Data Discrepancies or “Why don’t my numbers match?”

In the Office of Institutional Research & Effectiveness (OIRE) we are often faced with investigating data discrepancies between: 1) two reports we have provided for different purposes, 2) a report we provided that does not match reports provided by other departments or that does not match data pulled from CU-SIS (Cognos). The most common reason that an OIRE report might have different numbers than you expect is that different definitions are sometimes used to answer different questions. While fall enrollment numbers might seem simple, there are many definitional/inclusionary criteria that could have a potentially large impact on what numbers are provided by OIRE. Examples of such criteria include: state funded enrollment vs. extended studies enrollment, primary majors vs. multiple majors, and enrolled majors vs. eligible to enroll majors. All of these are valid methods to count enrollment, but the definition/criteria determines the final number… which may not match a different definition employed for another purpose.

We face similar considerations with faculty and staff counts. A count of ‘faculty’ for a question about office space might include faculty administrators or student faculty but a question about how many employees we have of each ‘type’ for a federal or state standardized report by ‘type’ might count them in different categories. A question about who is teaching our courses will produce yet a different number, since not all faculty teach, and some teaching is done by people whose highest ranking appointment is staff or regular administration. Finally, because some departments support multiple academic programs and some academic programs are multi-departmental, questions framed around academic programs will produce different counts than questions framed around organizational or funding departments. Each of these approaches is valid, but each will result in a very different bottom line.

In OIRE we strive to be consistent in our reporting, but there are legitimate reasons for using different definitions/criteria for different purposes. This can result in very similar-looking reports that actually represent very different populations. Because of this we always include report definitions on everything we produce. That way, the individuals using those reports have an understanding of the criteria that were utilized to create them.

Finally, while OIRE personnel will be able to articulate their methodologies for OIRE reports, we have no way of knowing how counts produced outside of OIRE were derived. Without the exact and precise definitions, timeframes, and data sources that were used in those non-OIRE reports, we likely won’t be able to explain the other reports’ numbers (unless they include clear, specific definitions/criteria). Because of the documentation OIRE includes on all reports and the additional documentation we maintain internally, we will be able to articulate the inclusionary/exclusionary criteria employed in our reports. This will allow you to reach out to the provider of the data in non-OIRE reports for that same information for comparison purposes.