ACADEMIC ANALYTICS

2016 Database (AAD 2016) Methodology

Introduction

There are two levels of licensure to the Academic Analytics Database; Comparative/Details and Master License Agreement (MLA). Both of these modules provide different levels of access to the data and tools to uncover and analyze the data.

Content and Coverage

The Academic Analytics Database includes data on the fundamental areas of scholarly research activity and provides users with benchmarking and analysis capabilities. The database is organized by institutions, broad fields, departments, Ph.D. programs, and by individual faculty members. The database includes information on over 225,000 faculty members associated with more than 9,700 Ph.D. programs and 11,000 departments at 407 universities in the United States and abroad. The database is fully customizable, allowing users to create custom peer groups at different levels of analysis as well as custom groups of disciplines based on selected options.

The Academic Analytics Scholarly Research Index (SRI) was developed to express the level of research activity across disciplines. The Scholarly Research Index can be applied at different levels; the level of individual faculty, units (programs or departments) within a university, and the overall performance of universities. The Index uses metrics that are independent of discipline values and of the portfolio of disciplines at universities. The Index is matched to seven separate variables in six areas of activity:
Total Publications
Total Citations
Total Books
Total Grants
Total Grant Dollars
Total Awards
Total Conference Proceedings

Comparative/Details Database

The Comparative/Details Database is the base level of licensure. Clients subscribing at this level receive two views of national scholarly activity, comparative data which act as a base or ground and details data which is an overlay on top of comparative data, providing enumerations of the base data. The Comparative Data are released annually and reflect scholarly activity in academic research for a given time window – a rolling range of years. The database includes metrics on professional honors and awards, federal grants, book publications, journal articles, conference proceedings, and citations. The coverage period of the current Comparative Data is:

Articles:	2013 - 2016
Citations:	2012 - 2016
Conference Proceedings:	2013 - 2016
Books:	2007 - 2016
Grants:	2012 - 2016
Awards:	No Limit

The Details Data include everything in the Comparative Data plus greater details regarding the individual pieces of scholarly activity linked to faculty in the Comparative Data. Unlike the Comparative Data, which provide a locked snapshot of activity for a given range of years, the Details Data are updated daily and contain activity outside of the most current Comparative Data time window. The Details Data may be accessed on the downloads screen of our online portal and in faculty-specific portal tools such as Quintiles by Rank, the Career Progression chart, and the Unit Modeling tool. The maximum coverage timeframe within the Details Data is:

Articles:	2004 - present
Citations:	2004 - present
Conference Proceedings:	2008 - present
Books:	2004 - present
Grants:	2006 - present
Awards:	No Limit

The Details Data allow clients to drill down into the particulars of individual faculty research activity to see publication meta-data for journals, books, and conference proceedings in addition to detailed information on awards and grants. The Details Data contain over fifty separate downloadable files which are organized into four segments:

Administration Data
Comparative Data
Comparative Details Data
Full Data

Master License Agreement (MLA)

The Master License Agreement is the second level of licensure. Clients with a Master License Agreement have access to the entire Academic Analytics production data set. Access includes the entire Comparative and Details databases, as well as access to the larger data warehouse. Further, this level of access offers other tools such as Collaborations and Research Insight. Finally, custom analyses performed by Academic Analytics analysts are provided to answer specific client questions.

This methodology describes the organization and construction of the Comparative/Details Database. The following sections describe construction of the database (inclusion of faculty and organization of entities), the elements of the database (research activity), calculations of performance comparisons, weighting of metrics, and metric definitions.

Faculty Members and Academic Units in the Comparative Database and Inclusion Methodology

This section describes how people are grouped and filtered for inclusion in the AAD2016 comparative database. AAD2016 includes 225,000 scholars in 20,950 comparative groupings at 407 institutions in the United States.

Each individual faculty member is assigned an Academic Analytics Unique Identifier (AA-UID), descriptive attributes and unit affiliations linked to their institution. The AA-UID is carried with that person should they move from institution to institution in the database. Whenever possible, this ID number is cross-walked to an institution-assigned ID for each person to facilitate (or expedite) year-over-year updates and maintain accurate matches to publications, grants, etc. An example where this ID structure is helpful is the disambiguation of two individuals with exactly the same name, no middle initials, and where both are in the same college or school of an institution. Without a distinguishing identifier, it is easy to confuse their publications (especially if both are engaged in similar research).

In addition to attaching an institution-level ID, (Open Researcher and Contributor ID) ORCID IDs can be submitted and warehoused alongside AA-UIDs and employee IDs. A list of match candidates can be provided if an institution submits faculty groupings without ORCID and would like to coordinate effort to match it with AA-UID and internal employee IDs.

Descriptive attributes and unit affiliations are obtained or inferred through two processes:

Submission – Institutions provide information about their faculty members.
Collection – Academic Analytics staff compile information from publicly available web resources.

Unintended changes in the inclusion of faculty between product years may occur. Usually, this is due to one of the following situations.

An institution was collected, and a more complete faculty roster was identified online.

An institution previously submitted a faculty list and what was provided does not coincide with the information that is publicly available online (e.g., only core faculty were submitted for a PhD Program, however, joint faculty from other departments were then identified through the collection process).

If an institution is interested in looking at only "core" or primary appointment faculty, the department level data is recommended as only core members of the departments are represented. All affiliated faculty are grouped into units classified as PhD Programs.

Submission Process

While access to the data is subscription based, participation is not. Academic Analytics encourages institutional involvement to ensure that the groupings represented are as inclusive and accurate as possible. Hence, each year, a letter is sent to every Ph.D. granting institution in the U.S. requesting participation in the identification of people and structure of their university for the purpose of including that information in the comparative database. In the initial release of AAD2016, 53% of scholars in the database were received through institutional submissions.

The window for submitting faculty for AAD2016 was between November 1, 2016 and March 30, 2017. However, submissions and updates are accepted throughout the year in an effort to constantly refine the data.

We request faculty lists include all tenure/tenure-track faculty and non-TTT faculty who are expected to produce research as of November 1, 2016. This includes:

- All tenure/tenure-track faculty, including administrators, faculty who are on sabbatical, faculty on leave without pay but are expected to return, and those who haven't published recently.

- Individuals not paid by the institution but who are affiliated with the institution and are expected to produce research as a condition of the affiliation.

- Others whose job involves scholarly research and are expected to pursue grants and publish. This could include research-track faculty, emeritus faculty who continue to hold a research obligation or professors of practice engaged in research.

A submission template is provided to institutions illustrating the preferred format of submitted faculty lists. In cases where an institution is unable to respond in the format requested, a list that is more convenient for the institution to produce will be accepted and processed into our system. Partial submissions are accepted in the event that an institution cannot easily provide all requested fields. Requested fields are assigned one of the following priorities.

Required – These fields are essential; without them, Academic Analytics is unable to process the submitted list. If unable to provide these fields, we ask that our request be directed to another office on campus.

