

The logo features the text "Cal Answers" in a bold, yellow, sans-serif font. The word "Cal" is significantly larger than "Answers". The text is set against a dark blue background that includes a faint, light-colored image of a classical building facade with columns and arches. A small white dot with a horizontal line extending to the right is positioned to the right of the word "Answers".

**Cal Answers**

***Cal Answers  
Analysis Training  
Part I***

***Creating Analyses in OBIEE***

University of California, Berkeley  
March 2012

## Table of Contents

|                                                           |    |
|-----------------------------------------------------------|----|
| Table of Contents .....                                   | 1  |
| Overview .....                                            | 2  |
| Getting Around OBIEE.....                                 | 2  |
| Cal Answers Home Page.....                                | 2  |
| Working with the File Catalog.....                        | 3  |
| Subject Areas .....                                       | 3  |
| Metadata.....                                             | 5  |
| Creating Analyses.....                                    | 5  |
| Columns .....                                             | 6  |
| Changing Order .....                                      | 6  |
| Sorting, Deleting, and Changing Properties .....          | 6  |
| Displaying Results .....                                  | 8  |
| Filters.....                                              | 8  |
| Creating/Editing Filters .....                            | 8  |
| AND/OR Filters .....                                      | 10 |
| Filter Options .....                                      | 11 |
| Using Saved Filters .....                                 | 11 |
| Formulas .....                                            | 12 |
| Simple Math and Additional Columns .....                  | 13 |
| Functions.....                                            | 13 |
| Bins.....                                                 | 14 |
| Advanced Analysis Features.....                           | 15 |
| Set Operations .....                                      | 15 |
| Advanced Tab.....                                         | 15 |
| Practice Analyses .....                                   | 16 |
| Degree Practice Analysis.....                             | 16 |
| Census Count Practice Analysis.....                       | 16 |
| Undergrad Cohort Graduation Rate Practice Analysis.....   | 17 |
| Physics and Statistics SAT Scores Practice Analysis ..... | 17 |

To report any problems with the Cal Answers Portal, Reports, or Dashboards:

- Send an e-mail to [calanswers-help@berkeley.edu](mailto:calanswers-help@berkeley.edu)
- Or
- Call the Help Desk at 642-8500, Monday – Friday, 8:00am – 5:00pm

## Overview

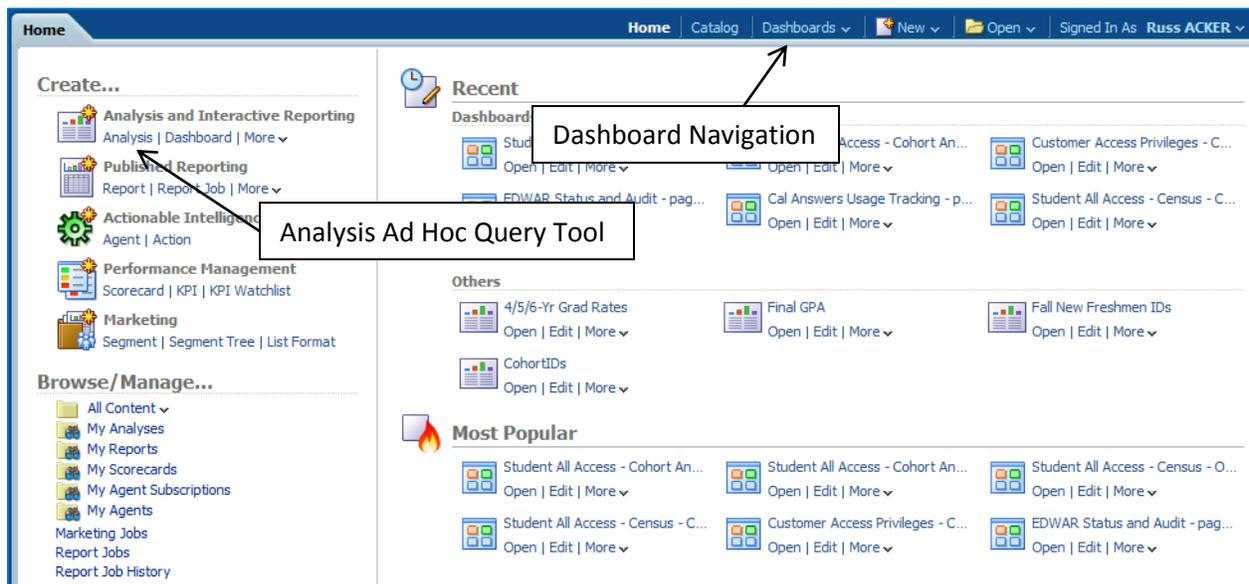
Cal Answers is the new reporting environment for UC Berkeley’s Enterprise Data Warehouse (EDW). It currently uses Oracle Business Intelligence Enterprise Edition (OBIEE) version 11g software.

This three-part training course will deal with using OBIEE, including how to write your own queries and create your own dashboard reports. Specifically, the sessions will cover:

1. Creating Analyses in OBIEE
2. Viewing Analysis Results in OBIEE
3. Advanced OBIEE – Dashboard Reports

## Getting Around OBIEE

When you log into OBIEE, you’ll first see the following home page. You’ll be using the links and dropdowns that appear in various sections of the screen to access dashboard reports and the Analysis query tool:

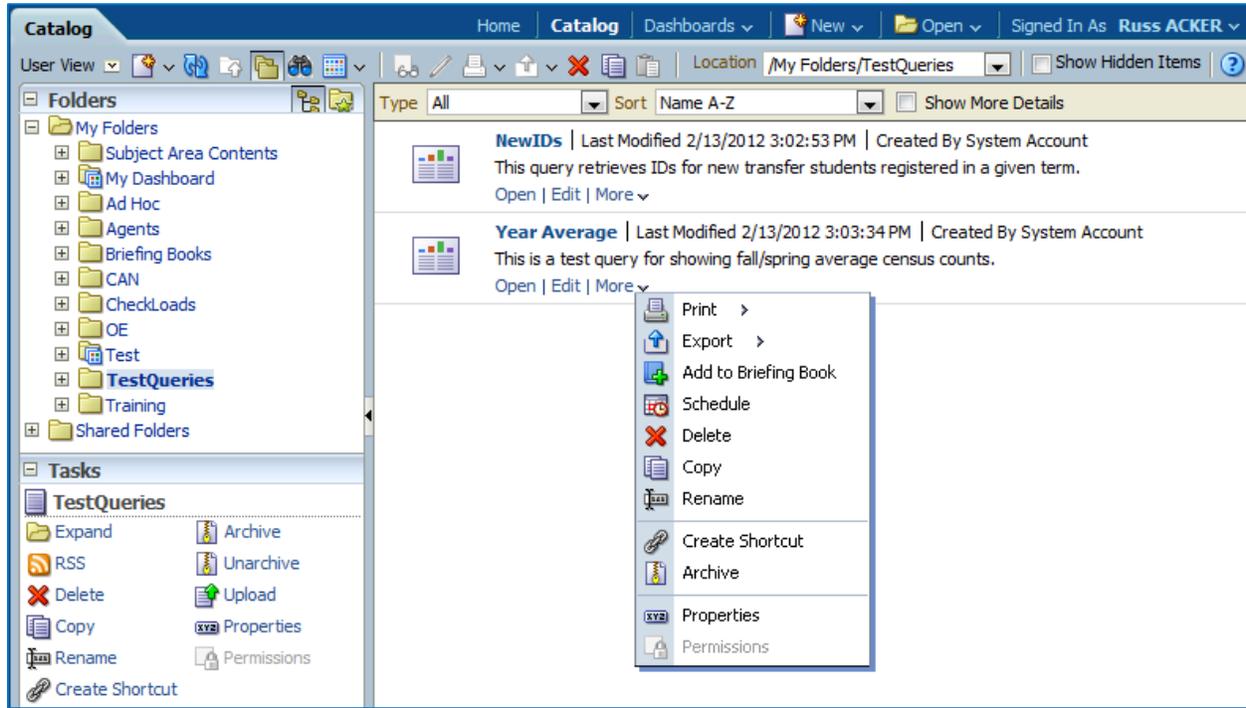


### **Cal Answers Home Page**

The Cal Answers home page shown above gives you access to everything in OBIEE. You can access dashboard reports, create new queries and dashboards, open existing queries and dashboards, and browse your files and folders, among many other things.

