Project-2-2-Copy1

December 7, 2018

```
In [1]: import pandas as pd
        import numpy as np
        import seaborn as sns
        import csv
        import json
        import collections
        import scholarly
        from itertools import product
        import nltk
        import spacy
        import matplotlib.pyplot as plt
        nlp = spacy.load('en_core_web_md')
```

1 Parse our data into CSV file

```
In [2]: def generate_citation_count(citation_dict, df):
            citation_list = zerolistmaker(df.shape[0])
            for key in citation_dict:
                citation_list[int(key)] = citation_dict[key]
            return citation_list
```

```
In [3]: def zerolistmaker(n):
            listofzeros = [0] * n
            return listofzeros
```

```
In [4]: def parse_data_into_dataframe(filename):
            row = []
            df = pd.DataFrame()
            refCount = 0
            title = None
            authors = None
            year = None
            publication_venue = None
            index = None
            abstract = None

            citation_dict = {}
```

```python
with open(filename) as file:
    content = {}
    for line in file:
        if(len(line) > 2):
            #* --- paperTitle
            #@ --- Authors
            #t ---- Year
            #c  --- publication venue
            #index 00---- index id of this paper
            #% ---- the id of references of this paper (there are multiple lines,
            #! --- Abstract
            value = line[2:-1]
            if(line[1] =='*'):
                title = value
            elif(line[1] =='@'):
                authors = value
            elif(line[1] == 't'):
                year = value
            elif(line[1] =='c'):
                publication_venue = value
            elif(line[1] =='i'):
                index = value[4:]
            elif(line[1] == '%'):
                if(value in citation_dict):
                    citation_dict[value] = citation_dict[value] + 1
                else:
                    citation_dict[value] = 1
            elif(line[1] == '!'):
                abstract = value
        if(line == '\n'):
            content = {'paperTitle': title, 'Authors' : authors, 'Year': year, 'pul
            row.append(content)
            title = None
            authors = None
            year = None
            publication_venue = None
            index = None
            abstract = None

df = pd.DataFrame(row)

df['citation_count'] = generate_citation_count(citation_dict, df)
return df
```

2 Concat our data from multiple data sets and write it into a CSV file.

```python
In [5]: def concat_data():
            df_1 = parse_data_into_dataframe('outputacm.txt')
            df_2 = parse_data_into_dataframe('citation-network2.txt')
            df_1 = df_1.drop(['index id'], axis=1)
            df_2 = df_2.drop(['index id'], axis=1)
            df_m = pd.concat([df_1, df_2], ignore_index=True)
            df_3 = parse_data_into_dataframe('V3.txt')
            df_3 = df_3.drop(['index id'], axis=1)
            df_m1 = pd.concat([df_m, df_3], ignore_index=True)
            return df_m1

        train_df = parse_data_into_dataframe('citation-network2.txt')
        train_df
```

```
Out[5]:                                            Abstract  \
        0                                             None
        1       Scholars continue to find that political parti...
        2                                             None
        3       In this paper, we explore the feasibility of u...
        4       Distributed filesystems are a typical solution...
        5       No question times are tough for many working e...
        6       The objective of the Collaboration Management ...
        7                                             None
        8       Instance based learning and clustering are pop...
        9       Dynamic memory management has been an importan...
        10                                            None
        11      In this paper we consider the problem of resyn...
        12                                            None
        13      An approach to solve constrained minimization ...
        14      The Surface Science Spectra journal is a spect...
        15      Up to 1995, all-star selection in the Ontario-...
        16      Traditionally, voice communication over the lo...
        17                                            None
        18      Let G = (V, E) be a graph and r : V  Z+. An r...
        19                                            None
        20                                            None
        21                                            None
        22      This paper describes the concept of GUI applic...
        23                                            None
        24      Modeling and simulation have taken a preponder...
        25      Given a rational parametrization of an algebra...
        26      We are building a biomedical information resou...
        27                                            None
        28                                            None
        29                                            None
        ...                                            ...
```

```
1397210  In this paper, the problem of bounded real lem...
1397211  Warranty service is getting important since it...
1397212  Web service composition is a very important wa...
1397213  The Internet is a complex structure arising fr...
1397214  Parallel symbolic computation has attracted co...
1397215  We give two applications of Nisan-Wigderson-ty...
1397216  Wandering the halls of a security conference i...
1397217  Supervised Machine Learning systems can induce...
1397218  This paper presents a novel approach to combin...
1397219  The ultimate success of microcomputer-based le...
1397220                                              None
1397221  In this paper we present persistent grid mappi...
1397222                                              None
1397223  Dear KV, I know you did a previous article whe...
1397224  Dynamic protocol recovery tries to recover a c...
1397225  Providing video streaming service over the Int...
1397226  A problem of stabilization of stochastic oscil...
1397227                                              None
1397228  Some computer architectures are inherently mor...
1397229  The mathematics of tangles has been applied to...
1397230  Keyword search in XML documents based on the n...
1397231  User-privacy in existing identity management s...
1397232  The EvoSpaces reverse-engineering tool represe...
1397233  Orthogonal moments such as Zernike moments and...
1397234  The new concept and method of imposing impreci...
1397235  The aim of this paper is to recognize upper-ca...
1397236  The currently observed exponentially increasin...
1397237  Motivation: Stable isotope labeling of small-m...
1397238  Shaped cathodes enhance key magnetron performa...
1397239  In this paper, we aim to identify what kinds o...
```

	Authors	Year	index id \
0	E. S. Cho,C. J. Kim,S. D. Kim,S. Y. Rhew	1998	0
1	Lori M. Weber,Alysha Loumakis,James Bergman	2003	1
2		2002	2
3	Choong-Gyoo Lim	2002	3
4	Jose Maria Perez,Felix Garcia,Jesus Carretero,...	2003	4
5	Jean Kumagai	2003	5
6	Marek Rusinkiewicz,Dimitrios Georgakopoulos	1999	6
7	Barton C. Massey,Evan Tick	1995	7
8	Jan Ramon	2002	8
9	Therapon Skotiniotis,Ji-en Morris Chang	2002	9
10	V. Martin,K. Schwan	1998	10
11	Peter W. Moo,Xiaolin Wu	1999	11
12	T. Aoki,T. Katayama	1998	12
13	Johannes Schropp	2003	13
14	Stephen W. Gaarenstroom	2003	14
15	Hasan Pirkul,David A. Schilling,W. J. Hurley	1998	15

```
16                Christo Frank,B. S. Manoj,C. Siva Ram Murthy  2002        16
17           Klaus Buchenrieder,Alexander Sedlmeier,Ch. Veith  1993        17
18                             Bill Jackson,Tibor Jordán  2003        18
19                                   Alex Alves Freitas  1998        19
20           Volker Krüger,Udo Mahlmeister,Gerald Sommer  1998        20
21                                   Paul Keith Baker  2002        21
22                             Andriy Panteleymonov  2002        22
23                                   James A. Larson  1984        23
24           Paul Bisgambiglia,Marielle Delhom,Jean-Françoi...  2002        24
25                                       Josef Schicho  2002        25
26                         L. Rodney Long,George R. Thoma  2000        26
27           Nicolas Anciaux,Christophe Bobineau,Luc Bougan...  2001        27
28           Franck Ravat,Marianne De Michiel,Gilles Zurfluh  1997        28
29                                                         1999        29
...                                                      ...       ...        ...
1397210              Gaomin Zhang,Yuanqing Xia,Peng Shi  2008   1397210
1397211   SangHyun Lee,SeongChae Seo,SoonJa Yeom,Kyungil...  2007   1397211
1397212      Gexin Li,Shuiguang Deng,Haijiang Xia,Chuan Lin  2007   1397212
1397213           Srinivasan Seetharaman,Mostafa Ammar  2006   1397213
1397214              Peter Nuth,Robert Halstead Jr.  1989   1397214
1397215           Boaz Barak,Shien Jin Ong,Salil Vadhan  2007   1397215
1397216                             Lynn Greiner  2008   1397216
1397217                             Alan Keller Gomes  2007   1397217
1397218              Chun-Jung Chen,Tien-Chi Chen  2007   1397218
1397219                       R. M. Aiken,L. Braun  1980   1397219
1397220                                              2004   1397220
1397221   Yotam Livny,Neta Sokolovsky,Tal Grinshpoun,Jih...  2008   1397221
1397222                                              2007   1397222
1397223                       George Neville-Neil  2007   1397223
1397224              Jochen Quante,Rainer Koschke  2007   1397224
1397225      Yang Guo,Kyoungwon Suh,Jim Kurose,Don Towsley  2008   1397225
1397226                             L. B. Ryashko  2007   1397226
1397227                                              2008   1397227
1397228              W. B. Dietz,L. Szewerenko  1979   1397228
1397229                             Isabel K. Darcy  2008   1397229
1397230           Yu Xu,Yannis Papakonstantinou  2007   1397230
1397231      Steven Gevers,Verslype Verslype,Bart De Decker  2007   1397231
1397232           Sazzadul Alam,Philippe Dugerdil  2007   1397232
1397233   V. Subbiah Bharathi,M. A. Leo Vijilious,L. Gan...  2007   1397233
1397234                             Hsien-Chung Wu  2007   1397234
1397235   Sarbajit Pal,Jhimli Mitra,Soumya Ghose,Paromit...  2007   1397235
1397236      Andreas Dietrich,Enrico Gobbetti,Sung-Eui Yoon  2007   1397236
1397237   Fangping Mu,Robert F. Williams,Clifford J. Unk...  2007   1397237
1397238   Timothy P. Fleming,Peter J. Mardahl,Lester Bow...  2007   1397238
1397239      Marianne Kukko,Nina Helander,Pasi Virtanen  2008   1397239

                                       paperTitle  \
0        Static and Dynamic Metrics for Effective Objec...
```

```
1         Who participates and why?: an analysis of citi...
2                                       Call for papers
3         Universal parametrization in constructing smoo...
4         Data Allocation and Load Balancing for Heterog...
5                  Employment opinion: reversal of fortune
6         From Coordination of Workflow and Group Activi...
7         Modes of Comprehension: Mode Analysis of Array...
8         Thesis: clustering and instance based learning...
9         Estimating internal memory fragmentation for J...
10        ILI: An Adaptive Infrastructure for Dynamic In...
11        Resynchronization Properties of Arithmetic Coding
12        Unification and Consistency Verification of Ob...
13        One- and multistep discretizations of index 2 ...
14        Surface Science Spectra: A Hybrid Journal-Data...
15        An Efficient, Objective Technique for Selectin...
16        Throughput Enhanced Wireless in Local Loop (TW...
17                    HW/SW Co-Design with PRAMs Using CoDES
18                        Non-separable detachments of graphs
19             On Objective Measures of Rule Surprisingness
20                    Attentive Face Detection and Recognition
21        Scientific papers: from web sites to e-communi...
22        Interoperable Thin Client Separation from GUI ...
23                              The Forms Pattern Language
24        An efficient and evolutionary hierarchical mod...
25               Simplification of surface parametrizations
26        Use of Shape Models to Search Digitized Spine ...
27                       PicoDBMS: Validation and Experience
28        Distributed Object Oriented Databases: An Allo...
29                                       External Referees
...                                                     ...
1397210   Technical communique: New bounded real lemma f...
1397211   A Study on Warning/Detection Degree of Warrant...
1397212   Automatic Service Composition Based on Process...
1397213   Characterizing and Mitigating Inter-domain Pol...
1397214   A study of LISP on a multiprocessor preliminar...
1397215                      Derandomization in Cryptography
1397216             Hacking your network's weakest link: you
1397217   Small Disjuncts Grouping by Rule Coverage and ...
1397218   Power System Stabilizer for Multi-Machine Usin...
1397219     Into the 80's with Microcomputer-Based Learning
1397220                                             AILS '04
1397221   A GPU persistent grid mapping for terrain rend...
1397222   Proceedings of the Ninth International Symposi...
1397223                                   The Next Big Thing
1397224                             Dynamic Protocol Recovery
1397225   DirectStream: A directory-based peer-to-peer v...
1397226   Stabilization of stochastically perturbed nonl...
1397227   Call for Papers: Mining Multimedia Streams in ...
```

```
1397228   Architectural Efficiency Measures: An Overview...
1397229        Modeling protein-DNA complexes with tangles
1397230     Efficient LCA based keyword search in xml data
1397231   Enhancing privacy in identity management systems
1397232   EvoSpaces Visualization Tool: Exploring Softwa...
1397233   Orthogonal Moments Based Texture Analysis of C...
1397234   Fuzzy linear regression model based on fuzzy s...
1397235   A Projection Based Statistical Approach for Ha...
1397236      Massive-Model Rendering Techniques: A Tutorial
1397237           Carbon-fate maps for metabolic reactions
1397238   Virtual Prototyping of Novel Cathode Designs f...
1397239   Knowledge Management in Renewing Software Deve...
```

	publication venue	citation_count
0	Proceedings of the Fifth Asia Pacific Software...	3
1	Social Science Computer Review	9
2	Sys Admin	0
3	Computer Aided Geometric Design	1
4	Proceedings of the 3st International Symposium...	0
5	IEEE Spectrum	0
6	Proceedings of the 1999 International Symposiu...	2
7	Proceedings of the 7th International Symposium...	0
8	AI Communications	0
9	Journal of Systems and Software	0
10	Proceedings of the International Conference on...	2
11	Proceedings of the Conference on Data Compression	2
12	Proceedings of the Fifth Asia Pacific Software...	0
13	Journal of Computational and Applied Mathematics	1
14	Computing in Science and Engineering	0
15	Interfaces	0
16	Proceedings of the 27th Annual IEEE Conference...	0
17	Proceedings of the 11th IFIP WG10.2 Internatio...	8
18	Journal of Combinatorial Theory Series B	2
19	Proceedings of the Second European Symposium o...	14
20	Mustererkennung 1998, 20. DAGM-Symposium	0
21	Technology and Health Care	0
22	Proceedings of the 6th European Conference on ...	0
23	Proceedings of the First International Confere...	1
24	Systems Analysis Modelling Simulation	0
25	Proceedings of the 2002 international symposiu...	2
26	CBMS	1
27	Very Large Data Bases	3
28	Proceedings of the 8th International Conferen...	0
29	Proceedings of the Sixth International Confere...	0
...	...	...
1397210	Automatica (Journal of IFAC)	1
1397211	ALPIT	0
1397212	Proceedings of the Third International Confere...	0

```
1397213  Proceedings of the Proceedings of the 2006 IEE...          2
1397214                          ACM SIGPLAN Lisp Pointers          0
1397215                             SIAM Journal on Computing        0
1397216                                            netWorker        0
1397217  Proceedings of the Seventh International Confe...          0
1397218  Proceedings of the Second International Confer...          0
1397219                                              Computer        0
1397220  Journal of Intelligent &amp; Fuzzy Systems: Ap...          0
1397221  The Visual Computer: International Journal of ...          0
1397222                                                SYNASC        0
1397223                                                 Queue        0
1397224  Proceedings of the 14th Working Conference on ...          1
1397225                              Computer Communications        1
1397226                           Automation and Remote Control     0
1397227  IEEE Transactions on Knowledge and Data Engine...          0
1397228                                              Computer        0
1397229       Computers &amp; Mathematics with Applications        0
1397230  Conference on Information and Knowledge Manage...          1
1397231  Proceedings of the 2007 ACM workshop on Privac...          1
1397232  Proceedings of the 14th Working Conference on ...          1
1397233                                                ICCIMA        0
1397234  Soft Computing - A Fusion of Foundations, Meth...          0
1397235                                                ICCIMA        0
1397236            IEEE Computer Graphics and Applications        1
1397237                                         Bioinformatics       0
1397238              Computing in Science and Engineering        0
1397239  Proceedings of the Proceedings of the 41st Ann...          0

[1397240 rows x 7 columns]
```

In [6]: `train_df.to_csv('train_new.csv', index=False)`

In [7]: `columns = ['Abstract', 'Authors', 'Year', 'index id', 'paperTitle', 'publication venue`

In [8]: `train_df = pd.read_csv('train_new.csv', nrows=10000)`

In [9]: `train_df`

```
Out[9]:                                               Abstract  \
        0                                                  NaN
        1      Scholars continue to find that political parti...
        2                                                  NaN
        3      In this paper, we explore the feasibility of u...
        4      Distributed filesystems are a typical solution...
        5      No question times are tough for many working e...
        6      The objective of the Collaboration Management ...
        7                                                  NaN
        8      Instance based learning and clustering are pop...
        9      Dynamic memory management has been an importan...
```

```
10                                          NaN
11    In this paper we consider the problem of resyn...
12                                          NaN
13    An approach to solve constrained minimization ...
14    The Surface Science Spectra journal is a spect...
15    Up to 1995, all-star selection in the Ontario-...
16    Traditionally, voice communication over the lo...
17                                          NaN
18    Let G = (V, E) be a graph and r : V  Z+. An r...
19                                          NaN
20                                          NaN
21                                          NaN
22    This paper describes the concept of GUI applic...
23                                          NaN
24    Modeling and simulation have taken a preponder...
25    Given a rational parametrization of an algebra...
26    We are building a biomedical information resou...
27                                          NaN
28                                          NaN
29                                          NaN
...                                         ...
9970  With promises of improved efficiencies, produc...
9971                                        NaN
9972                                        NaN
9973  This study explores the premise that knowledge...
9974  The paper considers the design of nearly perfe...
9975  Conventional fault relationships are mostly re...
9976  For nontrivial general (open) queueing network...
9977  In companies a large amount of information is ...
9978                                        NaN
9979                                        NaN
9980                                        NaN
9981                                        NaN
9982  Heilbronn's triangle problem asks for the leas...
9983  This article examines how the Internet transfo...
9984                                        NaN
9985  Two factors influencing the design anddevelopm...
9986  Collaborative decision making (CDM) is a joint...
9987  Constrained transmission lines are known to be...
9988                                        NaN
9989  Architectures that implement the Internet Prot...
9990  In many autonomous robot applications robots m...
9991  Traditional operating systems limit flexibilit...
9992                                        NaN
9993                                        NaN
9994  In this paper, we study a problem of schedulin...
9995                                        NaN
9996                                        NaN
```

```
9997                                                NaN
9998                                                NaN
9999   Interactive system design requires good descri...

                                               Authors  Year  index id  \
0                E. S. Cho,C. J. Kim,S. D. Kim,S. Y. Rhew  1998         0
1               Lori M. Weber,Alysha Loumakis,James Bergman  2003         1
2                                                  NaN  2002         2
3                                       Choong-Gyoo Lim  2002         3
4      Jose Maria Perez,Felix Garcia,Jesus Carretero,...  2003         4
5                                          Jean Kumagai  2003         5
6            Marek Rusinkiewicz,Dimitrios Georgakopoulos  1999         6
7                          Barton C. Massey,Evan Tick  1995         7
8                                            Jan Ramon  2002         8
9             Therapon Skotiniotis,Ji-en Morris Chang  2002         9
10                                   V. Martin,K. Schwan  1998        10
11                                 Peter W. Moo,Xiaolin Wu  1999        11
12                                   T. Aoki,T. Katayama  1998        12
13                                     Johannes Schropp  2003        13
14                                Stephen W. Gaarenstroom  2003        14
15          Hasan Pirkul,David A. Schilling,W. J. Hurley  1998        15
16          Christo Frank,B. S. Manoj,C. Siva Ram Murthy  2002        16
17    Klaus Buchenrieder,Alexander Sedlmeier,Ch. Veith  1993        17
18                               Bill Jackson,Tibor Jordán  2003        18
19                                   Alex Alves Freitas  1998        19
20          Volker Krüger,Udo Mahlmeister,Gerald Sommer  1998        20
21                                      Paul Keith Baker  2002        21
22                                  Andriy Panteleymonov  2002        22
23                                      James A. Larson  1984        23
24    Paul Bisgambiglia,Marielle Delhom,Jean-Françoi...  2002        24
25                                        Josef Schicho  2002        25
26                           L. Rodney Long,George R. Thoma  2000        26
27    Nicolas Anciaux,Christophe Bobineau,Luc Bougan...  2001        27
28      Franck Ravat,Marianne De Michiel,Gilles Zurfluh  1997        28
29                                                  NaN  1999        29
...                                                  ...   ...       ...
9970                                 William Granek P.E.  2000      9970
9971          Carles Farré,Ernest Teniente,Toni Urpí  1999      9971
9972          Christophe Loge,Valérie Gay,Eric Horlait  1994      9972
9973                                      Paul R. Carlile  2002      9973
9974  Fernando Cruz-Roldán,Ángel M. Bravo-Santos,Pil...  2003      9974
9975   Jwu E Chen,Chung Len Lee,Wen Zen Shen,Beyin Chen  2001      9975
9976                                        Armin Heindl  2003      9976
9977    Patrick Lambrix,Nahid Shahmehri,Niclas Wahlloef  1998      9977
9978  Motorola Paging Products Group Panelists: Jed ...  1997      9978
9979                      Christophe Lohou,Gilles Bertrand  2002      9979
9980  Anindo Bagchi,Edward F. Schmeichel,S. Louis Ha...  1994      9980
9981  B. Romaniuk,Michel Desvignes,J. Robiaille,Mari...  2001      9981
```

```
9982                       Tao Jiang,Ming Liy,Paul Vitányi  1999      9982
9983                       Tom Postmes,Suzanne Brunsting  2002      9983
9984             Marie-Pierre Gervais,Florin Muscutariu  2001      9984
9985                   D. A. Haverkamp,R. J. Richards  2002      9985
9986                                   Paul M. Carlson  2000      9986
9987                               Ziad Younes,Marija Ilic  1998      9987
9988                                         H. Tropf  1979      9988
9989  Andreas Dandalis,Viktor K. Prasanna,Jose D. P...  2000      9989
9990  Thorsten Schmitt,Michael Beetz,Robert Hanek,Se...  2002      9990
9991                   Patty Kostkova,Tim Wilkinson  2001      9991
9992                               Thomas Barkowsky  2002      9992
9993     Gerald M. Weinberg,Marie Benesh,James Bullock  2002      9993
9994          C.A. Glass,C.N. Potts,V.A. Strusevich  2001      9994
9995                               Michael Goodchild  2002      9995
9996  Jürgen Haas,Joachim Hornegger,Richard Huber,He...  1997      9996
9997                                   Roger M. Cooke  2003      9997
9998                                       Bill Trippe  2001      9998
9999                   Chris Phillips,Chris Scogings  2000      9999

                                              paperTitle  \
0     Static and Dynamic Metrics for Effective Objec...
1     Who participates and why?: an analysis of citi...
2                                         Call for papers
3     Universal parametrization in constructing smoo...
4     Data Allocation and Load Balancing for Heterog...
5               Employment opinion: reversal of fortune
6     From Coordination of Workflow and Group Activi...
7     Modes of Comprehension: Mode Analysis of Array...
8     Thesis: clustering and instance based learning...
9     Estimating internal memory fragmentation for J...
10    ILI: An Adaptive Infrastructure for Dynamic In...
11    Resynchronization Properties of Arithmetic Coding
12    Unification and Consistency Verification of Ob...
13    One- and multistep discretizations of index 2 ...
14    Surface Science Spectra: A Hybrid Journal-Data...
15    An Efficient, Objective Technique for Selectin...
16    Throughput Enhanced Wireless in Local Loop (TW...
17            HW/SW Co-Design with PRAMs Using CoDES
18                   Non-separable detachments of graphs
19        On Objective Measures of Rule Surprisingness
20            Attentive Face Detection and Recognition
21    Scientific papers: from web sites to e-communi...
22    Interoperable Thin Client Separation from GUI ...
23                               The Forms Pattern Language
24    An efficient and evolutionary hierarchical mod...
25            Simplification of surface parametrizations
26    Use of Shape Models to Search Digitized Spine ...
27                       PicoDBMS: Validation and Experience
```

```
28     Distributed Object Oriented Databases: An Allo...
29                              External Referees
...                                          ...
9970     E-Business, E-Commerce...Garbage In-Garbage Out
9971  The Constructive Method for Query Containment ...
9972  Computational Components for Synchronous Coope...
9973  A Pragmatic View of Knowledge and Boundaries: ...
9974  Design of multi-channel near-perfect-reconstru...
9975   Fanout Fault Analysis for Digital Logic Circuits
9976  Decomposition of general queueing networks wit...
9977  A Default Extension to Description Logics for ...
9978  Panel Chair: Tom Hilburn, Software Engineering...
9979  A New 3D 6-Subiteration Thinning Algorithm Bas...
9980         Gossiping with Multiple Sends and Receives
9981  Augmented Reality and Semi-automated Landmarki...
9982          The Expected Size of Heilbronn's Triangles
9983  Collective action in the age of the internet: ...
9984   Towards an ADL for Designing Agent-Based Systems
9985  Towards Safety Critical Middleware for Avionic...
9986  Exploiting the Opportunities of Collaborative ...
9987  Generation Strategies for Gaming Transmission ...
9988  Segmentation beim Erkennungsvorgang durch ein ...
9989  An Adaptive Cryptographic Engine for IPSec Arc...
9990  Watch their moves: applying probabilistic mult...
9991  Magnet: a virtual shared tuplespace resource m...
9992                                    State of the art
9993                                             Preface
9994  Scheduling Batches with Sequential Job Process...
9995              Geographic information system (GIS)
9996          Probabilistic Semantic Analysis of Speech
9997  Book review of "Elicitation of expert opinions...
9998            XML answers: it's time for an XML audit
9999  Task and Dialogue Modeling: Bridging the Divid...


                              publication venue  citation_count
0     Proceedings of the Fifth Asia Pacific Software...               3
1                     Social Science Computer Review               9
2                                        Sys Admin               0
3                     Computer Aided Geometric Design               1
4     Proceedings of the 3st International Symposium...               0
5                                     IEEE Spectrum               0
6     Proceedings of the 1999 International Symposiu...               2
7     Proceedings of the 7th International Symposium...               0
8                                 AI Communications               0
9                     Journal of Systems and Software               0
10    Proceedings of the International Conference on...               2
11    Proceedings of the Conference on Data Compression               2
12    Proceedings of the Fifth Asia Pacific Software...               0
```

13	Journal of Computational and Applied Mathematics	1
14	Computing in Science and Engineering	0
15	Interfaces	0
16	Proceedings of the 27th Annual IEEE Conference...	0
17	Proceedings of the 11th IFIP WG10.2 Internatio...	8
18	Journal of Combinatorial Theory Series B	2
19	Proceedings of the Second European Symposium o...	14
20	Mustererkennung 1998, 20. DAGM-Symposium	0
21	Technology and Health Care	0
22	Proceedings of the 6th European Conference on ...	0
23	Proceedings of the First International Confere...	1
24	Systems Analysis Modelling Simulation	0
25	Proceedings of the 2002 international symposiu...	2
26	CBMS	1
27	Very Large Data Bases	3
28	Proceedings of the 8th International Conferen...	0
29	Proceedings of the Sixth International Confere...	0
...	...	...
9970	Proceedings of the 1st Austin Workshop On Engi...	0
9971	Proceedings of the 10th International Conferen...	1
9972	Proceedings of the International COST 237 Work...	0
9973	Organization Science	27
9974	Signal Processing	0
9975	Proceedings of the 10th Anniversary Compendium...	0
9976	Performance Evaluation	3
9977	Proceedings of the Thirty-First Annual Hawaii ...	3
9978	Proceedings of the 10th Conference on Software...	0
9979	Proceedings of the 10th International Conferen...	6
9980	Proceedings of the 8th International Symposium...	0
9981	Proceedings of the 9th International Conferenc...	0
9982	Proceedings of the Fourteenth Annual IEEE Conf...	1
9983	Social Science Computer Review	4
9984	Revised Papers and Invited Contributions from ...	2
9985	Proceedings of the 27th Annual IEEE Conference...	1
9986	Transportation Science	0
9987	Proceedings of the Thirty-First Annual Hawaii ...	0
9988	Angewandte Szenenanalyse, DAGM Symposium	0
9989	Proceedings of the 2000 IEEE Symposium on Fiel...	7
9990	Eighteenth national conference on Artificial i...	3
9991	Virtual shared memory for distributed architec...	0
9992	Mental representation and processing of geogra...	0
9993	Roundtable on technical leadership: a SHAPE fo...	0
9994	INFORMS Journal on Computing	2
9995	Computer graphics companion	0
9996	Mustererkennung 1997, 19. DAGM-Symposium	0
9997	Fuzzy Sets and Systems	0
9998	Transform Magazine	0
9999	Proceedings of the First Australasian User Int...	1

```
        [10000 rows x 7 columns]
```

3   Get test data from Arxiv

```
In [10]: #Parse JSON data and create a df
         def parse_json_data_into_dataframe(filename, nrows):
             data = []
             count = 0
             with open(filename) as f:
                 for line in f:
                     if(count >= nrows):
                         break
                     data.append(json.loads(line))

             df = pd.DataFrame(data);
             return df;

In [11]: with open ('arxivData.json', 'rb') as f:
             data = json.load(f)
         #author = data['author']
         new_df = pd.DataFrame(data)
         author_list = []
         authors = new_df['author']
         authors
         new_df
```

```
Out[11]:                                          author  day          id  \
         0     [{'name': 'Ahmed Osman'}, {'name': 'Wojciech S...    1  1802.00209v1
         1     [{'name': 'Ji Young Lee'}, {'name': 'Franck De...   12  1603.03827v1
         2     [{'name': 'Iulian Vlad Serban'}, {'name': 'Tim...    2  1606.00776v2
         3     [{'name': 'Sebastian Ruder'}, {'name': 'Joachi...   23  1705.08142v2
         4     [{'name': 'Iulian V. Serban'}, {'name': 'Chinn...    7  1709.02349v2
         5     [{'name': 'Kelvin Guu'}, {'name': 'Tatsunori B...   26  1709.08878v1
         6     [{'name': 'Iulian V. Serban'}, {'name': 'Chinn...   20  1801.06700v1
         7     [{'name': 'Darko Brodic'}, {'name': 'Alessia A...   21  1609.06492v1
         8     [{'name': 'Mateusz Malinowski'}, {'name': 'Mar...    4  1610.01076v1
         9                        [{'name': 'Tony Beltramelli'}]   22  1705.07962v2
         10    [{'name': 'Fred Richardson'}, {'name': 'Dougla...    3  1504.00923v1
         11    [{'name': 'Hieu Pham'}, {'name': 'Melody Y. Gu...    9  1802.03268v2
         12    [{'name': 'Brenden M. Lake'}, {'name': 'Tomer ...    1  1604.00289v3
         13    [{'name': 'Hao Wang'}, {'name': 'Dit-Yan Yeung'}]    6  1604.01662v2
         14    [{'name': 'Tejas D. Kulkarni'}, {'name': 'Kart...   20  1604.06057v2
         15    [{'name': 'Deepak Pathak'}, {'name': 'Ross Gir...   19  1612.06370v2
         16    [{'name': 'Muhammad Ghifary'}, {'name': 'W. Ba...   21   1409.6041v1
         17    [{'name': 'Lionel Pigou'}, {'name': 'Aäron van...    5  1506.01911v3
         18    [{'name': 'Rakesh Achanta'}, {'name': 'Trevor ...   20  1509.05962v2
         19    [{'name': 'Jeff Donahue'}, {'name': 'Philipp K...   31  1605.09782v7
```

```
20                    [{'name': 'Zachary C. Lipton'}]    10   1606.03490v3
21    [{'name': 'Sahil Garg'}, {'name': 'Irina Rish'...   22   1701.06106v2
22      [{'name': 'Weifeng Ge'}, {'name': 'Yizhou Yu'}]   28   1702.08690v2
23    [{'name': 'Tanmay Gupta'}, {'name': 'Kevin Shi...    2   1704.00260v2
24    [{'name': 'Jan Hendrik Metzen'}, {'name': 'Mum...   19   1704.05712v3
25    [{'name': 'Quynh Nguyen'}, {'name': 'Matthias ...   26   1704.08045v2
26    [{'name': 'Chris Donahue'}, {'name': 'Zachary ...   22   1705.07904v3
27    [{'name': 'Mahesh Chandra Mukkamala'}, {'name'...   17   1706.05507v2
28    [{'name': 'Chunyuan Li'}, {'name': 'Hao Liu'},...    5   1709.01215v2
29    [{'name': 'Mateusz Buda'}, {'name': 'Atsuto Ma...   15   1710.05381v1
...                                                  ...  ...          ...
40970 [{'name': 'Smriti Bhagat'}, {'name': 'Udi Wein...   26   1311.6802v2
40971 [{'name': 'Kareem Amin'}, {'name': 'Afshin Ros...   26   1311.6838v1
40972 [{'name': 'Tianbao Yang'}, {'name': 'Shenghuo ...    4   1312.1031v2
40973 [{'name': 'Robert Kleinberg'}, {'name': 'Aleks...    4   1312.1277v2
40974 [{'name': 'Ferhat Özgür Çatak'}, {'name': 'Meh...   15   1312.4108v1
40975 [{'name': 'Gabriele Oliva'}, {'name': 'Roberto...   15   1312.4176v3
40976 [{'name': 'Yaqub Alwan'}, {'name': 'Zoran Cvet...   18   1312.5354v2
40977 [{'name': 'W. Liu'}, {'name': 'H. Liu'}, {'nam...   21   1312.6180v1
40978 [{'name': 'Luis Marujo'}, {'name': 'Anatole Ge...   23   1312.6597v2
40979 [{'name': 'Pranjal Awasthi'}, {'name': 'Maria-...   24   1312.6724v3
40980 [{'name': 'Ahmed K. Farahat'}, {'name': 'Ahmed...   24   1312.6838v1
40981 [{'name': 'Anupriya Gogna'}, {'name': 'Ankita ...   17   1312.6872v1
40982 [{'name': 'Prashanth L. A.'}, {'name': 'Abhran...   27   1312.7292v2
40983 [{'name': 'Christopher Genovese'}, {'name': 'M...   29   1312.7567v1
40984 [{'name': 'Andrey Bernstein'}, {'name': 'Nahum...   30   1312.7658v1
40985 [{'name': 'Andreas Veit'}, {'name': 'Christoph...    1   1404.0200v1
40986 [{'name': 'Peter Banda'}, {'name': 'Christof T...    2   1404.0427v2
40987 [{'name': 'Pokkuluri Kiran Sree'}, {'name': 'I...    2   1404.0453v1
40988 [{'name': 'Da Kuang'}, {'name': 'Alex Gittens'...    2   1404.0466v2
40989 [{'name': 'Pokkuluri Kiran Sree'}, {'name': 'I...    4   1404.1144v1
40990 [{'name': 'Joost Broekens'}, {'name': 'Tim Baa...    8   1404.2078v2
40991 [{'name': 'Kerstin Eder'}, {'name': 'Chris Har...    8   1404.2229v3
40992 [{'name': 'Lee-Ad Gottlieb'}, {'name': 'Aryeh ...   13   1404.3368v3
40993 [{'name': 'Amit Daniely'}, {'name': 'Shai Shal...   13   1404.3378v2
40994 [{'name': 'Karthik Raman'}, {'name': 'Thorsten...   14   1404.3656v1
40995 [{'name': 'Vitaly Feldman'}, {'name': 'Pravesh...   18   1404.4702v2
40996 [{'name': 'Orly Avner'}, {'name': 'Shie Mannor'}]   22   1404.5421v1
40997 [{'name': 'Ran Zhao'}, {'name': 'Deanna Needel...   22   1404.5899v1
40998 [{'name': 'Zongyan Huang'}, {'name': 'Matthew ...   25   1404.6369v1
40999 [{'name': 'Imen Trabelsi'}, {'name': 'Dorra Be...   27   1407.0380v1

                                              link  month  \
0       [{'rel': 'alternate', 'href': 'http://arxiv.or...     2
1       [{'rel': 'alternate', 'href': 'http://arxiv.or...     3
2       [{'rel': 'alternate', 'href': 'http://arxiv.or...     6
3       [{'rel': 'alternate', 'href': 'http://arxiv.or...     5
4       [{'rel': 'alternate', 'href': 'http://arxiv.or...     9
```

```
5       [{'rel': 'alternate', 'href': 'http://arxiv.or...         9
6       [{'rel': 'alternate', 'href': 'http://arxiv.or...         1
7       [{'rel': 'alternate', 'href': 'http://arxiv.or...         9
8       [{'rel': 'alternate', 'href': 'http://arxiv.or...        10
9       [{'rel': 'alternate', 'href': 'http://arxiv.or...         5
10      [{'rel': 'alternate', 'href': 'http://arxiv.or...         4
11      [{'rel': 'alternate', 'href': 'http://arxiv.or...         2
12      [{'rel': 'alternate', 'href': 'http://arxiv.or...         4
13      [{'rel': 'alternate', 'href': 'http://arxiv.or...         4
14      [{'rel': 'alternate', 'href': 'http://arxiv.or...         4
15      [{'rel': 'alternate', 'href': 'http://arxiv.or...        12
16      [{'rel': 'alternate', 'href': 'http://arxiv.or...         9
17      [{'rel': 'alternate', 'href': 'http://arxiv.or...         6
18      [{'rel': 'alternate', 'href': 'http://arxiv.or...         9
19      [{'rel': 'alternate', 'href': 'http://arxiv.or...         5
20      [{'rel': 'alternate', 'href': 'http://arxiv.or...         6
21      [{'rel': 'alternate', 'href': 'http://arxiv.or...         1
22      [{'rel': 'alternate', 'href': 'http://arxiv.or...         2
23      [{'rel': 'alternate', 'href': 'http://arxiv.or...         4
24      [{'rel': 'alternate', 'href': 'http://arxiv.or...         4
25      [{'rel': 'alternate', 'href': 'http://arxiv.or...         4
26      [{'rel': 'alternate', 'href': 'http://arxiv.or...         5
27      [{'rel': 'alternate', 'href': 'http://arxiv.or...         6
28      [{'rel': 'alternate', 'href': 'http://arxiv.or...         9
29      [{'rel': 'alternate', 'href': 'http://arxiv.or...        10
...                                                      ...       ...
40970   [{'rel': 'alternate', 'href': 'http://arxiv.or...        11
40971   [{'rel': 'alternate', 'href': 'http://arxiv.or...        11
40972   [{'rel': 'alternate', 'href': 'http://arxiv.or...        12
40973   [{'rel': 'alternate', 'href': 'http://arxiv.or...        12
40974   [{'rel': 'alternate', 'href': 'http://arxiv.or...        12
40975   [{'rel': 'alternate', 'href': 'http://arxiv.or...        12
40976   [{'rel': 'alternate', 'href': 'http://arxiv.or...        12
40977   [{'rel': 'alternate', 'href': 'http://arxiv.or...        12
40978   [{'rel': 'alternate', 'href': 'http://arxiv.or...        12
40979   [{'rel': 'alternate', 'href': 'http://arxiv.or...        12
40980   [{'rel': 'alternate', 'href': 'http://arxiv.or...        12
40981   [{'rel': 'alternate', 'href': 'http://arxiv.or...        12
40982   [{'rel': 'alternate', 'href': 'http://arxiv.or...        12
40983   [{'rel': 'alternate', 'href': 'http://arxiv.or...        12
40984   [{'rel': 'alternate', 'href': 'http://arxiv.or...        12
40985   [{'rel': 'alternate', 'href': 'http://arxiv.or...         4
40986   [{'rel': 'related', 'href': 'http://dx.doi.org...         4
40987   [{'rel': 'alternate', 'href': 'http://arxiv.or...         4
40988   [{'rel': 'alternate', 'href': 'http://arxiv.or...         4
40989   [{'rel': 'alternate', 'href': 'http://arxiv.or...         4
40990   [{'rel': 'alternate', 'href': 'http://arxiv.or...         4
40991   [{'rel': 'alternate', 'href': 'http://arxiv.or...         4
```

```
40992  [{'rel': 'alternate', 'href': 'http://arxiv.or...      4
40993  [{'rel': 'alternate', 'href': 'http://arxiv.or...      4
40994  [{'rel': 'alternate', 'href': 'http://arxiv.or...      4
40995  [{'rel': 'alternate', 'href': 'http://arxiv.or...      4
40996  [{'rel': 'alternate', 'href': 'http://arxiv.or...      4
40997  [{'rel': 'alternate', 'href': 'http://arxiv.or...      4
40998  [{'rel': 'related', 'href': 'http://dx.doi.org...      4
40999  [{'rel': 'alternate', 'href': 'http://arxiv.or...      6

                                           summary  \
0      We propose an architecture for VQA which utili...
1      Recent approaches based on artificial neural n...
2      We introduce the multiresolution recurrent neu...
3      Multi-task learning is motivated by the observ...
4      We present MILABOT: a deep reinforcement learn...
5      We propose a new generative model of sentences...
6      We present MILABOT: a deep reinforcement learn...
7      The paper introduces a new method for discrimi...
8      Together with the development of more accurate...
9      Transforming a graphical user interface screen...
10     Learned feature representations and sub-phonem...
11     We propose Efficient Neural Architecture Searc...
12     Recent progress in artificial intelligence (AI...
13     While perception tasks such as visual object r...
14     Learning goal-directed behavior in environment...
15     This paper presents a novel yet intuitive appr...
16     We propose a simple neural network model to de...
17     Recent studies have demonstrated the power of ...
18     In this paper, we address the task of Optical ...
19     The ability of the Generative Adversarial Netw...
20     Supervised machine learning models boast remar...
21     In this paper, we focus on online representati...
22     Deep neural networks require a large amount of...
23     An important goal of computer vision is to bui...
24     While deep learning is remarkably successful o...
25     While the optimization problem behind deep neu...
26     We propose a new algorithm for training genera...
27     Adaptive gradient methods have become recently...
28     We investigate the non-identifiability issues ...
29     In this study, we systematically investigate t...
...                                             ...
40970  Recommender systems leverage user demographic ...
40971  Inspired by real-time ad exchanges for online ...
40972  In \citep{Yangnips13}, the author presented di...
40973  In a multi-armed bandit problem, an online alg...
40974  Although Support Vector Machine (SVM) algorith...
40975  In this paper we provide a fully distributed i...
40976  We studied classification of human ECGs labell...
```