- *Person Name (both Last, First Mid and FML formats accepted)*
- *Academic Title (from HR or other system)*
- *Unit Name (Individual's primary academic unit affiliation or additional affiliation for which the individual has a disciplinary focus. In some cases, a person may not have a tenure home. Centers or research institutes may be submitted as their unit affiliation or as an additional affiliation. Individuals with a significant secondary focus may be submitted in more than one departmental unit.*

Highly Recommended – These fields, if not provided, are assigned/inferred by Academic Analytics. They have been separated from other recommended fields due to their increased impact on accuracy and processing time.

- *Person Identifier (from university system)*
- *Tenure Status (from HR or other system)*

Recommended – These fields, if not provided, are assigned/inferred by Academic Analytics through web research. By providing any of these fields, institutions help to ensure accuracy and improve processing time.

- *Faculty Rank Type (Professor, Associate, Assistant, Lecturer, Instructor, Other)*
- *Tenure Status Type (Tenured, TenureTrack, NotTenureTrack)[1]*
- *Faculty Type (Regular, Research, Clinical, Public, Librarian, Other)*
- *Is Administrator (university-level administrator, dean of a college, or associate dean)[2]*
- *Administrative Title*
- *Is Emeritus*
- *Is Primary Appointment (for individuals submitted with more than one unit)*
- *College*
- *PhD Program (where individuals may sit on a dissertation committee and mentor Ph.D. students)[3]*
- *Hire Date (from HR or other system, date of hire, not date of tenure)*
- *Terminal Degree (if multiple, include first obtained)*
- *Terminal Degree Year (if multiple, include first obtained)*
- *Terminal Degree Institution (if multiple, include first obtained)*
- *CIP Codes (to units and Ph.D. programs as classified in institution's system)[4]*

[1]*Note that in some cases NTT research track faculty are not easily defined or do not exist at a given institution.*

[2]*The expectation of research varies from institution to institution with regard to administrators. In some cases, this can impact unit placement in the comparative data.*

[3]*In cases where an institution cannot easily identify Ph.D. Program placement, Academic Analytics takes the list with departmental affiliations and infers Ph.D. Program placement which is later provided back to the institution for approval.*

[4]*Note that while CIP Codes are referenced in the creation of the Academic Analytics Taxonomy and are requested in the submission instructions, they are currently used for reference only. Reporting by CIP is not available.*

Not all requested fields are available in the comparative database. Some data remain incomplete because institutions have not submitted those fields. However, they are archived because they may prove useful for clients and Academic Analytics, particularly in the data matching processes.

<u>Useful Reference</u> – These fields are helpful in the data matching process but are not required to perform data matching.

- *Terminal Degree Field*
- *Publication Alias*
- *Prior Institution*
- *Gender*
- *ORCID*

The process by which a submission list is created varies from institution to institution. It also depends on the resources available. The following describes several different paths taken to compile a submission.

- *An HR file is sent to Academic Analytics followed by a review of Ph.D. program placement*
- *An HR file for departments is sent to Academic Analytics; the Graduate School compiles a list for Ph.D. Programs*
- *The previous year's submission list is distributed to Deans and/or Department Chairs for corrections and then compiled and reviewed*
- *An IR office identifies people using one or more systems on campus, assigns program affiliations and reaches out for feedback as needed to other offices*

Once a list is submitted, Academic Analytics uploads all individuals to the data warehouse and reviews the groupings for consistency with the other submitted and collected groupings. When questions arise processing a submission, they are sent to the Academic Solutions team who follow up with the submitting institution. This often prompts conversations about inclusion since research expectations by rank and appointment can vary. This is where working directly with the institution is preferable since the definitions of included faculty can be refined to a more granular level not possible through web collection.

Our approach is to cast a wide net and then build comparative groupings applicable to all institutions based on information available in the database. This is why some people who were included on a submitted list may not appear in the data provided back to the subscriber. A common example is clinical medicine departments. Data for these units are currently collected and stored in the warehouse; however, Academic Analytics cannot report back on clinical medicine departments because we do not yet have enough people grouped and matched to research activity metrics to build a useful national comparison.

People are classified based on criteria described in the section Collection Process which appears below. Units are organized into three categories: departments, Ph.D. programs, and other units. These categories are defined as follows:

Departments: these units can vary from institution to institution and include departments, divisions, schools (that do not contain departments), colleges (that do not contain departments), programs (degree granting units), and sometimes areas (in cases where they function as a department).

Ph.D. programs: this category includes every accredited Ph.D. program in the United States. In some cases, a program is separated into Ph.D. concentrations/specializations in an effort to provide more granular data (e.g., a Ph.D. in Business Administration with specializations in Accounting and Finance would be separated into two Ph.D. areas. In this case, data on Business as a combined effort of these two areas is available at the broad field level).

Other units: includes academic units that are not yet a part of the comparative data displayed to all subscribers (e.g., clinical medicine departments, research institutes, centers, labs, other program types, and custom groupings created by subscribers).

Once people are grouped and classified and academic units are assigned, a verification file is provided back to the institution for review of faculty placement, unit naming and classification. This allows the institution to make edits and submit them prior to finalizing the groupings that will appear in the comparative data released to all clients.

A cautionary note: There is a common misconception that all people who appear on submitted lists can be included in national benchmarking. While individuals who fall outside of these guidelines are still uploaded to the data warehouse and are available for custom groupings and the details data product, they are not included in the national comparative database. Data on these individuals are collected and stored with the idea that, over time, enough institutions will participate to open up new possibilities for comparisons. Based on client feedback, updates to the collection process will be reviewed and considered for future iterations of the database.

Maintaining standardized groupings across all institutions is a constant focus in database construction. Inclusion/exclusion decisions made by subscribing institutions are reviewed during the construction of each yearly database to ensure that the nuances of research expectation across all institutions is captured without skewing the data. This is most apparent in cases where certain types of individuals are not regularly included in submitted lists (e.g., if a small number of emeritus professors are included in a submitted list from one institution, and the following year the number of emeritus professors submitted doubles or triples, all emeritus professors are excluded by default and noted for later discussions with the client).

In the event that an institution decides to cease submission of a faculty list, person and unit groupings are carried over for one database year in the hope that the institution will reconsider and continue to submit. In some cases, there may be an extended delay in providing a list. For example, an institution submits a list for AAD2015 but is unable to submit a list for the 2016 database by the deadline of March 30 (2017). If a list is unavailable by the time AAD2016 is released, the 2015 faculty groupings are used to allow for additional time for an institution to provide a list. If a list is not submitted at all, the institution will be collected for AAD2017.

Collection Process

The first step of the collection process is to review a university's website for changes in college structure and Ph.D. programs. During this process, sources for faculty lists are identified and the best available list for a given unit is identified. In the case of departments, university directories, college directories and departmental faculty lists are reviewed to determine who to collect. Since definitive lists of individuals who sit on dissertation committees and mentor Ph.D. students are difficult to infer from website resources, the following hierarchy is employed to decide what source list to collect:

1. List with a header of Ph.D. program faculty
2. List from a graduate catalog/bulletin
3. List by research area from a departmental website
4. Department faculty list

Faculty members are then collected from the source lists identified during the structure review. Collection and inclusion of individual faculty are based on the person's title that appears on the collected list or their profile page when not available from the collected source list.

