Database Assignment Data Dictionary

Table of Contents

[Table vendor 1](#_Toc469325307)

[Table ugrant 2](#_Toc469325308)

[Table machine\_learning 2](#_Toc469325309)

[Table nsf\_award 3](#_Toc469325310)

[Table text\_analysis 4](#_Toc469325311)

[Table uc\_pay\_2011 4](#_Toc469325312)

# Table vendor

Table Description: University internal transaction data that describes the expenditures on a federal grants. Note that this information does not tell you what goods/services were bought, but tells you who was paid and how much was spent.

|  |  |  |
| --- | --- | --- |
| Field | Data Type | Description |
| index | Bigint | Unique integer primary key for each row in the table. |
| periodstartdate | date | Beginning date of the transaction |
| award\_id | Int | Public award ID |
| institutionid | varchar(200) | Business/vendor ID – unique to each vendor |
| paymentamount | double | Expenditure amount |
| university | varchar(45) | University acronym for each record of data |
| cfda | varchar(10) | Federal agency ID – maps to federal agency and federal agency programs |
| fipscode | varchar(6) | State county codes that map directly to the business/vendor zipcode |
| statecode | int(11) | State codes that map directly to the business/vendor zipcode |
| countycode | int(11) | County codes that map directly to the business/vendor zipcode |
| agency\_abbrev | varchar(30) | Abbreviation for the parent federal agency (except for NIH – NIH is used instead of HHS) |
| agency\_text | varchar(70) | Description of the agency – corresponds to agency field |
| sub\_agency\_text | varchar(100) | Description of the sub agency field (for example: National Institutes of Health is a sub agency of Health and Human Services) |

# Table ugrant

Table Description: University public grant information from NSF and NIH websites that describe the funded grant projects

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Field | Data Type | | Description | |
| index | | Bigint | | Unique integer primary key for each row in the table. |
| award\_id | | varchar(50) | | Public award ID |
| topic\_id | | int(11) | | LDA topic ID – corresponds to topic\_text |
| proportion | | double | | Proportion of best bit of a topic to a grant (decimal) |
| agency | | varchar(5) | | Federal agency source of the data |
| topic\_text | | text | | LDA topic text – describes the topic |

# Table machine\_learning

Table Description: table with information combined from vendor and ugrant, for use in machine learning models.

|  |  |  |
| --- | --- | --- |
| Field | Data Type | Description |
| index | Bigint | Unique integer primary key for each row in the table. |
| application\_id | varchar(20) | Unique identifier for each grant - Public application ID – only applicable to NIH grants |
| cfda\_code | varchar(10) | CFDA contains detailed program descriptions for 2,292 Federal assistance programs: <https://www.cfda.gov/> |
| year | int | Year in which grant was awarded |
| activity | Varchar | A 3-character code identifying the grant, contract, or intramural activity through which a project is supported. Here is a list of activity codes: http://grants.nih.gov/grants/funding/ac\_search\_results.htm |
| administering\_id | varchar(50) | Administering Institute or Center - A two-character code to designate the agency, NIH Institute, or Center administering the grant. See [definitions]: . |
| org\_name | Varchar | The name of the educational institution, research organization, business, or government agency receiving funding for the grant, contract, cooperative agreement, or intramural project. |
| org\_dept | varchar | The departmental affiliation of the contact principal investigator for a project, using a standardized categorization of departments. Names are available only for medical school departments. |
| topic\_id | int(11) | LDA topic ID – corresponds to topic\_text - Using text analysis techniques, a topic\_id was assigned to each grant. This topic\_id is a key for that topic. You can see what the topic contains by looking in the topiclda\_text table in umetricsgrants database. |
| study\_section | varchar(5) | A designator of the legislatively-mandated panel of subject matter experts that reviewed the research grant application for scientific and technical merit. |
| total\_cost | Int | Total project funding from all NIH Institute and Centers for a given fiscal year. |
| ed\_inst\_type | Varchar | Generic name for the grouping of components across an institution who has applied for or receives NIH funding. |
| arra\_funded | Boolean | “Y” indicates a project supported by funds appropriated through the American Recovery and Reinvestment Act of 2009. |

# Table nsf\_award

Table Description: TK

|  |  |  |
| --- | --- | --- |
| Field | Data Type | Description |
| index | Bigint | Unique integer primary key for each row in the table. |
| awardid | bigint | Public award ID |
| firstname | Text | First name |
| lastname | text | Last name |
| startdate | Text | Start date of award (format YYYY-MM-DD). |
| enddate | text | End date of award (format YYYY-MM-DD). |
| awardtitle | text | Title of award |
| awardeffectivedate | Text | Effective date of award (format YYYY-MM-DD) |
| awardexpirationdate | text | Award’s expiration date (format YYYY-MM-DD) |
| name | text | Name of institution receiving award. |
| cityname | text | Name of city in which institution is located. |
| zipcode | text | Zip code of address for grantee at institution. |
| phonenumber | Int | 10-digit phone number for institution. |
| streetaddress | text | Street address for grantee at institution. |
| countryname | Text | Country in which institution is located. |
| statename | Text | State in which institution is located. |
| statecode | Text | Code for state in which institution is located. |

# Table text\_analysis

Table Description: table with information combined from vendor and ugrant, for use in machine learning models.

|  |  |  |
| --- | --- | --- |
| Field | Data Type | Description |
| index | Bigint | Unique integer primary key for each row in the table. |
| application\_id | varchar(20) | Unique identifier for each grant - Public application ID – only applicable to NIH grants |
| abstract\_text | Text | Full text of abstract from application. |

# Table uc\_pay\_2011

Table Description: TK

|  |  |  |
| --- | --- | --- |
| Field | Data Type | Description |
| index | Bigint | Unique integer primary key for each row in the table. |
| id | bigint | ? |
| year | bigint | Year of pay record. |
| campus | Text | Campus where employee works. |
| name | Text | Employee name |
| title | Text | Employee’s title. |
| gross | double | Gross pay |
| base | Double | Base pay |
| overtime | double | Overtime pay |
| extra | bigint | ? |
| exclude | Bigint | ? |