17

```
40977  With the rapid advance of Internet technology ...
40978  In this work, we propose two stochastic archit...
40979  We study the design of interactive clustering ...
40980  In today's information systems, the availabili...
40981  In this paper we address the problem of recove...
40982  In this paper, we consider an intrusion detect...
40983  We derive nonparametric confidence intervals f...
40984  Approachability theory, introduced by Blackwel...
40985  The increasing use of renewable energy sources...
40986  The current biochemical information processing...
40987  This paper aims at providing a survey on the p...
40988  The dominant cost in solving least-square prob...
40989  Bioinformatics incorporates information regard...
40990  Understanding the affective, cognitive and beh...
40991  The success of the human-robot co-worker team ...
40992  We present the first sample compression algori...
40993  Using the recently developed framework of [Dan...
40994  MOOCs have the potential to revolutionize high...
40995  We study the complexity of learning and approx...
40996  We consider the problem of multiple users targ...
40997  In this paper, we compare and analyze clusteri...
40998  Cylindrical algebraic decomposition(CAD) is a ...
40999  Several speaker identification systems are giv...

                                                    tag  \
0     [{'term': 'cs.AI', 'scheme': 'http://arxiv.org...
1     [{'term': 'cs.CL', 'scheme': 'http://arxiv.org...
2     [{'term': 'cs.CL', 'scheme': 'http://arxiv.org...
3     [{'term': 'stat.ML', 'scheme': 'http://arxiv.o...
4     [{'term': 'cs.CL', 'scheme': 'http://arxiv.org...
5     [{'term': 'cs.CL', 'scheme': 'http://arxiv.org...
6     [{'term': 'cs.CL', 'scheme': 'http://arxiv.org...
7     [{'term': 'cs.CV', 'scheme': 'http://arxiv.org...
8     [{'term': 'cs.CV', 'scheme': 'http://arxiv.org...
9     [{'term': 'cs.LG', 'scheme': 'http://arxiv.org...
10    [{'term': 'cs.CL', 'scheme': 'http://arxiv.org...
11    [{'term': 'cs.LG', 'scheme': 'http://arxiv.org...
12    [{'term': 'cs.AI', 'scheme': 'http://arxiv.org...
13    [{'term': 'stat.ML', 'scheme': 'http://arxiv.o...
14    [{'term': 'cs.LG', 'scheme': 'http://arxiv.org...
15    [{'term': 'cs.CV', 'scheme': 'http://arxiv.org...
16    [{'term': 'cs.CV', 'scheme': 'http://arxiv.org...
17    [{'term': 'cs.CV', 'scheme': 'http://arxiv.org...
18    [{'term': 'stat.ML', 'scheme': 'http://arxiv.o...
19    [{'term': 'cs.LG', 'scheme': 'http://arxiv.org...
20    [{'term': 'cs.LG', 'scheme': 'http://arxiv.org...
21    [{'term': 'cs.LG', 'scheme': 'http://arxiv.org...
22    [{'term': 'cs.CV', 'scheme': 'http://arxiv.org...
```

```
23      [{'term': 'cs.CV', 'scheme': 'http://arxiv.org...
24      [{'term': 'stat.ML', 'scheme': 'http://arxiv.o...
25      [{'term': 'cs.LG', 'scheme': 'http://arxiv.org...
26      [{'term': 'cs.LG', 'scheme': 'http://arxiv.org...
27      [{'term': 'cs.LG', 'scheme': 'http://arxiv.org...
28      [{'term': 'stat.ML', 'scheme': 'http://arxiv.o...
29      [{'term': 'cs.CV', 'scheme': 'http://arxiv.org...
...                                                 ...
40970   [{'term': 'cs.LG', 'scheme': 'http://arxiv.org...
40971   [{'term': 'cs.LG', 'scheme': 'http://arxiv.org...
40972   [{'term': 'cs.DC', 'scheme': 'http://arxiv.org...
40973   [{'term': 'cs.DS', 'scheme': 'http://arxiv.org...
40974   [{'term': 'cs.LG', 'scheme': 'http://arxiv.org...
40975   [{'term': 'cs.LG', 'scheme': 'http://arxiv.org...
40976   [{'term': 'cs.CE', 'scheme': 'http://arxiv.org...
40977   [{'term': 'cs.LG', 'scheme': 'http://arxiv.org...
40978   [{'term': 'cs.LG', 'scheme': 'http://arxiv.org...
40979   [{'term': 'cs.DS', 'scheme': 'http://arxiv.org...
40980   [{'term': 'cs.DS', 'scheme': 'http://arxiv.org...
40981   [{'term': 'cs.NA', 'scheme': 'http://arxiv.org...
40982   [{'term': 'cs.SY', 'scheme': 'http://arxiv.org...
40983   [{'term': 'stat.ME', 'scheme': 'http://arxiv.o...
40984   [{'term': 'cs.LG', 'scheme': 'http://arxiv.org...
40985   [{'term': 'cs.LG', 'scheme': 'http://arxiv.org...
40986   [{'term': 'q-bio.MN', 'scheme': 'http://arxiv...
40987   [{'term': 'cs.CE', 'scheme': 'http://arxiv.org...
40988   [{'term': 'cs.LG', 'scheme': 'http://arxiv.org...
40989   [{'term': 'cs.CE', 'scheme': 'http://arxiv.org...
40990   [{'term': 'cs.LG', 'scheme': 'http://arxiv.org...
40991   [{'term': 'cs.RO', 'scheme': 'http://arxiv.org...
40992   [{'term': 'cs.LG', 'scheme': 'http://arxiv.org...
40993   [{'term': 'cs.LG', 'scheme': 'http://arxiv.org...
40994   [{'term': 'cs.LG', 'scheme': 'http://arxiv.org...
40995   [{'term': 'cs.LG', 'scheme': 'http://arxiv.org...
40996   [{'term': 'cs.LG', 'scheme': 'http://arxiv.org...
40997   [{'term': 'math.NA', 'scheme': 'http://arxiv.o...
40998   [{'term': 'cs.SC', 'scheme': 'http://arxiv.org...
40999   [{'term': 'cs.SD', 'scheme': 'http://arxiv.org...

                                             title  year
0      Dual Recurrent Attention Units for Visual Ques...  2018
1      Sequential Short-Text Classification with Recu...  2016
2      Multiresolution Recurrent Neural Networks: An ...  2016
3      Learning what to share between loosely related...  2017
4               A Deep Reinforcement Learning Chatbot  2017
5             Generating Sentences by Editing Prototypes  2017
6      A Deep Reinforcement Learning Chatbot (Short V...  2018
7      Document Image Coding and Clustering for Scrip...  2016
```

```
8        Tutorial on Answering Questions about Images w...    2016
9        pix2code: Generating Code from a Graphical Use...    2017
10       A Unified Deep Neural Network for Speaker and ...    2015
11       Efficient Neural Architecture Search via Param...    2018
12       Building Machines That Learn and Think Like Pe...    2016
13                 Towards Bayesian Deep Learning: A Survey    2016
14       Hierarchical Deep Reinforcement Learning: Inte...    2016
15              Learning Features by Watching Objects Move    2016
16       Domain Adaptive Neural Networks for Object Rec...    2014
17       Beyond Temporal Pooling: Recurrence and Tempor...    2015
18              Telugu OCR Framework using Deep Learning    2015
19                            Adversarial Feature Learning    2016
20                    The Mythos of Model Interpretability    2016
21       Neurogenesis-Inspired Dictionary Learning: Onl...    2017
22       Borrowing Treasures from the Wealthy: Deep Tra...    2017
23       Aligned Image-Word Representations Improve Ind...    2017
24       Universal Adversarial Perturbations Against Se...    2017
25       The loss surface of deep and wide neural networks    2017
26       Semantically Decomposing the Latent Spaces of ...    2017
27       Variants of RMSProp and Adagrad with Logarithm...    2017
28       ALICE: Towards Understanding Adversarial Learn...    2017
29       A systematic study of the class imbalance prob...    2017
...                                                    ...    ...
40970    Recommending with an Agenda: Active Learning o...    2013
40971    Learning Prices for Repeated Auctions with Str...    2013
40972    Analysis of Distributed Stochastic Dual Coordi...    2013
40973                    Bandits and Experts in Metric Spaces    2013
40974    A MapReduce based distributed SVM algorithm fo...    2013
40975                       Distributed k-means algorithm    2013
40976    Classification of Human Ventricular Arrhythmia...    2013
40977    Manifold regularized kernel logistic regressio...    2013
40978    Co-Multistage of Multiple Classifiers for Imba...    2013
40979         Local algorithms for interactive clustering    2013
40980    Greedy Column Subset Selection for Large-scale...    2013
40981                    Matrix recovery using Split Bregman    2013
40982    Two Timescale Convergent Q-learning for Sleep-...    2013
40983           Nonparametric Inference For Density Modes    2013
40984    Response-Based Approachability and its Applica...    2013
40985    Household Electricity Demand Forecasting -- Be...    2014
40986    Learning Two-input Linear and Nonlinear Analog...    2014
40987    Cellular Automata and Its Applications in Bioi...    2014
40988    piCholesky: Polynomial Interpolation of Multip...    2014
40989    AIS-MACA- Z: MACA based Clonal Classifier for ...    2014
40990    Optimistic Risk Perception in the Temporal Dif...    2014
40991    Towards the Safety of Human-in-the-Loop Roboti...    2014
40992    Near-optimal sample compression for nearest ne...    2014
40993    Complexity theoretic limitations on learning D...    2014
40994                    Methods for Ordinal Peer Grading    2014
```

```
40995   Nearly Tight Bounds on $\ell_1$ Approximation ...   2014
40996     Concurrent bandits and cognitive radio networks  2014
40997   A Comparison of Clustering and Missing Data Me...   2014
40998   Applying machine learning to the problem of ch...   2014
40999   A Multi Level Data Fusion Approach for Speaker...   2014

[41000 rows x 9 columns]
```

In [12]:
```python
final_list = []
citation_count_list = []
for entry in authors:
    a = entry
    all_authors = ""
    a = a[1:-1]
    a_list = a.split(",")
    for author in a_list:
        author_name = author[1:-1].split(":")
        if(len(author_name) > 1):
            author_name = author_name[1].replace("'","")
            all_authors += author_name + ","
    all_authors = all_authors[:-1]
    final_list.append(all_authors)
    citation_count_list.append(0)
```

In [13]:
```python
test1 = pd.DataFrame()
test1["authors"] = final_list
new_df["author"] = test1["authors"]
new_df = new_df.drop("link", axis = 1)
new_df = new_df.drop("tag", axis = 1)
new_df = new_df.drop("day", axis = 1)
new_df = new_df.drop("month", axis = 1)
new_df['Authors'] = new_df['summary']
new_df = new_df.drop("summary", axis = 1)
new_df.columns = ['Authors', 'index id', 'paperTitle', 'Year', 'Abstract']
test_df = new_df
```

In [14]:
```python
test_df['citation_count'] = citation_count_list
test_df.to_csv('test_new.csv', index=False)
test_df = pd.read_csv('test_new.csv', nrows=100)
```

In [15]:
```python
test_df
```

Out[15]:
```
                                            Authors      index id  \
0                      Ahmed Osman, Wojciech Samek   1802.00209v1
1                    Ji Young Lee, Franck Dernoncourt   1603.03827v1
2       Iulian Vlad Serban, Tim Klinger, Gerald Tesau...   1606.00776v2
3       Sebastian Ruder, Joachim Bingel, Isabelle Aug...   1705.08142v2
4       Iulian V. Serban, Chinnadhurai Sankar, Mathie...   1709.02349v2
5       Kelvin Guu, Tatsunori B. Hashimoto, Yonatan O...   1709.08878v1
```

```
6    Iulian V. Serban, Chinnadhurai Sankar, Mathie...  1801.06700v1
7    Darko Brodic, Alessia Amelio, Zoran N. Milivo...  1609.06492v1
8                     Mateusz Malinowski, Mario Fritz  1610.01076v1
9                                    Tony Beltramelli  1705.07962v2
10     Fred Richardson, Douglas Reynolds, Najim Dehak  1504.00923v1
11   Hieu Pham, Melody Y. Guan, Barret Zoph, Quoc ...  1802.03268v2
12   Brenden M. Lake, Tomer D. Ullman, Joshua B. T...  1604.00289v3
13                            Hao Wang, Dit-Yan Yeung  1604.01662v2
14   Tejas D. Kulkarni, Karthik R. Narasimhan, Ard...  1604.06057v2
15   Deepak Pathak, Ross Girshick, Piotr Dollár, T...  1612.06370v2
16   Muhammad Ghifary, W. Bastiaan Kleijn, Mengjie...   1409.6041v1
17   Lionel Pigou, Aäron van den Oord, Sander Diel...  1506.01911v3
18                      Rakesh Achanta, Trevor Hastie  1509.05962v2
19   Jeff Donahue, Philipp Krähenbühl, Trevor Darrell  1605.09782v7
20                                  Zachary C. Lipton  1606.03490v3
21   Sahil Garg, Irina Rish, Guillermo Cecchi, Aur...  1701.06106v2
22                             Weifeng Ge, Yizhou Yu   1702.08690v2
23   Tanmay Gupta, Kevin Shih, Saurabh Singh, Dere...  1704.00260v2
24   Jan Hendrik Metzen, Mummadi Chaithanya Kumar,...  1704.05712v3
25                        Quynh Nguyen, Matthias Hein  1704.08045v2
26   Chris Donahue, Zachary C. Lipton, Akshay Bals...  1705.07904v3
27          Mahesh Chandra Mukkamala, Matthias Hein   1706.05507v2
28   Chunyuan Li, Hao Liu, Changyou Chen, Yunchen ...  1709.01215v2
29    Mateusz Buda, Atsuto Maki, Maciej A. Mazurowski 1710.05381v1
..                                           ...           ...
70   Petr Baudi, Jan Pichl, Tomá Vyskoil, Jan ...   1603.06127v4
71   Jiatao Gu, Zhengdong Lu, Hang Li, Victor O. K...  1603.06393v3
72   Iulian Vlad Serban, Alberto García-Durán, Cag...  1603.06807v2
73   Chia-Wei Liu, Ryan Lowe, Iulian V. Serban, Mi...  1603.08023v2
74   Iulian Vlad Serban, Alessandro Sordoni, Ryan ...  1605.06069v3
75                                   Dirk Weissenborn  1606.03864v2
76                      Marc Dymetman, Chunyang Xiao  1607.02467v2
77     Ondrej Bajgar, Rudolf Kadlec, Jan Kleindienst  1610.00956v1
78   James Bradbury, Stephen Merity, Caiming Xiong...  1611.01576v2
79   Jakob N. Foerster, Justin Gilmer, Jan Chorows...  1611.09434v2
80   Micha Daniluk, Tim Rocktäschel, Johannes Wel...  1702.04521v1
81   Zhouhan Lin, Minwei Feng, Cicero Nogueira dos...  1703.03130v1
82   Samuel Rönnqvist, Niko Schenk, Christian Chia...  1704.08092v1
83   Lara J. Martin, Prithviraj Ammanabrolu, Xinyu...  1706.01331v3
84            Tong Wang, Xingdi Yuan, Adam Trischler  1706.01450v1
85   Wei Wen, Yuxiong He, Samyam Rajbhandari, Minj...  1709.05027v7
86     Huda Hakami, Danushka Bollegala, Hayashi Kohei  1709.06673v2
87   Zhengdong Lu, Haotian Cui, Xianggen Liu, Yuku...  1709.08853v4
88                                     Bin Bi, Hao Ma  1709.10204v2
89   Mirco Ravanelli, Philemon Brakel, Maurizio Om...  1710.00641v1
90   Baolin Peng, Xiujun Li, Jianfeng Gao, Jingjin...  1801.06176v1
91   Andrew L. Maas, Peng Qi, Ziang Xie, Awni Y. H...   1406.7806v2
92                            William Chan, Ian Lane  1504.01482v1
```

```
93                     David Krueger, Roland Memisevic   1511.08400v7
94     Noam Shazeer, Azalia Mirhoseini, Krzysztof Ma...   1701.06538v1
95      Yacine Jernite, Samuel R. Bowman, David Sontag   1705.00557v1
96                        Zhengyang Wang, Shuiwang Ji   1705.06824v1
97     Kyunghyun Cho, Bart van Merrienboer, Caglar G...   1406.1078v3
98              Zhiyuan Tang, Dong Wang, Zhiyong Zhang   1505.04630v5
99        Haim Sak, Andrew Senior, Françoise Beaufays   1402.1128v1

                                            paperTitle  Year  \
0    Dual Recurrent Attention Units for Visual Ques...  2018
1    Sequential Short-Text Classification with Recu...  2016
2    Multiresolution Recurrent Neural Networks: An ...  2016
3    Learning what to share between loosely related...  2017
4                  A Deep Reinforcement Learning Chatbot  2017
5             Generating Sentences by Editing Prototypes  2017
6    A Deep Reinforcement Learning Chatbot (Short V...  2018
7    Document Image Coding and Clustering for Scrip...  2016
8    Tutorial on Answering Questions about Images w...  2016
9    pix2code: Generating Code from a Graphical Use...  2017
10   A Unified Deep Neural Network for Speaker and ...  2015
11   Efficient Neural Architecture Search via Param...  2018
12   Building Machines That Learn and Think Like Pe...  2016
13              Towards Bayesian Deep Learning: A Survey  2016
14   Hierarchical Deep Reinforcement Learning: Inte...  2016
15           Learning Features by Watching Objects Move  2016
16   Domain Adaptive Neural Networks for Object Rec...  2014
17   Beyond Temporal Pooling: Recurrence and Tempor...  2015
18              Telugu OCR Framework using Deep Learning  2015
19                          Adversarial Feature Learning  2016
20                  The Mythos of Model Interpretability  2016
21   Neurogenesis-Inspired Dictionary Learning: Onl...  2017
22   Borrowing Treasures from the Wealthy: Deep Tra...  2017
23   Aligned Image-Word Representations Improve Ind...  2017
24   Universal Adversarial Perturbations Against Se...  2017
25   The loss surface of deep and wide neural networks  2017
26   Semantically Decomposing the Latent Spaces of ...  2017
27   Variants of RMSProp and Adagrad with Logarithm...  2017
28   ALICE: Towards Understanding Adversarial Learn...  2017
29   A systematic study of the class imbalance prob...  2017
..                                                 ...   ...
70   Sentence Pair Scoring: Towards Unified Framewo...  2016
71   Incorporating Copying Mechanism in Sequence-to...  2016
72   Generating Factoid Questions With Recurrent Ne...  2016
73   How NOT To Evaluate Your Dialogue System: An E...  2016
74   A Hierarchical Latent Variable Encoder-Decoder...  2016
75   Neural Associative Memory for Dual-Sequence Mo...  2016
76   Log-Linear RNNs: Towards Recurrent Neural Netw...  2016
77   Embracing data abundance: BookTest Dataset for...  2016
```

```
78                 Quasi-Recurrent Neural Networks  2016
79  Input Switched Affine Networks: An RNN Archite...  2016
80  Frustratingly Short Attention Spans in Neural ...  2017
81     A Structured Self-attentive Sentence Embedding  2017
82  A Recurrent Neural Model with Attention for th...  2017
83  Event Representations for Automated Story Gene...  2017
84  A Joint Model for Question Answering and Quest...  2017
85  Learning Intrinsic Sparse Structures within Lo...  2017
86  Why PairDiff works? -- A Mathematical Analysis...  2017
87  Object-oriented Neural Programming (OONP) for ...  2017
88  A Neural Comprehensive Ranker (NCR) for Open-D...  2017
89  Improving speech recognition by revising gated...  2017
90  Integrating planning for task-completion dialo...  2018
91  Building DNN Acoustic Models for Large Vocabul...  2014
92  Deep Recurrent Neural Networks for Acoustic Mo...  2015
93       Regularizing RNNs by Stabilizing Activations  2015
94  Outrageously Large Neural Networks: The Sparse...  2017
95  Discourse-Based Objectives for Fast Unsupervis...  2017
96  Learning Convolutional Text Representations fo...  2017
97  Learning Phrase Representations using RNN Enco...  2014
98  Recurrent Neural Network Training with Dark Kn...  2015
99  Long Short-Term Memory Based Recurrent Neural ...  2014


                                    Abstract  citation_count
0   We propose an architecture for VQA which utili...               0
1   Recent approaches based on artificial neural n...               0
2   We introduce the multiresolution recurrent neu...               0
3   Multi-task learning is motivated by the observ...               0
4   We present MILABOT: a deep reinforcement learn...               0
5   We propose a new generative model of sentences...               0
6   We present MILABOT: a deep reinforcement learn...               0
7   The paper introduces a new method for discrimi...               0
8   Together with the development of more accurate...               0
9   Transforming a graphical user interface screen...               0
10  Learned feature representations and sub-phonem...               0
11  We propose Efficient Neural Architecture Searc...               0
12  Recent progress in artificial intelligence (AI...               0
13  While perception tasks such as visual object r...               0
14  Learning goal-directed behavior in environment...               0
15  This paper presents a novel yet intuitive appr...               0
16  We propose a simple neural network model to de...               0
17  Recent studies have demonstrated the power of ...               0
18  In this paper, we address the task of Optical ...               0
19  The ability of the Generative Adversarial Netw...               0
20  Supervised machine learning models boast remar...               0
21  In this paper, we focus on online representati...               0
22  Deep neural networks require a large amount of...               0
23  An important goal of computer vision is to bui...               0
```

```
24   While deep learning is remarkably successful o...        0
25   While the optimization problem behind deep neu...        0
26   We propose a new algorithm for training genera...        0
27   Adaptive gradient methods have become recently...        0
28   We investigate the non-identifiability issues ...        0
29   In this study, we systematically investigate t...        0
..                                                      ...      ...
70   We review the task of Sentence Pair Scoring, p...        0
71   We address an important problem in sequence-to...        0
72   Over the past decade, large-scale supervised l...        0
73   We investigate evaluation metrics for dialogue...        0
74   Sequential data often possesses a hierarchical...        0
75   Many important NLP problems can be posed as du...        0
76   We introduce LL-RNNs (Log-Linear RNNs), an ext...        0
77   There is a practically unlimited amount of nat...        0
78   Recurrent neural networks are a powerful tool ...        0
79   There exist many problem domains where the int...        0
80   Neural language models predict the next token ...        0
81   This paper proposes a new model for extracting...        0
82   We introduce an attention-based Bi-LSTM for Ch...        0
83   Automated story generation is the problem of a...        0
84   We propose a generative machine comprehension ...        0
85   Model compression is significant for the wide ...        0
86   Representing the semantic relations that exist...        0
87   We propose Object-oriented Neural Programming ...        0
88   This paper proposes a novel neural machine rea...        0
89   Speech recognition is largely taking advantage...        0
90   Training a task-completion dialogue agent with...        0
91   Deep neural networks (DNNs) are now a central ...        0
92   We present a novel deep Recurrent Neural Netwo...        0
93   We stabilize the activations of Recurrent Neur...        0
94   The capacity of a neural network to absorb inf...        0
95   This work presents a novel objective function ...        0
96   Visual question answering is a recently propos...        0
97   In this paper, we propose a novel neural netwo...        0
98   Recurrent neural networks (RNNs), particularly...        0
99   Long Short-Term Memory (LSTM) is a recurrent n...        0

[100 rows x 6 columns]
```

In [16]: # drop venue column and fill NAN values with N/A
```python
train_df = train_df.drop(['publication venue'], axis=1)
train_df['Authors'].fillna('N/A', inplace=True)
train_df['Abstract'].fillna('N/A', inplace=True)
train_df['Year'].fillna(0.0, inplace=True)
train_df['Year'] = train_df['Year'].astype(np.int64)

test_df['Authors'].fillna('N/A', inplace=True)
```

```
        test_df['Abstract'].fillna('N/A', inplace=True)
        test_df['Year'].fillna(0.0, inplace=True)
        test_df['Year'] = test_df['Year'].astype(np.int64)
```

In [17]: train_df.head(12)

Out[17]:   Abstract \
        0   N/A
        1   Scholars continue to find that political parti...
        2   N/A
        3   In this paper, we explore the feasibility of u...
        4   Distributed filesystems are a typical solution...
        5   No question times are tough for many working e...
        6   The objective of the Collaboration Management ...
        7   N/A
        8   Instance based learning and clustering are pop...
        9   Dynamic memory management has been an importan...
        10   N/A
        11 In this paper we consider the problem of resyn...

                                                   Authors Year index id \
        0   E. S. Cho,C. J. Kim,S. D. Kim,S. Y. Rhew 1998   0
        1   Lori M. Weber,Alysha Loumakis,James Bergman 2003   1
        2   N/A 2002   2
        3   Choong-Gyoo Lim 2002   3
        4   Jose Maria Perez,Felix Garcia,Jesus Carretero,... 2003   4
        5   Jean Kumagai 2003   5
        6   Marek Rusinkiewicz,Dimitrios Georgakopoulos 1999   6
        7   Barton C. Massey,Evan Tick 1995   7
        8   Jan Ramon 2002   8
        9   Therapon Skotiniotis,Ji-en Morris Chang 2002   9
        10   V. Martin,K. Schwan 1998   10
        11   Peter W. Moo,Xiaolin Wu 1999   11

                                                 paperTitle citation_count
        0   Static and Dynamic Metrics for Effective Objec...   3
        1   Who participates and why?: an analysis of citi...   9
        2   Call for papers   0
        3   Universal parametrization in constructing smoo...   1
        4   Data Allocation and Load Balancing for Heterog...   0
        5   Employment opinion: reversal of fortune   0
        6   From Coordination of Workflow and Group Activi...   2
        7   Modes of Comprehension: Mode Analysis of Array...   0
        8   Thesis: clustering and instance based learning...   0
        9   Estimating internal memory fragmentation for J...   0
        10 ILI: An Adaptive Infrastructure for Dynamic In...   2
        11 Resynchronization Properties of Arithmetic Coding   2

In [18]: test_df
```

```
Out[18]:                                              Authors       index id  \
        0                      Ahmed Osman, Wojciech Samek    1802.00209v1
        1                     Ji Young Lee, Franck Dernoncourt    1603.03827v1
        2     Iulian Vlad Serban, Tim Klinger, Gerald Tesau...    1606.00776v2
        3     Sebastian Ruder, Joachim Bingel, Isabelle Aug...    1705.08142v2
        4     Iulian V. Serban, Chinnadhurai Sankar, Mathie...    1709.02349v2
        5     Kelvin Guu, Tatsunori B. Hashimoto, Yonatan O...    1709.08878v1
        6     Iulian V. Serban, Chinnadhurai Sankar, Mathie...    1801.06700v1
        7     Darko Brodic, Alessia Amelio, Zoran N. Milivo...    1609.06492v1
        8                      Mateusz Malinowski, Mario Fritz    1610.01076v1
        9                                      Tony Beltramelli    1705.07962v2
        10      Fred Richardson, Douglas Reynolds, Najim Dehak    1504.00923v1
        11    Hieu Pham, Melody Y. Guan, Barret Zoph, Quoc ...    1802.03268v2
        12    Brenden M. Lake, Tomer D. Ullman, Joshua B. T...    1604.00289v3
        13                              Hao Wang, Dit-Yan Yeung    1604.01662v2
        14    Tejas D. Kulkarni, Karthik R. Narasimhan, Ard...    1604.06057v2
        15    Deepak Pathak, Ross Girshick, Piotr Dollár, T...    1612.06370v2
        16    Muhammad Ghifary, W. Bastiaan Kleijn, Mengjie...     1409.6041v1
        17    Lionel Pigou, Aäron van den Oord, Sander Diel...    1506.01911v3
        18                        Rakesh Achanta, Trevor Hastie    1509.05962v2
        19    Jeff Donahue, Philipp Krähenbühl, Trevor Darrell    1605.09782v7
        20                                   Zachary C. Lipton    1606.03490v3
        21    Sahil Garg, Irina Rish, Guillermo Cecchi, Aur...    1701.06106v2
        22                              Weifeng Ge, Yizhou Yu    1702.08690v2
        23    Tanmay Gupta, Kevin Shih, Saurabh Singh, Dere...    1704.00260v2
        24    Jan Hendrik Metzen, Mummadi Chaithanya Kumar,...    1704.05712v3
        25                          Quynh Nguyen, Matthias Hein    1704.08045v2
        26    Chris Donahue, Zachary C. Lipton, Akshay Bals...    1705.07904v3
        27         Mahesh Chandra Mukkamala, Matthias Hein    1706.05507v2
        28    Chunyuan Li, Hao Liu, Changyou Chen, Yunchen ...    1709.01215v2
        29     Mateusz Buda, Atsuto Maki, Maciej A. Mazurowski    1710.05381v1
        ..                                                ...             ...
        70    Petr Baudi, Jan Pichl, Tomá Vyskoil, Jan ...   1603.06127v4
        71    Jiatao Gu, Zhengdong Lu, Hang Li, Victor O. K...    1603.06393v3
        72    Iulian Vlad Serban, Alberto García-Durán, Cag...    1603.06807v2
        73    Chia-Wei Liu, Ryan Lowe, Iulian V. Serban, Mi...    1603.08023v2
        74    Iulian Vlad Serban, Alessandro Sordoni, Ryan ...    1605.06069v3
        75                                   Dirk Weissenborn    1606.03864v2
        76                       Marc Dymetman, Chunyang Xiao    1607.02467v2
        77      Ondrej Bajgar, Rudolf Kadlec, Jan Kleindienst    1610.00956v1
        78    James Bradbury, Stephen Merity, Caiming Xiong...    1611.01576v2
        79    Jakob N. Foerster, Justin Gilmer, Jan Chorows...    1611.09434v2
        80    Micha Daniluk, Tim Rocktäschel, Johannes Wel...    1702.04521v1
        81    Zhouhan Lin, Minwei Feng, Cicero Nogueira dos...    1703.03130v1
        82    Samuel Rönnqvist, Niko Schenk, Christian Chia...    1704.08092v1
        83    Lara J. Martin, Prithviraj Ammanabrolu, Xinyu...    1706.01331v3
        84            Tong Wang, Xingdi Yuan, Adam Trischler    1706.01450v1
        85    Wei Wen, Yuxiong He, Samyam Rajbhandari, Minj...    1709.05027v7
```