Units that are classified as "other unit" are not collected or populated with individuals. Types of units currently classified as "other unit" include:

Clinical Medicine Departments
Centers/Institutes
Laboratories
Libraries
Regional Campus Units

Individuals with the following titles are collected, assigned an AAU-ID, uploaded to the data warehouse and grouped in a department and/or Ph.D. program:

Professor
Associate Professor
Assistant Professor

(Administrators, Distinguished, Endowed, "Named", Regular, Research, Regents, Trustee, University prefixes are all included.)

Based on the information available online, a person is assigned to their primary department and any Ph.D. program where they hold a primary, secondary or joint appointment. Additional appointments are not assigned to people in the department level data. However, there are times when a primary appointment is difficult to identify and a person is represented equally in more than one department. 93% of people represented in the AAD2016 comparative data are assigned to a single department. It is assumed that individuals with a joint or secondary appointment are actively mentoring Ph.D. students and sitting on dissertation committees and that a significant portion of their research is expected in that additional discipline. Professors with the following title types are not collected (However, if they are submitted, AAU-ID and unit assignments are made and an exclusion flag is placed on the person so that they will not appear in the comparative database, but can appear in custom reporting or in Faculty Details data feeds):

Adjunct	*Practice*
Clinical	*Public Service*
Emeritus	*Teaching/Instructional*
Extension	*Visiting*
In-Residence	*Cooperative and Courtesy appointments*
Part-Time	

Additional title types not collected:

Curator	*Musician/Artist/Conductor*
Fellow	*Nurse Practitioner*
Instructor	*Post-Doc*
Lab Technician	*Research Associate*
Lecturer	*Scholar*
Librarian	*Scientist*

In general, people who hold non-collected titles are not research active faculty members and do not sit on dissertation committees. Hence, they are not included in national benchmarking. Since definitions can vary from institution-to-institution, if a subscribing institution feels that an individual who holds an excluded title type should be included, it is possible to make adjustments to inclusion based on more specific criteria defined by the client institution. An example would be an Emeritus Professor who is still research active and working with Ph.D. students. Academic Analytics does not collect all emeritus faculty in the country. If a client advises to include some individuals, that is permitted and considered an exception.

Taxonomic structure

Each entity is classified into our taxonomy of 171 disciplines. While the taxonomy has three levels of aggregation, reports are generated at four levels. The most specific level reports on individual faculty activity. The first and second taxonomic level reports provide comparative data at the unit level (program or department) and on a broader scale. The third level is the broad field level. Faculty names and records of research activity are de-duplicated at each level of analysis for academic units that roll up to the same broad field, so that a professor who is a member of both the *Sociology* and *Political Science* units at level one of the taxonomy appears only once in the *Social and Behavioral Sciences* broad field at level 3. The table below shows the classification of three sample disciplines.

Level 1	Level 2	Level 3
Cell Biology	Biological Sciences	Biological and Biomedical Sciences
Statistics	Mathematical Sciences	Physical and Mathematical Sciences
Anthropology	Social Sciences	Social and Behavioral Sciences

Each Ph.D. program in the United States is assigned one or more disciplinary classifications to identify the discipline(s) the program addresses. Level one categories in the Academic Analytics taxonomy are defined to include at least 20 programs in each category. Where groups are smaller than 20 programs, they are aggregated to the next level of the taxonomy to provide the most specific classification possible while maintaining a sufficiently large number of programs to facilitate statistical comparison. Level one programs are then grouped into level 2 categories that allow comparisons, for example, of activity between universities with finely divided programs in the biomedical sciences with those having broadly based "umbrella" programs. Finally, level 3 of the taxonomy presents 11 broad areas of scholarly endeavor that match those typically used by the National Research Council (NRC), federal funding agencies, etc. (e.g., social and behavioral sciences, engineering, etc.). Comparisons of faculty research activity may be made at any of these levels.

Our taxonomy is based on the National Center for Education Statistics (NCES) Classification of Instructional Programs (CIP) code classifications, with some adjustments to include disciplines not represented with a six-digit CIP code and some omissions where 20 or more PhD Programs could not be identified. While CIP codes are requested in the submission template, reports by CIP code are currently unavailable due to the difficulty in making assignments for un-submitted faculty and the inconsistency across institutions in their assignment of CIP codes to their faculty.

Description of data elements

AAD2016 includes publications (books, journal articles, and conference proceedings), citations (to journal articles and conference proceedings), federal research grants, and professional honors and awards matched to individuals with varying coverage periods described below. Academic Analytics continuously looks to expand the descriptive data matched to faculty; hence, feedback on the following sections is welcomed and appreciated.

Journal articles

The 2016 comparative data includes over two million peer-reviewed articles published in 2013, 2014, 2015, and 2016 matched to scholars in over 28,000 warehoused journals. Publication data are derived from CrossRef, an association of scholarly publishers that develops shared infrastructure to support more effective scholarly communications. It effects linkages through CrossRef Digital Object Identifiers (CrossRef DOI), which are tagged to article metadata supplied by the participating publishers.

All publishers that assign a digital object identifier (DOI) to article content and submit to CrossRef are included. The completeness and accuracy of authors collected from CrossRef relies on the information that is submitted to CrossRef. While a DOI is a requirement for inclusion, it is possible for a journal that assigns DOIs to its content to not appear in the Academic Analytics warehouse in cases where a publisher has opted out of sharing content with third-party members or due to the timing of their submission to CrossRef. If a publisher submits incomplete information, we inherit the omission. For example, some publishers only submit an article's first author. As issues are identified, corrections are pursued to the extent that resources allow. Academic Analytics encourages faculty to request that publishers participate with CrossRef DOI.

Author affiliations listed in the bibliographic record are not separated and linked to each individual author, thus it is currently not possible to create custom reports that uses author affiliations. Author position may be known for a subset of journals, but is not considered in the comparative data due to the incompleteness of those data.

To prevent "double-counting" co-authored journal publications when two or more authors are in the same academic program, each of them is credited with having written a publication. However, if two or more authors of the same article are in the same Ph.D. program or academic unit, this publication counts only once toward the number of unique publications for the academic unit (likewise, the citations garnered by this publication are counted only once for the program, although each of the co-authors is credited as having written a publication that has garnered these citations when faculty members are considered individually).

Matching authors to scholars is an elaborate process that relies greatly on a manual validation effort. Probability scores help reviewers prioritize verification order and are based on keywords in abstracts/titles, co-authorship, publication patterns, statements of affiliation, and other descriptive metadata. Unscored match candidates are also reviewed in an effort to create the most complete article profile possible using articles with a DOI. The primary focus is to not attach false matches to individuals, this makes common names the most difficult to complete since there is often over 10,000 potential unscored articles to a single person. Scoring methods are under constant refinement to bring more probable match candidates to the top of the research queue. A small percentage of very highly scored articles are matched without manual review, this process is done with great caution to remain in line with the primary focus of not creating false author-to-scholar links.