**Working with the File Catalog**

The file catalog view allows you to manipulate and open saved objects in either your own “My Folders” area or in a shared section. To access this view, either click the “Catalog” link in the dark blue menu bar, or select one of the links in the “Browse/Manage...” section.



In the Catalog view, for any object listed, you can open, edit, or perform a variety of other actions accessed through the “More” dropdown shown above. Note in particular that one of the properties that you can save with an object is a description (hint, hint).

**Subject Areas**

When you begin creating a new Analysis from the home page (by using either the “New” menu in the dark blue menu bar, or choosing “Analysis” under the “Create...” section), you’ll have to choose a data subject area to use for the query. Currently, there are eight student data subject areas, as shown to the right.

The first four areas, which all begin with “Student Counts...” are the “easy” subject areas, in which a lot of the difficulties involved in filtering and figuring out which fields to use are already done for you. On the other hand, these subject areas each contain a relatively limited set of fields, in the range of a few dozen (they don’t have student IDs or other restricted fields). When you’re just learning to write your own queries, these make an excellent place to start.

The other four, which all begin with “Students...” are the full, proceed-at-your-own-risk subject areas. These each have



hundreds of fields available for you to use, but you have to handle all of your own filtering, which can be a little harder than it might seem. You can, of course, make use of saved filters in these subject areas, which we'll discuss later.

Here are some brief descriptions of the kinds of data in each subject area. Note that each of the "Student Counts..." summary area is just a subset of one of the full "Students..." areas:

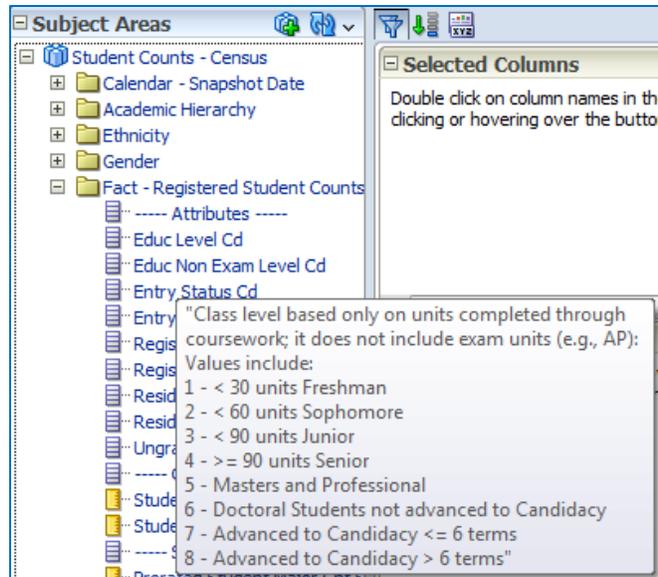
- **Students – Registered:** This subject area contains data on all students who have been accepted and are eligible to register in a given semester. IST will soon begin taking weekly snapshots of this data, but there are currently only two snapshots per semester: the census count and end-of-term (EOT) data. The census snapshot gives us our official 15<sup>th</sup> day count of students, while the EOT snapshot happens after grades are recorded.
  - **Student Counts – Census:** This is a subset of registered students, limited only to those who appear in the official census (reported to UCOP and the state) each semester.
- **Students – Degrees:** This subject area contains data on all students receiving a degree or certificate in a given semester. This data currently gets loaded twice a year: the first load is for summer degrees (and certificates, but let's just assume I mean both from now on) only, and the second includes fall and spring data. Remember that an academic year officially consists of a summer, the following fall, and the following spring.
  - **Student Counts – Degrees:** This is a smaller set of degree fields, limited to those degrees that have not been rescinded.
- **Students – Undergrad Applicants:** This area has data on all undergraduate applicants to UC Berkeley. At this point, unfortunately, it's quite difficult to use. We are currently working on a project to make this subject area more useful and usable. For now, though, I try very hard to use the following subject area whenever possible.
  - **Student Counts – Undergrad Applicants:** This is a smaller set of undergrad applicant fields. There aren't, however, any additional filters specific to applicants in place. This subject area already applies the filters that you need for gender and ethnicity data.
- **Students – Undergraduate Cohorts:** This subject area contains a set of data, derived from the registered student and degree datasets, dealing with retention and graduation for entering undergraduate cohorts. So you can use this, for instance, to tell how many of the new freshmen who entered in fall 2005 returned in the following fall, or how long each took to graduate.
  - **Student Counts – Undergrad Cohorts:** This is a smaller set of undergrad cohort fields. There aren't any additional filters in place, except for those specific to gender and ethnicity data.

**Metadata**

One of the most difficult things about writing your own queries is knowing which fields to use. To some extent, this just requires practice, but there are a couple of things that can help.

First, as shown to the right, there are popup descriptions defined for some fields in the query tool. IST is working on expanding this to every field, but that might take a while, due to the number of fields in the database.

There is also an “Enterprise Data Warehouse Metadata” dashboard defined, although it’s only partially usable at this point. IST is also working on fixing some of the problems with that, but that could also take a while.



## Creating Analyses

Creating a new Analysis and selecting a subject area on the home page will start the OBIEE ad hoc query tool:

Build Query

View/Format Results

Define Prompts

Do Fancy Stuff

Column/Filter Selection Pane

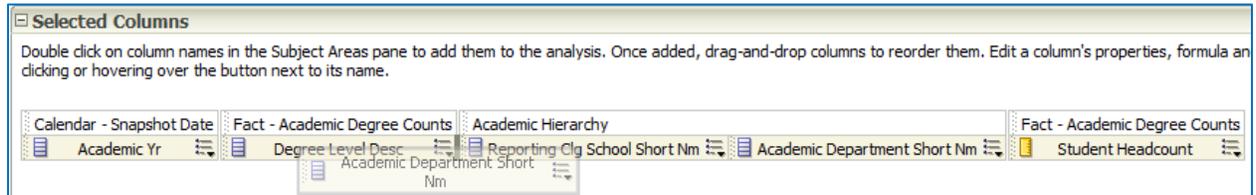
Query Columns and Filters

## Columns

The various fields available in a given subject area are stored in an expandable/collapsible tree. Click the plus signs (  ) to expand a branch, and the minus signs (  ) to collapse it. To bring a column into your query, either double-click it or drag it over into the “Selected Columns” area.

### Changing Order

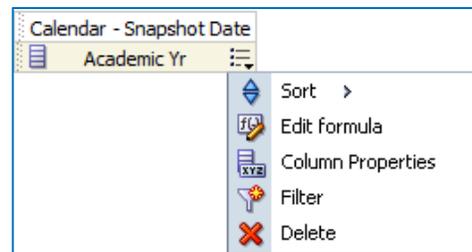
You can change the order of the fields in your query by simply dragging a field around. When you position the field in front of, behind, or in between other fields, you’ll see a vertical gray bar appear. At that point, you can drop the field in its new position.