```
86             Huda Hakami, Danushka Bollegala, Hayashi Kohei  1709.06673v2
87    Zhengdong Lu, Haotian Cui, Xianggen Liu, Yuku...  1709.08853v4
88                                          Bin Bi, Hao Ma  1709.10204v2
89    Mirco Ravanelli, Philemon Brakel, Maurizio Om...  1710.00641v1
90    Baolin Peng, Xiujun Li, Jianfeng Gao, Jingjin...  1801.06176v1
91    Andrew L. Maas, Peng Qi, Ziang Xie, Awni Y. H...   1406.7806v2
92                                William Chan, Ian Lane  1504.01482v1
93                      David Krueger, Roland Memisevic  1511.08400v7
94    Noam Shazeer, Azalia Mirhoseini, Krzysztof Ma...  1701.06538v1
95       Yacine Jernite, Samuel R. Bowman, David Sontag  1705.00557v1
96                          Zhengyang Wang, Shuiwang Ji  1705.06824v1
97    Kyunghyun Cho, Bart van Merrienboer, Caglar G...   1406.1078v3
98             Zhiyuan Tang, Dong Wang, Zhiyong Zhang  1505.04630v5
99       Haim Sak, Andrew Senior, Françoise Beaufays   1402.1128v1

                                            paperTitle  Year  \
0     Dual Recurrent Attention Units for Visual Ques...  2018
1     Sequential Short-Text Classification with Recu...  2016
2     Multiresolution Recurrent Neural Networks: An ...  2016
3     Learning what to share between loosely related...  2017
4               A Deep Reinforcement Learning Chatbot  2017
5          Generating Sentences by Editing Prototypes  2017
6     A Deep Reinforcement Learning Chatbot (Short V...  2018
7     Document Image Coding and Clustering for Scrip...  2016
8     Tutorial on Answering Questions about Images w...  2016
9     pix2code: Generating Code from a Graphical Use...  2017
10    A Unified Deep Neural Network for Speaker and ...  2015
11    Efficient Neural Architecture Search via Param...  2018
12    Building Machines That Learn and Think Like Pe...  2016
13             Towards Bayesian Deep Learning: A Survey  2016
14    Hierarchical Deep Reinforcement Learning: Inte...  2016
15          Learning Features by Watching Objects Move  2016
16    Domain Adaptive Neural Networks for Object Rec...  2014
17    Beyond Temporal Pooling: Recurrence and Tempor...  2015
18          Telugu OCR Framework using Deep Learning  2015
19                         Adversarial Feature Learning  2016
20               The Mythos of Model Interpretability  2016
21    Neurogenesis-Inspired Dictionary Learning: Onl...  2017
22    Borrowing Treasures from the Wealthy: Deep Tra...  2017
23    Aligned Image-Word Representations Improve Ind...  2017
24    Universal Adversarial Perturbations Against Se...  2017
25    The loss surface of deep and wide neural networks  2017
26    Semantically Decomposing the Latent Spaces of ...  2017
27    Variants of RMSProp and Adagrad with Logarithm...  2017
28    ALICE: Towards Understanding Adversarial Learn...  2017
29    A systematic study of the class imbalance prob...  2017
..                                                 ...   ...
70    Sentence Pair Scoring: Towards Unified Framewo...  2016
```

```
71  Incorporating Copying Mechanism in Sequence-to...    2016
72  Generating Factoid Questions With Recurrent Ne...    2016
73  How NOT To Evaluate Your Dialogue System: An E...    2016
74  A Hierarchical Latent Variable Encoder-Decoder...    2016
75  Neural Associative Memory for Dual-Sequence Mo...    2016
76  Log-Linear RNNs: Towards Recurrent Neural Netw...    2016
77  Embracing data abundance: BookTest Dataset for...    2016
78                       Quasi-Recurrent Neural Networks    2016
79  Input Switched Affine Networks: An RNN Archite...    2016
80  Frustratingly Short Attention Spans in Neural ...    2017
81      A Structured Self-attentive Sentence Embedding    2017
82  A Recurrent Neural Model with Attention for th...    2017
83  Event Representations for Automated Story Gene...    2017
84  A Joint Model for Question Answering and Quest...    2017
85  Learning Intrinsic Sparse Structures within Lo...    2017
86  Why PairDiff works? -- A Mathematical Analysis...    2017
87  Object-oriented Neural Programming (OONP) for ...    2017
88  A Neural Comprehensive Ranker (NCR) for Open-D...    2017
89  Improving speech recognition by revising gated...    2017
90  Integrating planning for task-completion dialo...    2018
91  Building DNN Acoustic Models for Large Vocabul...    2014
92  Deep Recurrent Neural Networks for Acoustic Mo...    2015
93      Regularizing RNNs by Stabilizing Activations    2015
94  Outrageously Large Neural Networks: The Sparse...    2017
95  Discourse-Based Objectives for Fast Unsupervis...    2017
96  Learning Convolutional Text Representations fo...    2017
97  Learning Phrase Representations using RNN Enco...    2014
98  Recurrent Neural Network Training with Dark Kn...    2015
99  Long Short-Term Memory Based Recurrent Neural ...    2014


                                         Abstract  citation_count
0   We propose an architecture for VQA which utili...              0
1   Recent approaches based on artificial neural n...              0
2   We introduce the multiresolution recurrent neu...              0
3   Multi-task learning is motivated by the observ...              0
4   We present MILABOT: a deep reinforcement learn...              0
5   We propose a new generative model of sentences...              0
6   We present MILABOT: a deep reinforcement learn...              0
7   The paper introduces a new method for discrimi...              0
8   Together with the development of more accurate...              0
9   Transforming a graphical user interface screen...              0
10  Learned feature representations and sub-phonem...              0
11  We propose Efficient Neural Architecture Searc...              0
12  Recent progress in artificial intelligence (AI...              0
13  While perception tasks such as visual object r...              0
14  Learning goal-directed behavior in environment...              0
15  This paper presents a novel yet intuitive appr...              0
16  We propose a simple neural network model to de...              0
```

```
17  Recent studies have demonstrated the power of ...          0
18  In this paper, we address the task of Optical ...          0
19  The ability of the Generative Adversarial Netw...          0
20  Supervised machine learning models boast remar...          0
21  In this paper, we focus on online representati...          0
22  Deep neural networks require a large amount of...          0
23  An important goal of computer vision is to bui...          0
24  While deep learning is remarkably successful o...          0
25  While the optimization problem behind deep neu...          0
26  We propose a new algorithm for training genera...          0
27  Adaptive gradient methods have become recently...          0
28  We investigate the non-identifiability issues ...          0
29  In this study, we systematically investigate t...          0
..                                                ...        ...
70  We review the task of Sentence Pair Scoring, p...          0
71  We address an important problem in sequence-to...          0
72  Over the past decade, large-scale supervised l...          0
73  We investigate evaluation metrics for dialogue...          0
74  Sequential data often possesses a hierarchical...          0
75  Many important NLP problems can be posed as du...          0
76  We introduce LL-RNNs (Log-Linear RNNs), an ext...          0
77  There is a practically unlimited amount of nat...          0
78  Recurrent neural networks are a powerful tool ...          0
79  There exist many problem domains where the int...          0
80  Neural language models predict the next token ...          0
81  This paper proposes a new model for extracting...          0
82  We introduce an attention-based Bi-LSTM for Ch...          0
83  Automated story generation is the problem of a...          0
84  We propose a generative machine comprehension ...          0
85  Model compression is significant for the wide ...          0
86  Representing the semantic relations that exist...          0
87  We propose Object-oriented Neural Programming ...          0
88  This paper proposes a novel neural machine rea...          0
89  Speech recognition is largely taking advantage...          0
90  Training a task-completion dialogue agent with...          0
91  Deep neural networks (DNNs) are now a central ...          0
92  We present a novel deep Recurrent Neural Netwo...          0
93  We stabilize the activations of Recurrent Neur...          0
94  The capacity of a neural network to absorb inf...          0
95  This work presents a novel objective function ...          0
96  Visual question answering is a recently propos...          0
97  In this paper, we propose a novel neural netwo...          0
98  Recurrent neural networks (RNNs), particularly...          0
99  Long Short-Term Memory (LSTM) is a recurrent n...          0

[100 rows x 6 columns]
```

4 Get Domain of each paper using Scumpy based on NER

```
In [19]: domain_list = ["Biomedical Research", "Chemistry", "Biology", "Economics", "Earth Scie
                        "Political", "Economics", "Literature", "Computer Science", "Software"
                        "Physics", "History", "Art", "Statistics", "Business", "Health", "Geome
```

```
In [20]: def get_domain(string):
             valid_domain = 'N/A'
             max_score = 0
             word1 = nlp(string)
             for domain in domain_list:
                 word2 = nlp(domain)
                 score = word1.similarity(word2)
                 if max_score < score:
                     max_score = score
                     valid_domain = domain

             return valid_domain
```

```
In [21]: def fetch_domain_for_each_paper(df):
             domain_list1 = []
             for index,row in df.iterrows():
                 author_domain = get_domain(row["paperTitle"])
                 domain_list1.append(author_domain)
             df["domain"] = domain_list1
             return df
```

```
In [22]: train_df = fetch_domain_for_each_paper(train_df)
```

```
In [23]: test_df = fetch_domain_for_each_paper(test_df)
```

```
In [24]: test_df
```

```
Out[24]:                                          Authors      index id  \
         0                 Ahmed Osman, Wojciech Samek  1802.00209v1
         1               Ji Young Lee, Franck Dernoncourt  1603.03827v1
         2    Iulian Vlad Serban, Tim Klinger, Gerald Tesau...  1606.00776v2
         3    Sebastian Ruder, Joachim Bingel, Isabelle Aug...  1705.08142v2
         4    Iulian V. Serban, Chinnadhurai Sankar, Mathie...  1709.02349v2
         5    Kelvin Guu, Tatsunori B. Hashimoto, Yonatan O...  1709.08878v1
         6    Iulian V. Serban, Chinnadhurai Sankar, Mathie...  1801.06700v1
         7    Darko Brodic, Alessia Amelio, Zoran N. Milivo...  1609.06492v1
         8                 Mateusz Malinowski, Mario Fritz  1610.01076v1
         9                              Tony Beltramelli  1705.07962v2
         10    Fred Richardson, Douglas Reynolds, Najim Dehak  1504.00923v1
         11   Hieu Pham, Melody Y. Guan, Barret Zoph, Quoc ...  1802.03268v2
         12   Brenden M. Lake, Tomer D. Ullman, Joshua B. T...  1604.00289v3
         13                          Hao Wang, Dit-Yan Yeung  1604.01662v2
         14   Tejas D. Kulkarni, Karthik R. Narasimhan, Ard...  1604.06057v2
```

```
15   Deepak Pathak, Ross Girshick, Piotr Dollár, T...   1612.06370v2
16   Muhammad Ghifary, W. Bastiaan Kleijn, Mengjie...    1409.6041v1
17   Lionel Pigou, Aäron van den Oord, Sander Diel...   1506.01911v3
18                         Rakesh Achanta, Trevor Hastie   1509.05962v2
19   Jeff Donahue, Philipp Krähenbühl, Trevor Darrell   1605.09782v7
20                                    Zachary C. Lipton   1606.03490v3
21   Sahil Garg, Irina Rish, Guillermo Cecchi, Aur...   1701.06106v2
22                                   Weifeng Ge, Yizhou Yu   1702.08690v2
23   Tanmay Gupta, Kevin Shih, Saurabh Singh, Dere...   1704.00260v2
24   Jan Hendrik Metzen, Mummadi Chaithanya Kumar,...   1704.05712v3
25                          Quynh Nguyen, Matthias Hein   1704.08045v2
26   Chris Donahue, Zachary C. Lipton, Akshay Bals...   1705.07904v3
27           Mahesh Chandra Mukkamala, Matthias Hein   1706.05507v2
28   Chunyuan Li, Hao Liu, Changyou Chen, Yunchen ...   1709.01215v2
29    Mateusz Buda, Atsuto Maki, Maciej A. Mazurowski   1710.05381v1
..                                              ...         ...
70   Petr Baudi, Jan Pichl, Tomá Vyskoil, Jan ...   1603.06127v4
71   Jiatao Gu, Zhengdong Lu, Hang Li, Victor O. K...   1603.06393v3
72   Iulian Vlad Serban, Alberto García-Durán, Cag...   1603.06807v2
73   Chia-Wei Liu, Ryan Lowe, Iulian V. Serban, Mi...   1603.08023v2
74   Iulian Vlad Serban, Alessandro Sordoni, Ryan ...   1605.06069v3
75                                    Dirk Weissenborn   1606.03864v2
76                         Marc Dymetman, Chunyang Xiao   1607.02467v2
77     Ondrej Bajgar, Rudolf Kadlec, Jan Kleindienst   1610.00956v1
78   James Bradbury, Stephen Merity, Caiming Xiong...   1611.01576v2
79   Jakob N. Foerster, Justin Gilmer, Jan Chorows...   1611.09434v2
80   Micha Daniluk, Tim Rocktäschel, Johannes Wel...   1702.04521v1
81   Zhouhan Lin, Minwei Feng, Cicero Nogueira dos...   1703.03130v1
82   Samuel Rönnqvist, Niko Schenk, Christian Chia...   1704.08092v1
83   Lara J. Martin, Prithviraj Ammanabrolu, Xinyu...   1706.01331v3
84           Tong Wang, Xingdi Yuan, Adam Trischler   1706.01450v1
85   Wei Wen, Yuxiong He, Samyam Rajbhandari, Minj...   1709.05027v7
86     Huda Hakami, Danushka Bollegala, Hayashi Kohei   1709.06673v2
87   Zhengdong Lu, Haotian Cui, Xianggen Liu, Yuku...   1709.08853v4
88                                       Bin Bi, Hao Ma   1709.10204v2
89   Mirco Ravanelli, Philemon Brakel, Maurizio Om...   1710.00641v1
90   Baolin Peng, Xiujun Li, Jianfeng Gao, Jingjin...   1801.06176v1
91   Andrew L. Maas, Peng Qi, Ziang Xie, Awni Y. H...    1406.7806v2
92                                   William Chan, Ian Lane   1504.01482v1
93               David Krueger, Roland Memisevic   1511.08400v7
94   Noam Shazeer, Azalia Mirhoseini, Krzysztof Ma...   1701.06538v1
95     Yacine Jernite, Samuel R. Bowman, David Sontag   1705.00557v1
96                         Zhengyang Wang, Shuiwang Ji   1705.06824v1
97   Kyunghyun Cho, Bart van Merrienboer, Caglar G...    1406.1078v3
98           Zhiyuan Tang, Dong Wang, Zhiyong Zhang   1505.04630v5
99     Haim Sak, Andrew Senior, Françoise Beaufays   1402.1128v1

                                          paperTitle  Year  \
```

```
0   Dual Recurrent Attention Units for Visual Ques...   2018
1   Sequential Short-Text Classification with Recu...   2016
2   Multiresolution Recurrent Neural Networks: An ...   2016
3   Learning what to share between loosely related...   2017
4                   A Deep Reinforcement Learning Chatbot   2017
5            Generating Sentences by Editing Prototypes   2017
6   A Deep Reinforcement Learning Chatbot (Short V...   2018
7   Document Image Coding and Clustering for Scrip...   2016
8   Tutorial on Answering Questions about Images w...   2016
9   pix2code: Generating Code from a Graphical Use...   2017
10  A Unified Deep Neural Network for Speaker and ...   2015
11  Efficient Neural Architecture Search via Param...   2018
12  Building Machines That Learn and Think Like Pe...   2016
13             Towards Bayesian Deep Learning: A Survey   2016
14  Hierarchical Deep Reinforcement Learning: Inte...   2016
15           Learning Features by Watching Objects Move   2016
16  Domain Adaptive Neural Networks for Object Rec...   2014
17  Beyond Temporal Pooling: Recurrence and Tempor...   2015
18            Telugu OCR Framework using Deep Learning   2015
19                          Adversarial Feature Learning   2016
20               The Mythos of Model Interpretability   2016
21  Neurogenesis-Inspired Dictionary Learning: Onl...   2017
22  Borrowing Treasures from the Wealthy: Deep Tra...   2017
23  Aligned Image-Word Representations Improve Ind...   2017
24  Universal Adversarial Perturbations Against Se...   2017
25  The loss surface of deep and wide neural networks   2017
26  Semantically Decomposing the Latent Spaces of ...   2017
27  Variants of RMSProp and Adagrad with Logarithm...   2017
28  ALICE: Towards Understanding Adversarial Learn...   2017
29  A systematic study of the class imbalance prob...   2017
..                                               ...   ...
70  Sentence Pair Scoring: Towards Unified Framewo...   2016
71  Incorporating Copying Mechanism in Sequence-to...   2016
72  Generating Factoid Questions With Recurrent Ne...   2016
73  How NOT To Evaluate Your Dialogue System: An E...   2016
74  A Hierarchical Latent Variable Encoder-Decoder...   2016
75  Neural Associative Memory for Dual-Sequence Mo...   2016
76  Log-Linear RNNs: Towards Recurrent Neural Netw...   2016
77  Embracing data abundance: BookTest Dataset for...   2016
78                    Quasi-Recurrent Neural Networks   2016
79  Input Switched Affine Networks: An RNN Archite...   2016
80  Frustratingly Short Attention Spans in Neural ...   2017
81    A Structured Self-attentive Sentence Embedding   2017
82  A Recurrent Neural Model with Attention for th...   2017
83  Event Representations for Automated Story Gene...   2017
84  A Joint Model for Question Answering and Quest...   2017
85  Learning Intrinsic Sparse Structures within Lo...   2017
86  Why PairDiff works? -- A Mathematical Analysis...   2017
```

```
87   Object-oriented Neural Programming (OONP) for ...   2017
88   A Neural Comprehensive Ranker (NCR) for Open-D...   2017
89   Improving speech recognition by revising gated...   2017
90   Integrating planning for task-completion dialo...   2018
91   Building DNN Acoustic Models for Large Vocabul...   2014
92   Deep Recurrent Neural Networks for Acoustic Mo...   2015
93        Regularizing RNNs by Stabilizing Activations   2015
94   Outrageously Large Neural Networks: The Sparse...   2017
95   Discourse-Based Objectives for Fast Unsupervis...   2017
96   Learning Convolutional Text Representations fo...   2017
97   Learning Phrase Representations using RNN Enco...   2014
98   Recurrent Neural Network Training with Dark Kn...   2015
99   Long Short-Term Memory Based Recurrent Neural ...   2014


                                    Abstract   citation_count  \
0    We propose an architecture for VQA which utili...              0
1    Recent approaches based on artificial neural n...              0
2    We introduce the multiresolution recurrent neu...              0
3    Multi-task learning is motivated by the observ...              0
4    We present MILABOT: a deep reinforcement learn...              0
5    We propose a new generative model of sentences...              0
6    We present MILABOT: a deep reinforcement learn...              0
7    The paper introduces a new method for discrimi...              0
8    Together with the development of more accurate...              0
9    Transforming a graphical user interface screen...              0
10   Learned feature representations and sub-phonem...              0
11   We propose Efficient Neural Architecture Searc...              0
12   Recent progress in artificial intelligence (AI...              0
13   While perception tasks such as visual object r...              0
14   Learning goal-directed behavior in environment...              0
15   This paper presents a novel yet intuitive appr...              0
16   We propose a simple neural network model to de...              0
17   Recent studies have demonstrated the power of ...              0
18   In this paper, we address the task of Optical ...              0
19   The ability of the Generative Adversarial Netw...              0
20   Supervised machine learning models boast remar...              0
21   In this paper, we focus on online representati...              0
22   Deep neural networks require a large amount of...              0
23   An important goal of computer vision is to bui...              0
24   While deep learning is remarkably successful o...              0
25   While the optimization problem behind deep neu...              0
26   We propose a new algorithm for training genera...              0
27   Adaptive gradient methods have become recently...              0
28   We investigate the non-identifiability issues ...              0
29   In this study, we systematically investigate t...              0
..                                         ...            ...
70   We review the task of Sentence Pair Scoring, p...              0
71   We address an important problem in sequence-to...              0
```

```
72  Over the past decade, large-scale supervised l...        0
73  We investigate evaluation metrics for dialogue...        0
74  Sequential data often possesses a hierarchical...        0
75  Many important NLP problems can be posed as du...        0
76  We introduce LL-RNNs (Log-Linear RNNs), an ext...       0
77  There is a practically unlimited amount of nat...        0
78  Recurrent neural networks are a powerful tool ...        0
79  There exist many problem domains where the int...        0
80  Neural language models predict the next token ...        0
81  This paper proposes a new model for extracting...        0
82  We introduce an attention-based Bi-LSTM for Ch...        0
83  Automated story generation is the problem of a...        0
84  We propose a generative machine comprehension ...        0
85  Model compression is significant for the wide ...        0
86  Representing the semantic relations that exist...        0
87  We propose Object-oriented Neural Programming ...        0
88  This paper proposes a novel neural machine rea...        0
89  Speech recognition is largely taking advantage...        0
90  Training a task-completion dialogue agent with...        0
91  Deep neural networks (DNNs) are now a central ...        0
92  We present a novel deep Recurrent Neural Netwo...        0
93  We stabilize the activations of Recurrent Neur...        0
94  The capacity of a neural network to absorb inf...        0
95  This work presents a novel objective function ...        0
96  Visual question answering is a recently propos...        0
97  In this paper, we propose a novel neural netwo...        0
98  Recurrent neural networks (RNNs), particularly...        0
99  Long Short-Term Memory (LSTM) is a recurrent n...        0

                  domain
0        Computer Science
1        Computer Science
2        Computer Science
3        Computer Science
4        Computer Science
5        Computer Science
6        Computer Science
7                Software
8        Computer Science
9                Software
10       Computer Science
11               Software
12       Computer Science
13       Computer Science
14           Neuroscience
15       Computer Science
16       Computer Science
17       Computer Science
```

```
18      Computer Science
19              Political
20                History
21      Computer Science
22      Computer Science
23      Computer Science
24              Political
25         Earth Science
26              Political
27         Earth Science
28      Computer Science
29      Computer Science
..                   ...
70      Computer Science
71      Computer Science
72      Computer Science
73      Computer Science
74              Software
75      Computer Science
76      Computer Science
77      Computer Science
78  Biomedical Research
79      Computer Science
80      Computer Science
81              Business
82              Political
83      Computer Science
84      Computer Science
85      Computer Science
86      Computer Science
87      Computer Science
88      Computer Science
89                Health
90              Business
91      Computer Science
92      Computer Science
93              Software
94      Computer Science
95      Computer Science
96      Computer Science
97              Software
98      Computer Science
99      Computer Science

[100 rows x 7 columns]

In [25]: paper_domain_dict = {}
         for index, rows in train_df.iterrows():
```

```
        domain = rows["domain"]
        if domain in paper_domain_dict:
            paper_domain_dict[domain] += 1
        else:
            paper_domain_dict[domain] = 1
```

5 Generate Author Table

```
In [26]: author_papers = {}
         author_citations = {}
         author_citations_list = {}
         author_domain_list = {}

         for index,row in train_df.iterrows():
             authors = row["Authors"]
             authors_list = authors.split(',')
             for author in authors_list:
                 if author in author_papers:
                     author_papers[author] += 1
                 else:
                     author_papers[author] = 1
                 if author in author_citations:
                     author_citations[author] += row["citation_count"]
                 else:
                     author_citations[author] = row["citation_count"]
                 if author in author_citations_list:
                     author_citations_list[author].append(row["citation_count"])
                 else:
                     author_citations_list[author] = [row["citation_count"]]
                 if author not in author_domain_list:
                     author_domain_list[author] = row["domain"]

In [27]: def compute_h_index(citations):
             hIndex = 0
             citations.sort(reverse=True)
             for i in range(len(citations)):
                 if citations[i] >= i + 1:
                     hIndex = i+1
             return hIndex

In [28]: def compute_i10_index(citations):
             count = 0;
             for i in range(len(citations)):
                 if citations[i] >= 10:
                     count = count + 1
             return count;

In [29]: def compute_g_index(citations, avg_citations):
             count = 0;
```

```
            for i in range(len(citations)):
                if citations[i] >= avg_citations:
                    count = count + 1
            return count;

In [30]: def generate_author_table():
             authors_list = []
             citation_count_list = []
             paper_count_list = []
             h_index_list = []
             g_index_list = []
             i10_index_list = []
             domain_list = []
             for key in author_citations:
                 authors_list.append(key)
                 citation_count_list.append(author_citations[key])
                 h_index_list.append(compute_h_index(author_citations_list[key]))
                 i10_index_list.append(compute_i10_index(author_citations_list[key]))
                 g_index_list.append(compute_g_index(author_citations_list[key], author_citati
                 paper_count_list.append(author_papers[key])
                 domain_list.append(author_domain_list[key])

             df_new = pd.DataFrame()
             df_new["author_name"] = authors_list
             df_new["citation_count"] = citation_count_list
             df_new["paper_count"] = paper_count_list
             df_new["h_index"] = h_index_list
             df_new["i10_index"] = i10_index_list
             df_new["g_index"] = g_index_list
             df_new["domain"] = domain_list
             return df_new

         df_n = generate_author_table()

In [31]: df_n = df_n.sort_values(by='h_index', ascending=False)
         df_n["average_citations"] = df_n["citation_count"]/df_n["paper_count"]

In [32]: score_list = []
         for index, rows in df_n.iterrows():
             count = rows['paper_count']
             domain = rows["domain"]
             total_count = paper_domain_dict[domain]
             score_list.append(count / total_count)
         df_n['domain_Score'] = score_list
         df_n
```

```
Out[32]:                    author_name  citation_count  paper_count  h_index  \
         4578        Christoph Weidenbach              93            5        5
         4073                 Amir Pnueli              32            6        4
```

3834	Ming C. Lin	51	4	4
814	Harald Ganzinger	29	6	4
6058	Frank Pfenning	150	7	4
3026	George L. Nemhauser	62	6	4
815	Andrei Voronkov	42	10	4
5885	Salil Vadhan	43	4	4
8114	Marc Joye	29	4	4
2701	David Pisinger	46	4	4
3040	Michael J. Black	53	4	4
2702	Paolo Toth	44	5	4
3320	David Blaauw	77	5	4
4068	Joachim Weickert	19	5	4
5192	Yuxin Chen	23	4	3
1119	B. Maggs	11	3	3
6363	Douglas W. Oard	23	3	3
2123	Moshe Y. Vardi	22	3	3
3773	Gerth Stølting Brodal	26	4	3
11803	Peter F. Patel-Schneider	34	4	3
1595	Robert K. Brayton	29	3	3
11802	Ian Horrocks	37	4	3
2881	Gilles Bertrand	20	5	3
8743	Jiawei Han	19	4	3
423	Rolf Ernst	45	5	3
6105	I. Walukiewicz	31	4	3
2163	S. Akamatsu	73	3	3
11797	Iliano Cervesato	40	3	3
2418	Cynthia Barnhart	28	3	3
2804	M. Rebaudengo	20	6	3
...	...	...	...	...
12225	Khanh Nguyen-Phi	0	1	0
12226	Hans Weinrichterr	0	1	0
12228	Elena V. Zudilova	0	1	0
12229	Larry Greenemeier	0	2	0
12230	Anthony Lauder	0	1	0
12231	J. Hong	0	1	0
12232	D. Kim	0	1	0
12233	Chisa Sumitomo	0	1	0
12208	Bruce Kogut	0	1	0
12206	Marcelo Salinas	0	1	0
12181	Harris G. Makatsoris	0	1	0
12205	Eduardo S. Schwartz	0	1	0
4344	Ishfaq Ahmad	0	2	0
12184	Vincent Y. Lum	0	1	0
12185	F. W. Lee	0	1	0
12186	S. K. Fong	0	1	0
4342	Magdi M. A. Morsi	0	1	0
12188	Travis E. Doom	0	2	0
12189	Anthony S. Wojcik	0	1	0

12190	Moon-Jung Chung	0	1	0
12191	M. Chang	0	1	0
12192	P. Chou	0	1	0
12193	H. Essafi	0	1	0
12194	C. Mazzoni	0	1	0
12195	P. Julien	0	1	0
12196	O. Jamet	0	1	0
4341	Teade Punter	0	1	0
4336	Paul Syverson	0	1	0
12204	Gonzalo Cortazar	0	1	0
10099	Ron Anderson	0	1	0

	i10_index	g_index	domain	average_citations	domain_Score
4578	4	3	Software	18.600000	0.004955
4073	1	2	Computer Science	5.333333	0.001162
3834	4	2	Computer Science	12.750000	0.000775
814	0	3	Political	4.833333	0.022642
6058	3	2	Geometry	21.428571	0.012658
3026	3	2	Business	10.333333	0.005720
815	2	4	Political	4.200000	0.037736
5885	1	1	Computer Science	10.750000	0.000775
8114	0	2	Geometry	7.250000	0.007233
2701	3	2	Computer Science	11.500000	0.000775
3040	3	2	Computer Science	13.250000	0.000775
2702	3	3	Computer Science	8.800000	0.000968
3320	2	2	Software	15.400000	0.004955
4068	0	4	Political	3.800000	0.018868
5192	1	2	Business	5.750000	0.003813
1119	0	2	Software	3.666667	0.002973
6363	1	1	History	7.666667	0.010417
2123	1	1	Political	7.333333	0.011321
3773	1	2	Geometry	6.500000	0.007233
11803	2	2	Software	8.500000	0.003964
1595	1	1	Computer Science	9.666667	0.000581
11802	2	2	Software	9.250000	0.003964
2881	1	2	Computer Science	4.000000	0.000968
8743	0	2	Computer Science	4.750000	0.000775
423	1	1	Computer Science	9.000000	0.000968
6105	1	2	Computer Science	7.750000	0.000775
2163	2	2	Computer Science	24.333333	0.000581
11797	2	2	Computer Science	13.333333	0.000581
2418	1	1	Business	9.333333	0.002860
2804	0	2	Computer Science	3.333333	0.001162
...	...	...	...	...	...
12225	0	1	Computer Science	0.000000	0.000194
12226	0	1	Computer Science	0.000000	0.000194
12228	0	1	Computer Science	0.000000	0.000194
12229	0	2	Software	0.000000	0.001982

```
12230         0         1          History      0.000000      0.003472
12231         0         1         Software      0.000000      0.000991
12232         0         1         Software      0.000000      0.000991
12233         0         1  Computer Science      0.000000      0.000194
12208         0         1         Business      0.000000      0.000953
12206         0         1         Business      0.000000      0.000953
12181         0         1         Business      0.000000      0.000953
12205         0         1         Business      0.000000      0.000953
4344          0         2  Computer Science      0.000000      0.000387
12184         0         1         Software      0.000000      0.000991
12185         0         1         Software      0.000000      0.000991
12186         0         1         Software      0.000000      0.000991
4342          0         1  Computer Science      0.000000      0.000194
12188         0         2  Computer Science      0.000000      0.000387
12189         0         1  Computer Science      0.000000      0.000194
12190         0         1  Computer Science      0.000000      0.000194
12191         0         1  Computer Science      0.000000      0.000194
12192         0         1  Computer Science      0.000000      0.000194
12193         0         1  Computer Science      0.000000      0.000194
12194         0         1  Computer Science      0.000000      0.000194
12195         0         1  Computer Science      0.000000      0.000194
12196         0         1  Computer Science      0.000000      0.000194
4341          0         1         Software      0.000000      0.000991
4336          0         1  Computer Science      0.000000      0.000194
12204         0         1         Business      0.000000      0.000953
10099         0         1  Computer Science      0.000000      0.000194

[20198 rows x 9 columns]
```

6 Calculate Author Metric

```
In [33]: df_n["SVS - index"] = 0.3*df_n["average_citations"] + 0.3*df_n["paper_count"] + 0.5*d

In [34]: df_n = df_n.sort_values(by='SVS - index', ascending=False)

In [35]: df_n

Out[35]:                    author_name  citation_count  paper_count  h_index  i10_index  \
         10599     Sepandar D. Kamvar             258            1        1          1
         10600     Mario T. Schlosser             258            1        1          1
         323     Hector Garcia-Molina             265            2        2          1
         7                        N/A               1          501        1          0
         13554         Alon Y. Halevy             168            1        1          1
         19674            Kushal Dave             122            1        1          1
         14666           Matteo Frigo             111            1        1          1
         14667     Charles E. Leiserson           111            1        1          1
         14668          Harald Prokop            111            1        1          1
         14669     Sridhar Ramachandran           111            1        1          1
```

5301	Guillaume Brat	110	1	1	1
5300	Klaus Havelund	110	1	1	1
5299	Willem Visser	110	1	1	1
5302	SeungJoon Park	110	1	1	1
19676	David M. Pennock	130	2	2	1
19675	Steve Lawrence	130	2	2	1
4432	Hantao Zhang	148	5	3	1
6058	Frank Pfenning	150	7	4	3
15726	Carsten Schürmann	102	1	1	1
532	Gavin Lowe	131	3	3	3
9672	Markus P. J. Fromherz	98	1	1	1
9671	Ying Zhang	98	1	1	1
9670	Wheeler Ruml	98	1	1	1
10660	Sebastian Thrun	132	4	3	2
14641	H. V. Jagadish	116	2	2	2
17337	N. Wiwatwattana	93	1	1	1
17331	A. Chapman	93	1	1	1
17332	L. V. S. Lakshmanan	93	1	1	1
17333	A. Nierman	93	1	1	1
17334	S. Paparizos	93	1	1	1
...	...	...	...	...	...
9938	Zhenghui Lin	0	1	0	0
9937	Haifeng Zhou	0	1	0	0
5599	Eui Park	0	1	0	0
5600	D. Strickland	0	1	0	0
9978	M. Laws	0	1	0	0
9877	Raymon Paul	0	1	0	0
9771	A. K. Noor	0	1	0	0
9770	Gary B. Parker	0	1	0	0
2103	Masato Nose	0	1	0	0
9820	Clara Bagnasco	0	1	0	0
9875	Husni Fahmi	0	1	0	0
9874	Sahra Sedigh	0	1	0	0
9873	Shahab Baqai	0	1	0	0
9872	Ahmed Bashandy	0	1	0	0
9871	Fujiwa Kato	0	1	0	0
9867	Jaedong Yang	0	1	0	0
9866	Hoyoung Kim	0	1	0	0
9865	Hyungjeong Yang	0	1	0	0
2104	Hideki Sawada	0	1	0	0
9750	S. Roth-Koch	0	1	0	0
5655	O. Wigertz	0	1	0	0
5656	András Méhes	0	1	0	0
5690	P. L. Palmer	0	1	0	0
5691	N. Fatemi-Ghomi	0	1	0	0
9763	Jennifer S. Lund	0	1	0	0
9762	Niall McLoughlin	0	1	0	0
9761	Martin Stetter	0	1	0	0

9760	Ingo Schiel	0	1	0	0
5697	Bernard Faure	0	1	0	0
10099	Ron Anderson	0	1	0	0

	g_index	domain	average_citations	domain_Score	SVS - index
10599	1	Software	258.000000	0.000991	207.200248
10600	1	Software	258.000000	0.000991	207.200248
323	1	Computer Science	132.500000	0.000387	173.350097
7	1	Computer Science	0.001996	0.097037	151.074858
13554	1	Computer Science	168.000000	0.000194	135.200048
19674	1	History	122.000000	0.003472	98.400868
14666	1	Microprocessor	111.000000	0.006173	89.601543
14667	1	Microprocessor	111.000000	0.006173	89.601543
14668	1	Microprocessor	111.000000	0.006173	89.601543
14669	1	Microprocessor	111.000000	0.006173	89.601543
5301	1	Computer Science	110.000000	0.000194	88.800048
5300	1	Computer Science	110.000000	0.000194	88.800048
5299	1	Computer Science	110.000000	0.000194	88.800048
5302	1	Computer Science	110.000000	0.000194	88.800048
19676	1	History	65.000000	0.006944	85.601736
19675	1	History	65.000000	0.006944	85.601736
4432	1	Geometry	29.600000	0.009042	84.882260
6058	2	Geometry	21.428571	0.012658	84.781736
15726	1	Computer Science	102.000000	0.000194	82.400048
532	1	Computer Science	43.666667	0.000581	80.500145
9672	1	Computer Science	98.000000	0.000194	79.200048
9671	1	Computer Science	98.000000	0.000194	79.200048
9670	1	Computer Science	98.000000	0.000194	79.200048
10660	2	Business	33.000000	0.003813	78.100953
14641	1	History	58.000000	0.006944	76.751736
17337	1	Software	93.000000	0.000991	75.200248
17331	1	Software	93.000000	0.000991	75.200248
17332	1	Software	93.000000	0.000991	75.200248
17333	1	Software	93.000000	0.000991	75.200248
17334	1	Software	93.000000	0.000991	75.200248
...	...	...	...	...	...
9938	1	Computer Science	0.000000	0.000194	0.550048
9937	1	Computer Science	0.000000	0.000194	0.550048
5599	1	Computer Science	0.000000	0.000194	0.550048
5600	1	Computer Science	0.000000	0.000194	0.550048
9978	1	Computer Science	0.000000	0.000194	0.550048
9877	1	Computer Science	0.000000	0.000194	0.550048
9771	1	Computer Science	0.000000	0.000194	0.550048
9770	1	Computer Science	0.000000	0.000194	0.550048
2103	1	Computer Science	0.000000	0.000194	0.550048
9820	1	Computer Science	0.000000	0.000194	0.550048
9875	1	Computer Science	0.000000	0.000194	0.550048
9874	1	Computer Science	0.000000	0.000194	0.550048