Since it is often the case that all entries from a journal issue are assigned a DOI, Academic Analytics classifies DOIs to exclude non-article content. This additional content is warehoused and available through the details product or custom reporting. The article classifications currently assigned include:

1. Article
2. Book Review
3. Other
   a. Announcements and Acknowledgements
   b. Awards and Recognition
   c. Corrections and Retractions
   d. Discussions and Forums
   e. Editorials and Commentaries
   f. From the Editor
   g. In Memoriam
   h. Interviews
   i. Introductions and Conclusions
   j. Letters to the Editor/Author
   k. Meeting Reports/Notes
   l. Posters
   m. Replies and Correspondence
   n. Reviews (Non-Book, such as film or other media reviews)
   o. Tables of Contents and Issue Summaries/Notes
   p. Other

It is assumed that the majority of journals participating in DOI are peer-reviewed. However, there is no indication that exists in the CrossRef data that identifies a journal as peer-reviewed. Unless we are pointed to a specific journal that is not peer-reviewed it is counted by default. Some publications, such as professional magazines, do participate in DOI and are warehoused but excluded from the comparative data. Academic Analytics understands that non-peer-reviewed content exists in the literatures of most disciplines with varying overall "importance" to specific disciplines. Non-DOI article content is not warehoused by Academic Analytics, but is under consideration as a future collection project pending resource availability.

A complete list of journal coverage is available on the client portal main page under "Home → Documents."

Conference proceedings

Conference proceedings published in 2013, 2014, 2015 and 2016 are included in the database and are matched to over 58,000 individual authors. Like articles, additional content is available through the details product and through custom reporting. The following disciplines in engineering and physical/mathematical science devote a substantial amount of their overall publishing activity to the publication and dissemination of conference proceedings.

Aerospace Engineering
Applied Mathematics
Applied Physics
Astronomy and Astrophysics
Biomedical Engineering
Civil Engineering
Computational Sciences
Computer and Information Sciences, various
Computer Engineering

Computer Science
Electrical Engineering
Engineering Mechanics
Engineering, General
Engineering, various
Environmental Engineering
Geological and Mining Engineering
Industrial Engineering
Information Science/Studies
Information Technology/Information Systems
Management Information Systems
Materials Engineering
Materials Science and Engineering
Mechanical Engineering
Nuclear Engineering
Operations Research
Physics, General
Systems Engineering

Proceedings that are assigned DOIs and submitted to CrossRef are included. The lag time between publication and appearance of publication records in CrossRef is greater than with journals. Hence, completeness of matches to scholars in subsequent releases of AAD2016 will increase as more data become available. It is estimated that 80% of conferences expected to submit proceedings with DOIs will be represented in the first release of AAD2016. Data matching processes are similar to articles, but conference proceedings present additional difficulties in that it is often the case that only the first author's affiliation is listed, and abstracts often are not available.

Though proceedings are not weighted in all disciplines, they are matched to all scholars. These numbers will appear in Faculty Counts and in the Faculty Details downloads. Some disciplines will have limited coverage based on conference participation with CrossRef.

A complete list of conference proceedings is available on the client portal main page under Home > Documents

Citations

Citations to articles and proceedings from 2012, 2013, 2014, 2015 and 2016 are included in the database. There are over 31 million citations included in the database. For articles published five years ago, five years have elapsed during which citations may have accrued (for the 2016 database, an article published in 2012 could have been cited in 2012, 2013, 2014, 2015 and 2016). An article published in 2013 could have been cited in 2013, 2014, 2015 or 2016 (four years of citations). An article published in 2014 could have been cited for 3 years (2014, 2015 or 2016); and so on. In sum, we have five years of citations to articles that are 5 years old, 4 years of citations to articles that are 4 years old, 3 years of citations to articles that are 3 years old, etc. Citations to articles and proceedings published prior to 2013 are not included. Self-citations are currently included in the database.

The CrossRef citation-linking network is the source for citations. According to their website (as of September 2017) "Our citation-linking network today covers over 91 million registered content records (journal articles, book chapters, data, theses, technical reports) from thousands of scholarly and professional publishers around the globe." Citations to journals and conference proceedings in "other content items" are included in the citation count. Currently, we do not have the ability to separate citations based on source.

Books

The Academic Analytics data warehouse includes book publication data obtained from Baker & Taylor and The British Library. A book title is reported once per author/editor/translator; all published works are weighted equally (distinctions between authors, editors and translators are available in the Academic Analytics database portal). Series editors are not included in the books metrics and do not appear in the datafeeds from our providers. As with journal articles, co-authored/edited books duplication is removed as the data are aggregated to broader categories.

There are over 143,500 authors, co-authors, editors, co-editors and translators of books published in 2007-2016 (inclusive) matched to 59,000 faculty members in the 2016 database. Introductions, forewords, afterwords, and citations in and to books are not currently captured. If a book has been published in multiple printings during the ten-year period under consideration, only one instance is included. If a book has been updated with each publication, all editions/volumes are included. If a book was originally published before 2007 and was published in a different format within the observed period, it is included (i.e., a book published originally in 2001 and released as a paperback in 2014 would be included).

Books are included and weighted in the comparative data for the following broad subject areas:
- Business
- Education
- Family, Consumer and Human Sciences
- Health Professions Sciences
- Humanities
- Social and Behavioral Sciences

Though books are not weighted in all disciplines, they are matched to all scholars. These numbers will appear in Faculty Counts and in the Faculty Details downloads. Books are not weighted based on publisher, but publishers are attached to the book records and are available through the Academic Analytics database portal.

A comprehensive source for book chapter data has not been identified and thus chapter authors are not represented in the comparative data. There has been effort to warehouse chapter authors

and thus some may appear in a details feed, but it is not likely that the data will be comprehensive enough for inclusion in the near future.

Matching books to authors/editors/translators is less difficult than journal articles in the sense that there are less possible records to match. However, they are more difficult in that they do not have all the additional descriptive data such as author institutional affiliation at time of publication. Academic Analytics uses a combination of name uniqueness and book subject to level one classification to present reviewers with more probable matches. Common names involve a more intensive review.

Academic Analytics classifies books into one of the following categories:

1. Book
2. Text Book
3. Edited Volume
4. Study Guide (not counted as a book in the comparative data)
5. Journal (not counted as a book in the comparative data)
6. Proceedings (not counted as a book in the comparative data)
7. Case Book (not counted as a book in the comparative data)

The classification of books and inclusion methodology are still in development; hence, feedback is encouraged. It is possible that these categories may be revised in the future.

Grants

The 2016 database includes grants data from 13 federal agencies matched to the principal investigator at the lead institution. There are over 167,000 grants matched which total over $35 billion in annualized competitive grant funding. Funding is attached to individuals, so a grant received while at one institution is carried to the next institution as people move between universities. Grants in active status during the 2012-2016 calendars year based on the project or budget start/end dates as of June 2017 are represented in the comparative database. Grants data are obtained through a combination of online search engines and through Freedom of Information Act (FOIA) requests.

