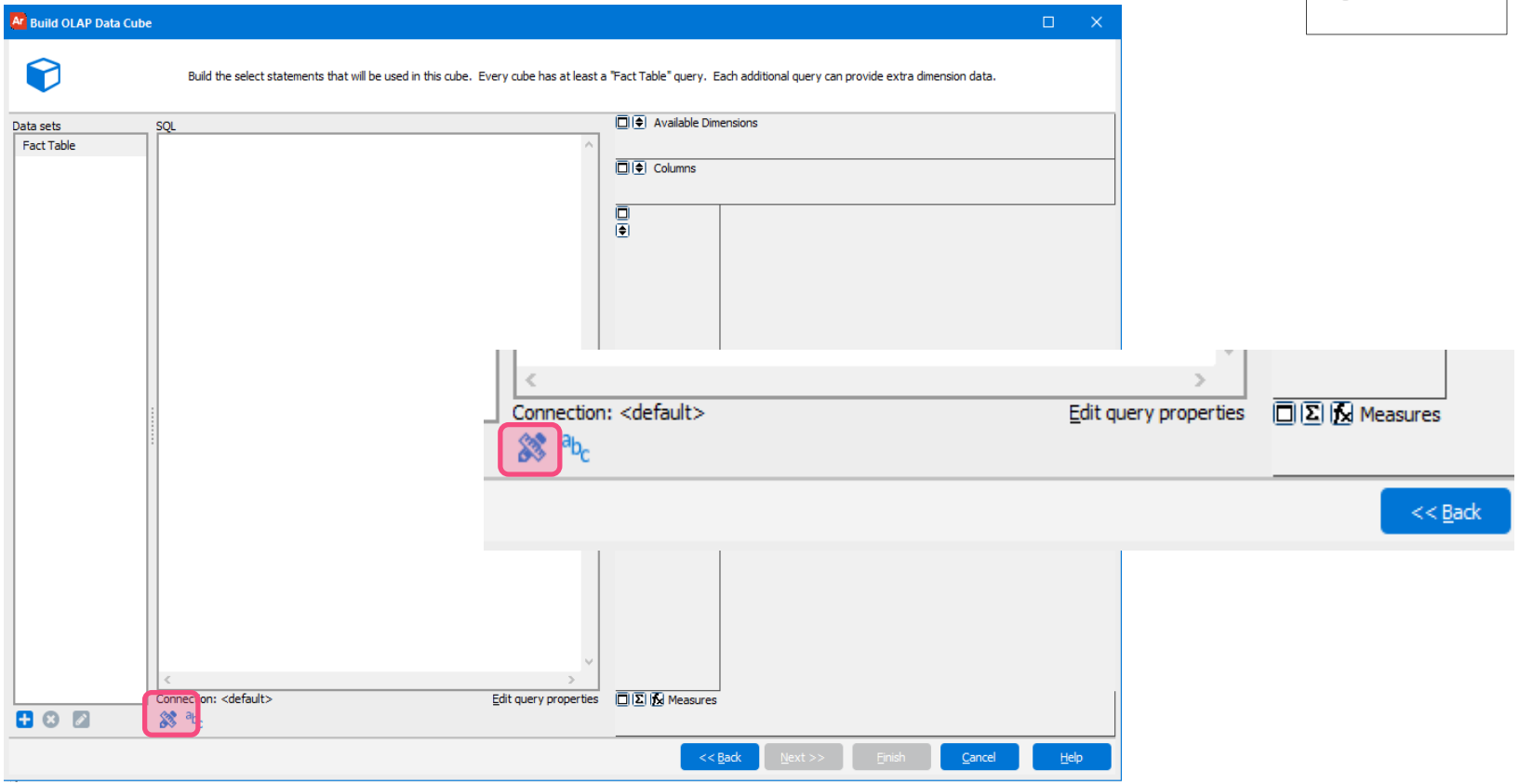


Building an Argos OLAP example



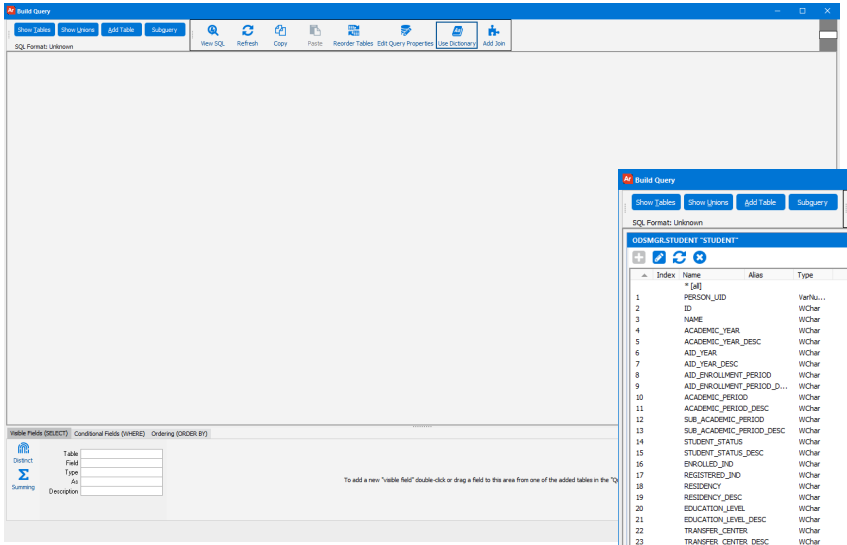
Building an Argos OLAP example

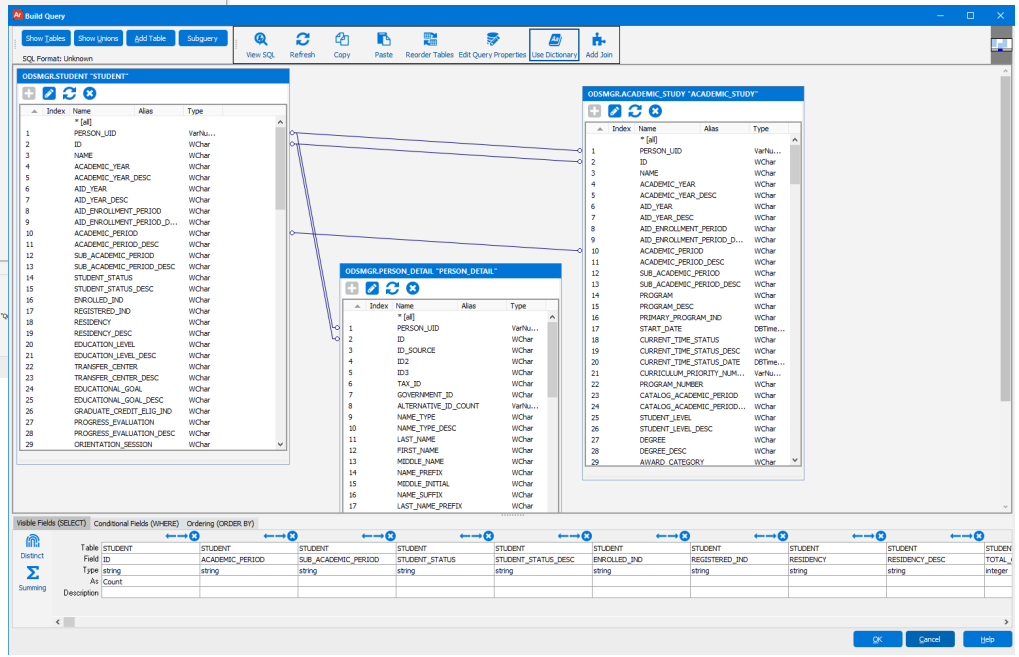
evisions conference
#Evisions19



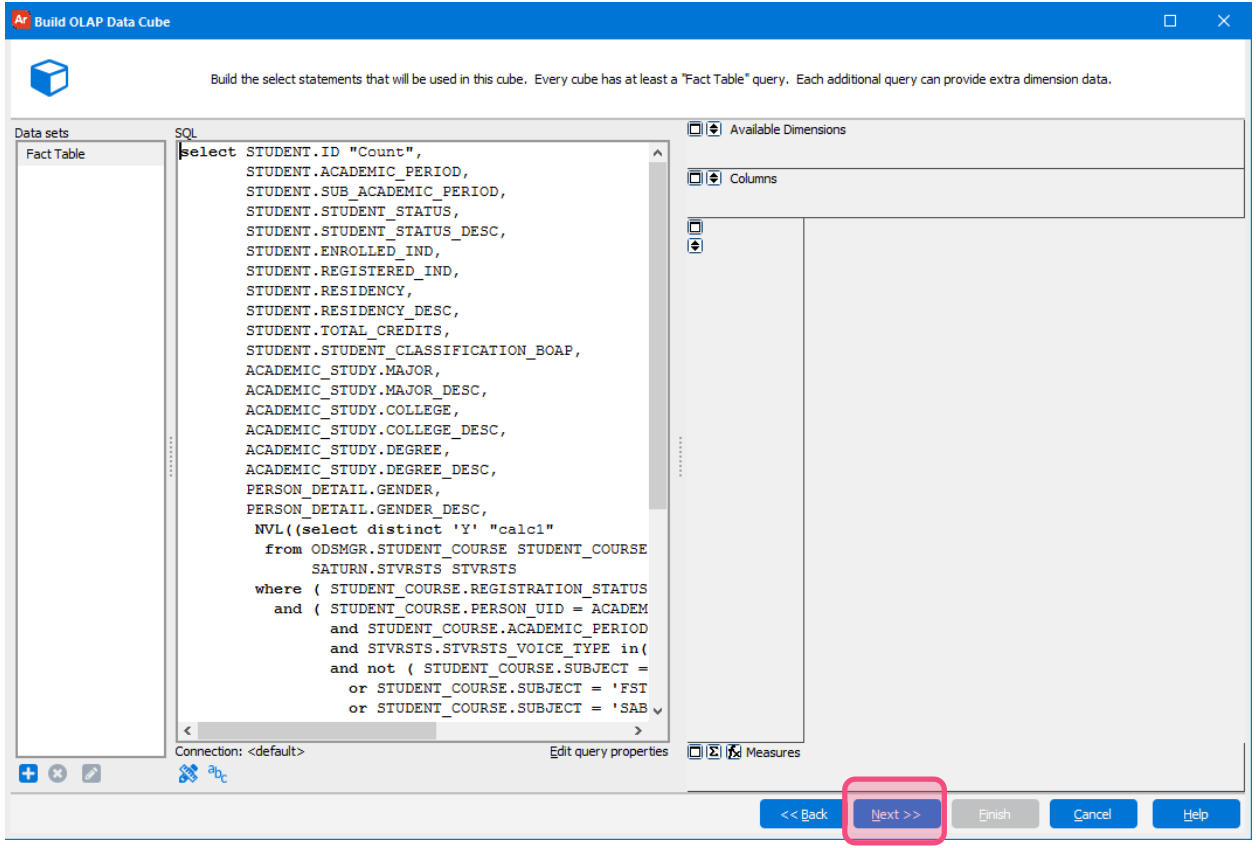
Building an Argos OLAP example

evisions conference
#Evisions19



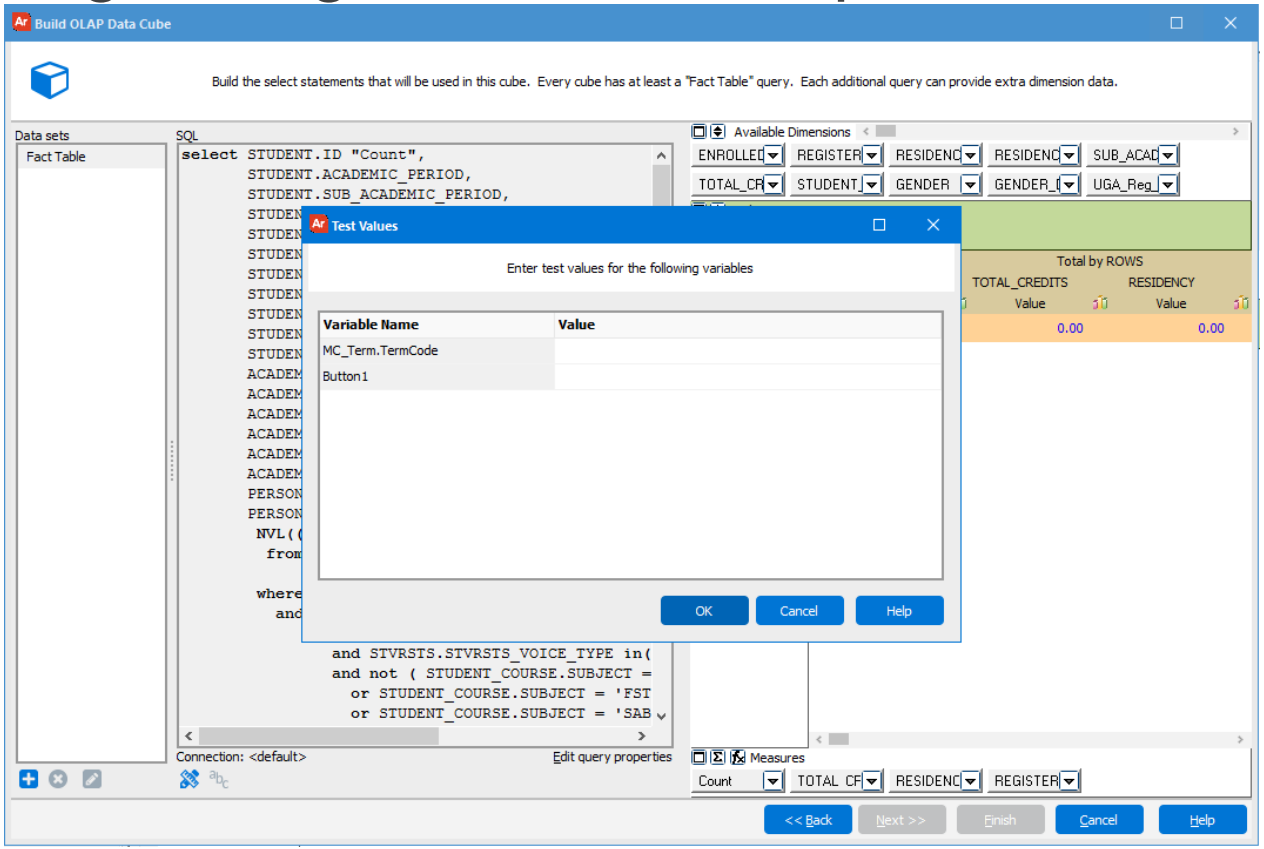


Building an Argos OLAP example



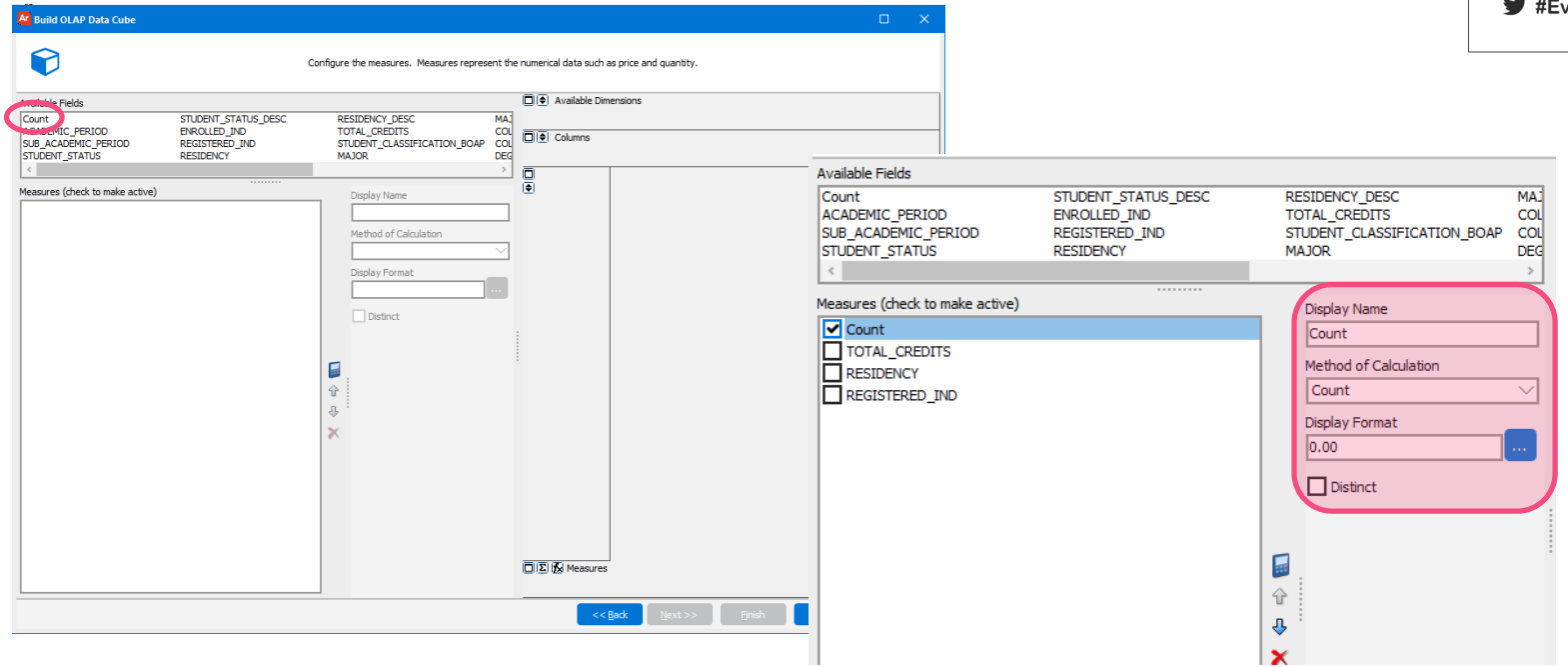
Building an Argos OLAP example

evisions conference
#Evisions19



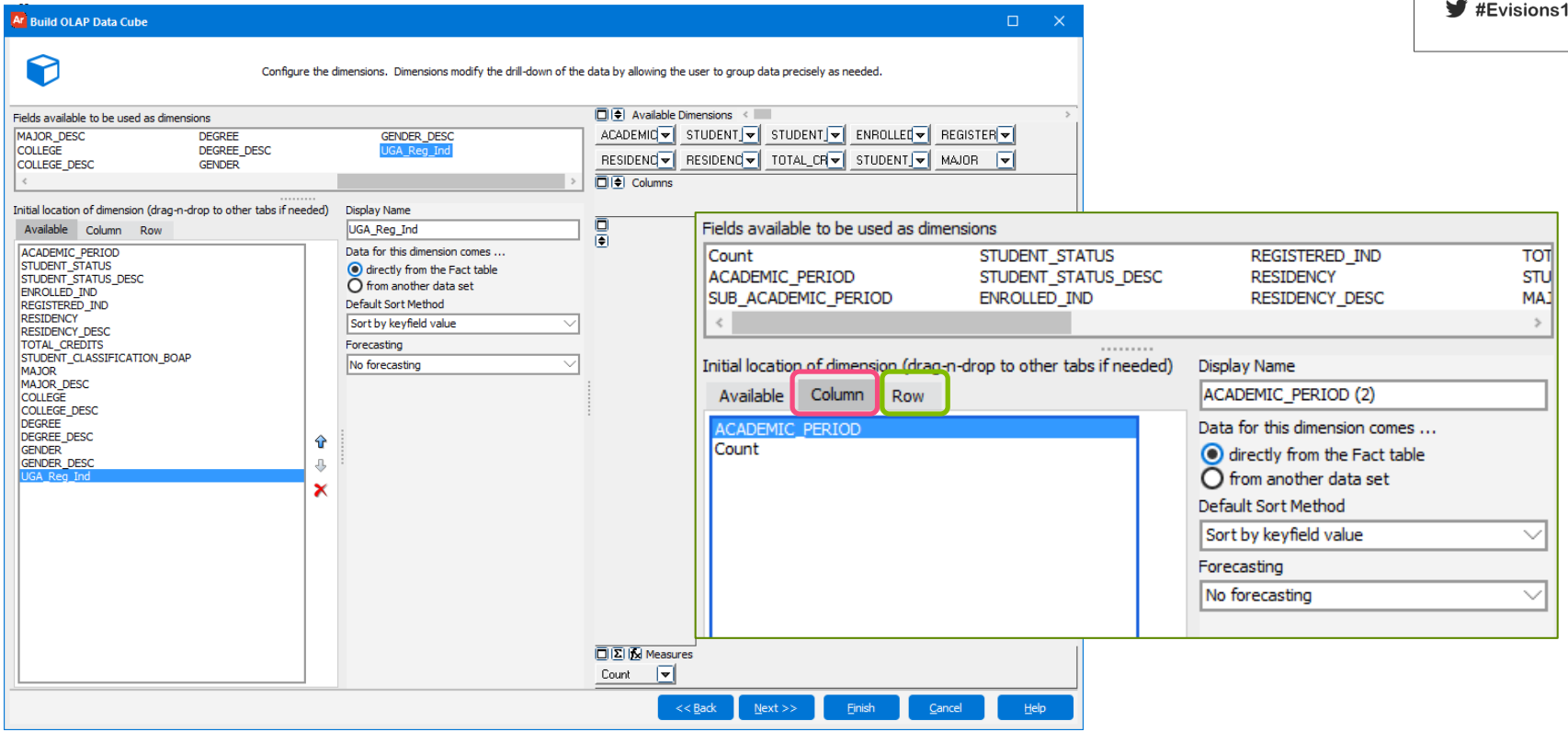
Building an Argos OLAP example

evisions conference
#Evisions19

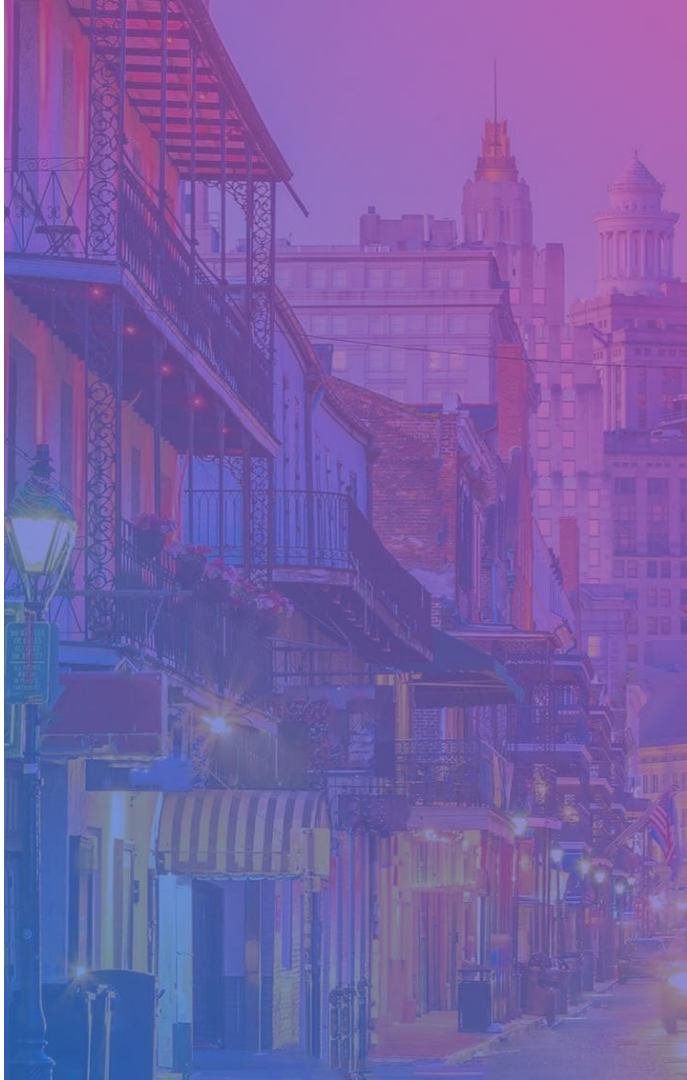


Building an Argos OLAP example

evisions conference
#Evisions19



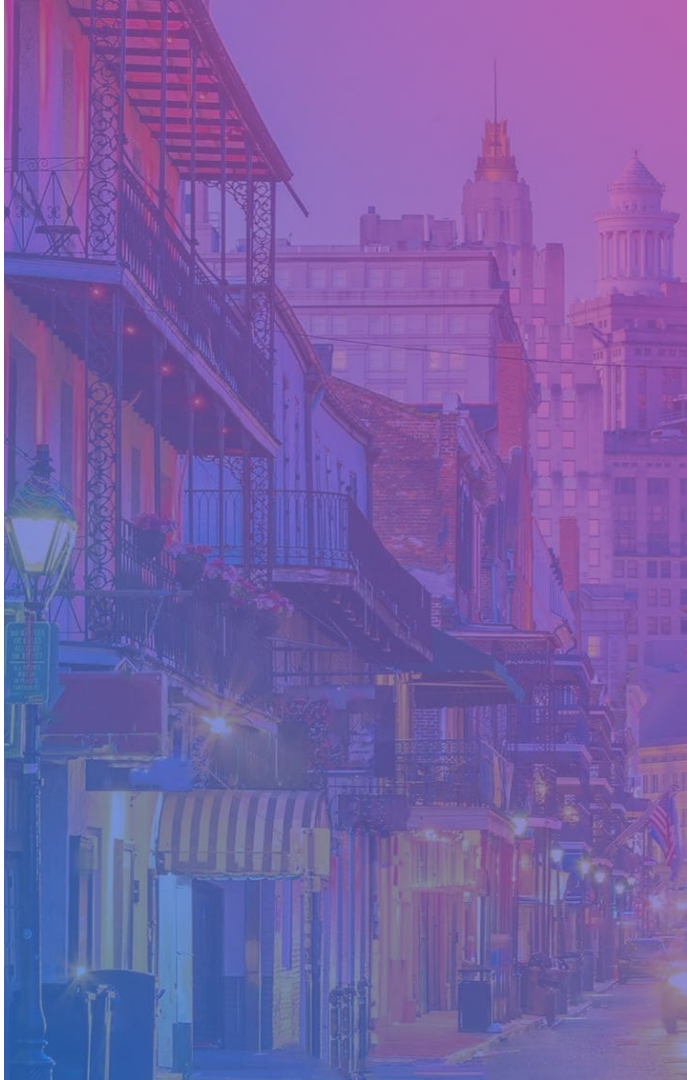
Common Pitfalls



Pitfalls

- There is a learning curve for using OLAPs
  - Understanding vocabulary and proper use (all the techniques we've demoed)
- Counts are off
  - Cartesians, bad queries, 'overlapping' data
- Granularity Problems
  - A special case of Cartesians.

Tips and Tricks