```
9873      1  Computer Science       0.000000    0.000194    0.550048
9872      1  Computer Science       0.000000    0.000194    0.550048
9871      1  Computer Science       0.000000    0.000194    0.550048
9867      1  Computer Science       0.000000    0.000194    0.550048
9866      1  Computer Science       0.000000    0.000194    0.550048
9865      1  Computer Science       0.000000    0.000194    0.550048
2104      1  Computer Science       0.000000    0.000194    0.550048
9750      1  Computer Science       0.000000    0.000194    0.550048
5655      1  Computer Science       0.000000    0.000194    0.550048
5656      1  Computer Science       0.000000    0.000194    0.550048
5690      1  Computer Science       0.000000    0.000194    0.550048
5691      1  Computer Science       0.000000    0.000194    0.550048
9763      1  Computer Science       0.000000    0.000194    0.550048
9762      1  Computer Science       0.000000    0.000194    0.550048
9761      1  Computer Science       0.000000    0.000194    0.550048
9760      1  Computer Science       0.000000    0.000194    0.550048
5697      1  Computer Science       0.000000    0.000194    0.550048
10099     1  Computer Science       0.000000    0.000194    0.550048

[20198 rows x 10 columns]
```

7   Ques 1 : Top 100 ranked researchers from multiple disciplines

```
In [36]:  # You must identify top 100 ranked researchers from multiple disciplines based on you
          # and see how it stacks up against their respective h-index.

          df_n.head(500)

Out[36]:                   author_name  citation_count  paper_count  h_index  \
          10599      Sepandar D. Kamvar             258            1        1
          10600      Mario T. Schlosser            258            1        1
          323      Hector Garcia-Molina            265            2        2
          7                        N/A               1          501        1
          13554        Alon Y. Halevy             168            1        1
          19674           Kushal Dave             122            1        1
          14666          Matteo Frigo             111            1        1
          14667     Charles E. Leiserson           111            1        1
          14668         Harald Prokop             111            1        1
          14669     Sridhar Ramachandran           111            1        1
          5301         Guillaume Brat             110            1        1
          5300         Klaus Havelund             110            1        1
          5299          Willem Visser             110            1        1
          5302          SeungJoon Park            110            1        1
          19676       David M. Pennock            130            2        2
          19675        Steve Lawrence             130            2        2
          4432            Hantao Zhang            148            5        3
          6058           Frank Pfenning           150            7        4
```

15726	Carsten Schürmann	102	1	1
532	Gavin Lowe	131	3	3
9672	Markus P. J. Fromherz	98	1	1
9671	Ying Zhang	98	1	1
9670	Wheeler Ruml	98	1	1
10660	Sebastian Thrun	132	4	3
14641	H. V. Jagadish	116	2	2
17337	N. Wiwatwattana	93	1	1
17331	A. Chapman	93	1	1
17332	L. V. S. Lakshmanan	93	1	1
17333	A. Nierman	93	1	1
17334	S. Paparizos	93	1	1
...	...	...	...	...
4581	Mark Vandenbosch	29	1	1
181	Kate A. Smith	29	1	1
4580	John S. Hulland	29	1	1
4579	Christopher R. Plouffe	29	1	1
2856	Philip K. Chan	29	1	1
18794	Craig Zilles	29	1	1
18795	Gurindar Sohi	29	1	1
3812	Edward H. Adelson	29	1	1
3811	Yair Weiss	29	1	1
2855	Matthew V. Mahoney	29	1	1
10564	Don Towsley	35	2	2
16099	Peter Sturm	35	2	1
11802	Ian Horrocks	37	4	3
8438	Li Gong	34	2	2
10878	Mostafa H. Ammar	34	2	1
17318	Tamal K. Dey	28	1	1
17319	Samrat Goswami	28	1	1
925	Omkar Pendse	28	1	1
928	Chris Wallace	28	1	1
2160	Z. Zhang	34	2	2
3784	Thomas Eiter	33	2	2
420	Marek Jersak	33	2	2
9215	Matthias Ruhl	33	2	2
9080	Liam Paninski	27	1	1
14950	U. Amann	27	1	1
12111	Scott Weiss	27	1	1
10762	Abraham P. Punnen	27	1	1
10760	Özlem Ergun	27	1	1
14438	Alessandra Di Pierro	27	1	1
1191	Jason Leigh	27	1	1

	i10_index	g_index	domain	average_citations	domain_Score \
10599	1	1	Software	258.000000	0.000991
10600	1	1	Software	258.000000	0.000991
323	1	1	Computer Science	132.500000	0.000387

7	0	1	Computer Science	0.001996	0.097037
13554	1	1	Computer Science	168.000000	0.000194
19674	1	1	History	122.000000	0.003472
14666	1	1	Microprocessor	111.000000	0.006173
14667	1	1	Microprocessor	111.000000	0.006173
14668	1	1	Microprocessor	111.000000	0.006173
14669	1	1	Microprocessor	111.000000	0.006173
5301	1	1	Computer Science	110.000000	0.000194
5300	1	1	Computer Science	110.000000	0.000194
5299	1	1	Computer Science	110.000000	0.000194
5302	1	1	Computer Science	110.000000	0.000194
19676	1	1	History	65.000000	0.006944
19675	1	1	History	65.000000	0.006944
4432	1	1	Geometry	29.600000	0.009042
6058	3	2	Geometry	21.428571	0.012658
15726	1	1	Computer Science	102.000000	0.000194
532	3	1	Computer Science	43.666667	0.000581
9672	1	1	Computer Science	98.000000	0.000194
9671	1	1	Computer Science	98.000000	0.000194
9670	1	1	Computer Science	98.000000	0.000194
10660	2	2	Business	33.000000	0.003813
14641	2	1	History	58.000000	0.006944
17337	1	1	Software	93.000000	0.000991
17331	1	1	Software	93.000000	0.000991
17332	1	1	Software	93.000000	0.000991
17333	1	1	Software	93.000000	0.000991
17334	1	1	Software	93.000000	0.000991
...	...	...	...	...	...
4581	1	1	Computer Science	29.000000	0.000194
181	1	1	Computer Science	29.000000	0.000194
4580	1	1	Computer Science	29.000000	0.000194
4579	1	1	Computer Science	29.000000	0.000194
2856	1	1	Computer Science	29.000000	0.000194
18794	1	1	Computer Science	29.000000	0.000194
18795	1	1	Computer Science	29.000000	0.000194
3812	1	1	Computer Science	29.000000	0.000194
3811	1	1	Computer Science	29.000000	0.000194
2855	1	1	Computer Science	29.000000	0.000194
10564	1	1	Computer Science	17.500000	0.000387
16099	1	1	Computer Science	17.500000	0.000387
11802	2	2	Software	9.250000	0.003964
8438	2	1	Software	17.000000	0.001982
10878	1	1	Business	17.000000	0.001907
17318	1	1	Earth Science	28.000000	0.001754
17319	1	1	Earth Science	28.000000	0.001754
925	1	1	Software	28.000000	0.000991
928	1	1	Software	28.000000	0.000991
2160	1	1	Computer Science	17.000000	0.000387

3784	2	1	Computer Science	16.500000	0.000387
420	1	1	Computer Science	16.500000	0.000387
9215	1	1	Computer Science	16.500000	0.000387
9080	1	1	Statistics	27.000000	0.006494
14950	1	1	Software	27.000000	0.000991
12111	1	1	Software	27.000000	0.000991
10762	1	1	Business	27.000000	0.000953
10760	1	1	Business	27.000000	0.000953
14438	1	1	Computer Science	27.000000	0.000194
1191	1	1	Computer Science	27.000000	0.000194

	SVS - index
10599	207.200248
10600	207.200248
323	173.350097
7	151.074858
13554	135.200048
19674	98.400868
14666	89.601543
14667	89.601543
14668	89.601543
14669	89.601543
5301	88.800048
5300	88.800048
5299	88.800048
5302	88.800048
19676	85.601736
19675	85.601736
4432	84.882260
6058	84.781736
15726	82.400048
532	80.500145
9672	79.200048
9671	79.200048
9670	79.200048
10660	78.100953
14641	76.751736
17337	75.200248
17331	75.200248
17332	75.200248
17333	75.200248
17334	75.200248
...	...
4581	24.000048
181	24.000048
4580	24.000048
4579	24.000048
2856	24.000048

```
18794      24.000048
18795      24.000048
3812       24.000048
3811       24.000048
2855       24.000048
10564      23.850097
16099      23.850097
11802      23.475991
8438       23.450496
10878      23.200477
17318      23.200439
17319      23.200439
925        23.200248
928        23.200248
2160       23.200097
3784       22.800097
420        22.550097
9215       22.550097
9080       22.401623
14950      22.400248
12111      22.400248
10762      22.400238
10760      22.400238
14438      22.400048
1191       22.400048

[500 rows x 10 columns]
```

8   Ques 2

```
In [37]: # Using a citation network graph, devise a reach function which can identify the degr
         # which can either be localised in a domain or have a more global inter-disciplinary
```

```
In [38]: df_new = train_df.copy()
```

```
In [39]: df_new = df_new.groupby('Year')['domain'].value_counts().unstack().fillna(0)
```

```
In [40]: df_new = df_new.unstack()
```

```
In [41]: df_new = df_new.reset_index()
```

```
In [42]: df_new.columns = ['domain', 'Year', 'Count']
```

```
In [43]: score_list = []
         for index, rows in df_new.iterrows():
             domain = rows['domain']
             count = rows['Count']
             total_count = paper_domain_dict[domain]
```

```
        score_list.append(count / total_count)
    df_new['domain_Score'] = score_list
    df_new
```

Out[43]:   domain Year Count domain_Score
         0   Art 1963   0.0   0.000000
         1   Art 1972   0.0   0.000000
         2   Art 1973   0.0   0.000000
         3   Art 1974   2.0   0.023256
         4   Art 1975   1.0   0.011628
         5   Art 1976   0.0   0.000000
         6   Art 1977   0.0   0.000000
         7   Art 1978   0.0   0.000000
         8   Art 1979   0.0   0.000000
         9   Art 1980   1.0   0.011628
         10   Art 1981   1.0   0.011628
         11   Art 1982   0.0   0.000000
         12   Art 1983   0.0   0.000000
         13   Art 1984   1.0   0.011628
         14   Art 1985   1.0   0.011628
         15   Art 1986   2.0   0.023256
         16   Art 1987   3.0   0.034884
         17   Art 1988   4.0   0.046512
         18   Art 1989   3.0   0.034884
         19   Art 1990   8.0   0.093023
         20   Art 1991   5.0   0.058140
         21   Art 1992   4.0   0.046512
         22   Art 1993   0.0   0.000000
         23   Art 1994   1.0   0.011628
         24   Art 1995   2.0   0.023256
         25   Art 1996   3.0   0.034884
         26   Art 1997   2.0   0.023256
         27   Art 1998   9.0   0.104651
         28   Art 1999   4.0   0.046512
         29   Art 2000   3.0   0.034884
         ..   ..
         630 Statistics 1974   0.0   0.000000
         631 Statistics 1975   0.0   0.000000
         632 Statistics 1976   0.0   0.000000
         633 Statistics 1977   2.0   0.012987
         634 Statistics 1978   0.0   0.000000
         635 Statistics 1979   1.0   0.006494
         636 Statistics 1980   1.0   0.006494
         637 Statistics 1981   0.0   0.000000
         638 Statistics 1982   0.0   0.000000
         639 Statistics 1983   0.0   0.000000
         640 Statistics 1984   0.0   0.000000
         641 Statistics 1985   0.0   0.000000
```

```
642  Statistics  1986   0.0      0.000000
643  Statistics  1987   1.0      0.006494
644  Statistics  1988   0.0      0.000000
645  Statistics  1989   2.0      0.012987
646  Statistics  1990   1.0      0.006494
647  Statistics  1991   1.0      0.006494
648  Statistics  1992   1.0      0.006494
649  Statistics  1993   3.0      0.019481
650  Statistics  1994   2.0      0.012987
651  Statistics  1995   7.0      0.045455
652  Statistics  1996   6.0      0.038961
653  Statistics  1997   6.0      0.038961
654  Statistics  1998   7.0      0.045455
655  Statistics  1999  13.0      0.084416
656  Statistics  2000  12.0      0.077922
657  Statistics  2001   7.0      0.045455
658  Statistics  2002  38.0      0.246753
659  Statistics  2003  43.0      0.279221

[660 rows x 4 columns]
```

9 Reach function

```python
In [44]: # Get citations count of paper present in test set
         def generate_citation_count_from_google_scholar(df):
             citation_list = []
             for index,row in df.iterrows():
                 print(index)
                 citations = 0
                 title = row['paperTitle']
                 search_query = scholarly.search_pubs_query(title)
                 val = next(search_query, "N/A")
                 if val != 'N/A' and 'title' in val.bib.keys():
                     title1 = val.bib['title']
                     if title1[0:3] == title[0:3]:
                         if hasattr(val, 'citedby'):
                             citations = val.citedby
                 citation_list.append(citations)
             df["citation_count"] = citation_list
             return df

         test_df = generate_citation_count_from_google_scholar(test_df)

0
1
2
3
```

4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51

52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99

```
In [46]: test_df.to_csv('test_new_arxiv.csv', index=False)

In [48]: test_df

Out[48]:                                           Authors       index id  \
         0                    Ahmed Osman, Wojciech Samek    1802.00209v1
         1                     Ji Young Lee, Franck Dernoncourt    1603.03827v1
         2    Iulian Vlad Serban, Tim Klinger, Gerald Tesau...    1606.00776v2
         3    Sebastian Ruder, Joachim Bingel, Isabelle Aug...    1705.08142v2
         4    Iulian V. Serban, Chinnadhurai Sankar, Mathie...    1709.02349v2
         5    Kelvin Guu, Tatsunori B. Hashimoto, Yonatan O...    1709.08878v1
         6    Iulian V. Serban, Chinnadhurai Sankar, Mathie...    1801.06700v1
         7    Darko Brodic, Alessia Amelio, Zoran N. Milivo...    1609.06492v1
         8                   Mateusz Malinowski, Mario Fritz    1610.01076v1
         9                                  Tony Beltramelli    1705.07962v2
         10     Fred Richardson, Douglas Reynolds, Najim Dehak    1504.00923v1
         11   Hieu Pham, Melody Y. Guan, Barret Zoph, Quoc ...    1802.03268v2
         12   Brenden M. Lake, Tomer D. Ullman, Joshua B. T...    1604.00289v3
         13                            Hao Wang, Dit-Yan Yeung    1604.01662v2
         14   Tejas D. Kulkarni, Karthik R. Narasimhan, Ard...    1604.06057v2
         15   Deepak Pathak, Ross Girshick, Piotr Dollár, T...    1612.06370v2
         16   Muhammad Ghifary, W. Bastiaan Kleijn, Mengjie...     1409.6041v1
         17   Lionel Pigou, Aäron van den Oord, Sander Diel...    1506.01911v3
         18                   Rakesh Achanta, Trevor Hastie    1509.05962v2
         19   Jeff Donahue, Philipp Krähenbühl, Trevor Darrell    1605.09782v7
         20                             Zachary C. Lipton    1606.03490v3
         21   Sahil Garg, Irina Rish, Guillermo Cecchi, Aur...    1701.06106v2
         22                             Weifeng Ge, Yizhou Yu    1702.08690v2
         23   Tanmay Gupta, Kevin Shih, Saurabh Singh, Dere...    1704.00260v2
         24   Jan Hendrik Metzen, Mummadi Chaithanya Kumar,...    1704.05712v3
         25                    Quynh Nguyen, Matthias Hein    1704.08045v2
         26   Chris Donahue, Zachary C. Lipton, Akshay Bals...    1705.07904v3
         27          Mahesh Chandra Mukkamala, Matthias Hein    1706.05507v2
         28   Chunyuan Li, Hao Liu, Changyou Chen, Yunchen ...    1709.01215v2
         29    Mateusz Buda, Atsuto Maki, Maciej A. Mazurowski    1710.05381v1
         ..                                             ...             ...
         70   Petr Baudi, Jan Pichl, Tomá Vyskoil, Jan ...   1603.06127v4
         71   Jiatao Gu, Zhengdong Lu, Hang Li, Victor O. K...    1603.06393v3
         72   Iulian Vlad Serban, Alberto García-Durán, Cag...    1603.06807v2
         73   Chia-Wei Liu, Ryan Lowe, Iulian V. Serban, Mi...    1603.08023v2
         74   Iulian Vlad Serban, Alessandro Sordoni, Ryan ...    1605.06069v3
         75                             Dirk Weissenborn    1606.03864v2
         76                   Marc Dymetman, Chunyang Xiao    1607.02467v2
         77     Ondrej Bajgar, Rudolf Kadlec, Jan Kleindienst    1610.00956v1
         78   James Bradbury, Stephen Merity, Caiming Xiong...    1611.01576v2
         79   Jakob N. Foerster, Justin Gilmer, Jan Chorows...    1611.09434v2
```

```
80   Micha Daniluk, Tim Rocktäschel, Johannes Wel...   1702.04521v1
81   Zhouhan Lin, Minwei Feng, Cicero Nogueira dos...   1703.03130v1
82   Samuel Rönnqvist, Niko Schenk, Christian Chia...   1704.08092v1
83   Lara J. Martin, Prithviraj Ammanabrolu, Xinyu...   1706.01331v3
84            Tong Wang, Xingdi Yuan, Adam Trischler   1706.01450v1
85   Wei Wen, Yuxiong He, Samyam Rajbhandari, Minj...   1709.05027v7
86     Huda Hakami, Danushka Bollegala, Hayashi Kohei   1709.06673v2
87   Zhengdong Lu, Haotian Cui, Xianggen Liu, Yuku...   1709.08853v4
88                                      Bin Bi, Hao Ma   1709.10204v2
89   Mirco Ravanelli, Philemon Brakel, Maurizio Om...   1710.00641v1
90   Baolin Peng, Xiujun Li, Jianfeng Gao, Jingjin...   1801.06176v1
91   Andrew L. Maas, Peng Qi, Ziang Xie, Awni Y. H...    1406.7806v2
92                            William Chan, Ian Lane   1504.01482v1
93                  David Krueger, Roland Memisevic   1511.08400v7
94   Noam Shazeer, Azalia Mirhoseini, Krzysztof Ma...   1701.06538v1
95     Yacine Jernite, Samuel R. Bowman, David Sontag   1705.00557v1
96                  Zhengyang Wang, Shuiwang Ji   1705.06824v1
97   Kyunghyun Cho, Bart van Merrienboer, Caglar G...    1406.1078v3
98            Zhiyuan Tang, Dong Wang, Zhiyong Zhang   1505.04630v5
99      Haim Sak, Andrew Senior, Françoise Beaufays   1402.1128v1

                                      paperTitle   Year  \
0    Dual Recurrent Attention Units for Visual Ques...   2018
1    Sequential Short-Text Classification with Recu...   2016
2    Multiresolution Recurrent Neural Networks: An ...   2016
3    Learning what to share between loosely related...   2017
4              A Deep Reinforcement Learning Chatbot   2017
5          Generating Sentences by Editing Prototypes   2017
6    A Deep Reinforcement Learning Chatbot (Short V...   2018
7    Document Image Coding and Clustering for Scrip...   2016
8    Tutorial on Answering Questions about Images w...   2016
9    pix2code: Generating Code from a Graphical Use...   2017
10   A Unified Deep Neural Network for Speaker and ...   2015
11   Efficient Neural Architecture Search via Param...   2018
12   Building Machines That Learn and Think Like Pe...   2016
13           Towards Bayesian Deep Learning: A Survey   2016
14   Hierarchical Deep Reinforcement Learning: Inte...   2016
15          Learning Features by Watching Objects Move   2016
16   Domain Adaptive Neural Networks for Object Rec...   2014
17   Beyond Temporal Pooling: Recurrence and Tempor...   2015
18           Telugu OCR Framework using Deep Learning   2015
19                       Adversarial Feature Learning   2016
20              The Mythos of Model Interpretability   2016
21   Neurogenesis-Inspired Dictionary Learning: Onl...   2017
22   Borrowing Treasures from the Wealthy: Deep Tra...   2017
23   Aligned Image-Word Representations Improve Ind...   2017
24   Universal Adversarial Perturbations Against Se...   2017
25   The loss surface of deep and wide neural networks   2017
```

```
26  Semantically Decomposing the Latent Spaces of ...  2017
27  Variants of RMSProp and Adagrad with Logarithm...  2017
28  ALICE: Towards Understanding Adversarial Learn...  2017
29  A systematic study of the class imbalance prob...  2017
..                                                 ...  ...
70  Sentence Pair Scoring: Towards Unified Framewo...  2016
71  Incorporating Copying Mechanism in Sequence-to...  2016
72  Generating Factoid Questions With Recurrent Ne...  2016
73  How NOT To Evaluate Your Dialogue System: An E...  2016
74  A Hierarchical Latent Variable Encoder-Decoder...  2016
75  Neural Associative Memory for Dual-Sequence Mo...  2016
76  Log-Linear RNNs: Towards Recurrent Neural Netw...  2016
77  Embracing data abundance: BookTest Dataset for...  2016
78                      Quasi-Recurrent Neural Networks  2016
79  Input Switched Affine Networks: An RNN Archite...  2016
80  Frustratingly Short Attention Spans in Neural ...  2017
81     A Structured Self-attentive Sentence Embedding  2017
82  A Recurrent Neural Model with Attention for th...  2017
83  Event Representations for Automated Story Gene...  2017
84  A Joint Model for Question Answering and Quest...  2017
85  Learning Intrinsic Sparse Structures within Lo...  2017
86  Why PairDiff works? -- A Mathematical Analysis...  2017
87  Object-oriented Neural Programming (OONP) for ...  2017
88  A Neural Comprehensive Ranker (NCR) for Open-D...  2017
89  Improving speech recognition by revising gated...  2017
90  Integrating planning for task-completion dialo...  2018
91  Building DNN Acoustic Models for Large Vocabul...  2014
92  Deep Recurrent Neural Networks for Acoustic Mo...  2015
93       Regularizing RNNs by Stabilizing Activations  2015
94  Outrageously Large Neural Networks: The Sparse...  2017
95  Discourse-Based Objectives for Fast Unsupervis...  2017
96  Learning Convolutional Text Representations fo...  2017
97  Learning Phrase Representations using RNN Enco...  2014
98  Recurrent Neural Network Training with Dark Kn...  2015
99  Long Short-Term Memory Based Recurrent Neural ...  2014


                                       Abstract  citation_count  \
0   We propose an architecture for VQA which utili...               1
1   Recent approaches based on artificial neural n...             105
2   We introduce the multiresolution recurrent neu...              59
3   Multi-task learning is motivated by the observ...               9
4   We present MILABOT: a deep reinforcement learn...               0
5   We propose a new generative model of sentences...              36
6   We present MILABOT: a deep reinforcement learn...               2
7   The paper introduces a new method for discrimi...               4
8   Together with the development of more accurate...               1
9   Transforming a graphical user interface screen...              26
10  Learned feature representations and sub-phonem...               0
```

```
11   We propose Efficient Neural Architecture Searc...      55
12   Recent progress in artificial intelligence (AI...     346
13   While perception tasks such as visual object r...      23
14   Learning goal-directed behavior in environment...     215
15   This paper presents a novel yet intuitive appr...      95
16   We propose a simple neural network model to de...      48
17   Recent studies have demonstrated the power of ...      56
18   In this paper, we address the task of Optical ...      10
19   The ability of the Generative Adversarial Netw...     357
20   Supervised machine learning models boast remar...     346
21   In this paper, we focus on online representati...       2
22   Deep neural networks require a large amount of...      34
23   An important goal of computer vision is to bui...       7
24   While deep learning is remarkably successful o...      14
25   While the optimization problem behind deep neu...      70
26   We propose a new algorithm for training genera...      25
27   Adaptive gradient methods have become recently...      13
28   We investigate the non-identifiability issues ...       0
29   In this study, we systematically investigate t...      55
..                                            ...          ...
70   We review the task of Sentence Pair Scoring, p...      19
71   We address an important problem in sequence-to...       0
72   Over the past decade, large-scale supervised l...       0
73   We investigate evaluation metrics for dialogue...       0
74   Sequential data often possesses a hierarchical...       0
75   Many important NLP problems can be posed as du...       0
76   We introduce LL-RNNs (Log-Linear RNNs), an ext...       0
77   There is a practically unlimited amount of nat...       0
78   Recurrent neural networks are a powerful tool ...       0
79   There exist many problem domains where the int...       0
80   Neural language models predict the next token ...       0
81   This paper proposes a new model for extracting...       0
82   We introduce an attention-based Bi-LSTM for Ch...       0
83   Automated story generation is the problem of a...       0
84   We propose a generative machine comprehension ...       0
85   Model compression is significant for the wide ...       0
86   Representing the semantic relations that exist...       0
87   We propose Object-oriented Neural Programming ...       0
88   This paper proposes a novel neural machine rea...       0
89   Speech recognition is largely taking advantage...       0
90   Training a task-completion dialogue agent with...       0
91   Deep neural networks (DNNs) are now a central ...       0
92   We present a novel deep Recurrent Neural Netwo...       0
93   We stabilize the activations of Recurrent Neur...       0
94   The capacity of a neural network to absorb inf...       0
95   This work presents a novel objective function ...       0
96   Visual question answering is a recently propos...       0
97   In this paper, we propose a novel neural netwo...       0
```

```
98  Recurrent neural networks (RNNs), particularly...          0
99  Long Short-Term Memory (LSTM) is a recurrent n...          0

                    domain
0       Computer Science
1       Computer Science
2       Computer Science
3       Computer Science
4       Computer Science
5       Computer Science
6       Computer Science
7               Software
8       Computer Science
9               Software
10      Computer Science
11              Software
12      Computer Science
13      Computer Science
14           Neuroscience
15      Computer Science
16      Computer Science
17      Computer Science
18      Computer Science
19              Political
20                History
21      Computer Science
22      Computer Science
23      Computer Science
24              Political
25          Earth Science
26              Political
27          Earth Science
28      Computer Science
29      Computer Science
..                    ...
70      Computer Science
71      Computer Science
72      Computer Science
73      Computer Science
74              Software
75      Computer Science
76      Computer Science
77      Computer Science
78  Biomedical Research
79      Computer Science
80      Computer Science
81              Business
82              Political
```

```
        83        Computer Science
        84        Computer Science
        85        Computer Science
        86        Computer Science
        87        Computer Science
        88        Computer Science
        89                  Health
        90                Business
        91        Computer Science
        92        Computer Science
        93                Software
        94        Computer Science
        95        Computer Science
        96        Computer Science
        97                Software
        98        Computer Science
        99        Computer Science

        [100 rows x 7 columns]

In [63]: domain_paper_list = []
         for index, rows in test_df.iterrows():
             authors = rows["Authors"]
             citations = rows["citation_count"]
             domain = rows["domain"]
             year = rows['Year']
             domain_list = df_new[df_new['domain'] == domain]
             print(domain_list)
             print(year)
             print(domain)
             domain_score = 0
             if not domain_list.empty:
                 if year is domain_list['Year'] == year:
                     domain_score = domain_list.loc[domain_list['Year'] == year, 'domain_Score
             domain_paper_list.append(domain_score)
         test_df['domain_score'] = domain_paper_list

             domain  Year  Count  domain_Score
165  Computer Science  1963    1.0      0.000194
166  Computer Science  1972    5.0      0.000968
167  Computer Science  1973    1.0      0.000194
168  Computer Science  1974   12.0      0.002324
169  Computer Science  1975    1.0      0.000194
170  Computer Science  1976    1.0      0.000194
171  Computer Science  1977    1.0      0.000194
172  Computer Science  1978    3.0      0.000581
173  Computer Science  1979    3.0      0.000581
174  Computer Science  1980    5.0      0.000968
```

```
175  Computer Science  1981    6.0      0.001162
176  Computer Science  1982    2.0      0.000387
177  Computer Science  1983   12.0      0.002324
178  Computer Science  1984    2.0      0.000387
179  Computer Science  1985   11.0      0.002131
180  Computer Science  1986   24.0      0.004648
181  Computer Science  1987   20.0      0.003874
182  Computer Science  1988   51.0      0.009878
183  Computer Science  1989   29.0      0.005617
184  Computer Science  1990   39.0      0.007554
185  Computer Science  1991   62.0      0.012009
186  Computer Science  1992   71.0      0.013752
187  Computer Science  1993   98.0      0.018981
188  Computer Science  1994  114.0      0.022080
189  Computer Science  1995  374.0      0.072439
190  Computer Science  1996  349.0      0.067596
191  Computer Science  1997  401.0      0.077668
192  Computer Science  1998  420.0      0.081348
193  Computer Science  1999  389.0      0.075344
194  Computer Science  2000  451.0      0.087352
195  Computer Science  2001  366.0      0.070889
196  Computer Science  2002  882.0      0.170831
197  Computer Science  2003  957.0      0.185357
2018
Computer Science
               domain  Year  Count  domain_Score
165  Computer Science  1963    1.0      0.000194
166  Computer Science  1972    5.0      0.000968
167  Computer Science  1973    1.0      0.000194
168  Computer Science  1974   12.0      0.002324
169  Computer Science  1975    1.0      0.000194
170  Computer Science  1976    1.0      0.000194
171  Computer Science  1977    1.0      0.000194
172  Computer Science  1978    3.0      0.000581
173  Computer Science  1979    3.0      0.000581
174  Computer Science  1980    5.0      0.000968
175  Computer Science  1981    6.0      0.001162
176  Computer Science  1982    2.0      0.000387
177  Computer Science  1983   12.0      0.002324
178  Computer Science  1984    2.0      0.000387
179  Computer Science  1985   11.0      0.002131
180  Computer Science  1986   24.0      0.004648
181  Computer Science  1987   20.0      0.003874
182  Computer Science  1988   51.0      0.009878
183  Computer Science  1989   29.0      0.005617
184  Computer Science  1990   39.0      0.007554
185  Computer Science  1991   62.0      0.012009
186  Computer Science  1992   71.0      0.013752
```

```
187  Computer Science  1993   98.0      0.018981
188  Computer Science  1994  114.0      0.022080
189  Computer Science  1995  374.0      0.072439
190  Computer Science  1996  349.0      0.067596
191  Computer Science  1997  401.0      0.077668
192  Computer Science  1998  420.0      0.081348
193  Computer Science  1999  389.0      0.075344
194  Computer Science  2000  451.0      0.087352
195  Computer Science  2001  366.0      0.070889
196  Computer Science  2002  882.0      0.170831
197  Computer Science  2003  957.0      0.185357
2016
Computer Science
                domain  Year  Count  domain_Score
165  Computer Science  1963    1.0      0.000194
166  Computer Science  1972    5.0      0.000968
167  Computer Science  1973    1.0      0.000194
168  Computer Science  1974   12.0      0.002324
169  Computer Science  1975    1.0      0.000194
170  Computer Science  1976    1.0      0.000194
171  Computer Science  1977    1.0      0.000194
172  Computer Science  1978    3.0      0.000581
173  Computer Science  1979    3.0      0.000581
174  Computer Science  1980    5.0      0.000968
175  Computer Science  1981    6.0      0.001162
176  Computer Science  1982    2.0      0.000387
177  Computer Science  1983   12.0      0.002324
178  Computer Science  1984    2.0      0.000387
179  Computer Science  1985   11.0      0.002131
180  Computer Science  1986   24.0      0.004648
181  Computer Science  1987   20.0      0.003874
182  Computer Science  1988   51.0      0.009878
183  Computer Science  1989   29.0      0.005617
184  Computer Science  1990   39.0      0.007554
185  Computer Science  1991   62.0      0.012009
186  Computer Science  1992   71.0      0.013752
187  Computer Science  1993   98.0      0.018981
188  Computer Science  1994  114.0      0.022080
189  Computer Science  1995  374.0      0.072439
190  Computer Science  1996  349.0      0.067596
191  Computer Science  1997  401.0      0.077668
192  Computer Science  1998  420.0      0.081348
193  Computer Science  1999  389.0      0.075344
194  Computer Science  2000  451.0      0.087352
195  Computer Science  2001  366.0      0.070889
196  Computer Science  2002  882.0      0.170831
197  Computer Science  2003  957.0      0.185357
2016
```

```
Computer Science
                  domain  Year  Count  domain_Score
165   Computer Science   1963    1.0       0.000194
166   Computer Science   1972    5.0       0.000968
167   Computer Science   1973    1.0       0.000194
168   Computer Science   1974   12.0       0.002324
169   Computer Science   1975    1.0       0.000194
170   Computer Science   1976    1.0       0.000194
171   Computer Science   1977    1.0       0.000194
172   Computer Science   1978    3.0       0.000581
173   Computer Science   1979    3.0       0.000581
174   Computer Science   1980    5.0       0.000968
175   Computer Science   1981    6.0       0.001162
176   Computer Science   1982    2.0       0.000387
177   Computer Science   1983   12.0       0.002324
178   Computer Science   1984    2.0       0.000387
179   Computer Science   1985   11.0       0.002131
180   Computer Science   1986   24.0       0.004648
181   Computer Science   1987   20.0       0.003874
182   Computer Science   1988   51.0       0.009878
183   Computer Science   1989   29.0       0.005617
184   Computer Science   1990   39.0       0.007554
185   Computer Science   1991   62.0       0.012009
186   Computer Science   1992   71.0       0.013752
187   Computer Science   1993   98.0       0.018981
188   Computer Science   1994  114.0       0.022080
189   Computer Science   1995  374.0       0.072439
190   Computer Science   1996  349.0       0.067596
191   Computer Science   1997  401.0       0.077668
192   Computer Science   1998  420.0       0.081348
193   Computer Science   1999  389.0       0.075344
194   Computer Science   2000  451.0       0.087352
195   Computer Science   2001  366.0       0.070889
196   Computer Science   2002  882.0       0.170831
197   Computer Science   2003  957.0       0.185357
2017
Computer Science
                  domain  Year  Count  domain_Score
165   Computer Science   1963    1.0       0.000194
166   Computer Science   1972    5.0       0.000968
167   Computer Science   1973    1.0       0.000194
168   Computer Science   1974   12.0       0.002324
169   Computer Science   1975    1.0       0.000194
170   Computer Science   1976    1.0       0.000194
171   Computer Science   1977    1.0       0.000194
172   Computer Science   1978    3.0       0.000581
173   Computer Science   1979    3.0       0.000581
174   Computer Science   1980    5.0       0.000968
```

```
175   Computer Science   1981     6.0        0.001162
176   Computer Science   1982     2.0        0.000387
177   Computer Science   1983    12.0        0.002324
178   Computer Science   1984     2.0        0.000387
179   Computer Science   1985    11.0        0.002131
180   Computer Science   1986    24.0        0.004648
181   Computer Science   1987    20.0        0.003874
182   Computer Science   1988    51.0        0.009878
183   Computer Science   1989    29.0        0.005617
184   Computer Science   1990    39.0        0.007554
185   Computer Science   1991    62.0        0.012009
186   Computer Science   1992    71.0        0.013752
187   Computer Science   1993    98.0        0.018981
188   Computer Science   1994   114.0        0.022080
189   Computer Science   1995   374.0        0.072439
190   Computer Science   1996   349.0        0.067596
191   Computer Science   1997   401.0        0.077668
192   Computer Science   1998   420.0        0.081348
193   Computer Science   1999   389.0        0.075344
194   Computer Science   2000   451.0        0.087352
195   Computer Science   2001   366.0        0.070889
196   Computer Science   2002   882.0        0.170831
197   Computer Science   2003   957.0        0.185357
2017
Computer Science
                domain   Year   Count   domain_Score
165   Computer Science   1963     1.0        0.000194
166   Computer Science   1972     5.0        0.000968
167   Computer Science   1973     1.0        0.000194
168   Computer Science   1974    12.0        0.002324
169   Computer Science   1975     1.0        0.000194
170   Computer Science   1976     1.0        0.000194
171   Computer Science   1977     1.0        0.000194
172   Computer Science   1978     3.0        0.000581
173   Computer Science   1979     3.0        0.000581
174   Computer Science   1980     5.0        0.000968
175   Computer Science   1981     6.0        0.001162
176   Computer Science   1982     2.0        0.000387
177   Computer Science   1983    12.0        0.002324
178   Computer Science   1984     2.0        0.000387
179   Computer Science   1985    11.0        0.002131
180   Computer Science   1986    24.0        0.004648
181   Computer Science   1987    20.0        0.003874
182   Computer Science   1988    51.0        0.009878
183   Computer Science   1989    29.0        0.005617
184   Computer Science   1990    39.0        0.007554
185   Computer Science   1991    62.0        0.012009
186   Computer Science   1992    71.0        0.013752
```

```
187   Computer Science   1993    98.0        0.018981
188   Computer Science   1994   114.0        0.022080
189   Computer Science   1995   374.0        0.072439
190   Computer Science   1996   349.0        0.067596
191   Computer Science   1997   401.0        0.077668
192   Computer Science   1998   420.0        0.081348
193   Computer Science   1999   389.0        0.075344
194   Computer Science   2000   451.0        0.087352
195   Computer Science   2001   366.0        0.070889
196   Computer Science   2002   882.0        0.170831
197   Computer Science   2003   957.0        0.185357
2017
Computer Science
                domain   Year   Count   domain_Score
165   Computer Science   1963     1.0        0.000194
166   Computer Science   1972     5.0        0.000968
167   Computer Science   1973     1.0        0.000194
168   Computer Science   1974    12.0        0.002324
169   Computer Science   1975     1.0        0.000194
170   Computer Science   1976     1.0        0.000194
171   Computer Science   1977     1.0        0.000194
172   Computer Science   1978     3.0        0.000581
173   Computer Science   1979     3.0        0.000581
174   Computer Science   1980     5.0        0.000968
175   Computer Science   1981     6.0        0.001162
176   Computer Science   1982     2.0        0.000387
177   Computer Science   1983    12.0        0.002324
178   Computer Science   1984     2.0        0.000387
179   Computer Science   1985    11.0        0.002131
180   Computer Science   1986    24.0        0.004648
181   Computer Science   1987    20.0        0.003874
182   Computer Science   1988    51.0        0.009878
183   Computer Science   1989    29.0        0.005617
184   Computer Science   1990    39.0        0.007554
185   Computer Science   1991    62.0        0.012009
186   Computer Science   1992    71.0        0.013752
187   Computer Science   1993    98.0        0.018981
188   Computer Science   1994   114.0        0.022080
189   Computer Science   1995   374.0        0.072439
190   Computer Science   1996   349.0        0.067596
191   Computer Science   1997   401.0        0.077668
192   Computer Science   1998   420.0        0.081348
193   Computer Science   1999   389.0        0.075344
194   Computer Science   2000   451.0        0.087352
195   Computer Science   2001   366.0        0.070889
196   Computer Science   2002   882.0        0.170831
197   Computer Science   2003   957.0        0.185357
2018
```