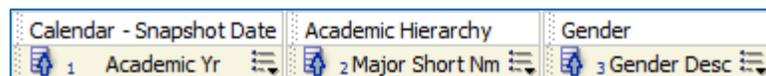
### Sorting, Deleting, and Changing Properties

Each field in your query has a dropdown menu on its right, with five options. We’ll cover “Sort,” “Column Properties,” and “Delete” now, and the other two choices in a later section.

First, you can apply or clear sorts by clicking the “Sort” option and then choosing the type of sort that you want. You can sort by multiple fields by simply turning on sorting for each of them.



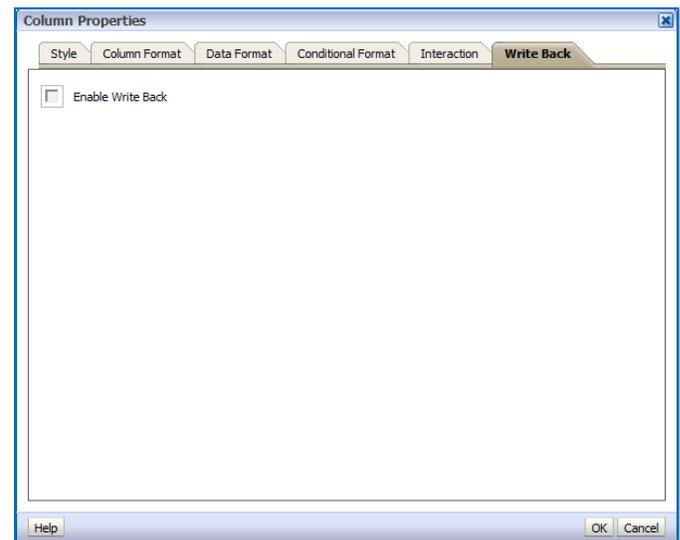
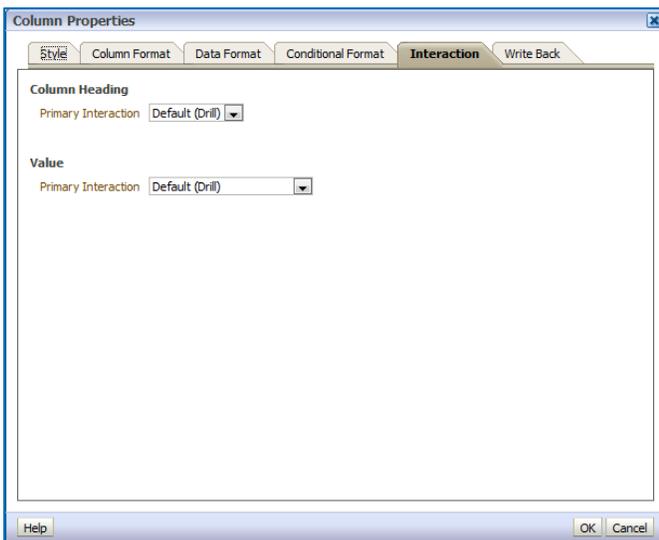
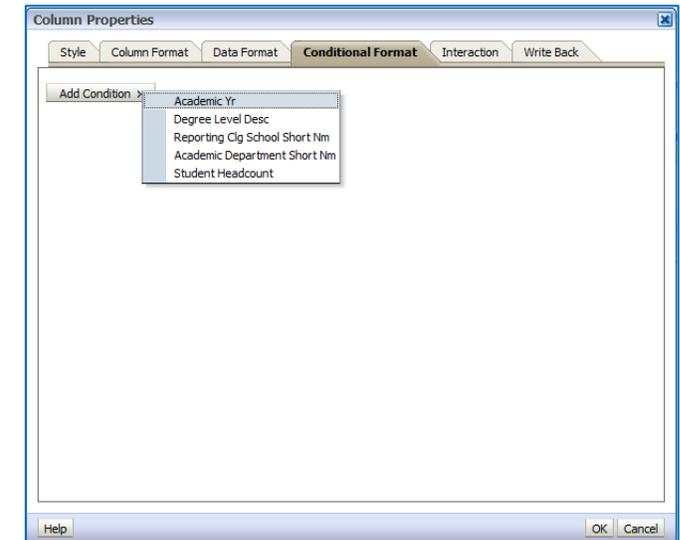
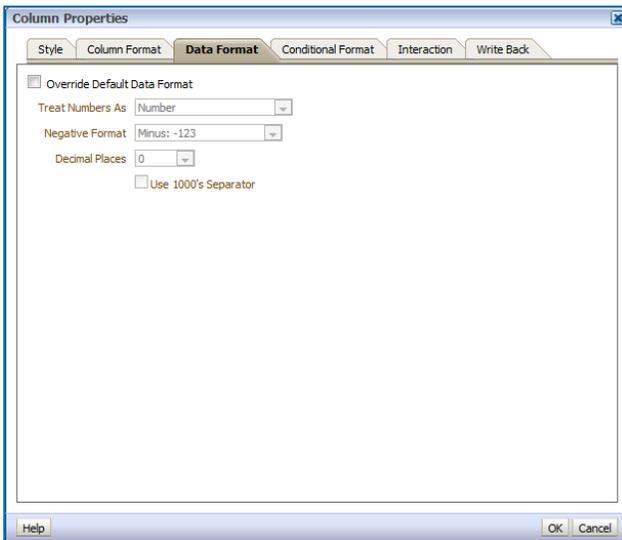
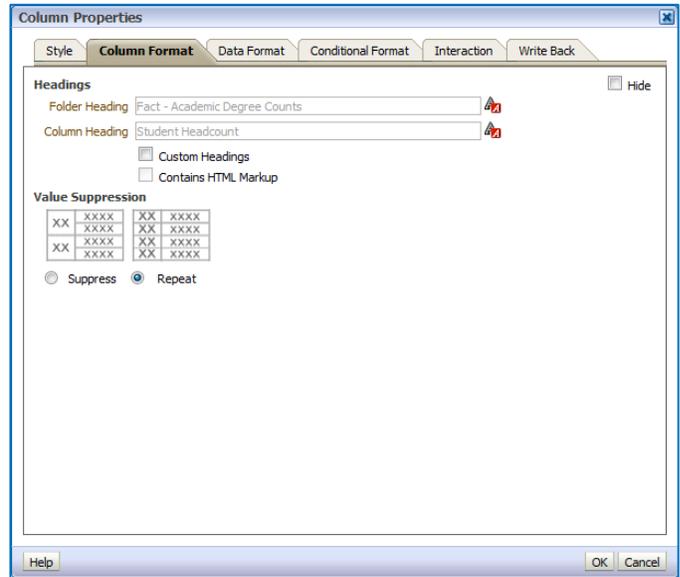
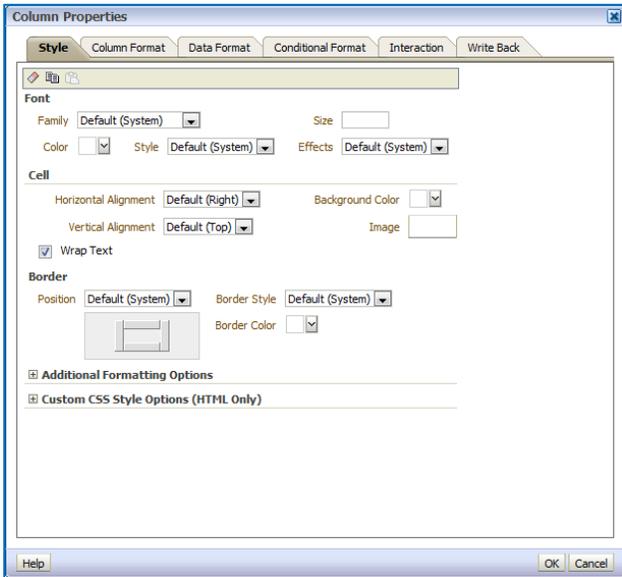
A small blue number will appear in front of each field name to let you know where it falls in the sort order.