Tip #1: Keep your data set focused

BEFORE

AFTER

evisions conference

#Evisions19

Tip #1: Keep your data set focused

Term:

DisplayTerm	TermCode
201402 - Spring 2014	201402
201405 - Summer 2014	201405
201408 - Fall 2014	201408
201502 - Spring 2015	201502
201505 - Summer 2015	201505
201508 - Fall 2015	201508

Available Dimensions

STUDENT_CLAS | COLLEGE | GENDER | GENDER_DESC | UGA_Reg_Ind | STUDENT_STAT | DEGREE_DESC | ACADEMIC_PER | STUDENT_STAT | RESIDENCY_DE | RESIDENCY_DE | MAJOR_DESC | COLLEGE_DESC

DEGREE | DEGREE_DESC | STUDENT_LEVE | STUDENT_LEVE | CAMPUS | CAMPUS_DESC | ACADEMIC_STA | ACADEMIC_STA | GRAD_YEAR_IN | GRAD_YEAR_IN | EXPECTED_GRA | MAJOR_DESC (| DEGREE (2)

Columns: ACADEMIC_PER | COLLEGE_DESC

STUDENT_STAT | RESIDENCY_DE

ACADEMIC_PERIOD (2) | COLLEGE_DESC (2) | STUDENT_STATUS_DESC (2) | RESIDENCY_DESC (3)

		201505							
		Biomed and Health Sci Inst	College of Agr and Env Science	College of Arts and Sciences	College of Business	College of Education	College of Engineering	College of Env and Design	
		ID	ID	ID	ID	ID	ID	ID	