Computer Science

	domain	Year	Count	domain_Score
594	Software	1963	0.0	0.000000
595	Software	1972	0.0	0.000000
596	Software	1973	0.0	0.000000
597	Software	1974	5.0	0.004955
598	Software	1975	1.0	0.000991
599	Software	1976	0.0	0.000000
600	Software	1977	1.0	0.000991
601	Software	1978	4.0	0.003964
602	Software	1979	2.0	0.001982
603	Software	1980	1.0	0.000991
604	Software	1981	1.0	0.000991
605	Software	1982	1.0	0.000991
606	Software	1983	0.0	0.000000
607	Software	1984	4.0	0.003964
608	Software	1985	1.0	0.000991
609	Software	1986	0.0	0.000000
610	Software	1987	4.0	0.003964
611	Software	1988	4.0	0.003964
612	Software	1989	3.0	0.002973
613	Software	1990	10.0	0.009911
614	Software	1991	14.0	0.013875
615	Software	1992	8.0	0.007929
616	Software	1993	18.0	0.017839
617	Software	1994	23.0	0.022795
618	Software	1995	50.0	0.049554
619	Software	1996	74.0	0.073340
620	Software	1997	69.0	0.068385
621	Software	1998	90.0	0.089197
622	Software	1999	94.0	0.093162
623	Software	2000	81.0	0.080278
624	Software	2001	76.0	0.075322
625	Software	2002	179.0	0.177403
626	Software	2003	191.0	0.189296

2016
Software

	domain	Year	Count	domain_Score
165	Computer Science	1963	1.0	0.000194
166	Computer Science	1972	5.0	0.000968
167	Computer Science	1973	1.0	0.000194
168	Computer Science	1974	12.0	0.002324
169	Computer Science	1975	1.0	0.000194
170	Computer Science	1976	1.0	0.000194
171	Computer Science	1977	1.0	0.000194
172	Computer Science	1978	3.0	0.000581
173	Computer Science	1979	3.0	0.000581
174	Computer Science	1980	5.0	0.000968

```
175  Computer Science  1981    6.0      0.001162
176  Computer Science  1982    2.0      0.000387
177  Computer Science  1983   12.0      0.002324
178  Computer Science  1984    2.0      0.000387
179  Computer Science  1985   11.0      0.002131
180  Computer Science  1986   24.0      0.004648
181  Computer Science  1987   20.0      0.003874
182  Computer Science  1988   51.0      0.009878
183  Computer Science  1989   29.0      0.005617
184  Computer Science  1990   39.0      0.007554
185  Computer Science  1991   62.0      0.012009
186  Computer Science  1992   71.0      0.013752
187  Computer Science  1993   98.0      0.018981
188  Computer Science  1994  114.0      0.022080
189  Computer Science  1995  374.0      0.072439
190  Computer Science  1996  349.0      0.067596
191  Computer Science  1997  401.0      0.077668
192  Computer Science  1998  420.0      0.081348
193  Computer Science  1999  389.0      0.075344
194  Computer Science  2000  451.0      0.087352
195  Computer Science  2001  366.0      0.070889
196  Computer Science  2002  882.0      0.170831
197  Computer Science  2003  957.0      0.185357
2016
Computer Science
       domain  Year  Count  domain_Score
594  Software  1963    0.0      0.000000
595  Software  1972    0.0      0.000000
596  Software  1973    0.0      0.000000
597  Software  1974    5.0      0.004955
598  Software  1975    1.0      0.000991
599  Software  1976    0.0      0.000000
600  Software  1977    1.0      0.000991
601  Software  1978    4.0      0.003964
602  Software  1979    2.0      0.001982
603  Software  1980    1.0      0.000991
604  Software  1981    1.0      0.000991
605  Software  1982    1.0      0.000991
606  Software  1983    0.0      0.000000
607  Software  1984    4.0      0.003964
608  Software  1985    1.0      0.000991
609  Software  1986    0.0      0.000000
610  Software  1987    4.0      0.003964
611  Software  1988    4.0      0.003964
612  Software  1989    3.0      0.002973
613  Software  1990   10.0      0.009911
614  Software  1991   14.0      0.013875
615  Software  1992    8.0      0.007929
```

```
616  Software  1993   18.0        0.017839
617  Software  1994   23.0        0.022795
618  Software  1995   50.0        0.049554
619  Software  1996   74.0        0.073340
620  Software  1997   69.0        0.068385
621  Software  1998   90.0        0.089197
622  Software  1999   94.0        0.093162
623  Software  2000   81.0        0.080278
624  Software  2001   76.0        0.075322
625  Software  2002  179.0        0.177403
626  Software  2003  191.0        0.189296
2017
Software
              domain  Year  Count   domain_Score
165  Computer Science  1963    1.0       0.000194
166  Computer Science  1972    5.0       0.000968
167  Computer Science  1973    1.0       0.000194
168  Computer Science  1974   12.0       0.002324
169  Computer Science  1975    1.0       0.000194
170  Computer Science  1976    1.0       0.000194
171  Computer Science  1977    1.0       0.000194
172  Computer Science  1978    3.0       0.000581
173  Computer Science  1979    3.0       0.000581
174  Computer Science  1980    5.0       0.000968
175  Computer Science  1981    6.0       0.001162
176  Computer Science  1982    2.0       0.000387
177  Computer Science  1983   12.0       0.002324
178  Computer Science  1984    2.0       0.000387
179  Computer Science  1985   11.0       0.002131
180  Computer Science  1986   24.0       0.004648
181  Computer Science  1987   20.0       0.003874
182  Computer Science  1988   51.0       0.009878
183  Computer Science  1989   29.0       0.005617
184  Computer Science  1990   39.0       0.007554
185  Computer Science  1991   62.0       0.012009
186  Computer Science  1992   71.0       0.013752
187  Computer Science  1993   98.0       0.018981
188  Computer Science  1994  114.0       0.022080
189  Computer Science  1995  374.0       0.072439
190  Computer Science  1996  349.0       0.067596
191  Computer Science  1997  401.0       0.077668
192  Computer Science  1998  420.0       0.081348
193  Computer Science  1999  389.0       0.075344
194  Computer Science  2000  451.0       0.087352
195  Computer Science  2001  366.0       0.070889
196  Computer Science  2002  882.0       0.170831
197  Computer Science  2003  957.0       0.185357
2015
```

Computer Science

	domain	Year	Count	domain_Score
594	Software	1963	0.0	0.000000
595	Software	1972	0.0	0.000000
596	Software	1973	0.0	0.000000
597	Software	1974	5.0	0.004955
598	Software	1975	1.0	0.000991
599	Software	1976	0.0	0.000000
600	Software	1977	1.0	0.000991
601	Software	1978	4.0	0.003964
602	Software	1979	2.0	0.001982
603	Software	1980	1.0	0.000991
604	Software	1981	1.0	0.000991
605	Software	1982	1.0	0.000991
606	Software	1983	0.0	0.000000
607	Software	1984	4.0	0.003964
608	Software	1985	1.0	0.000991
609	Software	1986	0.0	0.000000
610	Software	1987	4.0	0.003964
611	Software	1988	4.0	0.003964
612	Software	1989	3.0	0.002973
613	Software	1990	10.0	0.009911
614	Software	1991	14.0	0.013875
615	Software	1992	8.0	0.007929
616	Software	1993	18.0	0.017839
617	Software	1994	23.0	0.022795
618	Software	1995	50.0	0.049554
619	Software	1996	74.0	0.073340
620	Software	1997	69.0	0.068385
621	Software	1998	90.0	0.089197
622	Software	1999	94.0	0.093162
623	Software	2000	81.0	0.080278
624	Software	2001	76.0	0.075322
625	Software	2002	179.0	0.177403
626	Software	2003	191.0	0.189296

2018
Software

	domain	Year	Count	domain_Score
165	Computer Science	1963	1.0	0.000194
166	Computer Science	1972	5.0	0.000968
167	Computer Science	1973	1.0	0.000194
168	Computer Science	1974	12.0	0.002324
169	Computer Science	1975	1.0	0.000194
170	Computer Science	1976	1.0	0.000194
171	Computer Science	1977	1.0	0.000194
172	Computer Science	1978	3.0	0.000581
173	Computer Science	1979	3.0	0.000581
174	Computer Science	1980	5.0	0.000968

```
175    Computer Science  1981    6.0       0.001162
176    Computer Science  1982    2.0       0.000387
177    Computer Science  1983   12.0       0.002324
178    Computer Science  1984    2.0       0.000387
179    Computer Science  1985   11.0       0.002131
180    Computer Science  1986   24.0       0.004648
181    Computer Science  1987   20.0       0.003874
182    Computer Science  1988   51.0       0.009878
183    Computer Science  1989   29.0       0.005617
184    Computer Science  1990   39.0       0.007554
185    Computer Science  1991   62.0       0.012009
186    Computer Science  1992   71.0       0.013752
187    Computer Science  1993   98.0       0.018981
188    Computer Science  1994  114.0       0.022080
189    Computer Science  1995  374.0       0.072439
190    Computer Science  1996  349.0       0.067596
191    Computer Science  1997  401.0       0.077668
192    Computer Science  1998  420.0       0.081348
193    Computer Science  1999  389.0       0.075344
194    Computer Science  2000  451.0       0.087352
195    Computer Science  2001  366.0       0.070889
196    Computer Science  2002  882.0       0.170831
197    Computer Science  2003  957.0       0.185357
2016
Computer Science
                 domain  Year  Count  domain_Score
165    Computer Science  1963    1.0       0.000194
166    Computer Science  1972    5.0       0.000968
167    Computer Science  1973    1.0       0.000194
168    Computer Science  1974   12.0       0.002324
169    Computer Science  1975    1.0       0.000194
170    Computer Science  1976    1.0       0.000194
171    Computer Science  1977    1.0       0.000194
172    Computer Science  1978    3.0       0.000581
173    Computer Science  1979    3.0       0.000581
174    Computer Science  1980    5.0       0.000968
175    Computer Science  1981    6.0       0.001162
176    Computer Science  1982    2.0       0.000387
177    Computer Science  1983   12.0       0.002324
178    Computer Science  1984    2.0       0.000387
179    Computer Science  1985   11.0       0.002131
180    Computer Science  1986   24.0       0.004648
181    Computer Science  1987   20.0       0.003874
182    Computer Science  1988   51.0       0.009878
183    Computer Science  1989   29.0       0.005617
184    Computer Science  1990   39.0       0.007554
185    Computer Science  1991   62.0       0.012009
186    Computer Science  1992   71.0       0.013752
```

```
187  Computer Science  1993   98.0       0.018981
188  Computer Science  1994  114.0       0.022080
189  Computer Science  1995  374.0       0.072439
190  Computer Science  1996  349.0       0.067596
191  Computer Science  1997  401.0       0.077668
192  Computer Science  1998  420.0       0.081348
193  Computer Science  1999  389.0       0.075344
194  Computer Science  2000  451.0       0.087352
195  Computer Science  2001  366.0       0.070889
196  Computer Science  2002  882.0       0.170831
197  Computer Science  2003  957.0       0.185357
2016
Computer Science
           domain  Year  Count   domain_Score
495  Neuroscience  1963    0.0       0.000000
496  Neuroscience  1972    0.0       0.000000
497  Neuroscience  1973    0.0       0.000000
498  Neuroscience  1974    1.0       0.076923
499  Neuroscience  1975    0.0       0.000000
500  Neuroscience  1976    0.0       0.000000
501  Neuroscience  1977    0.0       0.000000
502  Neuroscience  1978    1.0       0.076923
503  Neuroscience  1979    0.0       0.000000
504  Neuroscience  1980    1.0       0.076923
505  Neuroscience  1981    0.0       0.000000
506  Neuroscience  1982    0.0       0.000000
507  Neuroscience  1983    0.0       0.000000
508  Neuroscience  1984    0.0       0.000000
509  Neuroscience  1985    1.0       0.076923
510  Neuroscience  1986    0.0       0.000000
511  Neuroscience  1987    0.0       0.000000
512  Neuroscience  1988    0.0       0.000000
513  Neuroscience  1989    0.0       0.000000
514  Neuroscience  1990    0.0       0.000000
515  Neuroscience  1991    0.0       0.000000
516  Neuroscience  1992    0.0       0.000000
517  Neuroscience  1993    0.0       0.000000
518  Neuroscience  1994    0.0       0.000000
519  Neuroscience  1995    0.0       0.000000
520  Neuroscience  1996    0.0       0.000000
521  Neuroscience  1997    0.0       0.000000
522  Neuroscience  1998    2.0       0.153846
523  Neuroscience  1999    1.0       0.076923
524  Neuroscience  2000    1.0       0.076923
525  Neuroscience  2001    0.0       0.000000
526  Neuroscience  2002    3.0       0.230769
527  Neuroscience  2003    2.0       0.153846
2016
```

Neuroscience

	domain	Year	Count	domain_Score
165	Computer Science	1963	1.0	0.000194
166	Computer Science	1972	5.0	0.000968
167	Computer Science	1973	1.0	0.000194
168	Computer Science	1974	12.0	0.002324
169	Computer Science	1975	1.0	0.000194
170	Computer Science	1976	1.0	0.000194
171	Computer Science	1977	1.0	0.000194
172	Computer Science	1978	3.0	0.000581
173	Computer Science	1979	3.0	0.000581
174	Computer Science	1980	5.0	0.000968
175	Computer Science	1981	6.0	0.001162
176	Computer Science	1982	2.0	0.000387
177	Computer Science	1983	12.0	0.002324
178	Computer Science	1984	2.0	0.000387
179	Computer Science	1985	11.0	0.002131
180	Computer Science	1986	24.0	0.004648
181	Computer Science	1987	20.0	0.003874
182	Computer Science	1988	51.0	0.009878
183	Computer Science	1989	29.0	0.005617
184	Computer Science	1990	39.0	0.007554
185	Computer Science	1991	62.0	0.012009
186	Computer Science	1992	71.0	0.013752
187	Computer Science	1993	98.0	0.018981
188	Computer Science	1994	114.0	0.022080
189	Computer Science	1995	374.0	0.072439
190	Computer Science	1996	349.0	0.067596
191	Computer Science	1997	401.0	0.077668
192	Computer Science	1998	420.0	0.081348
193	Computer Science	1999	389.0	0.075344
194	Computer Science	2000	451.0	0.087352
195	Computer Science	2001	366.0	0.070889
196	Computer Science	2002	882.0	0.170831
197	Computer Science	2003	957.0	0.185357

2016
Computer Science

	domain	Year	Count	domain_Score
165	Computer Science	1963	1.0	0.000194
166	Computer Science	1972	5.0	0.000968
167	Computer Science	1973	1.0	0.000194
168	Computer Science	1974	12.0	0.002324
169	Computer Science	1975	1.0	0.000194
170	Computer Science	1976	1.0	0.000194
171	Computer Science	1977	1.0	0.000194
172	Computer Science	1978	3.0	0.000581
173	Computer Science	1979	3.0	0.000581
174	Computer Science	1980	5.0	0.000968

```
175  Computer Science  1981    6.0       0.001162
176  Computer Science  1982    2.0       0.000387
177  Computer Science  1983   12.0       0.002324
178  Computer Science  1984    2.0       0.000387
179  Computer Science  1985   11.0       0.002131
180  Computer Science  1986   24.0       0.004648
181  Computer Science  1987   20.0       0.003874
182  Computer Science  1988   51.0       0.009878
183  Computer Science  1989   29.0       0.005617
184  Computer Science  1990   39.0       0.007554
185  Computer Science  1991   62.0       0.012009
186  Computer Science  1992   71.0       0.013752
187  Computer Science  1993   98.0       0.018981
188  Computer Science  1994  114.0       0.022080
189  Computer Science  1995  374.0       0.072439
190  Computer Science  1996  349.0       0.067596
191  Computer Science  1997  401.0       0.077668
192  Computer Science  1998  420.0       0.081348
193  Computer Science  1999  389.0       0.075344
194  Computer Science  2000  451.0       0.087352
195  Computer Science  2001  366.0       0.070889
196  Computer Science  2002  882.0       0.170831
197  Computer Science  2003  957.0       0.185357
2014
Computer Science
                domain  Year  Count  domain_Score
165  Computer Science  1963    1.0       0.000194
166  Computer Science  1972    5.0       0.000968
167  Computer Science  1973    1.0       0.000194
168  Computer Science  1974   12.0       0.002324
169  Computer Science  1975    1.0       0.000194
170  Computer Science  1976    1.0       0.000194
171  Computer Science  1977    1.0       0.000194
172  Computer Science  1978    3.0       0.000581
173  Computer Science  1979    3.0       0.000581
174  Computer Science  1980    5.0       0.000968
175  Computer Science  1981    6.0       0.001162
176  Computer Science  1982    2.0       0.000387
177  Computer Science  1983   12.0       0.002324
178  Computer Science  1984    2.0       0.000387
179  Computer Science  1985   11.0       0.002131
180  Computer Science  1986   24.0       0.004648
181  Computer Science  1987   20.0       0.003874
182  Computer Science  1988   51.0       0.009878
183  Computer Science  1989   29.0       0.005617
184  Computer Science  1990   39.0       0.007554
185  Computer Science  1991   62.0       0.012009
186  Computer Science  1992   71.0       0.013752
```

```
187   Computer Science   1993    98.0       0.018981
188   Computer Science   1994   114.0       0.022080
189   Computer Science   1995   374.0       0.072439
190   Computer Science   1996   349.0       0.067596
191   Computer Science   1997   401.0       0.077668
192   Computer Science   1998   420.0       0.081348
193   Computer Science   1999   389.0       0.075344
194   Computer Science   2000   451.0       0.087352
195   Computer Science   2001   366.0       0.070889
196   Computer Science   2002   882.0       0.170831
197   Computer Science   2003   957.0       0.185357
2015
Computer Science
                domain   Year   Count   domain_Score
165   Computer Science   1963     1.0       0.000194
166   Computer Science   1972     5.0       0.000968
167   Computer Science   1973     1.0       0.000194
168   Computer Science   1974    12.0       0.002324
169   Computer Science   1975     1.0       0.000194
170   Computer Science   1976     1.0       0.000194
171   Computer Science   1977     1.0       0.000194
172   Computer Science   1978     3.0       0.000581
173   Computer Science   1979     3.0       0.000581
174   Computer Science   1980     5.0       0.000968
175   Computer Science   1981     6.0       0.001162
176   Computer Science   1982     2.0       0.000387
177   Computer Science   1983    12.0       0.002324
178   Computer Science   1984     2.0       0.000387
179   Computer Science   1985    11.0       0.002131
180   Computer Science   1986    24.0       0.004648
181   Computer Science   1987    20.0       0.003874
182   Computer Science   1988    51.0       0.009878
183   Computer Science   1989    29.0       0.005617
184   Computer Science   1990    39.0       0.007554
185   Computer Science   1991    62.0       0.012009
186   Computer Science   1992    71.0       0.013752
187   Computer Science   1993    98.0       0.018981
188   Computer Science   1994   114.0       0.022080
189   Computer Science   1995   374.0       0.072439
190   Computer Science   1996   349.0       0.067596
191   Computer Science   1997   401.0       0.077668
192   Computer Science   1998   420.0       0.081348
193   Computer Science   1999   389.0       0.075344
194   Computer Science   2000   451.0       0.087352
195   Computer Science   2001   366.0       0.070889
196   Computer Science   2002   882.0       0.170831
197   Computer Science   2003   957.0       0.185357
2015
```

Computer Science

	domain	Year	Count	domain_Score
561	Political	1963	0.0	0.000000
562	Political	1972	0.0	0.000000
563	Political	1973	0.0	0.000000
564	Political	1974	0.0	0.000000
565	Political	1975	0.0	0.000000
566	Political	1976	0.0	0.000000
567	Political	1977	0.0	0.000000
568	Political	1978	0.0	0.000000
569	Political	1979	0.0	0.000000
570	Political	1980	0.0	0.000000
571	Political	1981	1.0	0.003774
572	Political	1982	0.0	0.000000
573	Political	1983	0.0	0.000000
574	Political	1984	1.0	0.003774
575	Political	1985	0.0	0.000000
576	Political	1986	0.0	0.000000
577	Political	1987	1.0	0.003774
578	Political	1988	1.0	0.003774
579	Political	1989	0.0	0.000000
580	Political	1990	5.0	0.018868
581	Political	1991	2.0	0.007547
582	Political	1992	3.0	0.011321
583	Political	1993	1.0	0.003774
584	Political	1994	5.0	0.018868
585	Political	1995	11.0	0.041509
586	Political	1996	20.0	0.075472
587	Political	1997	11.0	0.041509
588	Political	1998	18.0	0.067925
589	Political	1999	26.0	0.098113
590	Political	2000	28.0	0.105660
591	Political	2001	21.0	0.079245
592	Political	2002	49.0	0.184906
593	Political	2003	61.0	0.230189

2016
Political

	domain	Year	Count	domain_Score
330	History	1963	0.0	0.000000
331	History	1972	0.0	0.000000
332	History	1973	0.0	0.000000
333	History	1974	0.0	0.000000
334	History	1975	0.0	0.000000
335	History	1976	0.0	0.000000
336	History	1977	0.0	0.000000
337	History	1978	0.0	0.000000
338	History	1979	0.0	0.000000
339	History	1980	0.0	0.000000

```
340   History  1981    2.0        0.006944
341   History  1982    0.0        0.000000
342   History  1983    1.0        0.003472
343   History  1984    1.0        0.003472
344   History  1985    0.0        0.000000
345   History  1986    3.0        0.010417
346   History  1987    2.0        0.006944
347   History  1988    2.0        0.006944
348   History  1989    1.0        0.003472
349   History  1990    1.0        0.003472
350   History  1991    4.0        0.013889
351   History  1992    3.0        0.010417
352   History  1993    2.0        0.006944
353   History  1994    4.0        0.013889
354   History  1995   12.0        0.041667
355   History  1996   19.0        0.065972
356   History  1997   16.0        0.055556
357   History  1998   20.0        0.069444
358   History  1999   17.0        0.059028
359   History  2000   19.0        0.065972
360   History  2001   31.0        0.107639
361   History  2002   67.0        0.232639
362   History  2003   61.0        0.211806
2016
History
               domain   Year   Count   domain_Score
165   Computer Science  1963    1.0        0.000194
166   Computer Science  1972    5.0        0.000968
167   Computer Science  1973    1.0        0.000194
168   Computer Science  1974   12.0        0.002324
169   Computer Science  1975    1.0        0.000194
170   Computer Science  1976    1.0        0.000194
171   Computer Science  1977    1.0        0.000194
172   Computer Science  1978    3.0        0.000581
173   Computer Science  1979    3.0        0.000581
174   Computer Science  1980    5.0        0.000968
175   Computer Science  1981    6.0        0.001162
176   Computer Science  1982    2.0        0.000387
177   Computer Science  1983   12.0        0.002324
178   Computer Science  1984    2.0        0.000387
179   Computer Science  1985   11.0        0.002131
180   Computer Science  1986   24.0        0.004648
181   Computer Science  1987   20.0        0.003874
182   Computer Science  1988   51.0        0.009878
183   Computer Science  1989   29.0        0.005617
184   Computer Science  1990   39.0        0.007554
185   Computer Science  1991   62.0        0.012009
186   Computer Science  1992   71.0        0.013752
```

```
187  Computer Science  1993   98.0      0.018981
188  Computer Science  1994  114.0      0.022080
189  Computer Science  1995  374.0      0.072439
190  Computer Science  1996  349.0      0.067596
191  Computer Science  1997  401.0      0.077668
192  Computer Science  1998  420.0      0.081348
193  Computer Science  1999  389.0      0.075344
194  Computer Science  2000  451.0      0.087352
195  Computer Science  2001  366.0      0.070889
196  Computer Science  2002  882.0      0.170831
197  Computer Science  2003  957.0      0.185357
2017
Computer Science
                domain  Year  Count  domain_Score
165  Computer Science  1963    1.0      0.000194
166  Computer Science  1972    5.0      0.000968
167  Computer Science  1973    1.0      0.000194
168  Computer Science  1974   12.0      0.002324
169  Computer Science  1975    1.0      0.000194
170  Computer Science  1976    1.0      0.000194
171  Computer Science  1977    1.0      0.000194
172  Computer Science  1978    3.0      0.000581
173  Computer Science  1979    3.0      0.000581
174  Computer Science  1980    5.0      0.000968
175  Computer Science  1981    6.0      0.001162
176  Computer Science  1982    2.0      0.000387
177  Computer Science  1983   12.0      0.002324
178  Computer Science  1984    2.0      0.000387
179  Computer Science  1985   11.0      0.002131
180  Computer Science  1986   24.0      0.004648
181  Computer Science  1987   20.0      0.003874
182  Computer Science  1988   51.0      0.009878
183  Computer Science  1989   29.0      0.005617
184  Computer Science  1990   39.0      0.007554
185  Computer Science  1991   62.0      0.012009
186  Computer Science  1992   71.0      0.013752
187  Computer Science  1993   98.0      0.018981
188  Computer Science  1994  114.0      0.022080
189  Computer Science  1995  374.0      0.072439
190  Computer Science  1996  349.0      0.067596
191  Computer Science  1997  401.0      0.077668
192  Computer Science  1998  420.0      0.081348
193  Computer Science  1999  389.0      0.075344
194  Computer Science  2000  451.0      0.087352
195  Computer Science  2001  366.0      0.070889
196  Computer Science  2002  882.0      0.170831
197  Computer Science  2003  957.0      0.185357
2017
```

Computer Science

```
              domain  Year   Count   domain_Score
165  Computer Science  1963    1.0       0.000194
166  Computer Science  1972    5.0       0.000968
167  Computer Science  1973    1.0       0.000194
168  Computer Science  1974   12.0       0.002324
169  Computer Science  1975    1.0       0.000194
170  Computer Science  1976    1.0       0.000194
171  Computer Science  1977    1.0       0.000194
172  Computer Science  1978    3.0       0.000581
173  Computer Science  1979    3.0       0.000581
174  Computer Science  1980    5.0       0.000968
175  Computer Science  1981    6.0       0.001162
176  Computer Science  1982    2.0       0.000387
177  Computer Science  1983   12.0       0.002324
178  Computer Science  1984    2.0       0.000387
179  Computer Science  1985   11.0       0.002131
180  Computer Science  1986   24.0       0.004648
181  Computer Science  1987   20.0       0.003874
182  Computer Science  1988   51.0       0.009878
183  Computer Science  1989   29.0       0.005617
184  Computer Science  1990   39.0       0.007554
185  Computer Science  1991   62.0       0.012009
186  Computer Science  1992   71.0       0.013752
187  Computer Science  1993   98.0       0.018981
188  Computer Science  1994  114.0       0.022080
189  Computer Science  1995  374.0       0.072439
190  Computer Science  1996  349.0       0.067596
191  Computer Science  1997  401.0       0.077668
192  Computer Science  1998  420.0       0.081348
193  Computer Science  1999  389.0       0.075344
194  Computer Science  2000  451.0       0.087352
195  Computer Science  2001  366.0       0.070889
196  Computer Science  2002  882.0       0.170831
197  Computer Science  2003  957.0       0.185357
2017
Computer Science
          domain  Year  Count   domain_Score
561  Political  1963    0.0       0.000000
562  Political  1972    0.0       0.000000
563  Political  1973    0.0       0.000000
564  Political  1974    0.0       0.000000
565  Political  1975    0.0       0.000000
566  Political  1976    0.0       0.000000
567  Political  1977    0.0       0.000000
568  Political  1978    0.0       0.000000
569  Political  1979    0.0       0.000000
570  Political  1980    0.0       0.000000
```

```
571  Political  1981    1.0       0.003774
572  Political  1982    0.0       0.000000
573  Political  1983    0.0       0.000000
574  Political  1984    1.0       0.003774
575  Political  1985    0.0       0.000000
576  Political  1986    0.0       0.000000
577  Political  1987    1.0       0.003774
578  Political  1988    1.0       0.003774
579  Political  1989    0.0       0.000000
580  Political  1990    5.0       0.018868
581  Political  1991    2.0       0.007547
582  Political  1992    3.0       0.011321
583  Political  1993    1.0       0.003774
584  Political  1994    5.0       0.018868
585  Political  1995   11.0       0.041509
586  Political  1996   20.0       0.075472
587  Political  1997   11.0       0.041509
588  Political  1998   18.0       0.067925
589  Political  1999   26.0       0.098113
590  Political  2000   28.0       0.105660
591  Political  2001   21.0       0.079245
592  Political  2002   49.0       0.184906
593  Political  2003   61.0       0.230189
2017
Political
             domain  Year  Count  domain_Score
198  Earth Science   1963    0.0       0.000000
199  Earth Science   1972    0.0       0.000000
200  Earth Science   1973    1.0       0.001754
201  Earth Science   1974    6.0       0.010526
202  Earth Science   1975    0.0       0.000000
203  Earth Science   1976    0.0       0.000000
204  Earth Science   1977    3.0       0.005263
205  Earth Science   1978    2.0       0.003509
206  Earth Science   1979    2.0       0.003509
207  Earth Science   1980    4.0       0.007018
208  Earth Science   1981    3.0       0.005263
209  Earth Science   1982    0.0       0.000000
210  Earth Science   1983    0.0       0.000000
211  Earth Science   1984    5.0       0.008772
212  Earth Science   1985    4.0       0.007018
213  Earth Science   1986    3.0       0.005263
214  Earth Science   1987    8.0       0.014035
215  Earth Science   1988    8.0       0.014035
216  Earth Science   1989    4.0       0.007018
217  Earth Science   1990   14.0       0.024561
218  Earth Science   1991    5.0       0.008772
219  Earth Science   1992    9.0       0.015789
```

```
220  Earth Science  1993   17.0      0.029825
221  Earth Science  1994    9.0      0.015789
222  Earth Science  1995   23.0      0.040351
223  Earth Science  1996   28.0      0.049123
224  Earth Science  1997   41.0      0.071930
225  Earth Science  1998   35.0      0.061404
226  Earth Science  1999   27.0      0.047368
227  Earth Science  2000   23.0      0.040351
228  Earth Science  2001   27.0      0.047368
229  Earth Science  2002  105.0      0.184211
230  Earth Science  2003  154.0      0.270175
2017
Earth Science
        domain  Year  Count  domain_Score
561  Political  1963    0.0      0.000000
562  Political  1972    0.0      0.000000
563  Political  1973    0.0      0.000000
564  Political  1974    0.0      0.000000
565  Political  1975    0.0      0.000000
566  Political  1976    0.0      0.000000
567  Political  1977    0.0      0.000000
568  Political  1978    0.0      0.000000
569  Political  1979    0.0      0.000000
570  Political  1980    0.0      0.000000
571  Political  1981    1.0      0.003774
572  Political  1982    0.0      0.000000
573  Political  1983    0.0      0.000000
574  Political  1984    1.0      0.003774
575  Political  1985    0.0      0.000000
576  Political  1986    0.0      0.000000
577  Political  1987    1.0      0.003774
578  Political  1988    1.0      0.003774
579  Political  1989    0.0      0.000000
580  Political  1990    5.0      0.018868
581  Political  1991    2.0      0.007547
582  Political  1992    3.0      0.011321
583  Political  1993    1.0      0.003774
584  Political  1994    5.0      0.018868
585  Political  1995   11.0      0.041509
586  Political  1996   20.0      0.075472
587  Political  1997   11.0      0.041509
588  Political  1998   18.0      0.067925
589  Political  1999   26.0      0.098113
590  Political  2000   28.0      0.105660
591  Political  2001   21.0      0.079245
592  Political  2002   49.0      0.184906
593  Political  2003   61.0      0.230189
2017
```