Data are presented as annualized amounts based on the awarded or estimated amount of the grant divided by the number of years derived from either the start/end dates of the project or of the budget period where transactions have been reported. So, a $300,000 grant that runs from 1/1/2012-12/31/2015 would be assigned an annualized amount of $100,000. When the full history of a grant is not known, funding is calculated based on the budget start and end dates rather than the project start and end dates.

A grant is counted as long as the duration is within our 5-year window. So, a grant that runs from 5/1/2012-5/1/2015 would be counted in the comparative data AAD2012 through AAD2019. If the end date is updated to account for extensions or transaction/budget periods we add then it would be included accordingly (e.g., an update to end date 5/1/2016 would put it into AAD2020). Annualized dollars would be the same, unless, transaction/date changes impact the dollars per year in which case there would be changes between product years. Indirect costs are included; breakdowns of direct vs indirect dollars are not available across all agencies, so they cannot be separated at this time.

For renewals and extensions (including no-cost), we update the total dollars and the start/end dates and re-calculate the dollars if the Federal Award Id remains the same. Renewals that are assigned a new ID by a federal agency are counted as a new grant. For example, NSF renewals compete with all other pending proposals and are then assigned a new Federal Award Id. Since we count grants based on the ID, we count that renewal as a new grant. No-cost extensions are

currently included as they have come through grants reporting as updates. Filtering dollars per year based on the original duration of the grant is currently not available.

Due to the dynamic nature of the online grant search systems, Academic Analytics will provide more frequent updates to the dollars and durations of the grants represented in AAD2016 as new information on transactions and extensions are available. This will impact the dollars per year variable between releases of the AAD2016 database. Fluctuation is expected to be ongoing as we work to keep the total dollars updated to a current representation of the grant. This will mean that transactions occurring in FY2016/FY2017 to grants active during the AAD2016 inclusion window will be added over the course of the AAD2016 release. In cases where the funding is consistent YOY to a project, the dollars per year will show little or no change between releases, however, it is often the case that there is an increase or decrease in funding over the lifespan of a grant which will cause changes between each release.

Cooperative agreements which are often competitive funding have been excluded from the database due to the lack of availability of data when the index was developed and government involvement in directing the research. Academic Analytics has warehoused over 53,000 cooperative agreements and matched them to faculty. Cooperative agreements will not be included by default to maintain consistency with previous releases; but, data are available in custom reports. The following table summarizes total cooperative agreement dollars loaded by agency as of December 2016. Coverage varies by year/by agency.

Agency	Total Dollars
HHS	52,457,206,362
NIH	51,729,675,573
NSF	23,585,195,493
DOE	14,862,391,681
NIH-SUB	12,811,987,564
NOAA	8,351,294,185
NASA	5,407,187,909
EPA	2,543,113,831
Army	1,448,563,445
DOED	191,987,574
Navy	99,870,499
FAA	67,721,117
USDA	2,945,729

Co-PIs have been warehoused and matched to scholars when the data are available. Co-PI's can be found on the download page as part of the detail data warehouse and displayed as part of Faculty Counts. Further, Co-PI relationships are displayed in the Collaborations tool. The following table shows total Co-PI's loaded to the warehouse summarized by start year of the grant. Note that the NIH contact PI is classified as PI; all other PI's are classified as Co-PI.

Agency	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	Total
NSF	237	7226	7995	10524	9660	7524	7553	7209	7038	7494	8329	80789
NIH	200	245	768	1996	1860	1997	2154	2108	2856	3027	3220	20431
USDA	1479	1312	2180	2488	2669	2316	1846	985	684	386	338	16683
NEH	4	4	5	7	47	47	35	34	43	37	52	315

NIH sub-grants have been loaded to the data warehouse and will begin to appear in the details download as they are matched, but are not currently included in the comparative data. The below table summarizes what has been loaded to the warehouse for NIH sub-grant.

Year	Total Dollars
2009	5,287,587,677
2010	4,603,136,943
2011	3,316,103,862
2012	3,949,179,742
2013	2,643,778,943
2014	2,895,840,208
2015	1,882,492,898
2016	897,226,126

Our classification system to standardize federal funding includes the following categories with coverage varying by institutions and year, discretionary grants being included in the default comparative data:

1. BlockGrant
2. CareerGrant
3. CenterGrant
4. CooperativeAgreement
5. Contract (No contracts have been loaded to the data warehouse as of 9/2016.)
6. DiscretionaryGrant
7. Fellowship
8. FormulaGrant
9. InteragencyAgreement
10. Other
11. TrainingGrant

Interest in custom grant inclusion can be discussed with Academic Analytics in an effort to produce useful reports

Funding agencies included in the 2016 database

National Aeronautics and Space Administration (NASA)
Department of Agriculture (USDA)*
National Endowment for the Arts (NEA)
Department of Commerce (DOC)*
Department of Defense (DOD)*
Department of Education (ED)
Environmental Protection Agency (EPA)
Department of Energy (DOE)*
Department of Health and Human Services (HHS)
    (Includes NIH)
Institute of Museum and Library Services (IMLS)
National Endowment for the Humanities (NEH)
National Science Foundation (NSF)
Department of Transportation (DOT)*

*Not all agency offices are represented in the database. For a complete list of included offices, see the detailed listing below.

Federal research funding included in the comparative data

NASA
Source: www.research.gov
Included: Research grants
Excluded: Cooperative Agreements

USDA
Source: http://cris.nifa.usda.gov and FOIA
Included: NRI competitive grants, other grants, USDA grants
Excluded: 3D grant, animal health, cooperative agreement, Evans-Allen, Hatch, McIntire-Stennis, other extension grant, RREA, SERD grant, small business grant, special grant, state, USDA contract, USDA cooperative agreement, USDA in-house

NEA
Source: http://apps.nea.gov/GrantSearch/
Included: Literature fellowships
Excluded: Funding to groups instead of individuals

DOC
Source: https://grantsonline.rdc.noaa.gov
Included: National Oceanic and Atmospheric Administration (NOAA) grants (note that the grants download available from NOAA does not distinguish PI from CoPI and thus all individuals on a grant are represented in the comparative data as a PI).
Excluded: Offices outside of NOAA

DOD
Source: FOIA and www.usaspending.gov
Included: Project grants from the Air Force Office of Scientific Research (AFOSR), Army Research Office (ARO), and Office of Naval Research (ONR), Congressionally Directed Medical Research Programs (Army-CDMRP).
The Department of Air Force did not submit updates for the 2016 database. Air Force grants data coverage remains 2009-2013.
Excluded: Other DOD offices. ARO (PI record grants data not yet obtained for FY2014-FY2016), AFOSR (PI record grants data not yet obtained for FY2014-FY2016)

ED
Source: http://ies.ed.gov and www.usaspending.gov
Included: Discretionary (awards funds on the basis of a competitive process)
Excluded: Formula (noncompetitive awards based on a predetermined formula). Due to limited data availability, some programs funded in FY2016 are not represented.