		Value	Value	Value	Value	Value	Value	Value	
Active		31881.00	24.00	1783.00	9747.00	6648.00	3864.00	1275.00	8.00
	Georgia Resident	26332.00	11.00	1457.00	7985.00	5659.00	3198.00	1150.00	189.00
	International Student	1536.00	1.00	140.00	543.00	222.00	185.00	55.00	40.00
	Non Georgia Resident	4013.00	12.00	186.00	1219.00	767.00	481.00	70.00	49.00
Graduated-Inactive		4866.00	0.00	204.00	1363.00	1080.00	609.00	55.00	46.00
	Georgia Resident	4065.00	0.00	169.00	1156.00	882.00	513.00	49.00	23.00
	International Student	182.00	0.00	12.00	53.00	34.00	30.00	3.00	5.00
	Non Georgia Resident	619.00	0.00	23.00	154.00	164.00	66.00	3.00	18.00
Inactive		1394.00	0.00	77.00	524.00	176.00	168.00	47.00	9.00
	Georgia Resident	1084.00	0.00	65.00	456.00	145.00	135.00	44.00	7.00
	International Student	115.00	0.00	4.00	16.00	8.00	3.00	2.00	0.00
	Non Georgia Resident	195.00	0.00	8.00	52.00	23.00	30.00	1.00	2.00
No Show		133.00	0.00	3.00	40.00	19.00	38.00	1.00	1.00
	Georgia Resident	103.00	0.00	1.00	36.00	13.00	27.00	1.00	0.00
Total by COLUMNS		38296.00	24.00	2069.00	11685.00	7927.00	4679.00	1379.00	334.00