Political

	domain	Year	Count	domain_Score
198	Earth Science	1963	0.0	0.000000
199	Earth Science	1972	0.0	0.000000
200	Earth Science	1973	1.0	0.001754
201	Earth Science	1974	6.0	0.010526
202	Earth Science	1975	0.0	0.000000
203	Earth Science	1976	0.0	0.000000
204	Earth Science	1977	3.0	0.005263
205	Earth Science	1978	2.0	0.003509
206	Earth Science	1979	2.0	0.003509
207	Earth Science	1980	4.0	0.007018
208	Earth Science	1981	3.0	0.005263
209	Earth Science	1982	0.0	0.000000
210	Earth Science	1983	0.0	0.000000
211	Earth Science	1984	5.0	0.008772
212	Earth Science	1985	4.0	0.007018
213	Earth Science	1986	3.0	0.005263
214	Earth Science	1987	8.0	0.014035
215	Earth Science	1988	8.0	0.014035
216	Earth Science	1989	4.0	0.007018
217	Earth Science	1990	14.0	0.024561
218	Earth Science	1991	5.0	0.008772
219	Earth Science	1992	9.0	0.015789
220	Earth Science	1993	17.0	0.029825
221	Earth Science	1994	9.0	0.015789
222	Earth Science	1995	23.0	0.040351
223	Earth Science	1996	28.0	0.049123
224	Earth Science	1997	41.0	0.071930
225	Earth Science	1998	35.0	0.061404
226	Earth Science	1999	27.0	0.047368
227	Earth Science	2000	23.0	0.040351
228	Earth Science	2001	27.0	0.047368
229	Earth Science	2002	105.0	0.184211
230	Earth Science	2003	154.0	0.270175

2017
Earth Science

	domain	Year	Count	domain_Score
165	Computer Science	1963	1.0	0.000194
166	Computer Science	1972	5.0	0.000968
167	Computer Science	1973	1.0	0.000194
168	Computer Science	1974	12.0	0.002324
169	Computer Science	1975	1.0	0.000194
170	Computer Science	1976	1.0	0.000194
171	Computer Science	1977	1.0	0.000194
172	Computer Science	1978	3.0	0.000581
173	Computer Science	1979	3.0	0.000581
174	Computer Science	1980	5.0	0.000968

```
175  Computer Science  1981    6.0       0.001162
176  Computer Science  1982    2.0       0.000387
177  Computer Science  1983   12.0       0.002324
178  Computer Science  1984    2.0       0.000387
179  Computer Science  1985   11.0       0.002131
180  Computer Science  1986   24.0       0.004648
181  Computer Science  1987   20.0       0.003874
182  Computer Science  1988   51.0       0.009878
183  Computer Science  1989   29.0       0.005617
184  Computer Science  1990   39.0       0.007554
185  Computer Science  1991   62.0       0.012009
186  Computer Science  1992   71.0       0.013752
187  Computer Science  1993   98.0       0.018981
188  Computer Science  1994  114.0       0.022080
189  Computer Science  1995  374.0       0.072439
190  Computer Science  1996  349.0       0.067596
191  Computer Science  1997  401.0       0.077668
192  Computer Science  1998  420.0       0.081348
193  Computer Science  1999  389.0       0.075344
194  Computer Science  2000  451.0       0.087352
195  Computer Science  2001  366.0       0.070889
196  Computer Science  2002  882.0       0.170831
197  Computer Science  2003  957.0       0.185357
2017
Computer Science
              domain  Year  Count  domain_Score
165  Computer Science  1963    1.0       0.000194
166  Computer Science  1972    5.0       0.000968
167  Computer Science  1973    1.0       0.000194
168  Computer Science  1974   12.0       0.002324
169  Computer Science  1975    1.0       0.000194
170  Computer Science  1976    1.0       0.000194
171  Computer Science  1977    1.0       0.000194
172  Computer Science  1978    3.0       0.000581
173  Computer Science  1979    3.0       0.000581
174  Computer Science  1980    5.0       0.000968
175  Computer Science  1981    6.0       0.001162
176  Computer Science  1982    2.0       0.000387
177  Computer Science  1983   12.0       0.002324
178  Computer Science  1984    2.0       0.000387
179  Computer Science  1985   11.0       0.002131
180  Computer Science  1986   24.0       0.004648
181  Computer Science  1987   20.0       0.003874
182  Computer Science  1988   51.0       0.009878
183  Computer Science  1989   29.0       0.005617
184  Computer Science  1990   39.0       0.007554
185  Computer Science  1991   62.0       0.012009
186  Computer Science  1992   71.0       0.013752
```

```
187  Computer Science  1993   98.0       0.018981
188  Computer Science  1994  114.0       0.022080
189  Computer Science  1995  374.0       0.072439
190  Computer Science  1996  349.0       0.067596
191  Computer Science  1997  401.0       0.077668
192  Computer Science  1998  420.0       0.081348
193  Computer Science  1999  389.0       0.075344
194  Computer Science  2000  451.0       0.087352
195  Computer Science  2001  366.0       0.070889
196  Computer Science  2002  882.0       0.170831
197  Computer Science  2003  957.0       0.185357
2017
Computer Science
                domain  Year  Count  domain_Score
165  Computer Science  1963    1.0       0.000194
166  Computer Science  1972    5.0       0.000968
167  Computer Science  1973    1.0       0.000194
168  Computer Science  1974   12.0       0.002324
169  Computer Science  1975    1.0       0.000194
170  Computer Science  1976    1.0       0.000194
171  Computer Science  1977    1.0       0.000194
172  Computer Science  1978    3.0       0.000581
173  Computer Science  1979    3.0       0.000581
174  Computer Science  1980    5.0       0.000968
175  Computer Science  1981    6.0       0.001162
176  Computer Science  1982    2.0       0.000387
177  Computer Science  1983   12.0       0.002324
178  Computer Science  1984    2.0       0.000387
179  Computer Science  1985   11.0       0.002131
180  Computer Science  1986   24.0       0.004648
181  Computer Science  1987   20.0       0.003874
182  Computer Science  1988   51.0       0.009878
183  Computer Science  1989   29.0       0.005617
184  Computer Science  1990   39.0       0.007554
185  Computer Science  1991   62.0       0.012009
186  Computer Science  1992   71.0       0.013752
187  Computer Science  1993   98.0       0.018981
188  Computer Science  1994  114.0       0.022080
189  Computer Science  1995  374.0       0.072439
190  Computer Science  1996  349.0       0.067596
191  Computer Science  1997  401.0       0.077668
192  Computer Science  1998  420.0       0.081348
193  Computer Science  1999  389.0       0.075344
194  Computer Science  2000  451.0       0.087352
195  Computer Science  2001  366.0       0.070889
196  Computer Science  2002  882.0       0.170831
197  Computer Science  2003  957.0       0.185357
2017
```

Computer Science

	domain	Year	Count	domain_Score
165	Computer Science	1963	1.0	0.000194
166	Computer Science	1972	5.0	0.000968
167	Computer Science	1973	1.0	0.000194
168	Computer Science	1974	12.0	0.002324
169	Computer Science	1975	1.0	0.000194
170	Computer Science	1976	1.0	0.000194
171	Computer Science	1977	1.0	0.000194
172	Computer Science	1978	3.0	0.000581
173	Computer Science	1979	3.0	0.000581
174	Computer Science	1980	5.0	0.000968
175	Computer Science	1981	6.0	0.001162
176	Computer Science	1982	2.0	0.000387
177	Computer Science	1983	12.0	0.002324
178	Computer Science	1984	2.0	0.000387
179	Computer Science	1985	11.0	0.002131
180	Computer Science	1986	24.0	0.004648
181	Computer Science	1987	20.0	0.003874
182	Computer Science	1988	51.0	0.009878
183	Computer Science	1989	29.0	0.005617
184	Computer Science	1990	39.0	0.007554
185	Computer Science	1991	62.0	0.012009
186	Computer Science	1992	71.0	0.013752
187	Computer Science	1993	98.0	0.018981
188	Computer Science	1994	114.0	0.022080
189	Computer Science	1995	374.0	0.072439
190	Computer Science	1996	349.0	0.067596
191	Computer Science	1997	401.0	0.077668
192	Computer Science	1998	420.0	0.081348
193	Computer Science	1999	389.0	0.075344
194	Computer Science	2000	451.0	0.087352
195	Computer Science	2001	366.0	0.070889
196	Computer Science	2002	882.0	0.170831
197	Computer Science	2003	957.0	0.185357

2018
Computer Science

	domain	Year	Count	domain_Score
165	Computer Science	1963	1.0	0.000194
166	Computer Science	1972	5.0	0.000968
167	Computer Science	1973	1.0	0.000194
168	Computer Science	1974	12.0	0.002324
169	Computer Science	1975	1.0	0.000194
170	Computer Science	1976	1.0	0.000194
171	Computer Science	1977	1.0	0.000194
172	Computer Science	1978	3.0	0.000581
173	Computer Science	1979	3.0	0.000581
174	Computer Science	1980	5.0	0.000968

```
175  Computer Science  1981    6.0      0.001162
176  Computer Science  1982    2.0      0.000387
177  Computer Science  1983   12.0      0.002324
178  Computer Science  1984    2.0      0.000387
179  Computer Science  1985   11.0      0.002131
180  Computer Science  1986   24.0      0.004648
181  Computer Science  1987   20.0      0.003874
182  Computer Science  1988   51.0      0.009878
183  Computer Science  1989   29.0      0.005617
184  Computer Science  1990   39.0      0.007554
185  Computer Science  1991   62.0      0.012009
186  Computer Science  1992   71.0      0.013752
187  Computer Science  1993   98.0      0.018981
188  Computer Science  1994  114.0      0.022080
189  Computer Science  1995  374.0      0.072439
190  Computer Science  1996  349.0      0.067596
191  Computer Science  1997  401.0      0.077668
192  Computer Science  1998  420.0      0.081348
193  Computer Science  1999  389.0      0.075344
194  Computer Science  2000  451.0      0.087352
195  Computer Science  2001  366.0      0.070889
196  Computer Science  2002  882.0      0.170831
197  Computer Science  2003  957.0      0.185357
2018
Computer Science
                domain  Year  Count  domain_Score
165  Computer Science  1963    1.0      0.000194
166  Computer Science  1972    5.0      0.000968
167  Computer Science  1973    1.0      0.000194
168  Computer Science  1974   12.0      0.002324
169  Computer Science  1975    1.0      0.000194
170  Computer Science  1976    1.0      0.000194
171  Computer Science  1977    1.0      0.000194
172  Computer Science  1978    3.0      0.000581
173  Computer Science  1979    3.0      0.000581
174  Computer Science  1980    5.0      0.000968
175  Computer Science  1981    6.0      0.001162
176  Computer Science  1982    2.0      0.000387
177  Computer Science  1983   12.0      0.002324
178  Computer Science  1984    2.0      0.000387
179  Computer Science  1985   11.0      0.002131
180  Computer Science  1986   24.0      0.004648
181  Computer Science  1987   20.0      0.003874
182  Computer Science  1988   51.0      0.009878
183  Computer Science  1989   29.0      0.005617
184  Computer Science  1990   39.0      0.007554
185  Computer Science  1991   62.0      0.012009
186  Computer Science  1992   71.0      0.013752
```

```
187  Computer Science  1993   98.0      0.018981
188  Computer Science  1994  114.0      0.022080
189  Computer Science  1995  374.0      0.072439
190  Computer Science  1996  349.0      0.067596
191  Computer Science  1997  401.0      0.077668
192  Computer Science  1998  420.0      0.081348
193  Computer Science  1999  389.0      0.075344
194  Computer Science  2000  451.0      0.087352
195  Computer Science  2001  366.0      0.070889
196  Computer Science  2002  882.0      0.170831
197  Computer Science  2003  957.0      0.185357
2015
Computer Science
                domain  Year  Count  domain_Score
165  Computer Science  1963    1.0      0.000194
166  Computer Science  1972    5.0      0.000968
167  Computer Science  1973    1.0      0.000194
168  Computer Science  1974   12.0      0.002324
169  Computer Science  1975    1.0      0.000194
170  Computer Science  1976    1.0      0.000194
171  Computer Science  1977    1.0      0.000194
172  Computer Science  1978    3.0      0.000581
173  Computer Science  1979    3.0      0.000581
174  Computer Science  1980    5.0      0.000968
175  Computer Science  1981    6.0      0.001162
176  Computer Science  1982    2.0      0.000387
177  Computer Science  1983   12.0      0.002324
178  Computer Science  1984    2.0      0.000387
179  Computer Science  1985   11.0      0.002131
180  Computer Science  1986   24.0      0.004648
181  Computer Science  1987   20.0      0.003874
182  Computer Science  1988   51.0      0.009878
183  Computer Science  1989   29.0      0.005617
184  Computer Science  1990   39.0      0.007554
185  Computer Science  1991   62.0      0.012009
186  Computer Science  1992   71.0      0.013752
187  Computer Science  1993   98.0      0.018981
188  Computer Science  1994  114.0      0.022080
189  Computer Science  1995  374.0      0.072439
190  Computer Science  1996  349.0      0.067596
191  Computer Science  1997  401.0      0.077668
192  Computer Science  1998  420.0      0.081348
193  Computer Science  1999  389.0      0.075344
194  Computer Science  2000  451.0      0.087352
195  Computer Science  2001  366.0      0.070889
196  Computer Science  2002  882.0      0.170831
197  Computer Science  2003  957.0      0.185357
2016
```

Computer Science

	domain	Year	Count	domain_Score
330	History	1963	0.0	0.000000
331	History	1972	0.0	0.000000
332	History	1973	0.0	0.000000
333	History	1974	0.0	0.000000
334	History	1975	0.0	0.000000
335	History	1976	0.0	0.000000
336	History	1977	0.0	0.000000
337	History	1978	0.0	0.000000
338	History	1979	0.0	0.000000
339	History	1980	0.0	0.000000
340	History	1981	2.0	0.006944
341	History	1982	0.0	0.000000
342	History	1983	1.0	0.003472
343	History	1984	1.0	0.003472
344	History	1985	0.0	0.000000
345	History	1986	3.0	0.010417
346	History	1987	2.0	0.006944
347	History	1988	2.0	0.006944
348	History	1989	1.0	0.003472
349	History	1990	1.0	0.003472
350	History	1991	4.0	0.013889
351	History	1992	3.0	0.010417
352	History	1993	2.0	0.006944
353	History	1994	4.0	0.013889
354	History	1995	12.0	0.041667
355	History	1996	19.0	0.065972
356	History	1997	16.0	0.055556
357	History	1998	20.0	0.069444
358	History	1999	17.0	0.059028
359	History	2000	19.0	0.065972
360	History	2001	31.0	0.107639
361	History	2002	67.0	0.232639
362	History	2003	61.0	0.211806

2017
History

	domain	Year	Count	domain_Score
165	Computer Science	1963	1.0	0.000194
166	Computer Science	1972	5.0	0.000968
167	Computer Science	1973	1.0	0.000194
168	Computer Science	1974	12.0	0.002324
169	Computer Science	1975	1.0	0.000194
170	Computer Science	1976	1.0	0.000194
171	Computer Science	1977	1.0	0.000194
172	Computer Science	1978	3.0	0.000581
173	Computer Science	1979	3.0	0.000581
174	Computer Science	1980	5.0	0.000968

```
175   Computer Science  1981    6.0       0.001162
176   Computer Science  1982    2.0       0.000387
177   Computer Science  1983   12.0       0.002324
178   Computer Science  1984    2.0       0.000387
179   Computer Science  1985   11.0       0.002131
180   Computer Science  1986   24.0       0.004648
181   Computer Science  1987   20.0       0.003874
182   Computer Science  1988   51.0       0.009878
183   Computer Science  1989   29.0       0.005617
184   Computer Science  1990   39.0       0.007554
185   Computer Science  1991   62.0       0.012009
186   Computer Science  1992   71.0       0.013752
187   Computer Science  1993   98.0       0.018981
188   Computer Science  1994  114.0       0.022080
189   Computer Science  1995  374.0       0.072439
190   Computer Science  1996  349.0       0.067596
191   Computer Science  1997  401.0       0.077668
192   Computer Science  1998  420.0       0.081348
193   Computer Science  1999  389.0       0.075344
194   Computer Science  2000  451.0       0.087352
195   Computer Science  2001  366.0       0.070889
196   Computer Science  2002  882.0       0.170831
197   Computer Science  2003  957.0       0.185357
2016
Computer Science
                domain  Year  Count  domain_Score
165   Computer Science  1963    1.0       0.000194
166   Computer Science  1972    5.0       0.000968
167   Computer Science  1973    1.0       0.000194
168   Computer Science  1974   12.0       0.002324
169   Computer Science  1975    1.0       0.000194
170   Computer Science  1976    1.0       0.000194
171   Computer Science  1977    1.0       0.000194
172   Computer Science  1978    3.0       0.000581
173   Computer Science  1979    3.0       0.000581
174   Computer Science  1980    5.0       0.000968
175   Computer Science  1981    6.0       0.001162
176   Computer Science  1982    2.0       0.000387
177   Computer Science  1983   12.0       0.002324
178   Computer Science  1984    2.0       0.000387
179   Computer Science  1985   11.0       0.002131
180   Computer Science  1986   24.0       0.004648
181   Computer Science  1987   20.0       0.003874
182   Computer Science  1988   51.0       0.009878
183   Computer Science  1989   29.0       0.005617
184   Computer Science  1990   39.0       0.007554
185   Computer Science  1991   62.0       0.012009
186   Computer Science  1992   71.0       0.013752
```

```
187   Computer Science   1993    98.0      0.018981
188   Computer Science   1994   114.0      0.022080
189   Computer Science   1995   374.0      0.072439
190   Computer Science   1996   349.0      0.067596
191   Computer Science   1997   401.0      0.077668
192   Computer Science   1998   420.0      0.081348
193   Computer Science   1999   389.0      0.075344
194   Computer Science   2000   451.0      0.087352
195   Computer Science   2001   366.0      0.070889
196   Computer Science   2002   882.0      0.170831
197   Computer Science   2003   957.0      0.185357
2016
Computer Science
                domain  Year  Count  domain_Score
165   Computer Science   1963    1.0      0.000194
166   Computer Science   1972    5.0      0.000968
167   Computer Science   1973    1.0      0.000194
168   Computer Science   1974   12.0      0.002324
169   Computer Science   1975    1.0      0.000194
170   Computer Science   1976    1.0      0.000194
171   Computer Science   1977    1.0      0.000194
172   Computer Science   1978    3.0      0.000581
173   Computer Science   1979    3.0      0.000581
174   Computer Science   1980    5.0      0.000968
175   Computer Science   1981    6.0      0.001162
176   Computer Science   1982    2.0      0.000387
177   Computer Science   1983   12.0      0.002324
178   Computer Science   1984    2.0      0.000387
179   Computer Science   1985   11.0      0.002131
180   Computer Science   1986   24.0      0.004648
181   Computer Science   1987   20.0      0.003874
182   Computer Science   1988   51.0      0.009878
183   Computer Science   1989   29.0      0.005617
184   Computer Science   1990   39.0      0.007554
185   Computer Science   1991   62.0      0.012009
186   Computer Science   1992   71.0      0.013752
187   Computer Science   1993   98.0      0.018981
188   Computer Science   1994  114.0      0.022080
189   Computer Science   1995  374.0      0.072439
190   Computer Science   1996  349.0      0.067596
191   Computer Science   1997  401.0      0.077668
192   Computer Science   1998  420.0      0.081348
193   Computer Science   1999  389.0      0.075344
194   Computer Science   2000  451.0      0.087352
195   Computer Science   2001  366.0      0.070889
196   Computer Science   2002  882.0      0.170831
197   Computer Science   2003  957.0      0.185357
2016
```

Computer Science

	domain	Year	Count	domain_Score
165	Computer Science	1963	1.0	0.000194
166	Computer Science	1972	5.0	0.000968
167	Computer Science	1973	1.0	0.000194
168	Computer Science	1974	12.0	0.002324
169	Computer Science	1975	1.0	0.000194
170	Computer Science	1976	1.0	0.000194
171	Computer Science	1977	1.0	0.000194
172	Computer Science	1978	3.0	0.000581
173	Computer Science	1979	3.0	0.000581
174	Computer Science	1980	5.0	0.000968
175	Computer Science	1981	6.0	0.001162
176	Computer Science	1982	2.0	0.000387
177	Computer Science	1983	12.0	0.002324
178	Computer Science	1984	2.0	0.000387
179	Computer Science	1985	11.0	0.002131
180	Computer Science	1986	24.0	0.004648
181	Computer Science	1987	20.0	0.003874
182	Computer Science	1988	51.0	0.009878
183	Computer Science	1989	29.0	0.005617
184	Computer Science	1990	39.0	0.007554
185	Computer Science	1991	62.0	0.012009
186	Computer Science	1992	71.0	0.013752
187	Computer Science	1993	98.0	0.018981
188	Computer Science	1994	114.0	0.022080
189	Computer Science	1995	374.0	0.072439
190	Computer Science	1996	349.0	0.067596
191	Computer Science	1997	401.0	0.077668
192	Computer Science	1998	420.0	0.081348
193	Computer Science	1999	389.0	0.075344
194	Computer Science	2000	451.0	0.087352
195	Computer Science	2001	366.0	0.070889
196	Computer Science	2002	882.0	0.170831
197	Computer Science	2003	957.0	0.185357

2016
Computer Science

	domain	Year	Count	domain_Score
165	Computer Science	1963	1.0	0.000194
166	Computer Science	1972	5.0	0.000968
167	Computer Science	1973	1.0	0.000194
168	Computer Science	1974	12.0	0.002324
169	Computer Science	1975	1.0	0.000194
170	Computer Science	1976	1.0	0.000194
171	Computer Science	1977	1.0	0.000194
172	Computer Science	1978	3.0	0.000581
173	Computer Science	1979	3.0	0.000581
174	Computer Science	1980	5.0	0.000968

```
175  Computer Science  1981    6.0        0.001162
176  Computer Science  1982    2.0        0.000387
177  Computer Science  1983   12.0        0.002324
178  Computer Science  1984    2.0        0.000387
179  Computer Science  1985   11.0        0.002131
180  Computer Science  1986   24.0        0.004648
181  Computer Science  1987   20.0        0.003874
182  Computer Science  1988   51.0        0.009878
183  Computer Science  1989   29.0        0.005617
184  Computer Science  1990   39.0        0.007554
185  Computer Science  1991   62.0        0.012009
186  Computer Science  1992   71.0        0.013752
187  Computer Science  1993   98.0        0.018981
188  Computer Science  1994  114.0        0.022080
189  Computer Science  1995  374.0        0.072439
190  Computer Science  1996  349.0        0.067596
191  Computer Science  1997  401.0        0.077668
192  Computer Science  1998  420.0        0.081348
193  Computer Science  1999  389.0        0.075344
194  Computer Science  2000  451.0        0.087352
195  Computer Science  2001  366.0        0.070889
196  Computer Science  2002  882.0        0.170831
197  Computer Science  2003  957.0        0.185357
2016
Computer Science
     domain  Year  Count  domain_Score
297  Health  1963    0.0      0.000000
298  Health  1972    0.0      0.000000
299  Health  1973    0.0      0.000000
300  Health  1974    0.0      0.000000
301  Health  1975    0.0      0.000000
302  Health  1976    1.0      0.006211
303  Health  1977    0.0      0.000000
304  Health  1978    0.0      0.000000
305  Health  1979    0.0      0.000000
306  Health  1980    0.0      0.000000
307  Health  1981    0.0      0.000000
308  Health  1982    0.0      0.000000
309  Health  1983    1.0      0.006211
310  Health  1984    2.0      0.012422
311  Health  1985    0.0      0.000000
312  Health  1986    1.0      0.006211
313  Health  1987    1.0      0.006211
314  Health  1988    0.0      0.000000
315  Health  1989    0.0      0.000000
316  Health  1990    0.0      0.000000
317  Health  1991    0.0      0.000000
318  Health  1992    0.0      0.000000
```

```
319  Health  1993    0.0        0.000000
320  Health  1994    2.0        0.012422
321  Health  1995   10.0        0.062112
322  Health  1996    7.0        0.043478
323  Health  1997    9.0        0.055901
324  Health  1998   12.0        0.074534
325  Health  1999    9.0        0.055901
326  Health  2000   13.0        0.080745
327  Health  2001   16.0        0.099379
328  Health  2002   45.0        0.279503
329  Health  2003   32.0        0.198758
2016
Health
               domain  Year  Count  domain_Score
165  Computer Science  1963    1.0      0.000194
166  Computer Science  1972    5.0      0.000968
167  Computer Science  1973    1.0      0.000194
168  Computer Science  1974   12.0      0.002324
169  Computer Science  1975    1.0      0.000194
170  Computer Science  1976    1.0      0.000194
171  Computer Science  1977    1.0      0.000194
172  Computer Science  1978    3.0      0.000581
173  Computer Science  1979    3.0      0.000581
174  Computer Science  1980    5.0      0.000968
175  Computer Science  1981    6.0      0.001162
176  Computer Science  1982    2.0      0.000387
177  Computer Science  1983   12.0      0.002324
178  Computer Science  1984    2.0      0.000387
179  Computer Science  1985   11.0      0.002131
180  Computer Science  1986   24.0      0.004648
181  Computer Science  1987   20.0      0.003874
182  Computer Science  1988   51.0      0.009878
183  Computer Science  1989   29.0      0.005617
184  Computer Science  1990   39.0      0.007554
185  Computer Science  1991   62.0      0.012009
186  Computer Science  1992   71.0      0.013752
187  Computer Science  1993   98.0      0.018981
188  Computer Science  1994  114.0      0.022080
189  Computer Science  1995  374.0      0.072439
190  Computer Science  1996  349.0      0.067596
191  Computer Science  1997  401.0      0.077668
192  Computer Science  1998  420.0      0.081348
193  Computer Science  1999  389.0      0.075344
194  Computer Science  2000  451.0      0.087352
195  Computer Science  2001  366.0      0.070889
196  Computer Science  2002  882.0      0.170831
197  Computer Science  2003  957.0      0.185357
2016
```

```
Computer Science
            domain  Year  Count  domain_Score
165  Computer Science  1963    1.0      0.000194
166  Computer Science  1972    5.0      0.000968
167  Computer Science  1973    1.0      0.000194
168  Computer Science  1974   12.0      0.002324
169  Computer Science  1975    1.0      0.000194
170  Computer Science  1976    1.0      0.000194
171  Computer Science  1977    1.0      0.000194
172  Computer Science  1978    3.0      0.000581
173  Computer Science  1979    3.0      0.000581
174  Computer Science  1980    5.0      0.000968
175  Computer Science  1981    6.0      0.001162
176  Computer Science  1982    2.0      0.000387
177  Computer Science  1983   12.0      0.002324
178  Computer Science  1984    2.0      0.000387
179  Computer Science  1985   11.0      0.002131
180  Computer Science  1986   24.0      0.004648
181  Computer Science  1987   20.0      0.003874
182  Computer Science  1988   51.0      0.009878
183  Computer Science  1989   29.0      0.005617
184  Computer Science  1990   39.0      0.007554
185  Computer Science  1991   62.0      0.012009
186  Computer Science  1992   71.0      0.013752
187  Computer Science  1993   98.0      0.018981
188  Computer Science  1994  114.0      0.022080
189  Computer Science  1995  374.0      0.072439
190  Computer Science  1996  349.0      0.067596
191  Computer Science  1997  401.0      0.077668
192  Computer Science  1998  420.0      0.081348
193  Computer Science  1999  389.0      0.075344
194  Computer Science  2000  451.0      0.087352
195  Computer Science  2001  366.0      0.070889
196  Computer Science  2002  882.0      0.170831
197  Computer Science  2003  957.0      0.185357
2017
Computer Science
            domain  Year  Count  domain_Score
165  Computer Science  1963    1.0      0.000194
166  Computer Science  1972    5.0      0.000968
167  Computer Science  1973    1.0      0.000194
168  Computer Science  1974   12.0      0.002324
169  Computer Science  1975    1.0      0.000194
170  Computer Science  1976    1.0      0.000194
171  Computer Science  1977    1.0      0.000194
172  Computer Science  1978    3.0      0.000581
173  Computer Science  1979    3.0      0.000581
174  Computer Science  1980    5.0      0.000968
```

```
175  Computer Science  1981    6.0       0.001162
176  Computer Science  1982    2.0       0.000387
177  Computer Science  1983   12.0       0.002324
178  Computer Science  1984    2.0       0.000387
179  Computer Science  1985   11.0       0.002131
180  Computer Science  1986   24.0       0.004648
181  Computer Science  1987   20.0       0.003874
182  Computer Science  1988   51.0       0.009878
183  Computer Science  1989   29.0       0.005617
184  Computer Science  1990   39.0       0.007554
185  Computer Science  1991   62.0       0.012009
186  Computer Science  1992   71.0       0.013752
187  Computer Science  1993   98.0       0.018981
188  Computer Science  1994  114.0       0.022080
189  Computer Science  1995  374.0       0.072439
190  Computer Science  1996  349.0       0.067596
191  Computer Science  1997  401.0       0.077668
192  Computer Science  1998  420.0       0.081348
193  Computer Science  1999  389.0       0.075344
194  Computer Science  2000  451.0       0.087352
195  Computer Science  2001  366.0       0.070889
196  Computer Science  2002  882.0       0.170831
197  Computer Science  2003  957.0       0.185357
2017
Computer Science
               domain  Year  Count  domain_Score
165  Computer Science  1963    1.0       0.000194
166  Computer Science  1972    5.0       0.000968
167  Computer Science  1973    1.0       0.000194
168  Computer Science  1974   12.0       0.002324
169  Computer Science  1975    1.0       0.000194
170  Computer Science  1976    1.0       0.000194
171  Computer Science  1977    1.0       0.000194
172  Computer Science  1978    3.0       0.000581
173  Computer Science  1979    3.0       0.000581
174  Computer Science  1980    5.0       0.000968
175  Computer Science  1981    6.0       0.001162
176  Computer Science  1982    2.0       0.000387
177  Computer Science  1983   12.0       0.002324
178  Computer Science  1984    2.0       0.000387
179  Computer Science  1985   11.0       0.002131
180  Computer Science  1986   24.0       0.004648
181  Computer Science  1987   20.0       0.003874
182  Computer Science  1988   51.0       0.009878
183  Computer Science  1989   29.0       0.005617
184  Computer Science  1990   39.0       0.007554
185  Computer Science  1991   62.0       0.012009
186  Computer Science  1992   71.0       0.013752
```

```
187  Computer Science  1993   98.0        0.018981
188  Computer Science  1994  114.0        0.022080
189  Computer Science  1995  374.0        0.072439
190  Computer Science  1996  349.0        0.067596
191  Computer Science  1997  401.0        0.077668
192  Computer Science  1998  420.0        0.081348
193  Computer Science  1999  389.0        0.075344
194  Computer Science  2000  451.0        0.087352
195  Computer Science  2001  366.0        0.070889
196  Computer Science  2002  882.0        0.170831
197  Computer Science  2003  957.0        0.185357
2017
Computer Science
                domain  Year  Count  domain_Score
165  Computer Science  1963    1.0        0.000194
166  Computer Science  1972    5.0        0.000968
167  Computer Science  1973    1.0        0.000194
168  Computer Science  1974   12.0        0.002324
169  Computer Science  1975    1.0        0.000194
170  Computer Science  1976    1.0        0.000194
171  Computer Science  1977    1.0        0.000194
172  Computer Science  1978    3.0        0.000581
173  Computer Science  1979    3.0        0.000581
174  Computer Science  1980    5.0        0.000968
175  Computer Science  1981    6.0        0.001162
176  Computer Science  1982    2.0        0.000387
177  Computer Science  1983   12.0        0.002324
178  Computer Science  1984    2.0        0.000387
179  Computer Science  1985   11.0        0.002131
180  Computer Science  1986   24.0        0.004648
181  Computer Science  1987   20.0        0.003874
182  Computer Science  1988   51.0        0.009878
183  Computer Science  1989   29.0        0.005617
184  Computer Science  1990   39.0        0.007554
185  Computer Science  1991   62.0        0.012009
186  Computer Science  1992   71.0        0.013752
187  Computer Science  1993   98.0        0.018981
188  Computer Science  1994  114.0        0.022080
189  Computer Science  1995  374.0        0.072439
190  Computer Science  1996  349.0        0.067596
191  Computer Science  1997  401.0        0.077668
192  Computer Science  1998  420.0        0.081348
193  Computer Science  1999  389.0        0.075344
194  Computer Science  2000  451.0        0.087352
195  Computer Science  2001  366.0        0.070889
196  Computer Science  2002  882.0        0.170831
197  Computer Science  2003  957.0        0.185357
2017
```

Computer Science

	domain	Year	Count	domain_Score
561	Political	1963	0.0	0.000000
562	Political	1972	0.0	0.000000
563	Political	1973	0.0	0.000000
564	Political	1974	0.0	0.000000
565	Political	1975	0.0	0.000000
566	Political	1976	0.0	0.000000
567	Political	1977	0.0	0.000000
568	Political	1978	0.0	0.000000
569	Political	1979	0.0	0.000000
570	Political	1980	0.0	0.000000
571	Political	1981	1.0	0.003774
572	Political	1982	0.0	0.000000
573	Political	1983	0.0	0.000000
574	Political	1984	1.0	0.003774
575	Political	1985	0.0	0.000000
576	Political	1986	0.0	0.000000
577	Political	1987	1.0	0.003774
578	Political	1988	1.0	0.003774
579	Political	1989	0.0	0.000000
580	Political	1990	5.0	0.018868
581	Political	1991	2.0	0.007547
582	Political	1992	3.0	0.011321
583	Political	1993	1.0	0.003774
584	Political	1994	5.0	0.018868
585	Political	1995	11.0	0.041509
586	Political	1996	20.0	0.075472
587	Political	1997	11.0	0.041509
588	Political	1998	18.0	0.067925
589	Political	1999	26.0	0.098113
590	Political	2000	28.0	0.105660
591	Political	2001	21.0	0.079245
592	Political	2002	49.0	0.184906
593	Political	2003	61.0	0.230189