EPA
Source: FOIA and www.usaspending.gov
Included: Project grants
Excluded: Other funding types (only competitive project grants were submitted by the EPA)

DOE
Source: http://science.energy.gov
Included: Research grants
Excluded: Cooperative Agreements, Non-Office of Science grants

HHS
Source: http://exporter.nih.gov/ and http://taggs.hhs.gov
Included: Discretionary grants, research project (R), and research-related project (S) types active
Excluded: research construction programs (C), training projects (D), fellowship programs (F), resource programs (G), research career programs (K), general clinical research center program (M), research program projects and centers (P), training programs (T), cooperative agreements (U)

IMLS
Source: FOIA
Included: Conservation Assessment Prog., Museums for America, 21st Century Museum Profession, Native American Native Hawaiian, Save America's Treasures, Museum Grants African American History and Culture, Native American Library Services, Museum Assessment Program, National Leadership Grants – Museums, SPARKS, Cong. Set asides, Libraries, Cong. Set asides, Museums, 21st Century Museum Professionals Program, Librarians for the 21st Century, National Leadership Grants, Conservation Program, Native American Native Hawaiian Museum Services Program

NEH
Source: https://securegrants.neh.gov
Included: Challenge grants, Digital Humanities grants, Education programs, Preservation and Access, Public programs, Research programs
Excluded: Funding to groups instead of individuals

NSF
Source: http://www.nsf.gov/awardsearch and http://www.research.gov
Included: Standard and continuing grants
Excluded: BOA/task order, contract, contract interagency agreement, cooperative agreement, fellowship, fixed price award, interagency agreement, and personnel agreement

DOT
Source: http://www.faa.gov
Included: Federal Aviation Administration (FAA) grants
Excluded: Offices outside of FAA

The grants market share screens in the portal also include grants matched to the following agencies, however all other comparative data are restricted to federal agencies described above. Adding funding from additional non-federal agencies is of interest but is not currently active due to resource restrictions. Inclusion of the following two agencies is available via custom reporting.

ACS
Source: http://www.cancer.org
Included: None (available in grants market share screens only)
Excluded: Research and research scholar grants

AHA
Source: http://my.americanheart.org
Included: None (available in grants market share screens only)
Excluded: Research grants

Federal Agency Offices

The following lists agency offices that are included in the comparative data. A full list of federal agencies with coverage periods is available under Home > Documents

NASA: Goddard Space Flight Center, Marshall Space Flight Center, Glenn Research Center, Johnson Space Center, NASA Shared Services Center, Langley Research Center, Ames Research Center, Kennedy Space Center, STEREO Science Center, and Dryden Flight Research Center

USDA: National Institute of Food and Agriculture (NIFA), Agricultural Research Service (ARS), Economic Research Service (ERS)

DOC: National Oceanic and Atmospheric Administration (NOAA)

DOD: Air Force Office of Scientific Research (AFOSR), Army Research Office (ARO) and Office of Naval Research (ONR)

ED: Institute of Education Sciences (IES), Office of English Language Acquisition (OELA), Office of Elementary and Secondary Education (OESE), Office of Federal Student Aid (OFSA), Office of Innovation and Improvement (OII), Office of Postsecondary Education (OPE), Office of Safe and Drug-Free Schools (OSDFS), Office of Special Education and Rehabilitative Services (OSERS), and Office of Vocational and Adult Education (OVAE)

DOE: Office of Science (SC)

HHS: Substance Abuse and Mental Health Services Administration (SAMHSA), Administration for Children and Families (ACF), Centers for Disease Control (CDC), Health Resources and Services Administration (HRSA), Indian Health Service (IHS), Office of the Secretary (DHHS/OS), Centers for Medicare and Medicaid Services (CMS), Administration on Aging (AOA), Food and Drug Administration (FDA), Agency for Healthcare Research and Quality (AHRQ), National Institute on Alcohol Abuse and Alcoholism (NIAAA), National Institute on Aging (NIA), National Institute of Allergy and Infectious Diseases (NIAID), National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS), Center for Drug Evaluation and Research (CDER - FDA/NIH), National Cancer Institute (NCI), National Institute on Drug Abuse (NIDA), National Institute on Deafness and Other Communication Disorders (NIDCD), National Institute of Dental and Craniofacial Research (NIDCR), National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Biomedical Imaging and Bioengineering (NIBIB), National Institute of Environmental Health Sciences (NIEHS), National Eye Institute (NEI), National Institute of General Medical Sciences (NIGMS), National Institute of Child Health and Human Development (NICHD), National Human Genome Research Institute (NHGRI), National Heart, Lung, and Blood Institute (NHLBI), National Library of Medicine (NLM), National Center on Minority Health and Health Disparities (NCMHD), National Institute of Mental Health (NIMH), National Institute of Nursing Research (NINR), National Institute of Neurological Disorders and Stroke (NINDS), National Center for Research Resources (NCRR), John E. Fogarty International Center for Advanced Study in the Health Sciences (FIC)

NSF: Center for Precision Metrology (CMM), Center for the Build Environment (CBE), Directorate for Computer and Information Science and Engineering (CSE), Directorate for Engineering (ENG), Directorate for Geosciences (GEO), Directorate for Mathematical & Physical Sciences (MPS), Division of Antarctic Sciences (ANT), Division of Arctic Sciences (ARC), Division of Astronomical Sciences (AST), Division of Atmospheric Sciences (ATM), Division of Behavioral and Cognitive Sciences (BCS), Division of Biological Infrastructure (DBI), Division of Chemistry (CHE), Division of Computer and Network Systems (CNS), Division of Computing and Communication Foundations (CCF), Division of Design and Manufacturing Innovation (DMI), Division of Earth Sciences (EAR), Division of Electrical, Communications and Cyber Systems (ECC), Division of Elementary, Secondary and Informal Education (ESI), Division of Emerging Frontiers (EF), Division of Engineering Education and Centers (EEC), Division of Environmental Biology (DEB), Division of Graduate Education (DGE), Division of Human Resource Development (HRD), Division of Information & Intelligent Systems (IIS), Division of Information Systems (DIS), Division of Integrative Organismal Systems (IOS), Division of Materials Research (DMR), Division of Mathematical Sciences (DMS), Division of Molecular and Cellular Biosciences (MCB), Division of Ocean Sciences (OCE), Division of Physics (PHY), Division of Research, Evaluation, and Communication (REC), Division of Social and Economic Sciences (SES), Division of Undergraduate Education (DUE), Division Static Template (DRL), Experimental Program to Stimulate Competitive Research (EPSC), Industrial Innovation and Partnerships Division (IIP), Office of Cyberinfrastructure (OCI), Office of Emerging Frontiers in Research and Innovation (EFR), Office of Information Systems (OIS), Office of Integrative Activities (OIA), Office of Polar Programs (OPP), Polar Environment, Health, and Safety Office (PEH), Science Resources Statistics (SRS), Social, Behavioral & Economic Sciences (SBE)

DOT: Federal Aviation Administration (FAA)

Awards

The 2016 database includes 6,600 honorific awards from 990 governing societies with 5,100 awards from 867 societies represented in the comparative database. The difference in the number of awards warehoused vs included in the comparative data is not based on restrictions, it is based on matches. There are 1,505 awards not currently matched to any scholars represented in the AAD2016 database. Several scenarios could account for this discrepancy. The award may be defunct and all recipients are outside of our scope. The award may have appeared on a list of requested additions but does not actually apply to faculty based on our faculty inclusion methodology. Some awards that appear in the master count include foreign societies/awards where there are currently no recipients in our scope of US scholars.