To remove a field from a query, just click the “Delete” option. Note that this only removes the field from the Selected Columns area. If you’ve already defined a filter for the field, that filter will stay there. You have to delete the filter separately if you no longer want it.

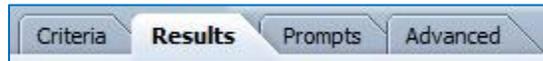
Finally, the “Column Properties” button opens up a dialog box where you can change numerous things about the appearance and functionality of the field (screenshots on next page):

- “Style” properties change the field’s formatting, including fonts, colors, borders, and alignment.
- “Column Format” properties change the field’s headings and value suppression settings.
- The “Data Format” properties let you change how the field values get displayed. With numeric fields, for instance, you can set the number of decimal places to display, choose to display a thousands separator, and define the format for negative numbers.
- The “Conditional Format” tab lets you set formatting when specific conditions are met.
- It’s unlikely that you’ll use the “Interaction” and “Write Back” tabs, but feel free to read up on them.



### Displaying Results

Displaying results for your query is easy. You just click on the “Results” tab:



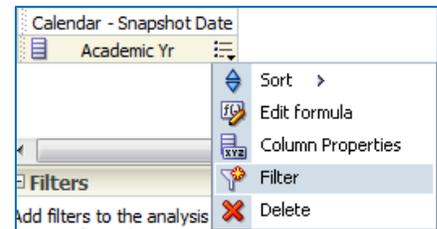
By default, you’ll see results in a simple table, but you have numerous options for viewing those results. We’ll cover these in detail in the next session.

### Filters

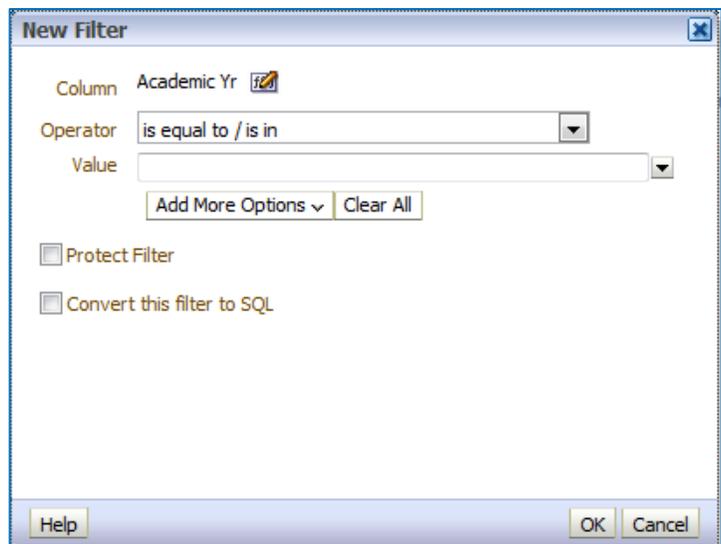
In almost all cases, you don’t really want your query to return data from every record in the database table. To limit the data that returns, you add filters to your queries. In OBIEE, you can either create these filters yourself, or use ones that either you or someone else has saved.

### Creating/Editing Filters

To create a new filter on a given field, click the funnel icon on the dropdown menu for that field in the “Selected Columns” area of your query. If you want to filter on a field, but don’t want the field to appear in the query results, click the “Filter” button on the right side of the “Filters” toolbar . This lets you add a field only as a filter. (You can also add the field the regular way, filter on it, and then delete it from the “Selected Columns” area. As noted earlier, deleting a field from the query doesn’t delete it from the filters.

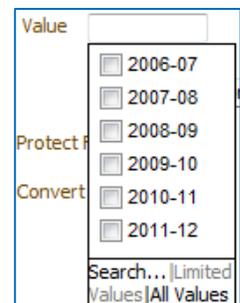


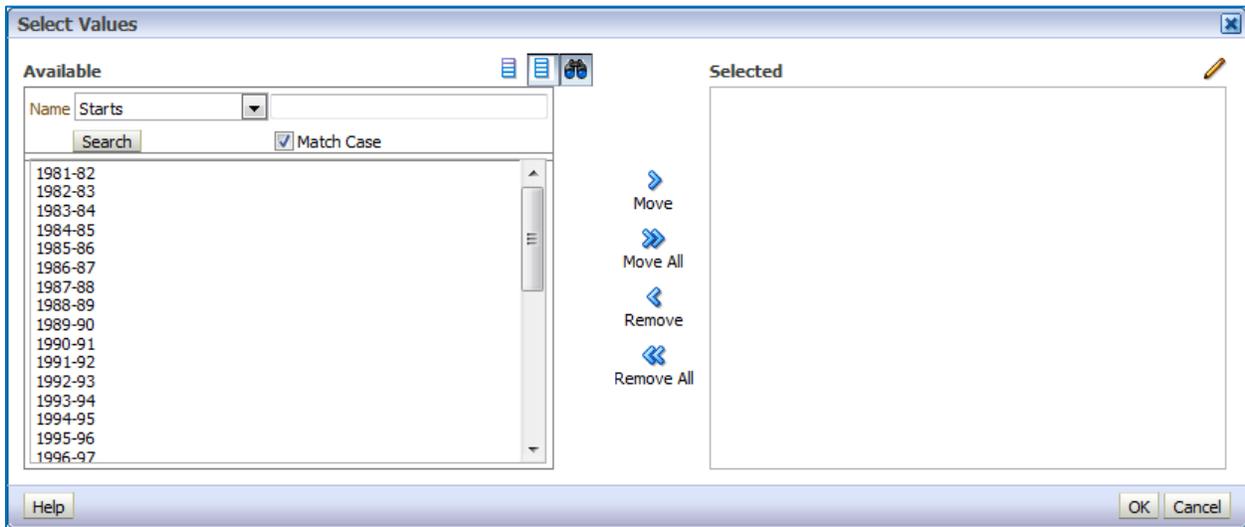
In either case, you will then see the New Filter dialog box, as shown to the right.



The “Operator” dropdown box lets you choose one of 20 different SQL operators to apply to this field, the default being “is equal to / is in.”

The “Value” edit box is where you provide the value to which this filter applies the chosen operator. You can type this in directly if you want (note that Oracle is case sensitive), but you can also choose values using the dropdown on the right side of the box. By default, this shows all valid values for that field that also meet all of the other filters defined in the query (this option, “Limited Values,” is grayed out to indicate that it’s the one being used). The “All Values” option will show every valid value for the field, in the same dropdown box. Finally, the “Search...” option lets you do more advanced searching for filter values, as shown in the screenshot on the next page.