2017
Political

	domain	Year	Count	domain_Score
66	Biomedical Research	1963	0.0	0.000000
67	Biomedical Research	1972	0.0	0.000000
68	Biomedical Research	1973	0.0	0.000000
69	Biomedical Research	1974	2.0	0.009050
70	Biomedical Research	1975	0.0	0.000000
71	Biomedical Research	1976	0.0	0.000000
72	Biomedical Research	1977	1.0	0.004525
73	Biomedical Research	1978	2.0	0.009050
74	Biomedical Research	1979	5.0	0.022624
75	Biomedical Research	1980	1.0	0.004525

```
76   Biomedical Research  1981    3.0        0.013575
77   Biomedical Research  1982    1.0        0.004525
78   Biomedical Research  1983    1.0        0.004525
79   Biomedical Research  1984    7.0        0.031674
80   Biomedical Research  1985    1.0        0.004525
81   Biomedical Research  1986    4.0        0.018100
82   Biomedical Research  1987    2.0        0.009050
83   Biomedical Research  1988    3.0        0.013575
84   Biomedical Research  1989    2.0        0.009050
85   Biomedical Research  1990    5.0        0.022624
86   Biomedical Research  1991    4.0        0.018100
87   Biomedical Research  1992    4.0        0.018100
88   Biomedical Research  1993    8.0        0.036199
89   Biomedical Research  1994    3.0        0.013575
90   Biomedical Research  1995   12.0        0.054299
91   Biomedical Research  1996   12.0        0.054299
92   Biomedical Research  1997   11.0        0.049774
93   Biomedical Research  1998   11.0        0.049774
94   Biomedical Research  1999    9.0        0.040724
95   Biomedical Research  2000   14.0        0.063348
96   Biomedical Research  2001   15.0        0.067873
97   Biomedical Research  2002   48.0        0.217195
98   Biomedical Research  2003   30.0        0.135747
2017
Biomedical Research
              domain  Year  Count  domain_Score
165  Computer Science  1963    1.0        0.000194
166  Computer Science  1972    5.0        0.000968
167  Computer Science  1973    1.0        0.000194
168  Computer Science  1974   12.0        0.002324
169  Computer Science  1975    1.0        0.000194
170  Computer Science  1976    1.0        0.000194
171  Computer Science  1977    1.0        0.000194
172  Computer Science  1978    3.0        0.000581
173  Computer Science  1979    3.0        0.000581
174  Computer Science  1980    5.0        0.000968
175  Computer Science  1981    6.0        0.001162
176  Computer Science  1982    2.0        0.000387
177  Computer Science  1983   12.0        0.002324
178  Computer Science  1984    2.0        0.000387
179  Computer Science  1985   11.0        0.002131
180  Computer Science  1986   24.0        0.004648
181  Computer Science  1987   20.0        0.003874
182  Computer Science  1988   51.0        0.009878
183  Computer Science  1989   29.0        0.005617
184  Computer Science  1990   39.0        0.007554
185  Computer Science  1991   62.0        0.012009
186  Computer Science  1992   71.0        0.013752
```

```
187  Computer Science  1993   98.0      0.018981
188  Computer Science  1994  114.0      0.022080
189  Computer Science  1995  374.0      0.072439
190  Computer Science  1996  349.0      0.067596
191  Computer Science  1997  401.0      0.077668
192  Computer Science  1998  420.0      0.081348
193  Computer Science  1999  389.0      0.075344
194  Computer Science  2000  451.0      0.087352
195  Computer Science  2001  366.0      0.070889
196  Computer Science  2002  882.0      0.170831
197  Computer Science  2003  957.0      0.185357
2017
Computer Science
                 domain  Year  Count  domain_Score
165  Computer Science  1963    1.0      0.000194
166  Computer Science  1972    5.0      0.000968
167  Computer Science  1973    1.0      0.000194
168  Computer Science  1974   12.0      0.002324
169  Computer Science  1975    1.0      0.000194
170  Computer Science  1976    1.0      0.000194
171  Computer Science  1977    1.0      0.000194
172  Computer Science  1978    3.0      0.000581
173  Computer Science  1979    3.0      0.000581
174  Computer Science  1980    5.0      0.000968
175  Computer Science  1981    6.0      0.001162
176  Computer Science  1982    2.0      0.000387
177  Computer Science  1983   12.0      0.002324
178  Computer Science  1984    2.0      0.000387
179  Computer Science  1985   11.0      0.002131
180  Computer Science  1986   24.0      0.004648
181  Computer Science  1987   20.0      0.003874
182  Computer Science  1988   51.0      0.009878
183  Computer Science  1989   29.0      0.005617
184  Computer Science  1990   39.0      0.007554
185  Computer Science  1991   62.0      0.012009
186  Computer Science  1992   71.0      0.013752
187  Computer Science  1993   98.0      0.018981
188  Computer Science  1994  114.0      0.022080
189  Computer Science  1995  374.0      0.072439
190  Computer Science  1996  349.0      0.067596
191  Computer Science  1997  401.0      0.077668
192  Computer Science  1998  420.0      0.081348
193  Computer Science  1999  389.0      0.075344
194  Computer Science  2000  451.0      0.087352
195  Computer Science  2001  366.0      0.070889
196  Computer Science  2002  882.0      0.170831
197  Computer Science  2003  957.0      0.185357
2017
```

Computer Science

```
                domain  Year   Count   domain_Score
165   Computer Science  1963    1.0      0.000194
166   Computer Science  1972    5.0      0.000968
167   Computer Science  1973    1.0      0.000194
168   Computer Science  1974   12.0      0.002324
169   Computer Science  1975    1.0      0.000194
170   Computer Science  1976    1.0      0.000194
171   Computer Science  1977    1.0      0.000194
172   Computer Science  1978    3.0      0.000581
173   Computer Science  1979    3.0      0.000581
174   Computer Science  1980    5.0      0.000968
175   Computer Science  1981    6.0      0.001162
176   Computer Science  1982    2.0      0.000387
177   Computer Science  1983   12.0      0.002324
178   Computer Science  1984    2.0      0.000387
179   Computer Science  1985   11.0      0.002131
180   Computer Science  1986   24.0      0.004648
181   Computer Science  1987   20.0      0.003874
182   Computer Science  1988   51.0      0.009878
183   Computer Science  1989   29.0      0.005617
184   Computer Science  1990   39.0      0.007554
185   Computer Science  1991   62.0      0.012009
186   Computer Science  1992   71.0      0.013752
187   Computer Science  1993   98.0      0.018981
188   Computer Science  1994  114.0      0.022080
189   Computer Science  1995  374.0      0.072439
190   Computer Science  1996  349.0      0.067596
191   Computer Science  1997  401.0      0.077668
192   Computer Science  1998  420.0      0.081348
193   Computer Science  1999  389.0      0.075344
194   Computer Science  2000  451.0      0.087352
195   Computer Science  2001  366.0      0.070889
196   Computer Science  2002  882.0      0.170831
197   Computer Science  2003  957.0      0.185357
2015
```

Computer Science

```
                domain  Year   Count   domain_Score
165   Computer Science  1963    1.0      0.000194
166   Computer Science  1972    5.0      0.000968
167   Computer Science  1973    1.0      0.000194
168   Computer Science  1974   12.0      0.002324
169   Computer Science  1975    1.0      0.000194
170   Computer Science  1976    1.0      0.000194
171   Computer Science  1977    1.0      0.000194
172   Computer Science  1978    3.0      0.000581
173   Computer Science  1979    3.0      0.000581
174   Computer Science  1980    5.0      0.000968
```

```
175  Computer Science  1981    6.0        0.001162
176  Computer Science  1982    2.0        0.000387
177  Computer Science  1983   12.0        0.002324
178  Computer Science  1984    2.0        0.000387
179  Computer Science  1985   11.0        0.002131
180  Computer Science  1986   24.0        0.004648
181  Computer Science  1987   20.0        0.003874
182  Computer Science  1988   51.0        0.009878
183  Computer Science  1989   29.0        0.005617
184  Computer Science  1990   39.0        0.007554
185  Computer Science  1991   62.0        0.012009
186  Computer Science  1992   71.0        0.013752
187  Computer Science  1993   98.0        0.018981
188  Computer Science  1994  114.0        0.022080
189  Computer Science  1995  374.0        0.072439
190  Computer Science  1996  349.0        0.067596
191  Computer Science  1997  401.0        0.077668
192  Computer Science  1998  420.0        0.081348
193  Computer Science  1999  389.0        0.075344
194  Computer Science  2000  451.0        0.087352
195  Computer Science  2001  366.0        0.070889
196  Computer Science  2002  882.0        0.170831
197  Computer Science  2003  957.0        0.185357
2017
Computer Science
       domain  Year  Count  domain_Score
594  Software  1963    0.0      0.000000
595  Software  1972    0.0      0.000000
596  Software  1973    0.0      0.000000
597  Software  1974    5.0      0.004955
598  Software  1975    1.0      0.000991
599  Software  1976    0.0      0.000000
600  Software  1977    1.0      0.000991
601  Software  1978    4.0      0.003964
602  Software  1979    2.0      0.001982
603  Software  1980    1.0      0.000991
604  Software  1981    1.0      0.000991
605  Software  1982    1.0      0.000991
606  Software  1983    0.0      0.000000
607  Software  1984    4.0      0.003964
608  Software  1985    1.0      0.000991
609  Software  1986    0.0      0.000000
610  Software  1987    4.0      0.003964
611  Software  1988    4.0      0.003964
612  Software  1989    3.0      0.002973
613  Software  1990   10.0      0.009911
614  Software  1991   14.0      0.013875
615  Software  1992    8.0      0.007929
```

```
616  Software  1993  18.0      0.017839
617  Software  1994  23.0      0.022795
618  Software  1995  50.0      0.049554
619  Software  1996  74.0      0.073340
620  Software  1997  69.0      0.068385
621  Software  1998  90.0      0.089197
622  Software  1999  94.0      0.093162
623  Software  2000  81.0      0.080278
624  Software  2001  76.0      0.075322
625  Software  2002  179.0     0.177403
626  Software  2003  191.0     0.189296
2016
Software
               domain  Year  Count  domain_Score
165  Computer Science  1963    1.0      0.000194
166  Computer Science  1972    5.0      0.000968
167  Computer Science  1973    1.0      0.000194
168  Computer Science  1974   12.0      0.002324
169  Computer Science  1975    1.0      0.000194
170  Computer Science  1976    1.0      0.000194
171  Computer Science  1977    1.0      0.000194
172  Computer Science  1978    3.0      0.000581
173  Computer Science  1979    3.0      0.000581
174  Computer Science  1980    5.0      0.000968
175  Computer Science  1981    6.0      0.001162
176  Computer Science  1982    2.0      0.000387
177  Computer Science  1983   12.0      0.002324
178  Computer Science  1984    2.0      0.000387
179  Computer Science  1985   11.0      0.002131
180  Computer Science  1986   24.0      0.004648
181  Computer Science  1987   20.0      0.003874
182  Computer Science  1988   51.0      0.009878
183  Computer Science  1989   29.0      0.005617
184  Computer Science  1990   39.0      0.007554
185  Computer Science  1991   62.0      0.012009
186  Computer Science  1992   71.0      0.013752
187  Computer Science  1993   98.0      0.018981
188  Computer Science  1994  114.0      0.022080
189  Computer Science  1995  374.0      0.072439
190  Computer Science  1996  349.0      0.067596
191  Computer Science  1997  401.0      0.077668
192  Computer Science  1998  420.0      0.081348
193  Computer Science  1999  389.0      0.075344
194  Computer Science  2000  451.0      0.087352
195  Computer Science  2001  366.0      0.070889
196  Computer Science  2002  882.0      0.170831
197  Computer Science  2003  957.0      0.185357
2016
```

Computer Science

	domain	Year	Count	domain_Score
594	Software	1963	0.0	0.000000
595	Software	1972	0.0	0.000000
596	Software	1973	0.0	0.000000
597	Software	1974	5.0	0.004955
598	Software	1975	1.0	0.000991
599	Software	1976	0.0	0.000000
600	Software	1977	1.0	0.000991
601	Software	1978	4.0	0.003964
602	Software	1979	2.0	0.001982
603	Software	1980	1.0	0.000991
604	Software	1981	1.0	0.000991
605	Software	1982	1.0	0.000991
606	Software	1983	0.0	0.000000
607	Software	1984	4.0	0.003964
608	Software	1985	1.0	0.000991
609	Software	1986	0.0	0.000000
610	Software	1987	4.0	0.003964
611	Software	1988	4.0	0.003964
612	Software	1989	3.0	0.002973
613	Software	1990	10.0	0.009911
614	Software	1991	14.0	0.013875
615	Software	1992	8.0	0.007929
616	Software	1993	18.0	0.017839
617	Software	1994	23.0	0.022795
618	Software	1995	50.0	0.049554
619	Software	1996	74.0	0.073340
620	Software	1997	69.0	0.068385
621	Software	1998	90.0	0.089197
622	Software	1999	94.0	0.093162
623	Software	2000	81.0	0.080278
624	Software	2001	76.0	0.075322
625	Software	2002	179.0	0.177403
626	Software	2003	191.0	0.189296

2014
Software

	domain	Year	Count	domain_Score
165	Computer Science	1963	1.0	0.000194
166	Computer Science	1972	5.0	0.000968
167	Computer Science	1973	1.0	0.000194
168	Computer Science	1974	12.0	0.002324
169	Computer Science	1975	1.0	0.000194
170	Computer Science	1976	1.0	0.000194
171	Computer Science	1977	1.0	0.000194
172	Computer Science	1978	3.0	0.000581
173	Computer Science	1979	3.0	0.000581
174	Computer Science	1980	5.0	0.000968

```
175   Computer Science   1981    6.0       0.001162
176   Computer Science   1982    2.0       0.000387
177   Computer Science   1983   12.0       0.002324
178   Computer Science   1984    2.0       0.000387
179   Computer Science   1985   11.0       0.002131
180   Computer Science   1986   24.0       0.004648
181   Computer Science   1987   20.0       0.003874
182   Computer Science   1988   51.0       0.009878
183   Computer Science   1989   29.0       0.005617
184   Computer Science   1990   39.0       0.007554
185   Computer Science   1991   62.0       0.012009
186   Computer Science   1992   71.0       0.013752
187   Computer Science   1993   98.0       0.018981
188   Computer Science   1994  114.0       0.022080
189   Computer Science   1995  374.0       0.072439
190   Computer Science   1996  349.0       0.067596
191   Computer Science   1997  401.0       0.077668
192   Computer Science   1998  420.0       0.081348
193   Computer Science   1999  389.0       0.075344
194   Computer Science   2000  451.0       0.087352
195   Computer Science   2001  366.0       0.070889
196   Computer Science   2002  882.0       0.170831
197   Computer Science   2003  957.0       0.185357
2016
Computer Science
               domain  Year  Count  domain_Score
165   Computer Science   1963    1.0       0.000194
166   Computer Science   1972    5.0       0.000968
167   Computer Science   1973    1.0       0.000194
168   Computer Science   1974   12.0       0.002324
169   Computer Science   1975    1.0       0.000194
170   Computer Science   1976    1.0       0.000194
171   Computer Science   1977    1.0       0.000194
172   Computer Science   1978    3.0       0.000581
173   Computer Science   1979    3.0       0.000581
174   Computer Science   1980    5.0       0.000968
175   Computer Science   1981    6.0       0.001162
176   Computer Science   1982    2.0       0.000387
177   Computer Science   1983   12.0       0.002324
178   Computer Science   1984    2.0       0.000387
179   Computer Science   1985   11.0       0.002131
180   Computer Science   1986   24.0       0.004648
181   Computer Science   1987   20.0       0.003874
182   Computer Science   1988   51.0       0.009878
183   Computer Science   1989   29.0       0.005617
184   Computer Science   1990   39.0       0.007554
185   Computer Science   1991   62.0       0.012009
186   Computer Science   1992   71.0       0.013752
```

```
187  Computer Science  1993   98.0        0.018981
188  Computer Science  1994  114.0        0.022080
189  Computer Science  1995  374.0        0.072439
190  Computer Science  1996  349.0        0.067596
191  Computer Science  1997  401.0        0.077668
192  Computer Science  1998  420.0        0.081348
193  Computer Science  1999  389.0        0.075344
194  Computer Science  2000  451.0        0.087352
195  Computer Science  2001  366.0        0.070889
196  Computer Science  2002  882.0        0.170831
197  Computer Science  2003  957.0        0.185357
2015
Computer Science
        domain  Year  Count  domain_Score
264  Geometry  1963    0.0      0.000000
265  Geometry  1972    0.0      0.000000
266  Geometry  1973    0.0      0.000000
267  Geometry  1974    1.0      0.001808
268  Geometry  1975    0.0      0.000000
269  Geometry  1976    0.0      0.000000
270  Geometry  1977    0.0      0.000000
271  Geometry  1978    0.0      0.000000
272  Geometry  1979    0.0      0.000000
273  Geometry  1980    0.0      0.000000
274  Geometry  1981    2.0      0.003617
275  Geometry  1982    0.0      0.000000
276  Geometry  1983    0.0      0.000000
277  Geometry  1984    2.0      0.003617
278  Geometry  1985    3.0      0.005425
279  Geometry  1986    5.0      0.009042
280  Geometry  1987    1.0      0.001808
281  Geometry  1988   10.0      0.018083
282  Geometry  1989    1.0      0.001808
283  Geometry  1990    4.0      0.007233
284  Geometry  1991    3.0      0.005425
285  Geometry  1992   12.0      0.021700
286  Geometry  1993   10.0      0.018083
287  Geometry  1994    8.0      0.014467
288  Geometry  1995   36.0      0.065099
289  Geometry  1996   53.0      0.095841
290  Geometry  1997   44.0      0.079566
291  Geometry  1998   27.0      0.048825
292  Geometry  1999   37.0      0.066908
293  Geometry  2000   34.0      0.061483
294  Geometry  2001   25.0      0.045208
295  Geometry  2002   92.0      0.166365
296  Geometry  2003  143.0      0.258590
2016
```

Geometry

```
         domain  Year   Count  domain_Score
165  Computer Science  1963    1.0      0.000194
166  Computer Science  1972    5.0      0.000968
167  Computer Science  1973    1.0      0.000194
168  Computer Science  1974   12.0      0.002324
169  Computer Science  1975    1.0      0.000194
170  Computer Science  1976    1.0      0.000194
171  Computer Science  1977    1.0      0.000194
172  Computer Science  1978    3.0      0.000581
173  Computer Science  1979    3.0      0.000581
174  Computer Science  1980    5.0      0.000968
175  Computer Science  1981    6.0      0.001162
176  Computer Science  1982    2.0      0.000387
177  Computer Science  1983   12.0      0.002324
178  Computer Science  1984    2.0      0.000387
179  Computer Science  1985   11.0      0.002131
180  Computer Science  1986   24.0      0.004648
181  Computer Science  1987   20.0      0.003874
182  Computer Science  1988   51.0      0.009878
183  Computer Science  1989   29.0      0.005617
184  Computer Science  1990   39.0      0.007554
185  Computer Science  1991   62.0      0.012009
186  Computer Science  1992   71.0      0.013752
187  Computer Science  1993   98.0      0.018981
188  Computer Science  1994  114.0      0.022080
189  Computer Science  1995  374.0      0.072439
190  Computer Science  1996  349.0      0.067596
191  Computer Science  1997  401.0      0.077668
192  Computer Science  1998  420.0      0.081348
193  Computer Science  1999  389.0      0.075344
194  Computer Science  2000  451.0      0.087352
195  Computer Science  2001  366.0      0.070889
196  Computer Science  2002  882.0      0.170831
197  Computer Science  2003  957.0      0.185357
```
2016
Computer Science
```
         domain  Year   Count  domain_Score
165  Computer Science  1963    1.0      0.000194
166  Computer Science  1972    5.0      0.000968
167  Computer Science  1973    1.0      0.000194
168  Computer Science  1974   12.0      0.002324
169  Computer Science  1975    1.0      0.000194
170  Computer Science  1976    1.0      0.000194
171  Computer Science  1977    1.0      0.000194
172  Computer Science  1978    3.0      0.000581
173  Computer Science  1979    3.0      0.000581
174  Computer Science  1980    5.0      0.000968
```

```
175  Computer Science  1981    6.0      0.001162
176  Computer Science  1982    2.0      0.000387
177  Computer Science  1983   12.0      0.002324
178  Computer Science  1984    2.0      0.000387
179  Computer Science  1985   11.0      0.002131
180  Computer Science  1986   24.0      0.004648
181  Computer Science  1987   20.0      0.003874
182  Computer Science  1988   51.0      0.009878
183  Computer Science  1989   29.0      0.005617
184  Computer Science  1990   39.0      0.007554
185  Computer Science  1991   62.0      0.012009
186  Computer Science  1992   71.0      0.013752
187  Computer Science  1993   98.0      0.018981
188  Computer Science  1994  114.0      0.022080
189  Computer Science  1995  374.0      0.072439
190  Computer Science  1996  349.0      0.067596
191  Computer Science  1997  401.0      0.077668
192  Computer Science  1998  420.0      0.081348
193  Computer Science  1999  389.0      0.075344
194  Computer Science  2000  451.0      0.087352
195  Computer Science  2001  366.0      0.070889
196  Computer Science  2002  882.0      0.170831
197  Computer Science  2003  957.0      0.185357
2015
Computer Science
                domain  Year  Count  domain_Score
165  Computer Science  1963    1.0      0.000194
166  Computer Science  1972    5.0      0.000968
167  Computer Science  1973    1.0      0.000194
168  Computer Science  1974   12.0      0.002324
169  Computer Science  1975    1.0      0.000194
170  Computer Science  1976    1.0      0.000194
171  Computer Science  1977    1.0      0.000194
172  Computer Science  1978    3.0      0.000581
173  Computer Science  1979    3.0      0.000581
174  Computer Science  1980    5.0      0.000968
175  Computer Science  1981    6.0      0.001162
176  Computer Science  1982    2.0      0.000387
177  Computer Science  1983   12.0      0.002324
178  Computer Science  1984    2.0      0.000387
179  Computer Science  1985   11.0      0.002131
180  Computer Science  1986   24.0      0.004648
181  Computer Science  1987   20.0      0.003874
182  Computer Science  1988   51.0      0.009878
183  Computer Science  1989   29.0      0.005617
184  Computer Science  1990   39.0      0.007554
185  Computer Science  1991   62.0      0.012009
186  Computer Science  1992   71.0      0.013752
```

```
187  Computer Science  1993   98.0      0.018981
188  Computer Science  1994  114.0      0.022080
189  Computer Science  1995  374.0      0.072439
190  Computer Science  1996  349.0      0.067596
191  Computer Science  1997  401.0      0.077668
192  Computer Science  1998  420.0      0.081348
193  Computer Science  1999  389.0      0.075344
194  Computer Science  2000  451.0      0.087352
195  Computer Science  2001  366.0      0.070889
196  Computer Science  2002  882.0      0.170831
197  Computer Science  2003  957.0      0.185357
2015
Computer Science
                domain  Year  Count  domain_Score
165  Computer Science  1963    1.0      0.000194
166  Computer Science  1972    5.0      0.000968
167  Computer Science  1973    1.0      0.000194
168  Computer Science  1974   12.0      0.002324
169  Computer Science  1975    1.0      0.000194
170  Computer Science  1976    1.0      0.000194
171  Computer Science  1977    1.0      0.000194
172  Computer Science  1978    3.0      0.000581
173  Computer Science  1979    3.0      0.000581
174  Computer Science  1980    5.0      0.000968
175  Computer Science  1981    6.0      0.001162
176  Computer Science  1982    2.0      0.000387
177  Computer Science  1983   12.0      0.002324
178  Computer Science  1984    2.0      0.000387
179  Computer Science  1985   11.0      0.002131
180  Computer Science  1986   24.0      0.004648
181  Computer Science  1987   20.0      0.003874
182  Computer Science  1988   51.0      0.009878
183  Computer Science  1989   29.0      0.005617
184  Computer Science  1990   39.0      0.007554
185  Computer Science  1991   62.0      0.012009
186  Computer Science  1992   71.0      0.013752
187  Computer Science  1993   98.0      0.018981
188  Computer Science  1994  114.0      0.022080
189  Computer Science  1995  374.0      0.072439
190  Computer Science  1996  349.0      0.067596
191  Computer Science  1997  401.0      0.077668
192  Computer Science  1998  420.0      0.081348
193  Computer Science  1999  389.0      0.075344
194  Computer Science  2000  451.0      0.087352
195  Computer Science  2001  366.0      0.070889
196  Computer Science  2002  882.0      0.170831
197  Computer Science  2003  957.0      0.185357
2015
```

Computer Science

	domain	Year	Count	domain_Score
165	Computer Science	1963	1.0	0.000194
166	Computer Science	1972	5.0	0.000968
167	Computer Science	1973	1.0	0.000194
168	Computer Science	1974	12.0	0.002324
169	Computer Science	1975	1.0	0.000194
170	Computer Science	1976	1.0	0.000194
171	Computer Science	1977	1.0	0.000194
172	Computer Science	1978	3.0	0.000581
173	Computer Science	1979	3.0	0.000581
174	Computer Science	1980	5.0	0.000968
175	Computer Science	1981	6.0	0.001162
176	Computer Science	1982	2.0	0.000387
177	Computer Science	1983	12.0	0.002324
178	Computer Science	1984	2.0	0.000387
179	Computer Science	1985	11.0	0.002131
180	Computer Science	1986	24.0	0.004648
181	Computer Science	1987	20.0	0.003874
182	Computer Science	1988	51.0	0.009878
183	Computer Science	1989	29.0	0.005617
184	Computer Science	1990	39.0	0.007554
185	Computer Science	1991	62.0	0.012009
186	Computer Science	1992	71.0	0.013752
187	Computer Science	1993	98.0	0.018981
188	Computer Science	1994	114.0	0.022080
189	Computer Science	1995	374.0	0.072439
190	Computer Science	1996	349.0	0.067596
191	Computer Science	1997	401.0	0.077668
192	Computer Science	1998	420.0	0.081348
193	Computer Science	1999	389.0	0.075344
194	Computer Science	2000	451.0	0.087352
195	Computer Science	2001	366.0	0.070889
196	Computer Science	2002	882.0	0.170831
197	Computer Science	2003	957.0	0.185357

2015
Computer Science

	domain	Year	Count	domain_Score
99	Business	1963	0.0	0.000000
100	Business	1972	0.0	0.000000
101	Business	1973	0.0	0.000000
102	Business	1974	1.0	0.000953
103	Business	1975	1.0	0.000953
104	Business	1976	0.0	0.000000
105	Business	1977	0.0	0.000000
106	Business	1978	1.0	0.000953
107	Business	1979	2.0	0.001907
108	Business	1980	2.0	0.001907

```
109  Business  1981    2.0        0.001907
110  Business  1982    1.0        0.000953
111  Business  1983    1.0        0.000953
112  Business  1984    1.0        0.000953
113  Business  1985    1.0        0.000953
114  Business  1986    1.0        0.000953
115  Business  1987    1.0        0.000953
116  Business  1988    2.0        0.001907
117  Business  1989    4.0        0.003813
118  Business  1990    4.0        0.003813
119  Business  1991   12.0        0.011439
120  Business  1992    8.0        0.007626
121  Business  1993   14.0        0.013346
122  Business  1994    8.0        0.007626
123  Business  1995   28.0        0.026692
124  Business  1996   51.0        0.048618
125  Business  1997   42.0        0.040038
126  Business  1998  108.0        0.102955
127  Business  1999   88.0        0.083889
128  Business  2000  105.0        0.100095
129  Business  2001   98.0        0.093422
130  Business  2002  202.0        0.192564
131  Business  2003  260.0        0.247855
2015
Business
               domain  Year  Count  domain_Score
165  Computer Science  1963    1.0      0.000194
166  Computer Science  1972    5.0      0.000968
167  Computer Science  1973    1.0      0.000194
168  Computer Science  1974   12.0      0.002324
169  Computer Science  1975    1.0      0.000194
170  Computer Science  1976    1.0      0.000194
171  Computer Science  1977    1.0      0.000194
172  Computer Science  1978    3.0      0.000581
173  Computer Science  1979    3.0      0.000581
174  Computer Science  1980    5.0      0.000968
175  Computer Science  1981    6.0      0.001162
176  Computer Science  1982    2.0      0.000387
177  Computer Science  1983   12.0      0.002324
178  Computer Science  1984    2.0      0.000387
179  Computer Science  1985   11.0      0.002131
180  Computer Science  1986   24.0      0.004648
181  Computer Science  1987   20.0      0.003874
182  Computer Science  1988   51.0      0.009878
183  Computer Science  1989   29.0      0.005617
184  Computer Science  1990   39.0      0.007554
185  Computer Science  1991   62.0      0.012009
186  Computer Science  1992   71.0      0.013752
```

```
187  Computer Science  1993    98.0        0.018981
188  Computer Science  1994   114.0        0.022080
189  Computer Science  1995   374.0        0.072439
190  Computer Science  1996   349.0        0.067596
191  Computer Science  1997   401.0        0.077668
192  Computer Science  1998   420.0        0.081348
193  Computer Science  1999   389.0        0.075344
194  Computer Science  2000   451.0        0.087352
195  Computer Science  2001   366.0        0.070889
196  Computer Science  2002   882.0        0.170831
197  Computer Science  2003   957.0        0.185357
2015
Computer Science
            domain  Year  Count  domain_Score
198  Earth Science  1963    0.0        0.000000
199  Earth Science  1972    0.0        0.000000
200  Earth Science  1973    1.0        0.001754
201  Earth Science  1974    6.0        0.010526
202  Earth Science  1975    0.0        0.000000
203  Earth Science  1976    0.0        0.000000
204  Earth Science  1977    3.0        0.005263
205  Earth Science  1978    2.0        0.003509
206  Earth Science  1979    2.0        0.003509
207  Earth Science  1980    4.0        0.007018
208  Earth Science  1981    3.0        0.005263
209  Earth Science  1982    0.0        0.000000
210  Earth Science  1983    0.0        0.000000
211  Earth Science  1984    5.0        0.008772
212  Earth Science  1985    4.0        0.007018
213  Earth Science  1986    3.0        0.005263
214  Earth Science  1987    8.0        0.014035
215  Earth Science  1988    8.0        0.014035
216  Earth Science  1989    4.0        0.007018
217  Earth Science  1990   14.0        0.024561
218  Earth Science  1991    5.0        0.008772
219  Earth Science  1992    9.0        0.015789
220  Earth Science  1993   17.0        0.029825
221  Earth Science  1994    9.0        0.015789
222  Earth Science  1995   23.0        0.040351
223  Earth Science  1996   28.0        0.049123
224  Earth Science  1997   41.0        0.071930
225  Earth Science  1998   35.0        0.061404
226  Earth Science  1999   27.0        0.047368
227  Earth Science  2000   23.0        0.040351
228  Earth Science  2001   27.0        0.047368
229  Earth Science  2002  105.0        0.184211
230  Earth Science  2003  154.0        0.270175
2015
```

Earth Science

	domain	Year	Count	domain_Score
165	Computer Science	1963	1.0	0.000194
166	Computer Science	1972	5.0	0.000968
167	Computer Science	1973	1.0	0.000194
168	Computer Science	1974	12.0	0.002324
169	Computer Science	1975	1.0	0.000194
170	Computer Science	1976	1.0	0.000194
171	Computer Science	1977	1.0	0.000194
172	Computer Science	1978	3.0	0.000581
173	Computer Science	1979	3.0	0.000581
174	Computer Science	1980	5.0	0.000968
175	Computer Science	1981	6.0	0.001162
176	Computer Science	1982	2.0	0.000387
177	Computer Science	1983	12.0	0.002324
178	Computer Science	1984	2.0	0.000387
179	Computer Science	1985	11.0	0.002131
180	Computer Science	1986	24.0	0.004648
181	Computer Science	1987	20.0	0.003874
182	Computer Science	1988	51.0	0.009878
183	Computer Science	1989	29.0	0.005617
184	Computer Science	1990	39.0	0.007554
185	Computer Science	1991	62.0	0.012009
186	Computer Science	1992	71.0	0.013752
187	Computer Science	1993	98.0	0.018981
188	Computer Science	1994	114.0	0.022080
189	Computer Science	1995	374.0	0.072439
190	Computer Science	1996	349.0	0.067596
191	Computer Science	1997	401.0	0.077668
192	Computer Science	1998	420.0	0.081348
193	Computer Science	1999	389.0	0.075344
194	Computer Science	2000	451.0	0.087352
195	Computer Science	2001	366.0	0.070889
196	Computer Science	2002	882.0	0.170831
197	Computer Science	2003	957.0	0.185357

2015
Computer Science

	domain	Year	Count	domain_Score
330	History	1963	0.0	0.000000
331	History	1972	0.0	0.000000
332	History	1973	0.0	0.000000
333	History	1974	0.0	0.000000
334	History	1975	0.0	0.000000
335	History	1976	0.0	0.000000
336	History	1977	0.0	0.000000
337	History	1978	0.0	0.000000
338	History	1979	0.0	0.000000
339	History	1980	0.0	0.000000

```
340   History   1981    2.0        0.006944
341   History   1982    0.0        0.000000
342   History   1983    1.0        0.003472
343   History   1984    1.0        0.003472
344   History   1985    0.0        0.000000
345   History   1986    3.0        0.010417
346   History   1987    2.0        0.006944
347   History   1988    2.0        0.006944
348   History   1989    1.0        0.003472
349   History   1990    1.0        0.003472
350   History   1991    4.0        0.013889
351   History   1992    3.0        0.010417
352   History   1993    2.0        0.006944
353   History   1994    4.0        0.013889
354   History   1995   12.0        0.041667
355   History   1996   19.0        0.065972
356   History   1997   16.0        0.055556
357   History   1998   20.0        0.069444
358   History   1999   17.0        0.059028
359   History   2000   19.0        0.065972
360   History   2001   31.0        0.107639
361   History   2002   67.0        0.232639
362   History   2003   61.0        0.211806
2015
History
                 domain   Year   Count   domain_Score
165   Computer Science   1963    1.0        0.000194
166   Computer Science   1972    5.0        0.000968
167   Computer Science   1973    1.0        0.000194
168   Computer Science   1974   12.0        0.002324
169   Computer Science   1975    1.0        0.000194
170   Computer Science   1976    1.0        0.000194
171   Computer Science   1977    1.0        0.000194
172   Computer Science   1978    3.0        0.000581
173   Computer Science   1979    3.0        0.000581
174   Computer Science   1980    5.0        0.000968
175   Computer Science   1981    6.0        0.001162
176   Computer Science   1982    2.0        0.000387
177   Computer Science   1983   12.0        0.002324
178   Computer Science   1984    2.0        0.000387
179   Computer Science   1985   11.0        0.002131
180   Computer Science   1986   24.0        0.004648
181   Computer Science   1987   20.0        0.003874
182   Computer Science   1988   51.0        0.009878
183   Computer Science   1989   29.0        0.005617
184   Computer Science   1990   39.0        0.007554
185   Computer Science   1991   62.0        0.012009
186   Computer Science   1992   71.0        0.013752
```

```
187  Computer Science  1993   98.0       0.018981
188  Computer Science  1994  114.0       0.022080
189  Computer Science  1995  374.0       0.072439
190  Computer Science  1996  349.0       0.067596
191  Computer Science  1997  401.0       0.077668
192  Computer Science  1998  420.0       0.081348
193  Computer Science  1999  389.0       0.075344
194  Computer Science  2000  451.0       0.087352
195  Computer Science  2001  366.0       0.070889
196  Computer Science  2002  882.0       0.170831
197  Computer Science  2003  957.0       0.185357
2015
Computer Science
                domain  Year  Count  domain_Score
165  Computer Science  1963    1.0       0.000194
166  Computer Science  1972    5.0       0.000968
167  Computer Science  1973    1.0       0.000194
168  Computer Science  1974   12.0       0.002324
169  Computer Science  1975    1.0       0.000194
170  Computer Science  1976    1.0       0.000194
171  Computer Science  1977    1.0       0.000194
172  Computer Science  1978    3.0       0.000581
173  Computer Science  1979    3.0       0.000581
174  Computer Science  1980    5.0       0.000968
175  Computer Science  1981    6.0       0.001162
176  Computer Science  1982    2.0       0.000387
177  Computer Science  1983   12.0       0.002324
178  Computer Science  1984    2.0       0.000387
179  Computer Science  1985   11.0       0.002131
180  Computer Science  1986   24.0       0.004648
181  Computer Science  1987   20.0       0.003874
182  Computer Science  1988   51.0       0.009878
183  Computer Science  1989   29.0       0.005617
184  Computer Science  1990   39.0       0.007554
185  Computer Science  1991   62.0       0.012009
186  Computer Science  1992   71.0       0.013752
187  Computer Science  1993   98.0       0.018981
188  Computer Science  1994  114.0       0.022080
189  Computer Science  1995  374.0       0.072439
190  Computer Science  1996  349.0       0.067596
191  Computer Science  1997  401.0       0.077668
192  Computer Science  1998  420.0       0.081348
193  Computer Science  1999  389.0       0.075344
194  Computer Science  2000  451.0       0.087352
195  Computer Science  2001  366.0       0.070889
196  Computer Science  2002  882.0       0.170831
197  Computer Science  2003  957.0       0.185357
2016
```