Measures: ID

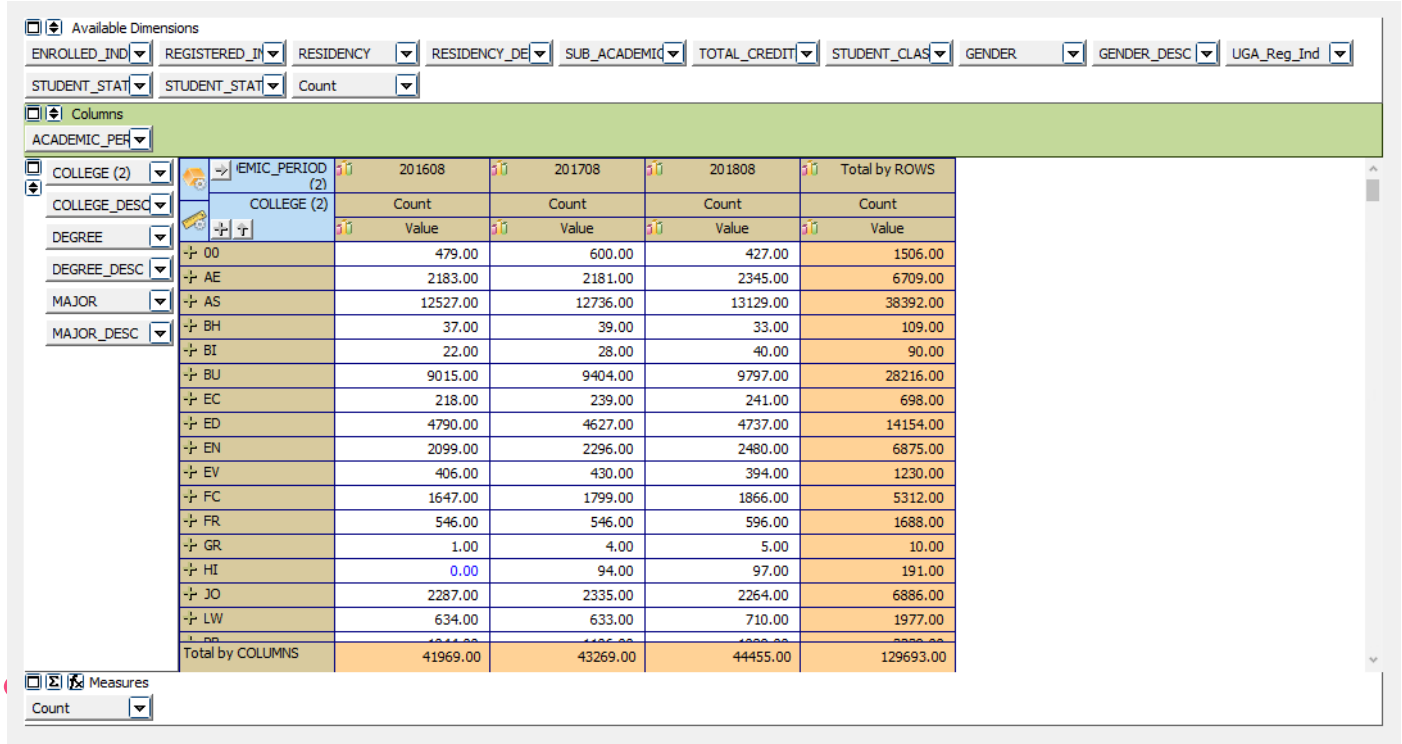
Tip #2: ‘Pre-process’ the data in your query.

Tip #3: Add/create your own Dimensions

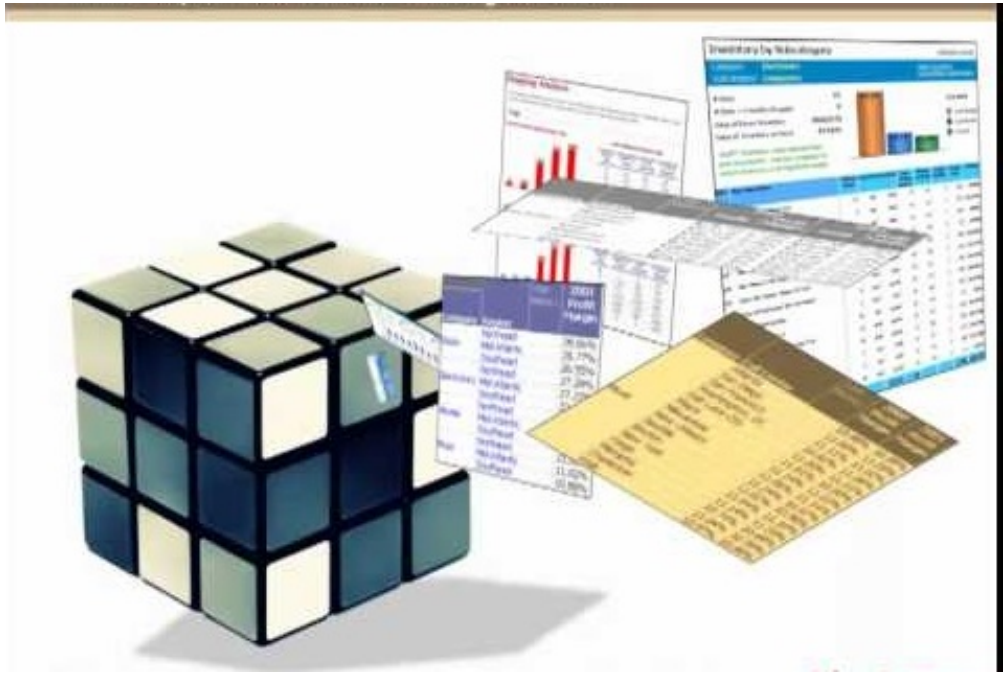
- Creating a custom dimension from an existing field.

- For example: Converting DATE field to year/Month.
  - TO_CHAR (SYSDATE, ‘mm-Month’) -> 01-January
  - TO_CHAR (ACTIVITY_DATE, ‘YYYY’) -> 2018

Tip #4: Hide unnecessary measures

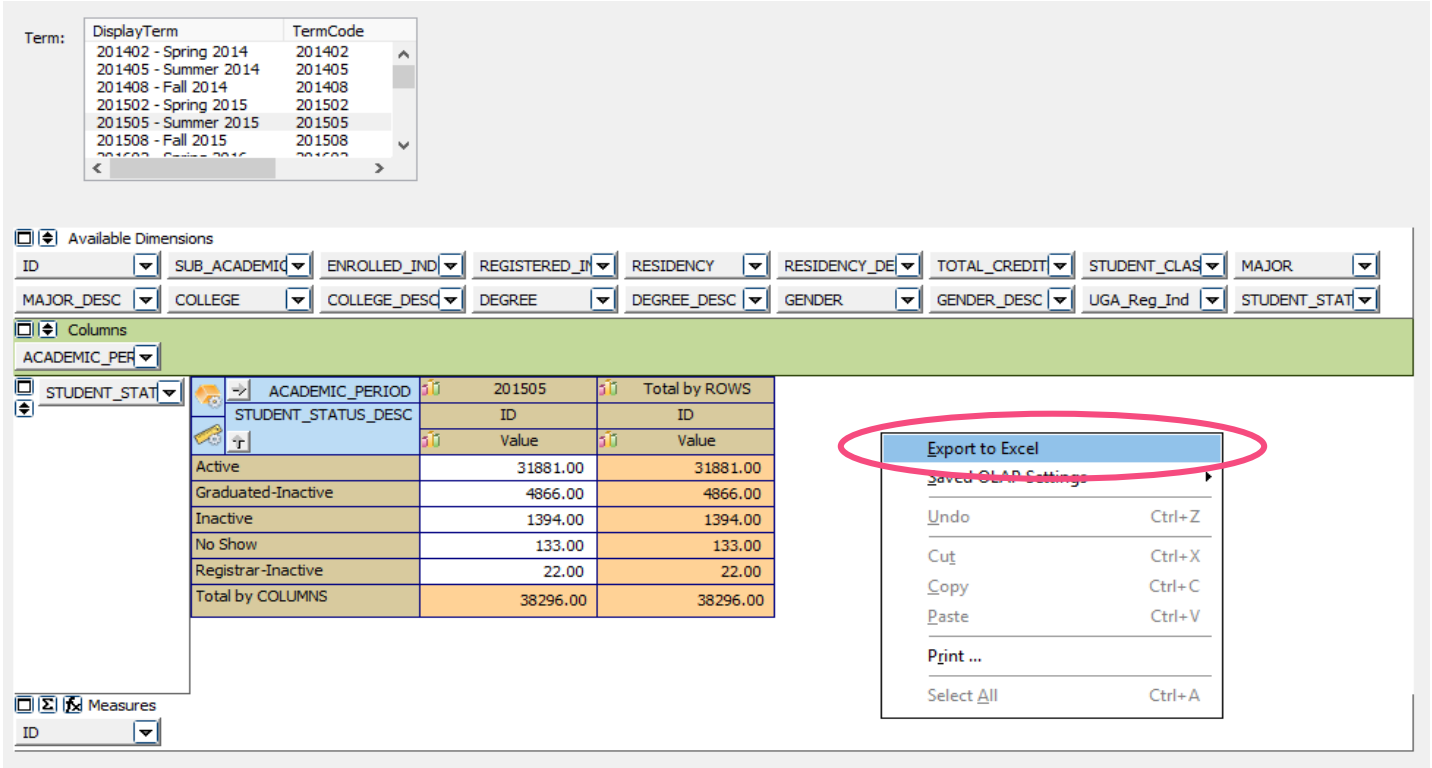


COLLEGE (2)	201608 Count Value	201708 Count Value	201808 Count Value	Total by ROWS Count Value
00	479.00	600.00	427.00	1506.00
AE	2183.00	2181.00	2345.00	6709.00
AS	12527.00	12736.00	13129.00	38392.00
BH	37.00	39.00	33.00	109.00
BI	22.00	28.00	40.00	90.00
BU	9015.00	9404.00	9797.00	28216.00
EC	218.00	239.00	241.00	698.00
ED	4790.00	4627.00	4737.00	14154.00
EN	2099.00	2296.00	2480.00	6875.00
EV	406.00	430.00	394.00	1230.00
FC	1647.00	1799.00	1866.00	5312.00
FR	546.00	546.00	596.00	1688.00
GR	1.00	4.00	5.00	10.00
HI	0.00	94.00	97.00	191.00
JO	2287.00	2335.00	2264.00	6886.00
LW	634.00	633.00	710.00	1977.00
Total by COLUMNS	41969.00	43269.00	44455.00	129693.00