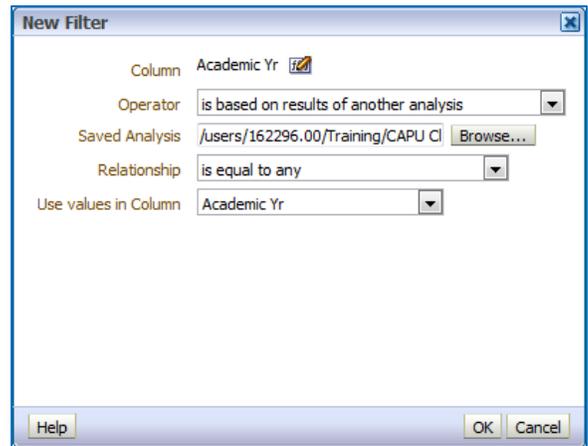
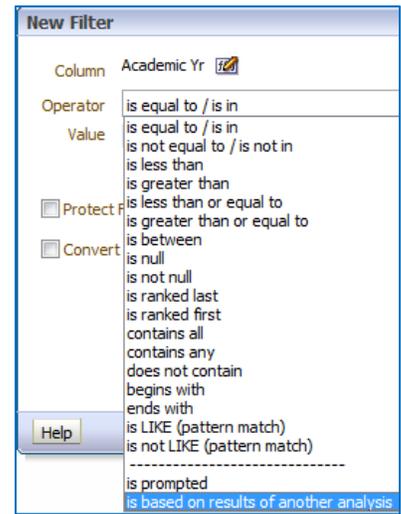


There are also two advanced filter options that you’ll probably use fairly often.

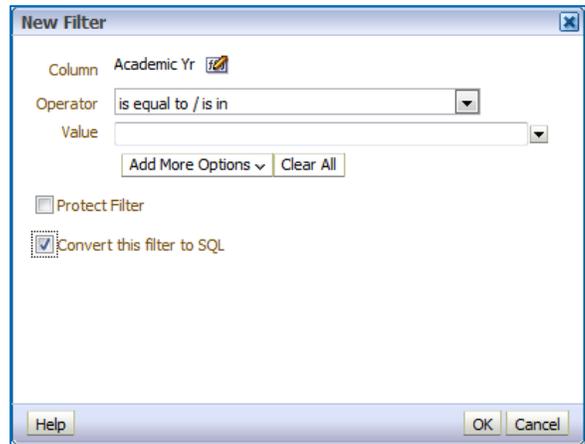
First, you can create a filter that uses the results of another query. You select this option at the very bottom of the “Operator” dropdown that we discussed earlier, as shown to the right. When you choose the “is based on results of another analysis” option, the dialog box shown at the lower right will appear.

In this dialog box, you choose the saved query, operator, and column that you want to use for filter values. When you execute this filtered query, OBIEE will first run the saved query, then show you only those results from the second query where the filter criteria are met.

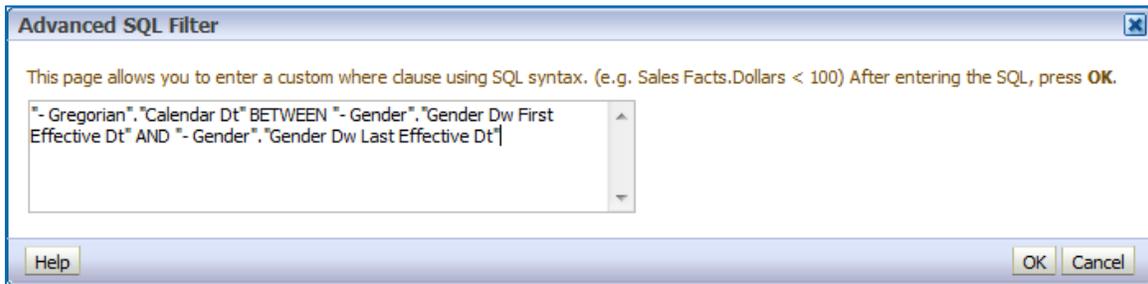
A good example of using this type of filter is when you want data in one subject area only for a particular set of students from another subject area. For instance, if you wanted to know SAT scores (available in the applicant subject area) for students who received degrees in a specific major (from the degree subject area), you could create a degree query that included student IDs, then use that to filter an applicant query.



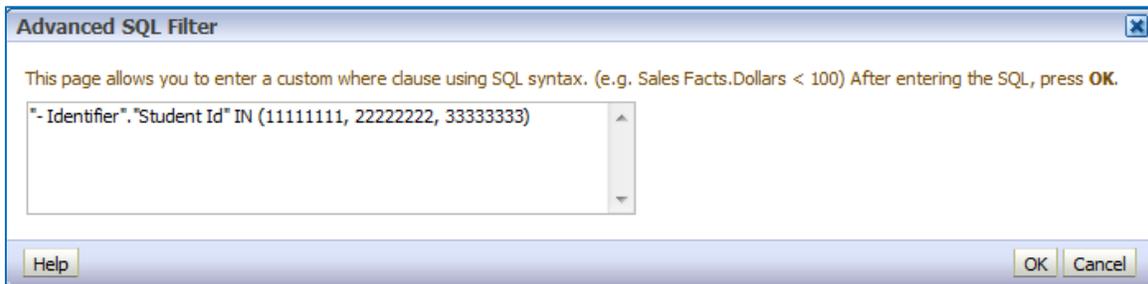
The second advanced option converts the filter to SQL so that you can make it as complex as you like. To convert a filter to SQL, simply select that option in the “New Filter” dialog box and click the “OK” button. This will display the “Advanced SQL Filter” dialog box, where you can enter the SQL.



An example of this kind of filter might be one where you’re looking for records in which one field’s value is between the values of two other fields (note that you have to type this in yourself; there are no shortcuts, picklists, or any other kind of help):

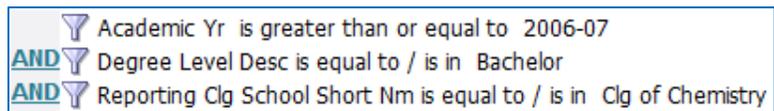


Another example of an advanced SQL filter is one that you may wind up doing a lot: examining data for a (possibly large, up to about 10,000) number of specific students for whom you have IDs. In that case, you can make the filter look something like this, with the student IDs pasted into an IN statement and separated by commas:

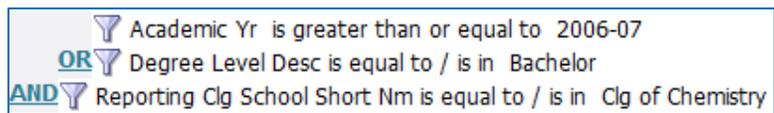


**AND/OR Filters**

When you create filters on multiple fields within the same query, OBIEE’s default is to add them as logical ANDs, as shown to the right. So every filter field must evaluate to TRUE for a particular record to appear in the results.



If you want certain fields to be ORed together instead, so that only one or the other needs to be TRUE, just click the AND between those two fields. Similarly, clicking an OR will convert it back to an AND. Note that indentation indicates field groupings (the equivalent of parentheses in an SQL statement).

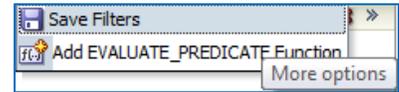


### Filter Options

Once you have a filter defined for a field, there are several actions that you can take on it. When you move your mouse over an existing filter, a toolbar pops up to the right with four options: 

The pencil button allows you to edit that filter, while the copy and paste buttons do just what you'd think. Finally, you can remove a particular filter by using the delete button at the end. If you want to remove all filters at once, use the double-delete button () on the right side of the "Filters" toolbar.

