Statistics on Authorship

# Statistics Implemented by Tim

I indentify all the papers published by one individual, the *primary author*, in a certain time period. This is 1999 or earlier, 2000-2003, 2004-2006, 2007-2009, and 2010 and later (defined by yearBoundary in ProcessSinglePublicationCSVList). I do not output results for the earliest or latest periods.

With this set of papers I further subdivide them into 4 types which are combinations of the following properties: *alphabetical* or *non-alphabetical* author order, *solo papers* (single author) or *team papers* (more than one author). Note that there are non non-alphabetical solo papers as I treat all single author papers as being alphabetical.

For all the papers of a given type I count how many first author, final author, second author and Penultimate authorships the primary author had. Note that I assign papers to these subcategaoriastions in teh order given. Thus if the primary author is the last author of two, it is counted as a last authorship not a second authorship.

I can also combine these sets and counts. For instance I produce results for all team papers, both alphabetically order and non-alphabetically ordered.

For each set of papers (published in one period and of other specified characteristics) I produce the following information. First the actual counts of first authorships, final authorships, second authorships, penultimate authorships, other positions, total number of papers. The last figure should of course equal the sum of the previous ones. I also produce a fractional number which are these counts divided by the total number of papers of all types published in that period.

I also provide the mean number of authors on all the papers in a period (so minimum is one if defined), and the mean team size, that is the mean number of authors on team papers (so this cannot be less than two if defined). I can provide other statistics on these distributions e.g. the median and quartile numbers.

Statistics requested by Pietro

1. number of papers written in teams (more than one author) [types 2,3]
2. fraction of papers written in teams (more than one author) [total number of types 2 & 3 divided by total number of all papers used]
3. number of solo-authored papers [types 0,1]
4. fraction of solo-authored papers [total number of types 0 and 1 divided by total number of all papers used]
5. mean team size (= average number of co-authors per paper across all papers with more than one author) [number of types 2 & 3 divided by total number of all papers used]
6. number of papers with more than one author in which co-authors are listed alphabetically [type 3]
7. fraction of papers with more than one author in which co-authors are listed alphabetically [number of type 3 divided by total number of all papers used]
8. number of papers with more than one author and in which co-authors are not listed alphabetically [type 2]
9. fraction of papers with more than one author and in which co-authors are not listed alphabetically [number of type 2 divided by total number of all papers used]
10. number of (multi-authored) papers in which the author is listed as first author
11. fraction of (multi-authored) papers in which the author is listed as first author
12. number of (multi-authored) papers in which co-authors are not listed alphabetically and the author is listed as first author
13. fraction of (multi-authored) papers in which co-authors are not listed alphabetically and the author is listed as first author
14. number of papers with three co-authors in which the author is listed as second author
15. fraction of papers with three co-authors in which the author is listed as second author
16. number of papers with three co-authors not listed alphabetically in which the author is listed as second author
17. fraction of papers with three co-authors not listed alphabetically in which the author is listed as second author
18. number of papers with more than three co-authors in which the author is listed as penultimate author
19. fraction of papers with more than three co-authors in which the author is listed as penultimate author
20. number of papers with more than three co-authors not listed alphabetically in which the author is listed as penultimate
21. author fraction of papers with more than three co-authors not listed alphabetically in which the author is listed as penultimate author
22. number of (multi-authored) papers in which the author is listed as last author
23. fraction of (multi-authored) papers in which the author is listed as last author
24. number of (multi-authored) papers in which co-authors are not listed alphabetically and the author is listed as last author
25. fraction of (multi-authored) papers in which co-authors are not listed alphabetically and the author is listed as last author