The general principle for the inclusion of an award in the database is that the award must be open to all people in a (sub-) discipline or to a large subset (i.e., age, gender) at the national and/or international level. State and local awards are not currently captured by Academic Analytics. Expansion of the awards list is ongoing and largely based on feedback from client institutions.

The following awards are also represented as grants and counted as both an award and a grant:
- NIH MERIT awards
- NSF CAREER awards
- NEH awards

2016 Calculations and Definitions

Calculating Program, Broad Field and Institutional Index Scores in the AAD 2016 Database

Person Based Calculations

For the AAD 2016 database, we have continued using a person-based comparative method introduced with the AAD 2012 database. We calculate the aggregate Scholarly Research Index based on the mean person-level scores of individual faculty members. That is, the SRI score for a program is the average of the individual SRI scores of the faculty in that program; the SRI score for a university will be the average SRI score of the faculty (de-duplicated) at that university. Using a "person based" SRI score has a number of advantages. First, it allows us to expand comparisons beyond units to groups of individuals aggregated under different rubrics. We are now able to compare research activity of individuals by faculty rank, comparing faculty holding different academic ranks and comparing faculty within a single academic rank. Further, we can look at all faculty in a discipline and view their research activity compared to years since terminal degree, viewing their career progression within the discipline. In addition, we have found that this method compensates for the "portfolio effect" found in the data for many institutions. For example, our previous unit-based method of calculation for an entire institution took the average of all the Ph.D. programs at that institution, which means a program with only 5 faculty members influenced the university Scholarly Research Index as much as a program with 100 faculty members. Now, the institutional Scholarly Research Index is the result of the average SRI scores of all individual faculty members, effectively eliminating this issue.

Metrics for Person-Based Scholarly Research Index (AAD 2016 Default Metrics)

Metrics for the person-based Scholarly Research Index are:

Total Publications	Total Grant Dollars
Total Citations	Total Awards
Total Books	Total Conference Proceedings
Total Grants	

By capturing national data on all faculty research activity within a specific discipline and attaching those data to individual faculty, Academic Analytics is able to organize and rank individual faculty based on their individual research activity within their discipline. We have organized faculty into 20% bands, or quintiles, for presentation of these data. A tool in our online portal provides a picture of research activity in the discipline through a description of each of the quintiles in terms of:

Average Citations	Average Books
Average Journal Articles	Average Number of Grants
Average Awards	Average Annual Grant Dollars

For clients at the Details level of licensure, Academic Analytics provides a second view of research activity by academic rank. For this view, all discipline faculty have been separated out by academic rank. They have been organized into five quintiles according to their personal scholarly SRI scores within their rank. The first quintile is the most active; the fifth quintile is the least active. Each academic rank is identified and the total number of individuals at that rank in the discipline is given. Each quintile is described in terms of:

Average Citations	Average Books
Average Journal Articles	Average Number of Grants
Average Awards	Average Annual Grant Dollars

Unit Aggregations

The Academic Analytics database is organized according to a taxonomy of 171 disciplines. Each program or department is classified into our taxonomy with one or more disciplinary classifications. Reports are generated at four levels of aggregation; the base level is the most specific, the individual faculty. The first and second level reports provide comparative data on a broader scale. The third level is the Broad Field. Faculty names and records of research activity are de-duplicated at each level of analysis for academic units that roll up to the same broad field, so that a professor who is a member of both the *Sociology* and *Political Science* units at level one of the taxonomy appears only once in the *Social and Behavioral Sciences* broad field at level 3. Thus, the SRI score of each unit or level of aggregation; program, department, Broad Field or institution is the average of the faculty who comprise that unit.

Taxonomy level 01, 02, and 03 are comprised of both volume and per capita metrics. Volume metrics are the de-duplicated faculty counts of research activity. Because volume metrics are often a function of how many faculty are in a unit (more faculty leads to higher totals) we derive per capita metrics such as, "per faculty" and "percentage of faculty." The per capita metrics reduce the faculty size effect and allow smaller units to compare to larger units. Unit level metrics are:

1. **Books metrics**
    Total books published
    Number of faculty with a book publication
        Percentage of faculty who have authored a book
    Books published per faculty member

2. **Journal publication metrics**
    Total journal articles
    Number of faculty with a journal article
        Percentage of faculty who have authored a journal article
    Journal articles published per faculty member
    Articles per author

3. **Conference proceedings metrics**
    Total conference proceedings
    Number of faculty with a conference proceeding
        Percentage of faculty who have a conference proceeding
    Conference proceedings per faculty member
    Conference proceeding per author

4. **Citations metrics**
    Total citations
    Number of faculty with a citation
        Percentage of faculty with a citation
    Citations per faculty member
    Citations per journal article
    Citations per author

5. **Federal funding metrics**
    Total federal grants
    Total federal grant dollars
    Number of faculty who have a federal grant
        Percentage of faculty who have won new and competitive federal research funding
    Federal grants per faculty member
    Federal grant dollars won per faculty member
    Dollars per grant

6. **Honorific awards metrics**
    Total honorific awards
    Number of faculty who have won an honorific award
       Percentage of faculty who have won an honorific award
    Honorific awards per faculty member

The following descriptive statistics are provided for each metric and are relevant to the comparison group chosen.

1. Rank
2. Z-score (metric value - μ of the metric value) / σ of the metric value
3. Percentile rank

University Level Aggregation

The institutional SRI score is the average of all individual faculty at the institution. Faculty are de-duplicated at each level of analysis for academic units that roll up to the same broad field.

Metrics at the university level are provided as de-duplicated volume totals and not as per capita metrics. Due to the diversity of metrics that are considered relevant or important among disciplines, the faculty denominator in the calculation becomes less meaningful when calculating per capita metrics.

Unit Based Calculation and Metrics

Historically, the Academic Analytics databases have contained two basic types of metrics: "Full data metrics" and "Per capita metrics." Full data metrics represent raw data on faculty research activity. This includes total counts of journal publications, citations, number of grants and grant dollars, awards, book publications, the number of conference proceedings published, and the number of faculty who have produced or received each of the above. For AAD 2011 and prior databases, these full data metrics were not used in their raw form to understand the research activity of a given program or department relative to other programs or departments, because they did not take into account the size of the program in terms of total number of faculty. Larger programs typically have more total publications, awards, books, and other forms of research activity simply because they have more people contributing to overall scholarship. Per capita metrics solved this problem by converting raw metrics into per capita and percentage form. For example, "total journal article publications" for a given program became "journal article publications per faculty member."