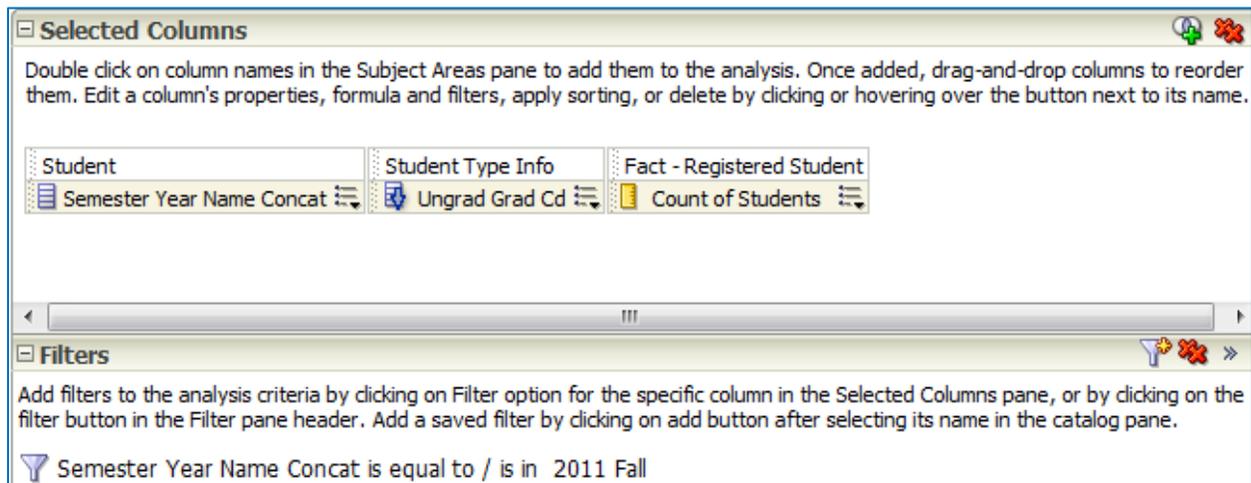
Finally, the double-arrow at the far right side of the toolbar gives you a couple of options, one of which is to save the filter so that you or others can reuse it in other queries.



### Using Saved Filters

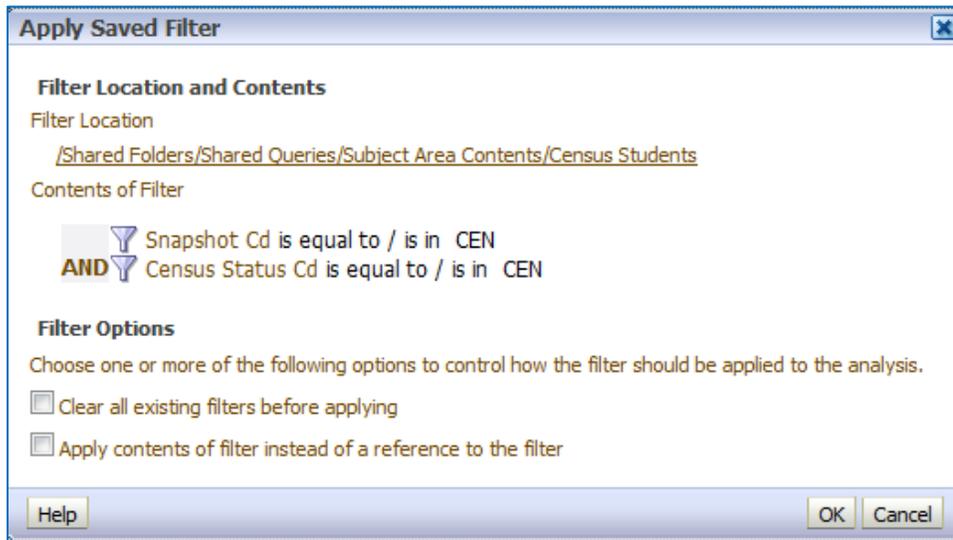
As noted above, you may find it useful to save complicated filters that could apply to many different queries, so that you don't have to keep recreating them. We've done this in the "Students – Registered" subject area, for instance, for common requests such as retrieving only census students, or showing gender and ethnicity as they were recorded during the time period at which you're looking.

For instance, running the following query would give incorrect results if you were trying to retrieve census counts, since it's not specifically limited to that group of students:



As noted above, however, we have a predefined census filter that you can use in any query. Shared saved filters appear in the Catalog pane on the lower left side of the screen, below the subject area's available fields, in the "Shared Folders > Shared Queries > Subject Area Contents > Students - Registered" folder. To use one, just double-click it. That will display the "Apply Saved Filter" dialog box.

This dialog shows you the location and definition of the saved filter, and gives you two options (you can select neither, one, or both):

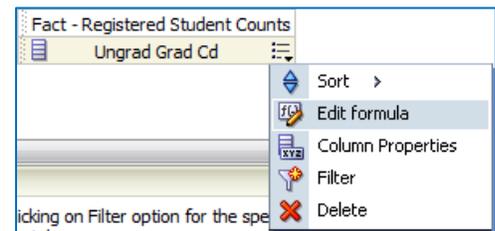


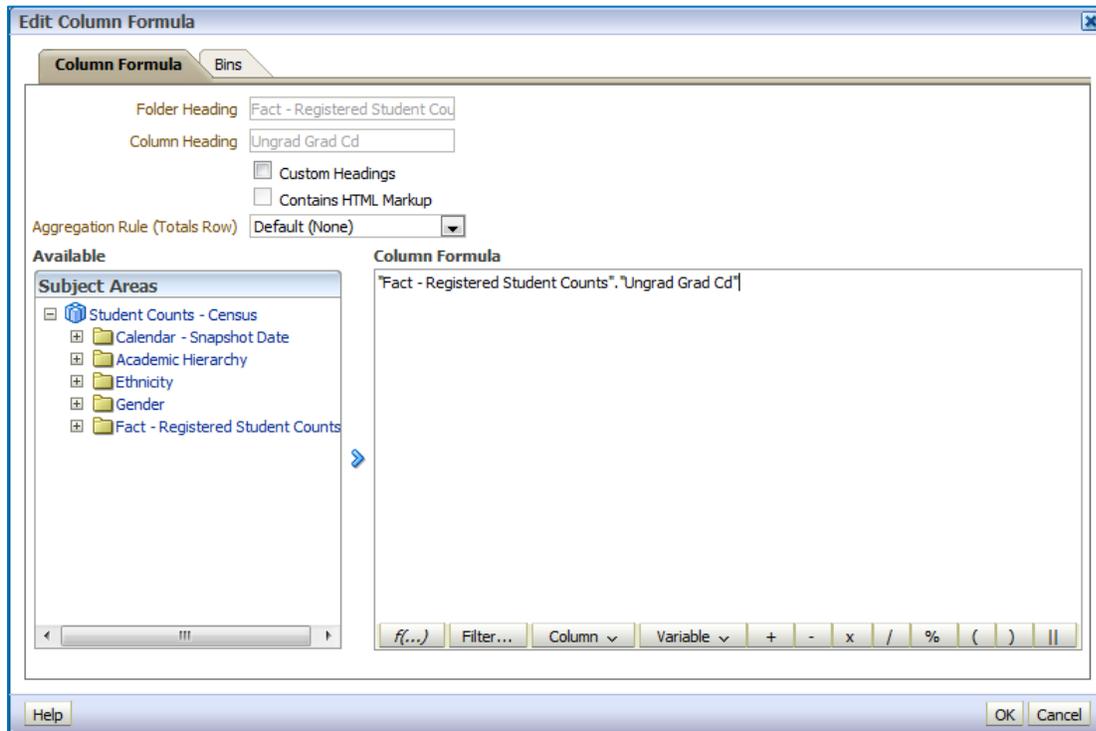
- Clear all existing filters before applying – This option will remove any existing filters that you’ve already defined for the query, replacing them with just the saved filter. If you don’t check this box, the saved filter will just get added to the bottom of any existing filter fields.
- Apply contents of filter instead of a reference to the filter – This option defines how the filter gets added to your query. If you don’t check this box, the filter will be added to the query as a reference only. This has a couple of advantages and one limitation. One advantage is that you can’t accidentally change the saved filter after adding it to your query. The other big advantage is that if the saved filter ever changes, all of the queries that you used it in will automatically pick up the new version. The main limitation is that you can’t change the definition of the filter in this specific query. If you do check this box, then the filter will get added to your query just like any other filter. You can then change it in any way that you want, but changes to the original saved filter won’t flow through to your query anymore.