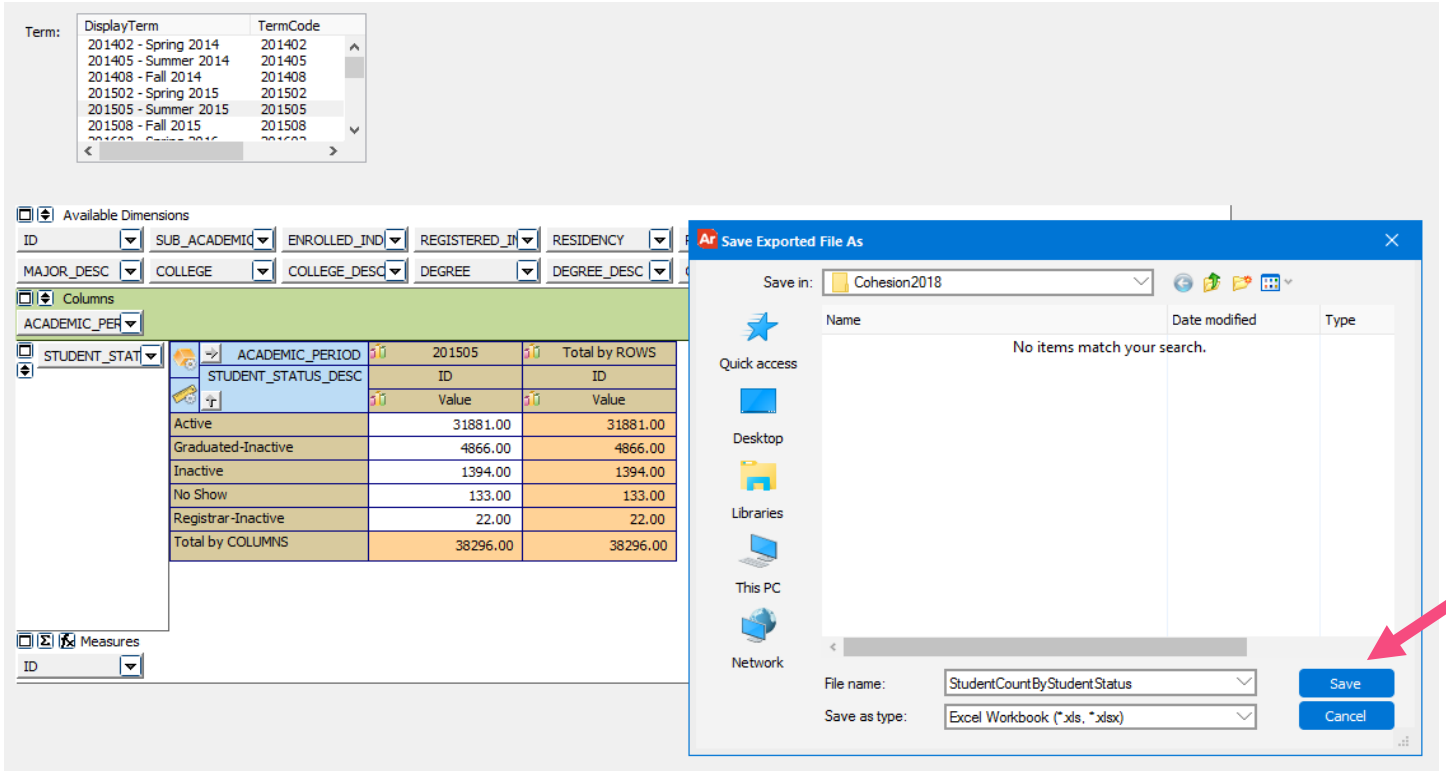


Tip #5: Extract the data

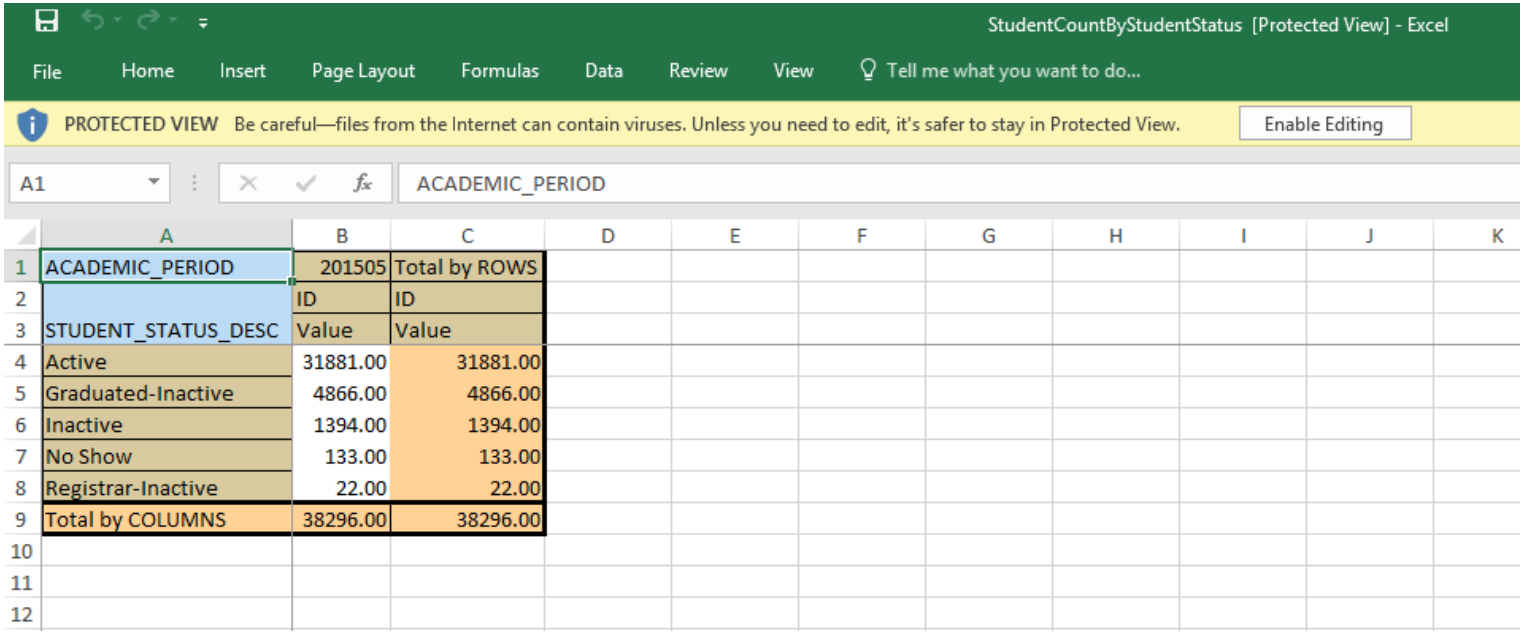
Extract to Excel – Step 1



Extract to Excel - Step 2



Extract to Excel – Step 3

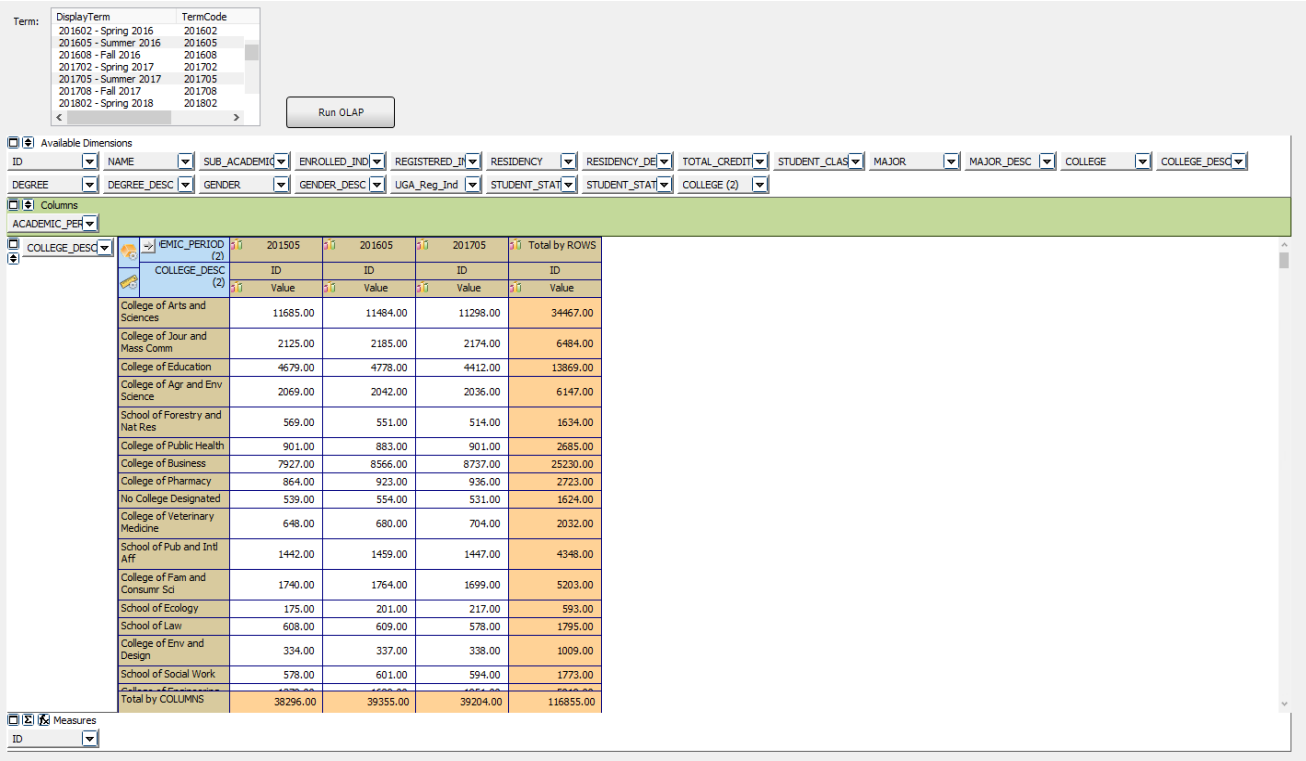


	A	B	C	D	E	F	G	H	I	J	K
1	ACADEMIC_PERIOD	201505	Total by ROWS								
2		ID	ID								
3	STUDENT_STATUS_DESC	Value	Value								
4	Active	31881.00	31881.00								
5	Graduated-Inactive	4866.00	4866.00								
6	Inactive	1394.00	1394.00								
7	No Show	133.00	133.00								
8	Registrar-Inactive	22.00	22.00								
9	Total by COLUMNS	38296.00	38296.00								
10											
11											
12											

