The Inside Track

The data are in two tab-separated text files:

pnas_papers.txt

Fields:

filename unique filename for each abstract
memb_edit scientist who Contributed, Communicated or Edited (for direct submissions) each paper
first_name first name from memb_edit
last_name last name from memb_edit
track publication track: Contributed, Communicated or Edited (direct submissions)
prearr 1 = paper had a prearranged editor; 0 = not
oa 1 = paper was open access; 0 = not
inaug 1 = paper was member's inaugural article; 0 = not
confl 1 = conflict of interest declared; 0 = no conflict of interest declared
confl_state conflict of interest statement
rec_date date paper received/sent for review
acc_date date paper approved for publication/Contributed/Communicated
online_date date of online publication
print_date date of print publication
authors list of authors
auth_count number of authors
title paper title
section* academy section
doi digital object identifier
pmid pubmed id
url url for abstract

* as given in metadata for paper. For the citation analysis, this field was edited to correspond to current academy sections: **Biophysics** is now **Biophysics and Computational Biology**; **Psychology** is now **Psychological and Cognitive Sciences**; **Geology** and **Geophysics** have been merged into **Earth Atmospheric and Planetary Sciences**.

pnas_cites.txt

Fields:

filename as above
cites_url url with CrossRef citations
cr_cites count of the citations at this url
timestamp time at which cr_cites was scraped from web