**Formulas**

The last query field option that we’ll cover lets you use that field in a mathematical formula of some sort. This can include everything from simply adding two fields together to implementing complex programming logic.

When you click the “Edit Formula” option, you’ll see the following “Edit Column Formula” dialog box:

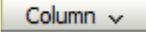




Although this dialog box has many options, there are several that you're unlikely to use, so we'll cover the most important ones in the following sections.

### Simple Math and Additional Columns

It's quite possible that your formula will just involve doing simple math operations or string concatenation involving two or more fields. The eight buttons on the right side of the row just beneath the "Column Formula" box  let you do exactly that. You can add, subtract, multiply, or divide two fields with the first four buttons. You can convert numeric values to percentages with the next, define operational order with the two parenthesis buttons, and concatenate two text values with the last button.

You can just type in the names of additional columns in the formula, or you can use the "Column" button  to choose them from a list (definitely the easier way to go).

### Functions

For more complicated formulas, you can click the "Insert Function" button  to see over 100 predefined functions in the following categories: aggregate, running aggregate, string, math, calendar/date, data type conversion, system information, conditional branching expressions, database functions, and time series calculations.

You can combine and nest these functions in a wide variety of configurations. Note that when you select one, some information about that function, including syntax and a brief description, appears at the bottom of the dialog box.

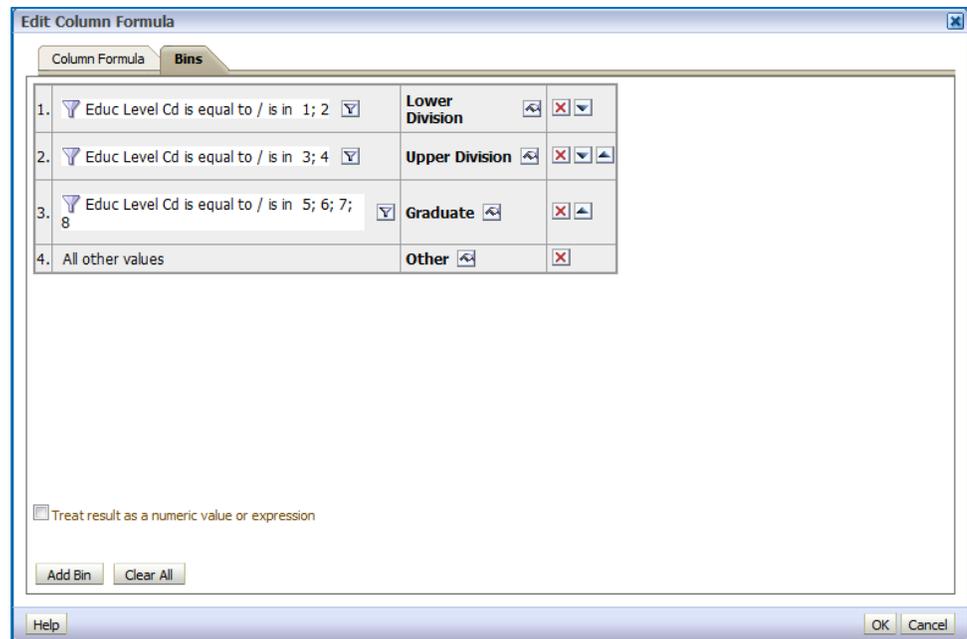
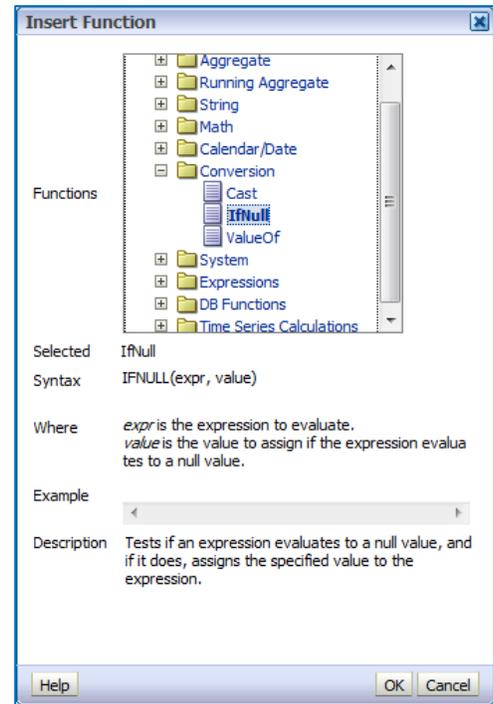
Finally, be aware that if the column name is highlighted in the “Column Formula” box of the previous dialog box, and you insert a function that includes an expression (“expr”) as part of its syntax, OBIEE will actually insert the column name into the function for you, saving you some work.

**Bins**

Notice that the “Edit Column Formula” dialog box has a second, very tiny tab called, “Bins.” You can use this tab to have OBIEE build a CASE statement for you in a formula (you can also create a CASE statement yourself by using one of the functions that appears under “Expressions” if you want). A CASE statement is a lot like an IF-THEN-ELSE statement, except that it’s a little easier to read when you have multiple options. Basically, it puts values in different buckets that you define.

Adding bins uses the same “Create/Edit Filter” dialog box that we’ve already seen. After selecting values, you’ll be asked to name the bin, which can be pretty much any text string. Note that there’s also a checkbox that lets you create a bin for any values that you haven’t specified; you’ll also have to name that bin.

Eventually, you’ll wind up something similar to the example on the right. When you run this query, the Educ\_Level\_Cd field, instead of showing values such as 1, 2, and 3, will show the names that you provided, such as “Lower Division” and “Upper Division.”



## Advanced Analysis Features

There are a couple of more advanced Analysis features available in OBIEE that you may find use for on rare occasions, one more so than the other. Before using either one, however, you should make sure that you understand how they work, since they can both easily mess up your query.

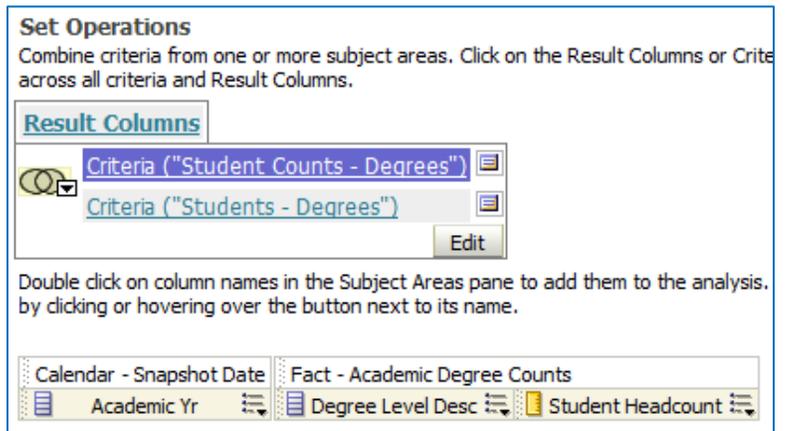
### Set Operations

If you need to combine results from multiple queries, using SQL set operators, you can click the  button that appears in the “Selected Columns” toolbar. You then select the subject area that contains the “similar request” data.