Tip #6: Using Percentages

Tip #6: Using Percentages

evisions conference
#Evisions19



Term:

DisplayTerm	TermCode
201602 - Spring 2016	201602
201605 - Summer 2016	201605
201608 - Fall 2016	201608
201702 - Spring 2017	201702
201705 - Summer 2017	201705
201708 - Fall 2017	201708
201802 - Spring 2018	201802

Run OLAP

Available Dimensions

ID | NAME | SUB_ACADEMIC | ENROLLED_IND | REGISTERED_IN | RESIDENCY | RESIDENCY_DE | TOTAL_CREDIT | STUDENT_CLAS | MAJOR | MAJOR_DESC | COLLEGE | COLLEGE_DESC

DEGREE | DEGREE_DESC | GENDER | GENDER_DESC | UGA_Reg_Ind | STUDENT_STAT | STUDENT_STAT | COLLEGE (2)

Columns

ACADEMIC_PER

COLLEGE_DESC

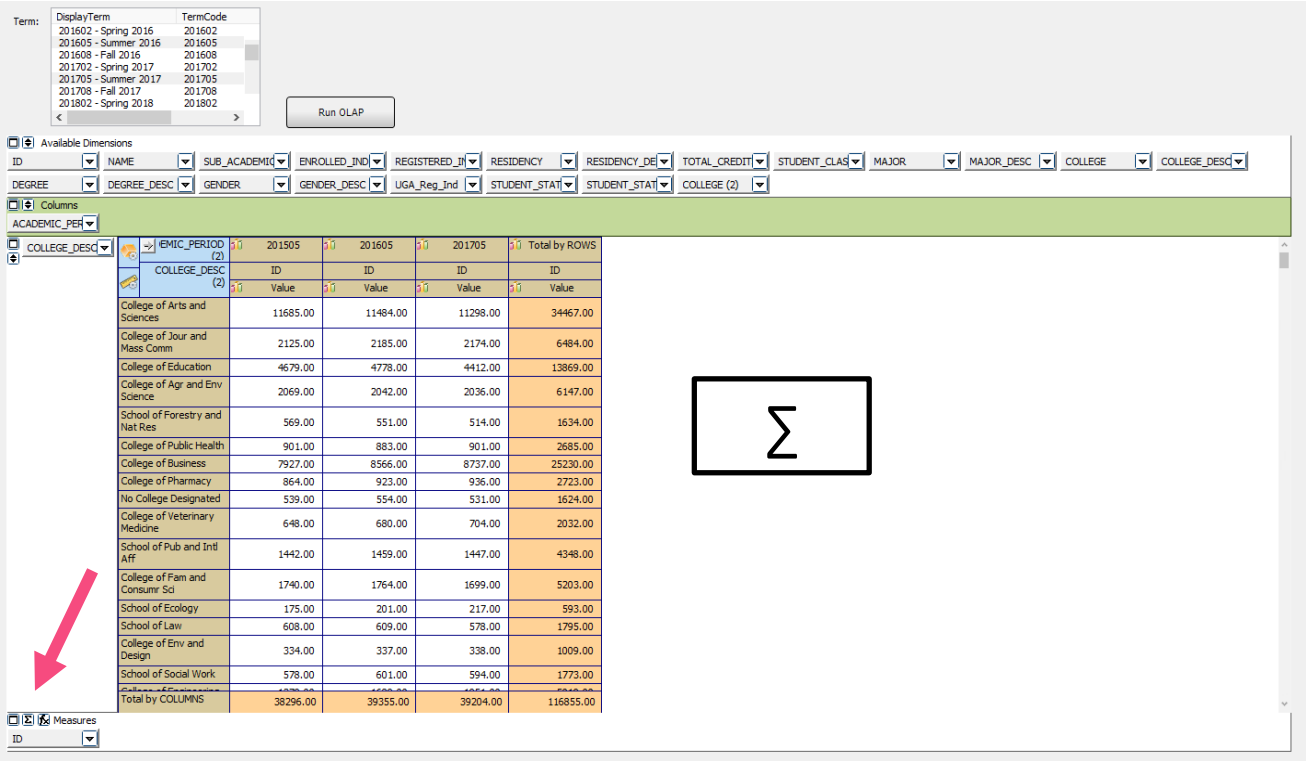
EMIC_PERIOD (2)	201505	201605	201705	Total by ROWS
COLLEGE_DESC (2)	ID	ID	ID	ID
	Value	Value	Value	Value
College of Arts and Sciences	11685.00	11484.00	11298.00	34467.00
College of Jour and Mass Comm	2125.00	2185.00	2174.00	6484.00
College of Education	4679.00	4778.00	4412.00	13869.00
College of Agr and Env Science	2069.00	2042.00	2036.00	6147.00
School of Forestry and Nat Res	569.00	551.00	514.00	1634.00
College of Public Health	901.00	883.00	901.00	2685.00
College of Business	7927.00	8566.00	8737.00	25230.00
College of Pharmacy	864.00	923.00	936.00	2723.00
No College Designated	539.00	554.00	531.00	1624.00
College of Veterinary Medicine	648.00	680.00	704.00	2032.00
School of Pub and Intl Aff	1442.00	1459.00	1447.00	4348.00
College of Fam and Consumr Sci	1740.00	1764.00	1699.00	5203.00
School of Ecology	175.00	201.00	217.00	593.00
School of Law	608.00	609.00	578.00	1795.00
College of Env and Design	334.00	337.00	338.00	1009.00
School of Social Work	578.00	601.00	594.00	1773.00
Total by COLUMNS	38296.00	39355.00	39204.00	116855.00

Measures

ID

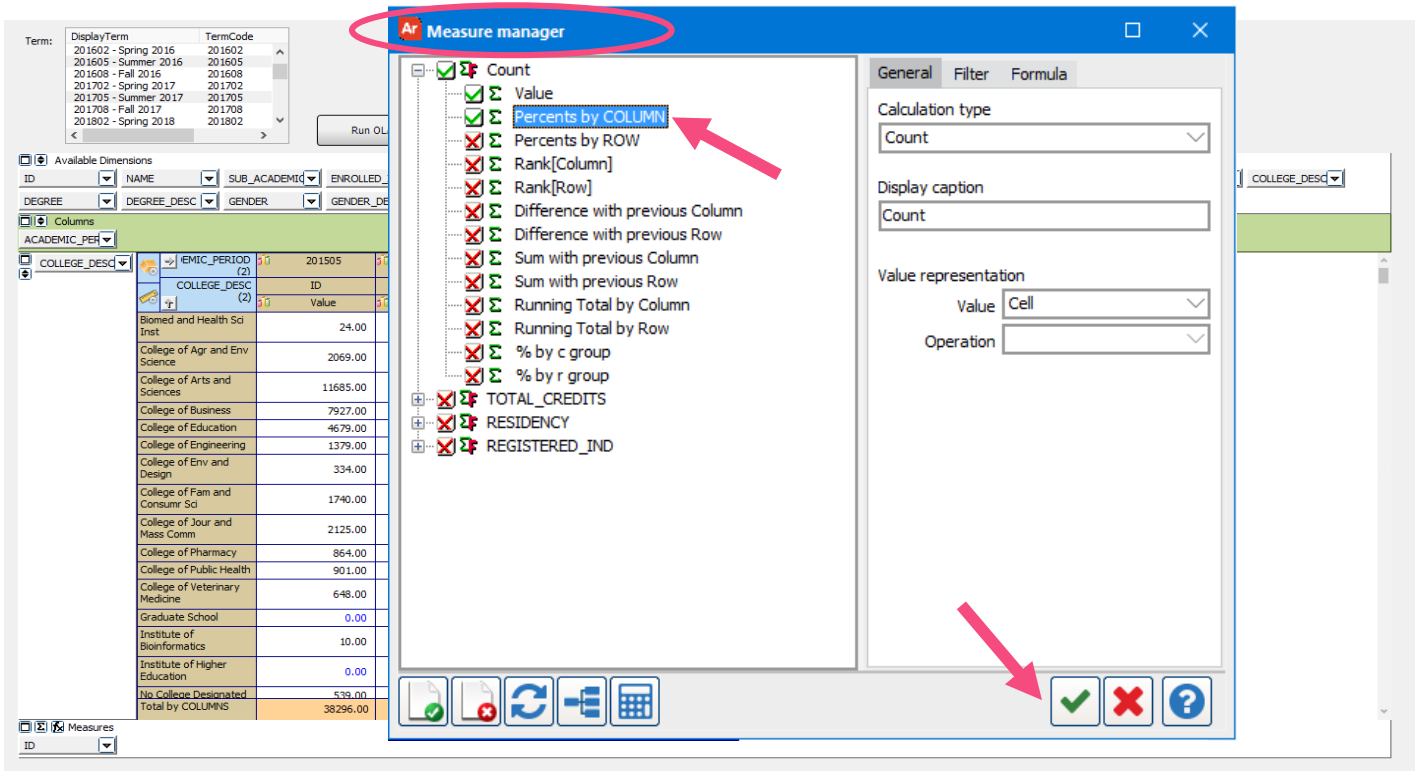
Tip #6: Using Percentages

evisions conference
#Evisions19



Tip #6: Using Percentages

evisions conference
#Evisions19



Tip #6: Using Percentages

evisions conference

#Evisions19

Testing "New DataBlock OLAP Shows Percentages"