In AAD 2011 and prior databases, the z-scores for each metric serve to facilitate comparisons between unlike data types (e.g., journal article publications versus honorific awards, or "Department of Chemistry" versus "Department of History"). A program's z-score for a given metric represents the research activity of this program relative to the average program in the same discipline. In this calculation method, z-scores are calculated by subtracting the average score of the metric for the entire discipline from the program's score, and then dividing this difference by the standard deviation for that metric in the discipline. The standard deviation is simply a measure of the overall variation of the scores in the discipline, and this division is a way to put all scores on the same scale. For example, if the average number of journal publications per faculty member in discipline "A" is 6.0, and the standard deviation for this discipline is 2.0, then a program with 8.0 publications per faculty has a z-score of 1.0; i.e., $[(8.0 – 6.0)/2.0] = 1.0$, or "one standard deviation about the mean." For more information on z-scores, please visit the following URL: http://en.wikipedia.org/wiki/Standard_score.

As you can see from the formula above, programs with scores above the national average have positive z-scores, and programs with scores below the average have negative z-scores. The absolute value of the z-score represents how far the program is from the comparative group average in either the positive or negative direction. Thus, a score of 2.0 represents a program that is very research active relative to the average program in its discipline (two standard deviations above the comparative group average), while a z-score of -2.0 represents a fairly inactive program (two standard deviations below the comparative group average). Z-scores are a convenient way to represent the relative standing of a program within its discipline, compared to standard or user-specified peer groups.

Z-scores are also useful for comparing the performance of a program across metrics. If we wanted to compare the research activity of a given program in terms of grant dollars to its activity in terms of journal publications, we would not want to look at grant dollars per faculty member and publications per faculty member, because these two metrics are on different scales (i.e., dollars and articles). For example, if a given program has $10,000 per faculty member in grant funding and 6.0 articles per faculty member in terms of journal publications, we are unsure about which score represents greater activity due to different scaling. We can, however, directly compare the program's z-scores. If the z-score for grant dollars is 1.0, and the z-score for publications is 2.0, then we know that the program is relatively more active in publishing journal articles than in obtaining grant money when compared to the rest of the comparative group.

The z-score representation is also advantageous because it allows for the construction of a single summary measure of the overall research activity of a program, relative to the rest of its discipline, by combining its scores on many metrics. Since z-scores are all on the same scale, we can combine a program's scores on all fifteen metrics to create the summary rank or percentile in the discipline.

Definitions of Metrics

In the following explanation of metrics 'μ' refers to the mean of the population while 'σ' refers to the standard deviation of the population.

Books Metrics
Percentage of Faculty with a Book Publication – pfacbp
> hasbook / numfac

Z-Score of Percentage of Faculty with a Book Publication – b1
> (pfacbp - μ of pfacbp) / σ of pfacbp

Rank on Percentage of Faculty with a Book Publication – Rb1
Book Publications Per Faculty - bppfac
> totbooks / numfac

Z-Score of Book Publications Per Faculty – b2
> (bppfac - μ of bppfac) / σ of bppfac

Rank on Book Publications per Faculty – Rb2

Publication Metrics

Percentage of Faculty with a Journal Publication – pfacjp
    hasjpub / numfac

Z-Score of Percentage of Faculty with a Journal Publication – p1
    (pfacjp - μ of pfacjp) / σ of bppfac

Rank on Percentage of Faculty with a Journal Publication – Rp1

Journal Publications per Faculty Member – jppfac
    totpubs / numfac

Z-Score of Journal Publications per Faculty Member – p2
    (jppfac - μ of jppfac) / σ of jppfac

Rank on Journal Publications per Faculty Member – Rp2

Conference Proceedings Metrics

Percentage of Faculty with a Conference Proceeding – percentFacultyConfProc
    HasConfProc / numfac

Z-Score of Percentage of Faculty with a Conference Proceedings –
zScore_percentFacultyConfProc
    (zScore_percentFacultyConfProc - μ of zScore_percentFacultyConfProc) / σ of
    percentFacultyConfProc

Rank on Percentage of Faculty with a Conference Proceedings –
rank_percentFacultyConfProc

Conference Proceedings per Faculty Member – perFacultyConfProc
    TotConfProc / numfac

Z-Score of Conference Proceedings per Faculty Member – zScore_perFacultyConfProc
    (zScore_perFacultyConfProc - μ of zScore_perFacultyConfProc) / σ of
zScore_perFacultyConfProc

Rank on Conference Proceedings per Faculty Member – rank_perFacultyConfProc

Citations Metrics

Percentage of Faculty with a Citation – pfacc
    hascit / numfac

Z-Score of Percentage of Faculty with a Citation – p3
    (pfacc - μ of pfacc) / σ of pfacc

Rank on Percentage of Faculty with a Citation – Rp3

Citations per Faculty Member – cpfac
    totcits / numfac

Z-Score of Citations per Faculty Member – p4
     (cpfac – μ of cpfac) / σ of cpfac

Rank on Citations per Faculty Member – Rp4

Citations per Publication – cpp
     totcits / totpubs

Z-Score of Citations per Publication – p5
     (cpp - μ of cpp) / σ of cpp

Rank on Citations per Publication – Rp5

Grant Metrics
Percentage of Faculty with a Grant – pfacgrnt
     hasgrnt / numfac

Z-Score of Percentage of Faculty with a Grant – g1
     (pfacgrnt - μ of pfacgrnt) / σ of pfacgrnt

Rank on Percentage of Faculty with a Grant – Rg1

Grants per Faculty Member – gpfac
     totgno / numfac

Z-Score of Grants per Faculty Member – g2
     (gpfac – μ of gpfac) / σ of gpfac

Rank on Grants per Faculty Member – Rg2

Grant Dollars Per Faculty Member – gdpfac
     totgdol / numfac

Z-Score of Grant Dollars Per Faculty Member – g3
     (gdpfac - μ of gdpfac) / σ of gdpfac

Rank on Grant Dollars Per Faculty Member – Rg3

Dollars per Grant – dolpgrnt
     totgdol / totgno

Z-Score of Dollars per Grant – g4
     (dolpgrnt - μ of dolpgrnt) / σ of dolpgrnt

Rank on Dollars per grant – Rg4

Award Metrics

Percentage of Faculty with an Award – pfacawd
hasawd / numfac

Z-Score of Percentage of Faculty with an Award – a2
(pfacawd- μ of pfacawd) / σ of pfacawd

Rank on Percentage of Faculty with an Award – Ra2

Awards per Faculty Member – awdpfac
Totawds / numfac

Z-Score of Awards per Faculty Member – a1
(awdpfac – μ of awdpfac) / σ of awdpfac

Rank on Awards per Faculty Member – Ra1

Weight Metrics

Weight of Awards Metrics in this Discipline – awdwt
Weight of Grants Metrics in this Discipline – grntwt
Weight of Books Metrics in this Discipline – bookswt
Weight of Citations Metrics in this Discipline – citswt
Weight of Journal Publication Metrics in this Discipline – pubswt
Weight of Conference Proceedings Metrics in this Discipline – confprocwt