Next, you’ll see a screen like that shown to the right. Here, you define the fields to be used in each of the queries whose results you’re combining.

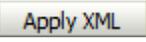
The Venn diagram  lets you choose which set operation you want to do: UNION, UNION ALL, INTERSECT, or MINUS.

When doing set operations, you must make sure that the number and data types of all columns match across all of the queries.



### Advanced Tab

One last thing to cover in the query designer is the “Advanced” tab. There are three sections here:

- **Analysis XML** – This is an XML representation of the analysis that you’ve built so far. You should never edit this text directly, since that will almost certainly cause problems. You can, however, use this to quickly share analyses with other people. Just create the analysis, then copy this XML and give it to the people you want to share it with. They can then create a new, empty analysis, go to the “Advanced” tab, replace the “Request XML” with the new version you gave them, and then click the  button. They will then have exactly the same analysis loaded as you did, including both the query itself and any results formatting that it might have.
- **SQL Issued** – This box shows the logical SQL that gets sent to the Oracle BI Server. Note that this is not the actual SQL that goes to the database, making it not really all that useful. It’s hard to think of a reason why you’d want to change this here.
- **Advanced SQL Clauses** – I’ve yet to come across a real need to use an Advanced SQL Clause, but you have the ability to enter those if you want.

## Practice Analyses

### Degree Practice Analysis

If you drill into the “Degree Counts” dashboard report in Cal Answers, you can see degree recipient headcounts, for unrescinded degrees at a given level, for all of the departments within a particular college or division. Pick any college/division that you want and duplicate the numbers with your own query (just get the right numbers and don’t worry about making the results look exactly the same; that’s the next class). For instance, here’s what your results would look like for bachelor’s degrees in the College of Chemistry:

| Academic Yr | Degree Level Desc | Reporting Clg School Short Nm | Academic Department Short Nm | Student Headcount |
|-------------|-------------------|-------------------------------|------------------------------|-------------------|
| 2007-08     | Bachelor          | Clg of Chemistry              | Chem & Biomolecular Eng      | 92                |
|             |                   |                               | Chemistry                    | 141               |
| 2008-09     | Bachelor          | Clg of Chemistry              | Chem & Biomolecular Eng      | 85                |
|             |                   |                               | Chemistry                    | 140               |
| 2009-10     | Bachelor          | Clg of Chemistry              | Chem & Biomolecular Eng      | 90                |
|             |                   |                               | Chemistry                    | 102               |
| 2010-11     | Bachelor          | Clg of Chemistry              | Chem & Biomolecular Eng      | 92                |
|             |                   |                               | Chemistry                    | 103               |
| 2011-12     | Bachelor          | Clg of Chemistry              | Chem & Biomolecular Eng      | 7                 |
|             |                   |                               | Chemistry                    | 10                |

For the easier way of doing this, use the “Student Counts – Degrees” subject area. For more of a challenge, use the “Students – Degrees” subject area.

When you’re done, save the query to your “My Folders.” You may want to create subfolders there to help you manage all the files that you’ll eventually have.

### Census Count Practice Analysis

The “Census by 2 Fields” dashboard report defaults to showing census headcounts by gender and 4-value ethnicity for the campus. Using the prompts at the top, you can filter the results down to a specific college/school or division. For instance, here are the results for the College of Chemistry:

|                |             |                                | Student Census Headcounts |            |             |            |             |            |
|----------------|-------------|--------------------------------|---------------------------|------------|-------------|------------|-------------|------------|
|                |             |                                | 2009 Spring               | 2009 Fall  | 2010 Spring | 2010 Fall  | 2011 Spring | 2011 Fall  |
| Student Level  | Gender Desc | Ucb Level 1 Ethnic Rollup Desc |                           |            |             |            |             |            |
| U              | Female      | Asian/Pacific Islander         | 166                       | 179        | 171         | 174        | 163         | 173        |
|                |             | International                  | 38                        | 41         | 43          | 47         | 40          | 57         |
|                |             | Underrepresented Minority      | 15                        | 18         | 13          | 13         | 8           | 13         |
|                |             | White/Other                    | 78                        | 81         | 74          | 71         | 66          | 61         |
|                | Male        | Asian/Pacific Islander         | 233                       | 257        | 235         | 261        | 251         | 293        |
|                |             | International                  | 52                        | 51         | 47          | 61         | 53          | 66         |
|                |             | Underrepresented Minority      | 24                        | 31         | 29          | 29         | 26          | 29         |
|                |             | White/Other                    | 136                       | 154        | 145         | 157        | 153         | 163        |
| <b>U Total</b> |             |                                | <b>742</b>                | <b>812</b> | <b>757</b>  | <b>813</b> | <b>760</b>  | <b>855</b> |

Choosing any college or division that you want, duplicate these counts (again, not the format). For the easier version, use the “Student Counts – Census” subject area. For a considerably more difficult version, use the “Students – Registered” subject area, and remember to use saved filters! Save the query when you’re done.

**Undergrad Cohort Graduation Rate Practice Analysis**

This one is a bit more difficult, so feel free to ask questions. The Fall 2004 entering cohort of new freshmen consisted of 3,674 students. Of those, the first declared major for 190 of them was in the College of Chemistry and in the College of Natural Resources for 276 more. Of those two groups, which one had the better 6-year graduation record?

For this query, you'll use the "Students-Undergraduate Cohorts" subject area. You'll want Fall 2004 data only, looking at applicant types of "4" (Freshman, HS Grad), where the first undergraduate major is in reporting colleges of Chemistry or Natural Resources. That should be enough to get you the 190 and 276 numbers. (Note that this just tells you they graduated from UC Berkeley. They may not have graduated with majors in those colleges, however.)

To get the 6-year graduation data, you'll add in a field called "Yrs\_To\_Graduation\_Nbr" and look for values between 0 and 6. That will tell you everyone who graduated within six years. Here's what I got:

Save your query!

| Semester Yr | Semester Nm | Applicant Type Cd Shrt Desc | Reporting Clg School Short Nm - First Undergrad Major | Count Of Students |
|-------------|-------------|-----------------------------|-------------------------------------------------------|-------------------|
| 2004        | Fall        | FRESHMAN, HS GRAD           | Clg of Chemistry                                      | 180               |
|             |             |                             | Clg of Natural Resources                              | 254               |

**Physics and Statistics SAT Scores Practice Analysis**

In 2010-11, 91 students received bachelor's degrees in Physics and 93 in Statistics. Which group had higher SAT Math scores?

For this, you'll actually need to write three queries. The first will retrieve the student IDs of Physics or Statistics bachelor's degree recipients in 2010-11. The second and third queries will then use those IDs to retrieve SAT Math scores (not everyone will have one) from applicant data for each set. (There are other ways to do this, but this seems like the easiest to me.)

I just downloaded the results to Excel and calculated the following:

|                              | Physics | Statistics |
|------------------------------|---------|------------|
| <b>Total Degrees</b>         | 91      | 93         |
| <b>Missing SAT Math</b>      | 31      | 34         |
| <b>Avg. SAT Math Score</b>   | 741     | 747        |
| <b>Median SAT Math Score</b> | 750     | 760        |