Term:

DisplayTerm	TermCode
201602 - Spring 2016	201602
201605 - Summer 2016	201605
201608 - Fall 2016	201608
201702 - Spring 2017	201702
201705 - Summer 2017	201705
201708 - Fall 2017	201708
201802 - Spring 2018	201802

Run OLAP

Available Dimensions

ID | NAME | SUB_ACADEMIC | ENROLLED_IND | REGISTERED_IN | RESIDENCY | RESIDENCY_DE | TOTAL_CREDIT | STUDENT_CLAS | MAJOR | MAJOR_DESC | COLLEGE | COLLEGE_DESC

DEGREE | DEGREE_DESC | GENDER | GENDER_DESC | UGA_Reg_Ind | STUDENT_STAT

Columns: ACADEMIC_PER

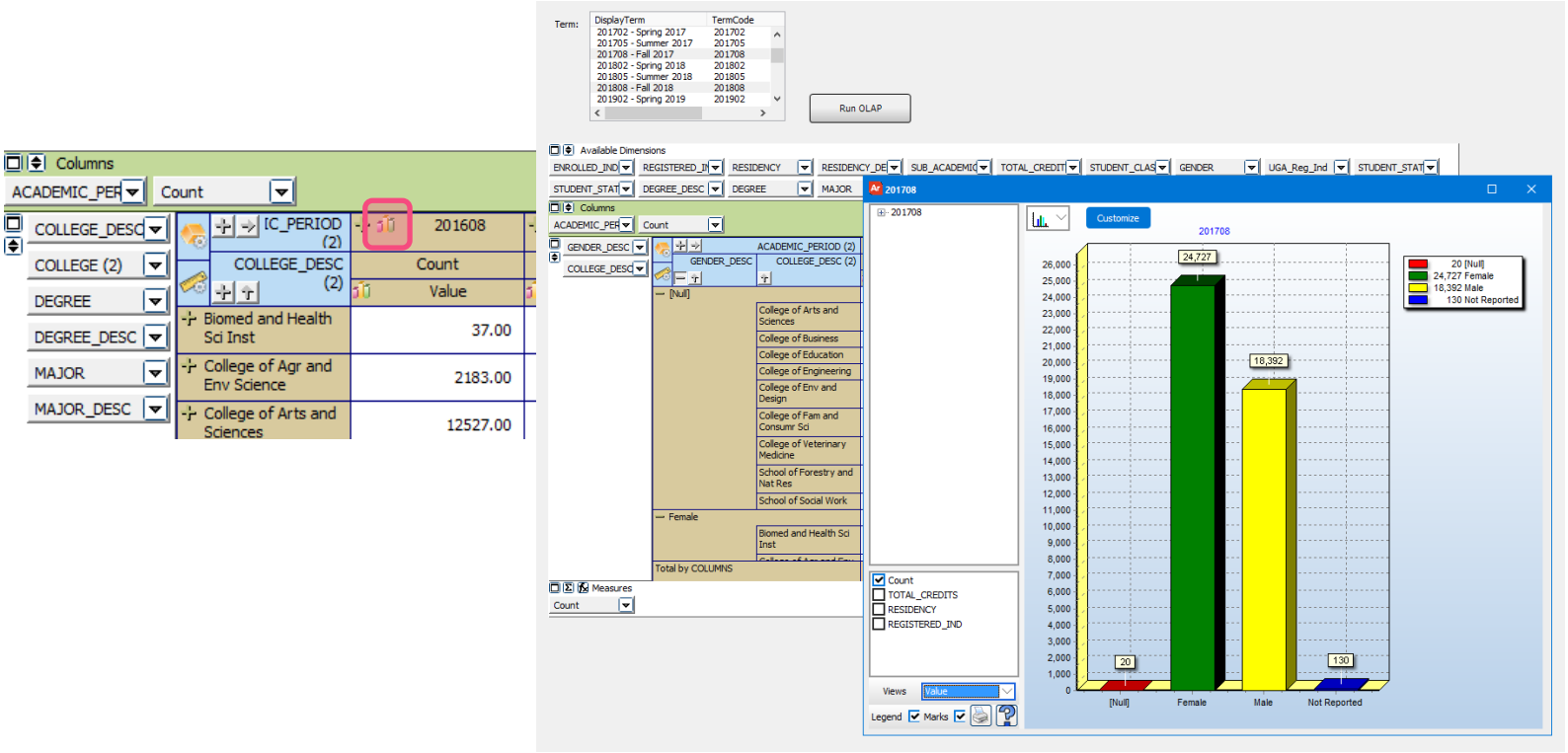
Rows: COLLEGE (2) / COLLEGE_DESC / STUDENT_STAT

ACADEMIC_PERIOD (2)			201505		201605		201705		Total by ROWS	
COLLEGE (2)	COLLEGE_DESC (2)	STUDENT_STATUS_...	ID		ID		ID		ID	
			Value	Percents by COLUMN	Value	Percents by COLUMN	Value	Percents by COLUMN	Value	Percents by COLUMN
00			539.00	1.41%	554.00	1.41%	531.00	1.35%	1624.00	1.39%
AE			2069.00	5.40%	2042.00	5.19%	2036.00	5.19%	6147.00	5.26%
AS	College of Arts and Sciences		11685.00	30.51%	11484.00	29.18%	11298.00	28.82%	34467.00	29.50%
		Active	9747.00	25.45%	9629.00	24.47%	9481.00	24.18%	28857.00	24.69%
		Graduated-Inactive	1363.00	3.56%	1306.00	3.32%	1304.00	3.33%	3973.00	3.40%
		Inactive	524.00	1.37%	508.00	1.29%	471.00	1.20%	1503.00	1.29%
		No Show	40.00	0.10%	41.00	0.10%	40.00	0.10%	121.00	0.10%
		Registrar-Inactive	11.00	0.03%	0.00	0.00%	2.00	0.01%	13.00	0.01%
BH			24.00	0.06%	34.00	0.09%	37.00	0.09%	95.00	0.08%
BI			10.00	0.03%	15.00	0.04%	22.00	0.06%	47.00	0.04%
BU			7927.00	20.70%	8566.00	21.77%	8737.00	22.29%	25230.00	21.59%
EC			175.00	0.46%	201.00	0.51%	217.00	0.55%	593.00	0.51%
ED			4679.00	12.22%	4778.00	12.14%	4412.00	11.25%	13869.00	11.87%
EN	College of Engineering		1379.00	3.60%	1689.00	4.29%	1951.00	4.98%	5019.00	4.30%
		Active	1275.00	3.33%	1563.00	3.97%	1715.00	4.37%	4553.00	3.90%
		Graduated-Inactive	55.00	0.14%	62.00	0.16%	165.00	0.42%	282.00	0.24%
		Inactive	47.00	0.12%	60.00	0.15%	64.00	0.16%	171.00	0.15%
		No Show	1.00	0.00%	4.00	0.01%	7.00	0.02%	12.00	0.01%
		Registrar-Inactive	1.00	0.00%	0.00	0.00%	0.00	0.00%	1.00	0.00%
EV			334.00	0.87%	337.00	0.86%	338.00	0.86%	1009.00	0.86%
FC			1740.00	4.54%	1764.00	4.48%	1699.00	4.33%	5203.00	4.45%
FR			569.00	1.49%	551.00	1.40%	514.00	1.31%	1634.00	1.40%
Total by COLUMNS			38296.00	100.00%	39355.00	100.00%	39204.00	100.00%	116855.00	100.00%

Measures: ID

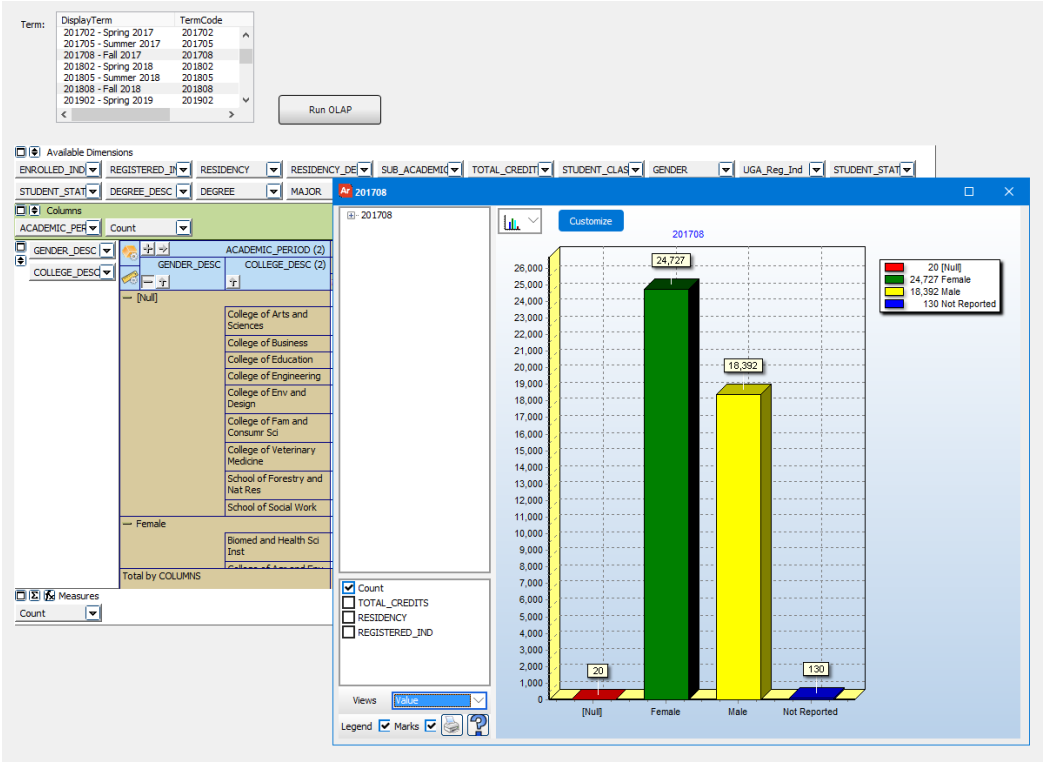
Tip #7: Charting from the OLAP

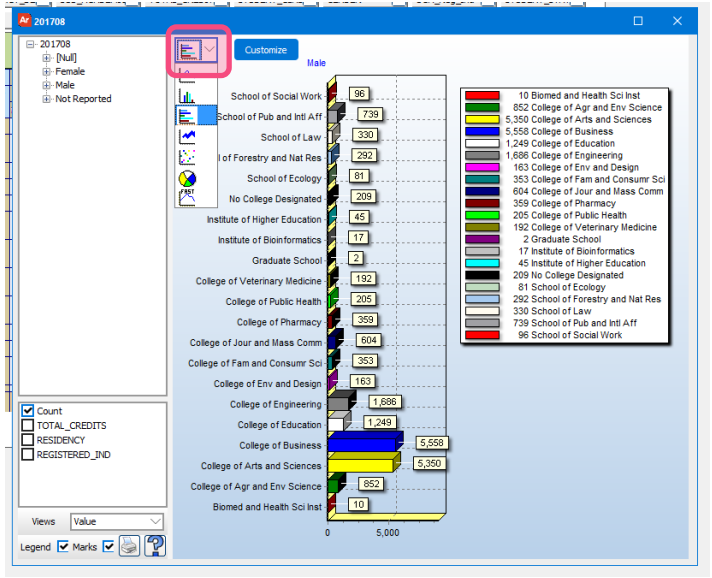
evisions conference
#Evisions19



Tip #7: Charting from the OLAP

evisions conference
#Evisions19

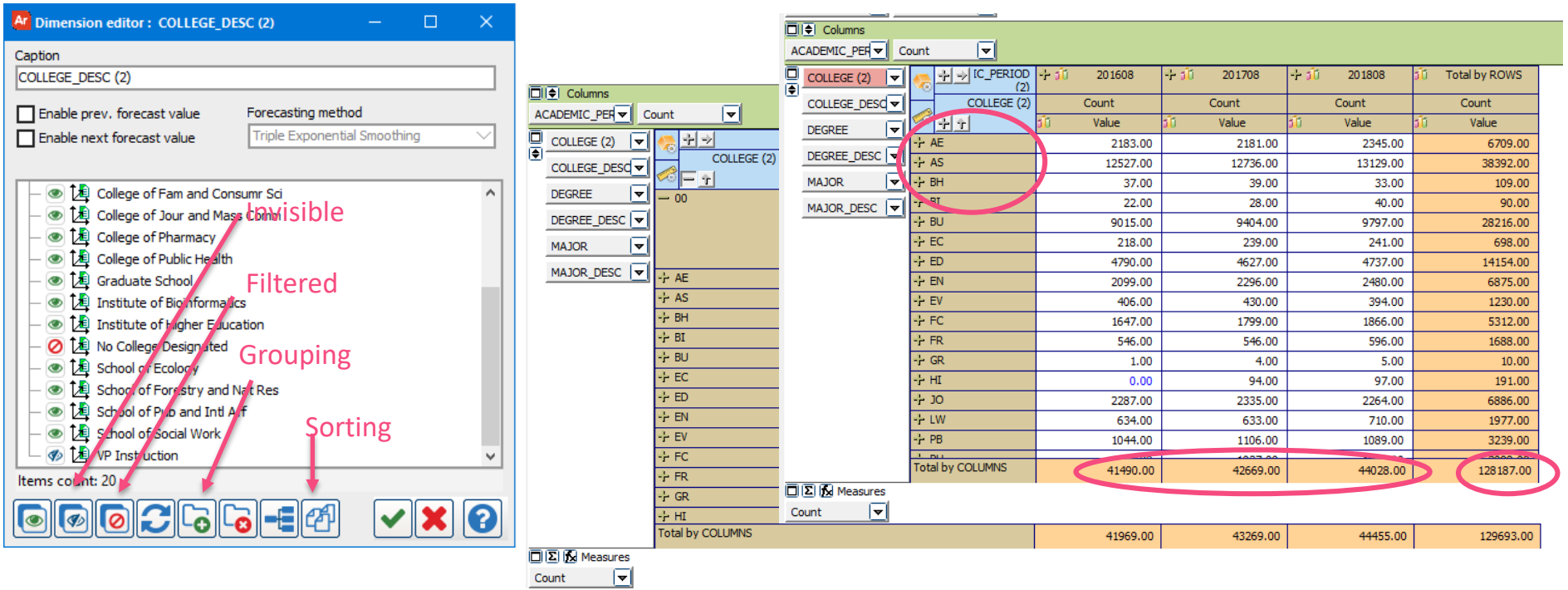




evisions conference

#Evisions19

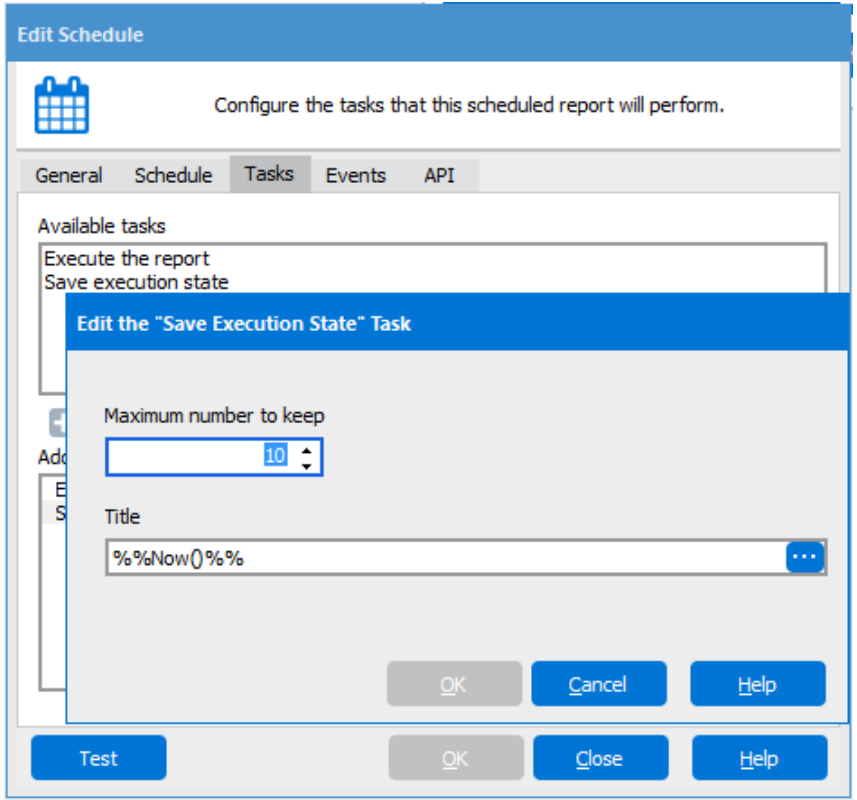
Tip #8: End User Filtering

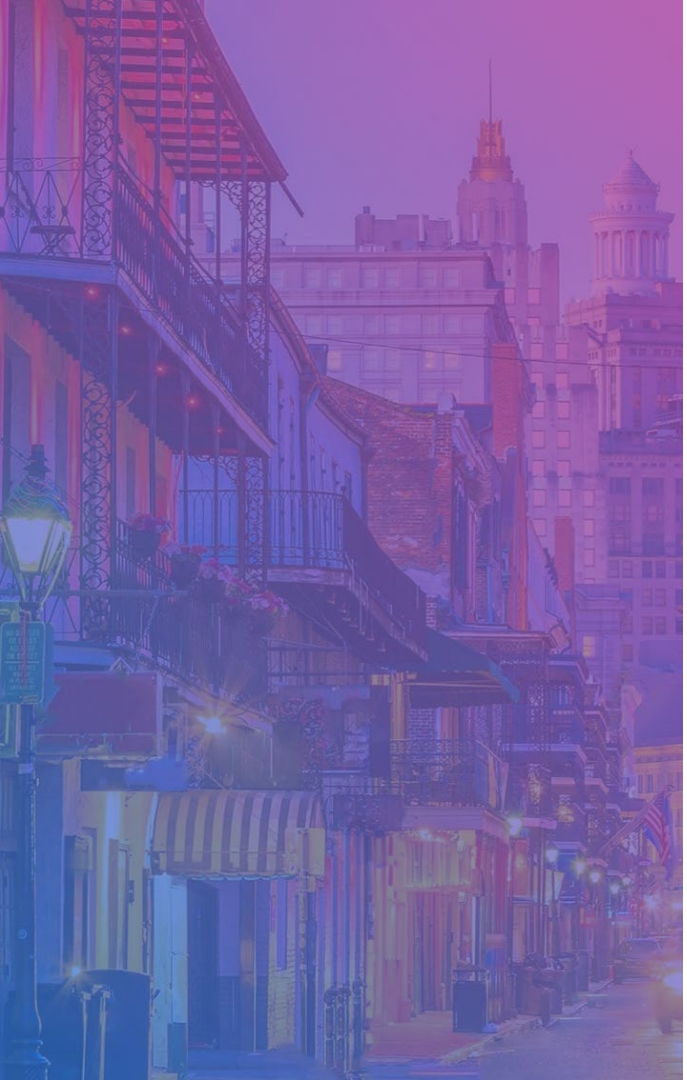


Tip #9: Saved States

A version of the OLAP can be saved for future use. These saved states can be created in the schedule

Warning!!





SUMMARY

evisions conference

#Evisions19

1 OLAP Overview

2 Argos Demonstration

3 Creating an OLAP

4 Common Pitfalls

5 Tips and Tricks

6 Summary

7 Q & A

Questions & Answers

THANK YOU!

Andrea Suarez Halow
suarez@uga.edu

John Tong
jtong@uga.edu

evisions conference

Presentation notes - order

1 Intro - both

7 Agenda - John

8 OLAP Overview - John

12 Argos Demonstration - Andrea

- 18 terms - Andrea
- 19 Dimensions - John
- 21 Measures - John

3Creating an OLAP - John

- 24 design - John
- 28 build - Andrea

38 Common Pitfalls- John

40 Tips and Tricks- John

- 1- 5 John
- 6-9 Andrea

59 Summary - Andrea

60 Q & A