Computer Science

```
                domain  Year   Count  domain_Score
165  Computer Science  1963    1.0       0.000194
166  Computer Science  1972    5.0       0.000968
167  Computer Science  1973    1.0       0.000194
168  Computer Science  1974   12.0       0.002324
169  Computer Science  1975    1.0       0.000194
170  Computer Science  1976    1.0       0.000194
171  Computer Science  1977    1.0       0.000194
172  Computer Science  1978    3.0       0.000581
173  Computer Science  1979    3.0       0.000581
174  Computer Science  1980    5.0       0.000968
175  Computer Science  1981    6.0       0.001162
176  Computer Science  1982    2.0       0.000387
177  Computer Science  1983   12.0       0.002324
178  Computer Science  1984    2.0       0.000387
179  Computer Science  1985   11.0       0.002131
180  Computer Science  1986   24.0       0.004648
181  Computer Science  1987   20.0       0.003874
182  Computer Science  1988   51.0       0.009878
183  Computer Science  1989   29.0       0.005617
184  Computer Science  1990   39.0       0.007554
185  Computer Science  1991   62.0       0.012009
186  Computer Science  1992   71.0       0.013752
187  Computer Science  1993   98.0       0.018981
188  Computer Science  1994  114.0       0.022080
189  Computer Science  1995  374.0       0.072439
190  Computer Science  1996  349.0       0.067596
191  Computer Science  1997  401.0       0.077668
192  Computer Science  1998  420.0       0.081348
193  Computer Science  1999  389.0       0.075344
194  Computer Science  2000  451.0       0.087352
195  Computer Science  2001  366.0       0.070889
196  Computer Science  2002  882.0       0.170831
197  Computer Science  2003  957.0       0.185357
2016
Computer Science
                domain  Year   Count  domain_Score
165  Computer Science  1963    1.0       0.000194
166  Computer Science  1972    5.0       0.000968
167  Computer Science  1973    1.0       0.000194
168  Computer Science  1974   12.0       0.002324
169  Computer Science  1975    1.0       0.000194
170  Computer Science  1976    1.0       0.000194
171  Computer Science  1977    1.0       0.000194
172  Computer Science  1978    3.0       0.000581
173  Computer Science  1979    3.0       0.000581
174  Computer Science  1980    5.0       0.000968
```

```
175  Computer Science  1981    6.0      0.001162
176  Computer Science  1982    2.0      0.000387
177  Computer Science  1983   12.0      0.002324
178  Computer Science  1984    2.0      0.000387
179  Computer Science  1985   11.0      0.002131
180  Computer Science  1986   24.0      0.004648
181  Computer Science  1987   20.0      0.003874
182  Computer Science  1988   51.0      0.009878
183  Computer Science  1989   29.0      0.005617
184  Computer Science  1990   39.0      0.007554
185  Computer Science  1991   62.0      0.012009
186  Computer Science  1992   71.0      0.013752
187  Computer Science  1993   98.0      0.018981
188  Computer Science  1994  114.0      0.022080
189  Computer Science  1995  374.0      0.072439
190  Computer Science  1996  349.0      0.067596
191  Computer Science  1997  401.0      0.077668
192  Computer Science  1998  420.0      0.081348
193  Computer Science  1999  389.0      0.075344
194  Computer Science  2000  451.0      0.087352
195  Computer Science  2001  366.0      0.070889
196  Computer Science  2002  882.0      0.170831
197  Computer Science  2003  957.0      0.185357
2016
Computer Science
                domain  Year  Count  domain_Score
165  Computer Science  1963    1.0      0.000194
166  Computer Science  1972    5.0      0.000968
167  Computer Science  1973    1.0      0.000194
168  Computer Science  1974   12.0      0.002324
169  Computer Science  1975    1.0      0.000194
170  Computer Science  1976    1.0      0.000194
171  Computer Science  1977    1.0      0.000194
172  Computer Science  1978    3.0      0.000581
173  Computer Science  1979    3.0      0.000581
174  Computer Science  1980    5.0      0.000968
175  Computer Science  1981    6.0      0.001162
176  Computer Science  1982    2.0      0.000387
177  Computer Science  1983   12.0      0.002324
178  Computer Science  1984    2.0      0.000387
179  Computer Science  1985   11.0      0.002131
180  Computer Science  1986   24.0      0.004648
181  Computer Science  1987   20.0      0.003874
182  Computer Science  1988   51.0      0.009878
183  Computer Science  1989   29.0      0.005617
184  Computer Science  1990   39.0      0.007554
185  Computer Science  1991   62.0      0.012009
186  Computer Science  1992   71.0      0.013752
```

```
187  Computer Science  1993   98.0       0.018981
188  Computer Science  1994  114.0       0.022080
189  Computer Science  1995  374.0       0.072439
190  Computer Science  1996  349.0       0.067596
191  Computer Science  1997  401.0       0.077668
192  Computer Science  1998  420.0       0.081348
193  Computer Science  1999  389.0       0.075344
194  Computer Science  2000  451.0       0.087352
195  Computer Science  2001  366.0       0.070889
196  Computer Science  2002  882.0       0.170831
197  Computer Science  2003  957.0       0.185357
2016
Computer Science
       domain   Year  Count  domain_Score
594  Software   1963    0.0      0.000000
595  Software   1972    0.0      0.000000
596  Software   1973    0.0      0.000000
597  Software   1974    5.0      0.004955
598  Software   1975    1.0      0.000991
599  Software   1976    0.0      0.000000
600  Software   1977    1.0      0.000991
601  Software   1978    4.0      0.003964
602  Software   1979    2.0      0.001982
603  Software   1980    1.0      0.000991
604  Software   1981    1.0      0.000991
605  Software   1982    1.0      0.000991
606  Software   1983    0.0      0.000000
607  Software   1984    4.0      0.003964
608  Software   1985    1.0      0.000991
609  Software   1986    0.0      0.000000
610  Software   1987    4.0      0.003964
611  Software   1988    4.0      0.003964
612  Software   1989    3.0      0.002973
613  Software   1990   10.0      0.009911
614  Software   1991   14.0      0.013875
615  Software   1992    8.0      0.007929
616  Software   1993   18.0      0.017839
617  Software   1994   23.0      0.022795
618  Software   1995   50.0      0.049554
619  Software   1996   74.0      0.073340
620  Software   1997   69.0      0.068385
621  Software   1998   90.0      0.089197
622  Software   1999   94.0      0.093162
623  Software   2000   81.0      0.080278
624  Software   2001   76.0      0.075322
625  Software   2002  179.0      0.177403
626  Software   2003  191.0      0.189296
2016
```

Software

	domain	Year	Count	domain_Score
165	Computer Science	1963	1.0	0.000194
166	Computer Science	1972	5.0	0.000968
167	Computer Science	1973	1.0	0.000194
168	Computer Science	1974	12.0	0.002324
169	Computer Science	1975	1.0	0.000194
170	Computer Science	1976	1.0	0.000194
171	Computer Science	1977	1.0	0.000194
172	Computer Science	1978	3.0	0.000581
173	Computer Science	1979	3.0	0.000581
174	Computer Science	1980	5.0	0.000968
175	Computer Science	1981	6.0	0.001162
176	Computer Science	1982	2.0	0.000387
177	Computer Science	1983	12.0	0.002324
178	Computer Science	1984	2.0	0.000387
179	Computer Science	1985	11.0	0.002131
180	Computer Science	1986	24.0	0.004648
181	Computer Science	1987	20.0	0.003874
182	Computer Science	1988	51.0	0.009878
183	Computer Science	1989	29.0	0.005617
184	Computer Science	1990	39.0	0.007554
185	Computer Science	1991	62.0	0.012009
186	Computer Science	1992	71.0	0.013752
187	Computer Science	1993	98.0	0.018981
188	Computer Science	1994	114.0	0.022080
189	Computer Science	1995	374.0	0.072439
190	Computer Science	1996	349.0	0.067596
191	Computer Science	1997	401.0	0.077668
192	Computer Science	1998	420.0	0.081348
193	Computer Science	1999	389.0	0.075344
194	Computer Science	2000	451.0	0.087352
195	Computer Science	2001	366.0	0.070889
196	Computer Science	2002	882.0	0.170831
197	Computer Science	2003	957.0	0.185357

2016
Computer Science

	domain	Year	Count	domain_Score
165	Computer Science	1963	1.0	0.000194
166	Computer Science	1972	5.0	0.000968
167	Computer Science	1973	1.0	0.000194
168	Computer Science	1974	12.0	0.002324
169	Computer Science	1975	1.0	0.000194
170	Computer Science	1976	1.0	0.000194
171	Computer Science	1977	1.0	0.000194
172	Computer Science	1978	3.0	0.000581
173	Computer Science	1979	3.0	0.000581
174	Computer Science	1980	5.0	0.000968

```
175   Computer Science  1981    6.0       0.001162
176   Computer Science  1982    2.0       0.000387
177   Computer Science  1983   12.0       0.002324
178   Computer Science  1984    2.0       0.000387
179   Computer Science  1985   11.0       0.002131
180   Computer Science  1986   24.0       0.004648
181   Computer Science  1987   20.0       0.003874
182   Computer Science  1988   51.0       0.009878
183   Computer Science  1989   29.0       0.005617
184   Computer Science  1990   39.0       0.007554
185   Computer Science  1991   62.0       0.012009
186   Computer Science  1992   71.0       0.013752
187   Computer Science  1993   98.0       0.018981
188   Computer Science  1994  114.0       0.022080
189   Computer Science  1995  374.0       0.072439
190   Computer Science  1996  349.0       0.067596
191   Computer Science  1997  401.0       0.077668
192   Computer Science  1998  420.0       0.081348
193   Computer Science  1999  389.0       0.075344
194   Computer Science  2000  451.0       0.087352
195   Computer Science  2001  366.0       0.070889
196   Computer Science  2002  882.0       0.170831
197   Computer Science  2003  957.0       0.185357
2016
Computer Science
              domain  Year  Count  domain_Score
165   Computer Science  1963    1.0       0.000194
166   Computer Science  1972    5.0       0.000968
167   Computer Science  1973    1.0       0.000194
168   Computer Science  1974   12.0       0.002324
169   Computer Science  1975    1.0       0.000194
170   Computer Science  1976    1.0       0.000194
171   Computer Science  1977    1.0       0.000194
172   Computer Science  1978    3.0       0.000581
173   Computer Science  1979    3.0       0.000581
174   Computer Science  1980    5.0       0.000968
175   Computer Science  1981    6.0       0.001162
176   Computer Science  1982    2.0       0.000387
177   Computer Science  1983   12.0       0.002324
178   Computer Science  1984    2.0       0.000387
179   Computer Science  1985   11.0       0.002131
180   Computer Science  1986   24.0       0.004648
181   Computer Science  1987   20.0       0.003874
182   Computer Science  1988   51.0       0.009878
183   Computer Science  1989   29.0       0.005617
184   Computer Science  1990   39.0       0.007554
185   Computer Science  1991   62.0       0.012009
186   Computer Science  1992   71.0       0.013752
```

```
187  Computer Science  1993   98.0       0.018981
188  Computer Science  1994  114.0       0.022080
189  Computer Science  1995  374.0       0.072439
190  Computer Science  1996  349.0       0.067596
191  Computer Science  1997  401.0       0.077668
192  Computer Science  1998  420.0       0.081348
193  Computer Science  1999  389.0       0.075344
194  Computer Science  2000  451.0       0.087352
195  Computer Science  2001  366.0       0.070889
196  Computer Science  2002  882.0       0.170831
197  Computer Science  2003  957.0       0.185357
2016
Computer Science
                 domain  Year  Count  domain_Score
66  Biomedical Research  1963    0.0      0.000000
67  Biomedical Research  1972    0.0      0.000000
68  Biomedical Research  1973    0.0      0.000000
69  Biomedical Research  1974    2.0      0.009050
70  Biomedical Research  1975    0.0      0.000000
71  Biomedical Research  1976    0.0      0.000000
72  Biomedical Research  1977    1.0      0.004525
73  Biomedical Research  1978    2.0      0.009050
74  Biomedical Research  1979    5.0      0.022624
75  Biomedical Research  1980    1.0      0.004525
76  Biomedical Research  1981    3.0      0.013575
77  Biomedical Research  1982    1.0      0.004525
78  Biomedical Research  1983    1.0      0.004525
79  Biomedical Research  1984    7.0      0.031674
80  Biomedical Research  1985    1.0      0.004525
81  Biomedical Research  1986    4.0      0.018100
82  Biomedical Research  1987    2.0      0.009050
83  Biomedical Research  1988    3.0      0.013575
84  Biomedical Research  1989    2.0      0.009050
85  Biomedical Research  1990    5.0      0.022624
86  Biomedical Research  1991    4.0      0.018100
87  Biomedical Research  1992    4.0      0.018100
88  Biomedical Research  1993    8.0      0.036199
89  Biomedical Research  1994    3.0      0.013575
90  Biomedical Research  1995   12.0      0.054299
91  Biomedical Research  1996   12.0      0.054299
92  Biomedical Research  1997   11.0      0.049774
93  Biomedical Research  1998   11.0      0.049774
94  Biomedical Research  1999    9.0      0.040724
95  Biomedical Research  2000   14.0      0.063348
96  Biomedical Research  2001   15.0      0.067873
97  Biomedical Research  2002   48.0      0.217195
98  Biomedical Research  2003   30.0      0.135747
2016
```

Biomedical Research

	domain	Year	Count	domain_Score
165	Computer Science	1963	1.0	0.000194
166	Computer Science	1972	5.0	0.000968
167	Computer Science	1973	1.0	0.000194
168	Computer Science	1974	12.0	0.002324
169	Computer Science	1975	1.0	0.000194
170	Computer Science	1976	1.0	0.000194
171	Computer Science	1977	1.0	0.000194
172	Computer Science	1978	3.0	0.000581
173	Computer Science	1979	3.0	0.000581
174	Computer Science	1980	5.0	0.000968
175	Computer Science	1981	6.0	0.001162
176	Computer Science	1982	2.0	0.000387
177	Computer Science	1983	12.0	0.002324
178	Computer Science	1984	2.0	0.000387
179	Computer Science	1985	11.0	0.002131
180	Computer Science	1986	24.0	0.004648
181	Computer Science	1987	20.0	0.003874
182	Computer Science	1988	51.0	0.009878
183	Computer Science	1989	29.0	0.005617
184	Computer Science	1990	39.0	0.007554
185	Computer Science	1991	62.0	0.012009
186	Computer Science	1992	71.0	0.013752
187	Computer Science	1993	98.0	0.018981
188	Computer Science	1994	114.0	0.022080
189	Computer Science	1995	374.0	0.072439
190	Computer Science	1996	349.0	0.067596
191	Computer Science	1997	401.0	0.077668
192	Computer Science	1998	420.0	0.081348
193	Computer Science	1999	389.0	0.075344
194	Computer Science	2000	451.0	0.087352
195	Computer Science	2001	366.0	0.070889
196	Computer Science	2002	882.0	0.170831
197	Computer Science	2003	957.0	0.185357

2016
Computer Science

	domain	Year	Count	domain_Score
165	Computer Science	1963	1.0	0.000194
166	Computer Science	1972	5.0	0.000968
167	Computer Science	1973	1.0	0.000194
168	Computer Science	1974	12.0	0.002324
169	Computer Science	1975	1.0	0.000194
170	Computer Science	1976	1.0	0.000194
171	Computer Science	1977	1.0	0.000194
172	Computer Science	1978	3.0	0.000581
173	Computer Science	1979	3.0	0.000581
174	Computer Science	1980	5.0	0.000968

```
175  Computer Science  1981    6.0      0.001162
176  Computer Science  1982    2.0      0.000387
177  Computer Science  1983   12.0      0.002324
178  Computer Science  1984    2.0      0.000387
179  Computer Science  1985   11.0      0.002131
180  Computer Science  1986   24.0      0.004648
181  Computer Science  1987   20.0      0.003874
182  Computer Science  1988   51.0      0.009878
183  Computer Science  1989   29.0      0.005617
184  Computer Science  1990   39.0      0.007554
185  Computer Science  1991   62.0      0.012009
186  Computer Science  1992   71.0      0.013752
187  Computer Science  1993   98.0      0.018981
188  Computer Science  1994  114.0      0.022080
189  Computer Science  1995  374.0      0.072439
190  Computer Science  1996  349.0      0.067596
191  Computer Science  1997  401.0      0.077668
192  Computer Science  1998  420.0      0.081348
193  Computer Science  1999  389.0      0.075344
194  Computer Science  2000  451.0      0.087352
195  Computer Science  2001  366.0      0.070889
196  Computer Science  2002  882.0      0.170831
197  Computer Science  2003  957.0      0.185357
2017
Computer Science
       domain  Year  Count  domain_Score
99    Business  1963    0.0      0.000000
100   Business  1972    0.0      0.000000
101   Business  1973    0.0      0.000000
102   Business  1974    1.0      0.000953
103   Business  1975    1.0      0.000953
104   Business  1976    0.0      0.000000
105   Business  1977    0.0      0.000000
106   Business  1978    1.0      0.000953
107   Business  1979    2.0      0.001907
108   Business  1980    2.0      0.001907
109   Business  1981    2.0      0.001907
110   Business  1982    1.0      0.000953
111   Business  1983    1.0      0.000953
112   Business  1984    1.0      0.000953
113   Business  1985    1.0      0.000953
114   Business  1986    1.0      0.000953
115   Business  1987    1.0      0.000953
116   Business  1988    2.0      0.001907
117   Business  1989    4.0      0.003813
118   Business  1990    4.0      0.003813
119   Business  1991   12.0      0.011439
120   Business  1992    8.0      0.007626
```

```
121  Business  1993   14.0        0.013346
122  Business  1994    8.0        0.007626
123  Business  1995   28.0        0.026692
124  Business  1996   51.0        0.048618
125  Business  1997   42.0        0.040038
126  Business  1998  108.0        0.102955
127  Business  1999   88.0        0.083889
128  Business  2000  105.0        0.100095
129  Business  2001   98.0        0.093422
130  Business  2002  202.0        0.192564
131  Business  2003  260.0        0.247855
2017
Business
        domain  Year  Count  domain_Score
561  Political  1963    0.0      0.000000
562  Political  1972    0.0      0.000000
563  Political  1973    0.0      0.000000
564  Political  1974    0.0      0.000000
565  Political  1975    0.0      0.000000
566  Political  1976    0.0      0.000000
567  Political  1977    0.0      0.000000
568  Political  1978    0.0      0.000000
569  Political  1979    0.0      0.000000
570  Political  1980    0.0      0.000000
571  Political  1981    1.0      0.003774
572  Political  1982    0.0      0.000000
573  Political  1983    0.0      0.000000
574  Political  1984    1.0      0.003774
575  Political  1985    0.0      0.000000
576  Political  1986    0.0      0.000000
577  Political  1987    1.0      0.003774
578  Political  1988    1.0      0.003774
579  Political  1989    0.0      0.000000
580  Political  1990    5.0      0.018868
581  Political  1991    2.0      0.007547
582  Political  1992    3.0      0.011321
583  Political  1993    1.0      0.003774
584  Political  1994    5.0      0.018868
585  Political  1995   11.0      0.041509
586  Political  1996   20.0      0.075472
587  Political  1997   11.0      0.041509
588  Political  1998   18.0      0.067925
589  Political  1999   26.0      0.098113
590  Political  2000   28.0      0.105660
591  Political  2001   21.0      0.079245
592  Political  2002   49.0      0.184906
593  Political  2003   61.0      0.230189
2017
```

Political

	domain	Year	Count	domain_Score
165	Computer Science	1963	1.0	0.000194
166	Computer Science	1972	5.0	0.000968
167	Computer Science	1973	1.0	0.000194
168	Computer Science	1974	12.0	0.002324
169	Computer Science	1975	1.0	0.000194
170	Computer Science	1976	1.0	0.000194
171	Computer Science	1977	1.0	0.000194
172	Computer Science	1978	3.0	0.000581
173	Computer Science	1979	3.0	0.000581
174	Computer Science	1980	5.0	0.000968
175	Computer Science	1981	6.0	0.001162
176	Computer Science	1982	2.0	0.000387
177	Computer Science	1983	12.0	0.002324
178	Computer Science	1984	2.0	0.000387
179	Computer Science	1985	11.0	0.002131
180	Computer Science	1986	24.0	0.004648
181	Computer Science	1987	20.0	0.003874
182	Computer Science	1988	51.0	0.009878
183	Computer Science	1989	29.0	0.005617
184	Computer Science	1990	39.0	0.007554
185	Computer Science	1991	62.0	0.012009
186	Computer Science	1992	71.0	0.013752
187	Computer Science	1993	98.0	0.018981
188	Computer Science	1994	114.0	0.022080
189	Computer Science	1995	374.0	0.072439
190	Computer Science	1996	349.0	0.067596
191	Computer Science	1997	401.0	0.077668
192	Computer Science	1998	420.0	0.081348
193	Computer Science	1999	389.0	0.075344
194	Computer Science	2000	451.0	0.087352
195	Computer Science	2001	366.0	0.070889
196	Computer Science	2002	882.0	0.170831
197	Computer Science	2003	957.0	0.185357

2017
Computer Science

	domain	Year	Count	domain_Score
165	Computer Science	1963	1.0	0.000194
166	Computer Science	1972	5.0	0.000968
167	Computer Science	1973	1.0	0.000194
168	Computer Science	1974	12.0	0.002324
169	Computer Science	1975	1.0	0.000194
170	Computer Science	1976	1.0	0.000194
171	Computer Science	1977	1.0	0.000194
172	Computer Science	1978	3.0	0.000581
173	Computer Science	1979	3.0	0.000581
174	Computer Science	1980	5.0	0.000968

```
175   Computer Science   1981     6.0         0.001162
176   Computer Science   1982     2.0         0.000387
177   Computer Science   1983    12.0         0.002324
178   Computer Science   1984     2.0         0.000387
179   Computer Science   1985    11.0         0.002131
180   Computer Science   1986    24.0         0.004648
181   Computer Science   1987    20.0         0.003874
182   Computer Science   1988    51.0         0.009878
183   Computer Science   1989    29.0         0.005617
184   Computer Science   1990    39.0         0.007554
185   Computer Science   1991    62.0         0.012009
186   Computer Science   1992    71.0         0.013752
187   Computer Science   1993    98.0         0.018981
188   Computer Science   1994   114.0         0.022080
189   Computer Science   1995   374.0         0.072439
190   Computer Science   1996   349.0         0.067596
191   Computer Science   1997   401.0         0.077668
192   Computer Science   1998   420.0         0.081348
193   Computer Science   1999   389.0         0.075344
194   Computer Science   2000   451.0         0.087352
195   Computer Science   2001   366.0         0.070889
196   Computer Science   2002   882.0         0.170831
197   Computer Science   2003   957.0         0.185357
2017
Computer Science
                domain   Year   Count   domain_Score
165   Computer Science   1963     1.0         0.000194
166   Computer Science   1972     5.0         0.000968
167   Computer Science   1973     1.0         0.000194
168   Computer Science   1974    12.0         0.002324
169   Computer Science   1975     1.0         0.000194
170   Computer Science   1976     1.0         0.000194
171   Computer Science   1977     1.0         0.000194
172   Computer Science   1978     3.0         0.000581
173   Computer Science   1979     3.0         0.000581
174   Computer Science   1980     5.0         0.000968
175   Computer Science   1981     6.0         0.001162
176   Computer Science   1982     2.0         0.000387
177   Computer Science   1983    12.0         0.002324
178   Computer Science   1984     2.0         0.000387
179   Computer Science   1985    11.0         0.002131
180   Computer Science   1986    24.0         0.004648
181   Computer Science   1987    20.0         0.003874
182   Computer Science   1988    51.0         0.009878
183   Computer Science   1989    29.0         0.005617
184   Computer Science   1990    39.0         0.007554
185   Computer Science   1991    62.0         0.012009
186   Computer Science   1992    71.0         0.013752
```

```
187  Computer Science  1993   98.0      0.018981
188  Computer Science  1994  114.0      0.022080
189  Computer Science  1995  374.0      0.072439
190  Computer Science  1996  349.0      0.067596
191  Computer Science  1997  401.0      0.077668
192  Computer Science  1998  420.0      0.081348
193  Computer Science  1999  389.0      0.075344
194  Computer Science  2000  451.0      0.087352
195  Computer Science  2001  366.0      0.070889
196  Computer Science  2002  882.0      0.170831
197  Computer Science  2003  957.0      0.185357
2017
Computer Science
                domain  Year  Count  domain_Score
165  Computer Science  1963    1.0      0.000194
166  Computer Science  1972    5.0      0.000968
167  Computer Science  1973    1.0      0.000194
168  Computer Science  1974   12.0      0.002324
169  Computer Science  1975    1.0      0.000194
170  Computer Science  1976    1.0      0.000194
171  Computer Science  1977    1.0      0.000194
172  Computer Science  1978    3.0      0.000581
173  Computer Science  1979    3.0      0.000581
174  Computer Science  1980    5.0      0.000968
175  Computer Science  1981    6.0      0.001162
176  Computer Science  1982    2.0      0.000387
177  Computer Science  1983   12.0      0.002324
178  Computer Science  1984    2.0      0.000387
179  Computer Science  1985   11.0      0.002131
180  Computer Science  1986   24.0      0.004648
181  Computer Science  1987   20.0      0.003874
182  Computer Science  1988   51.0      0.009878
183  Computer Science  1989   29.0      0.005617
184  Computer Science  1990   39.0      0.007554
185  Computer Science  1991   62.0      0.012009
186  Computer Science  1992   71.0      0.013752
187  Computer Science  1993   98.0      0.018981
188  Computer Science  1994  114.0      0.022080
189  Computer Science  1995  374.0      0.072439
190  Computer Science  1996  349.0      0.067596
191  Computer Science  1997  401.0      0.077668
192  Computer Science  1998  420.0      0.081348
193  Computer Science  1999  389.0      0.075344
194  Computer Science  2000  451.0      0.087352
195  Computer Science  2001  366.0      0.070889
196  Computer Science  2002  882.0      0.170831
197  Computer Science  2003  957.0      0.185357
2017
```

Computer Science

```
              domain  Year  Count  domain_Score
165  Computer Science  1963    1.0      0.000194
166  Computer Science  1972    5.0      0.000968
167  Computer Science  1973    1.0      0.000194
168  Computer Science  1974   12.0      0.002324
169  Computer Science  1975    1.0      0.000194
170  Computer Science  1976    1.0      0.000194
171  Computer Science  1977    1.0      0.000194
172  Computer Science  1978    3.0      0.000581
173  Computer Science  1979    3.0      0.000581
174  Computer Science  1980    5.0      0.000968
175  Computer Science  1981    6.0      0.001162
176  Computer Science  1982    2.0      0.000387
177  Computer Science  1983   12.0      0.002324
178  Computer Science  1984    2.0      0.000387
179  Computer Science  1985   11.0      0.002131
180  Computer Science  1986   24.0      0.004648
181  Computer Science  1987   20.0      0.003874
182  Computer Science  1988   51.0      0.009878
183  Computer Science  1989   29.0      0.005617
184  Computer Science  1990   39.0      0.007554
185  Computer Science  1991   62.0      0.012009
186  Computer Science  1992   71.0      0.013752
187  Computer Science  1993   98.0      0.018981
188  Computer Science  1994  114.0      0.022080
189  Computer Science  1995  374.0      0.072439
190  Computer Science  1996  349.0      0.067596
191  Computer Science  1997  401.0      0.077668
192  Computer Science  1998  420.0      0.081348
193  Computer Science  1999  389.0      0.075344
194  Computer Science  2000  451.0      0.087352
195  Computer Science  2001  366.0      0.070889
196  Computer Science  2002  882.0      0.170831
197  Computer Science  2003  957.0      0.185357
2017
Computer Science
              domain  Year  Count  domain_Score
165  Computer Science  1963    1.0      0.000194
166  Computer Science  1972    5.0      0.000968
167  Computer Science  1973    1.0      0.000194
168  Computer Science  1974   12.0      0.002324
169  Computer Science  1975    1.0      0.000194
170  Computer Science  1976    1.0      0.000194
171  Computer Science  1977    1.0      0.000194
172  Computer Science  1978    3.0      0.000581
173  Computer Science  1979    3.0      0.000581
174  Computer Science  1980    5.0      0.000968
```

```
175   Computer Science  1981    6.0      0.001162
176   Computer Science  1982    2.0      0.000387
177   Computer Science  1983   12.0      0.002324
178   Computer Science  1984    2.0      0.000387
179   Computer Science  1985   11.0      0.002131
180   Computer Science  1986   24.0      0.004648
181   Computer Science  1987   20.0      0.003874
182   Computer Science  1988   51.0      0.009878
183   Computer Science  1989   29.0      0.005617
184   Computer Science  1990   39.0      0.007554
185   Computer Science  1991   62.0      0.012009
186   Computer Science  1992   71.0      0.013752
187   Computer Science  1993   98.0      0.018981
188   Computer Science  1994  114.0      0.022080
189   Computer Science  1995  374.0      0.072439
190   Computer Science  1996  349.0      0.067596
191   Computer Science  1997  401.0      0.077668
192   Computer Science  1998  420.0      0.081348
193   Computer Science  1999  389.0      0.075344
194   Computer Science  2000  451.0      0.087352
195   Computer Science  2001  366.0      0.070889
196   Computer Science  2002  882.0      0.170831
197   Computer Science  2003  957.0      0.185357
2017
Computer Science
     domain  Year  Count  domain_Score
297  Health  1963   0.0      0.000000
298  Health  1972   0.0      0.000000
299  Health  1973   0.0      0.000000
300  Health  1974   0.0      0.000000
301  Health  1975   0.0      0.000000
302  Health  1976   1.0      0.006211
303  Health  1977   0.0      0.000000
304  Health  1978   0.0      0.000000
305  Health  1979   0.0      0.000000
306  Health  1980   0.0      0.000000
307  Health  1981   0.0      0.000000
308  Health  1982   0.0      0.000000
309  Health  1983   1.0      0.006211
310  Health  1984   2.0      0.012422
311  Health  1985   0.0      0.000000
312  Health  1986   1.0      0.006211
313  Health  1987   1.0      0.006211
314  Health  1988   0.0      0.000000
315  Health  1989   0.0      0.000000
316  Health  1990   0.0      0.000000
317  Health  1991   0.0      0.000000
318  Health  1992   0.0      0.000000
```

```
319   Health  1993     0.0       0.000000
320   Health  1994     2.0       0.012422
321   Health  1995    10.0       0.062112
322   Health  1996     7.0       0.043478
323   Health  1997     9.0       0.055901
324   Health  1998    12.0       0.074534
325   Health  1999     9.0       0.055901
326   Health  2000    13.0       0.080745
327   Health  2001    16.0       0.099379
328   Health  2002    45.0       0.279503
329   Health  2003    32.0       0.198758
2017
Health
        domain  Year  Count  domain_Score
99    Business  1963    0.0       0.000000
100   Business  1972    0.0       0.000000
101   Business  1973    0.0       0.000000
102   Business  1974    1.0       0.000953
103   Business  1975    1.0       0.000953
104   Business  1976    0.0       0.000000
105   Business  1977    0.0       0.000000
106   Business  1978    1.0       0.000953
107   Business  1979    2.0       0.001907
108   Business  1980    2.0       0.001907
109   Business  1981    2.0       0.001907
110   Business  1982    1.0       0.000953
111   Business  1983    1.0       0.000953
112   Business  1984    1.0       0.000953
113   Business  1985    1.0       0.000953
114   Business  1986    1.0       0.000953
115   Business  1987    1.0       0.000953
116   Business  1988    2.0       0.001907
117   Business  1989    4.0       0.003813
118   Business  1990    4.0       0.003813
119   Business  1991   12.0       0.011439
120   Business  1992    8.0       0.007626
121   Business  1993   14.0       0.013346
122   Business  1994    8.0       0.007626
123   Business  1995   28.0       0.026692
124   Business  1996   51.0       0.048618
125   Business  1997   42.0       0.040038
126   Business  1998  108.0       0.102955
127   Business  1999   88.0       0.083889
128   Business  2000  105.0       0.100095
129   Business  2001   98.0       0.093422
130   Business  2002  202.0       0.192564
131   Business  2003  260.0       0.247855
2018
```

Business

	domain	Year	Count	domain_Score
165	Computer Science	1963	1.0	0.000194
166	Computer Science	1972	5.0	0.000968
167	Computer Science	1973	1.0	0.000194
168	Computer Science	1974	12.0	0.002324
169	Computer Science	1975	1.0	0.000194
170	Computer Science	1976	1.0	0.000194
171	Computer Science	1977	1.0	0.000194
172	Computer Science	1978	3.0	0.000581
173	Computer Science	1979	3.0	0.000581
174	Computer Science	1980	5.0	0.000968
175	Computer Science	1981	6.0	0.001162
176	Computer Science	1982	2.0	0.000387
177	Computer Science	1983	12.0	0.002324
178	Computer Science	1984	2.0	0.000387
179	Computer Science	1985	11.0	0.002131
180	Computer Science	1986	24.0	0.004648
181	Computer Science	1987	20.0	0.003874
182	Computer Science	1988	51.0	0.009878
183	Computer Science	1989	29.0	0.005617
184	Computer Science	1990	39.0	0.007554
185	Computer Science	1991	62.0	0.012009
186	Computer Science	1992	71.0	0.013752
187	Computer Science	1993	98.0	0.018981
188	Computer Science	1994	114.0	0.022080
189	Computer Science	1995	374.0	0.072439
190	Computer Science	1996	349.0	0.067596
191	Computer Science	1997	401.0	0.077668
192	Computer Science	1998	420.0	0.081348
193	Computer Science	1999	389.0	0.075344
194	Computer Science	2000	451.0	0.087352
195	Computer Science	2001	366.0	0.070889
196	Computer Science	2002	882.0	0.170831
197	Computer Science	2003	957.0	0.185357

2014
Computer Science

	domain	Year	Count	domain_Score
165	Computer Science	1963	1.0	0.000194
166	Computer Science	1972	5.0	0.000968
167	Computer Science	1973	1.0	0.000194
168	Computer Science	1974	12.0	0.002324
169	Computer Science	1975	1.0	0.000194
170	Computer Science	1976	1.0	0.000194
171	Computer Science	1977	1.0	0.000194
172	Computer Science	1978	3.0	0.000581
173	Computer Science	1979	3.0	0.000581
174	Computer Science	1980	5.0	0.000968

```
175  Computer Science  1981    6.0        0.001162
176  Computer Science  1982    2.0        0.000387
177  Computer Science  1983   12.0        0.002324
178  Computer Science  1984    2.0        0.000387
179  Computer Science  1985   11.0        0.002131
180  Computer Science  1986   24.0        0.004648
181  Computer Science  1987   20.0        0.003874
182  Computer Science  1988   51.0        0.009878
183  Computer Science  1989   29.0        0.005617
184  Computer Science  1990   39.0        0.007554
185  Computer Science  1991   62.0        0.012009
186  Computer Science  1992   71.0        0.013752
187  Computer Science  1993   98.0        0.018981
188  Computer Science  1994  114.0        0.022080
189  Computer Science  1995  374.0        0.072439
190  Computer Science  1996  349.0        0.067596
191  Computer Science  1997  401.0        0.077668
192  Computer Science  1998  420.0        0.081348
193  Computer Science  1999  389.0        0.075344
194  Computer Science  2000  451.0        0.087352
195  Computer Science  2001  366.0        0.070889
196  Computer Science  2002  882.0        0.170831
197  Computer Science  2003  957.0        0.185357
2015
Computer Science
        domain  Year  Count  domain_Score
594  Software  1963    0.0        0.000000
595  Software  1972    0.0        0.000000
596  Software  1973    0.0        0.000000
597  Software  1974    5.0        0.004955
598  Software  1975    1.0        0.000991
599  Software  1976    0.0        0.000000
600  Software  1977    1.0        0.000991
601  Software  1978    4.0        0.003964
602  Software  1979    2.0        0.001982
603  Software  1980    1.0        0.000991
604  Software  1981    1.0        0.000991
605  Software  1982    1.0        0.000991
606  Software  1983    0.0        0.000000
607  Software  1984    4.0        0.003964
608  Software  1985    1.0        0.000991
609  Software  1986    0.0        0.000000
610  Software  1987    4.0        0.003964
611  Software  1988    4.0        0.003964
612  Software  1989    3.0        0.002973
613  Software  1990   10.0        0.009911
614  Software  1991   14.0        0.013875
615  Software  1992    8.0        0.007929
```

```
616   Software   1993   18.0        0.017839
617   Software   1994   23.0        0.022795
618   Software   1995   50.0        0.049554
619   Software   1996   74.0        0.073340
620   Software   1997   69.0        0.068385
621   Software   1998   90.0        0.089197
622   Software   1999   94.0        0.093162
623   Software   2000   81.0        0.080278
624   Software   2001   76.0        0.075322
625   Software   2002   179.0       0.177403
626   Software   2003   191.0       0.189296
2015
Software
                 domain   Year   Count   domain_Score
165   Computer Science   1963    1.0        0.000194
166   Computer Science   1972    5.0        0.000968
167   Computer Science   1973    1.0        0.000194
168   Computer Science   1974   12.0        0.002324
169   Computer Science   1975    1.0        0.000194
170   Computer Science   1976    1.0        0.000194
171   Computer Science   1977    1.0        0.000194
172   Computer Science   1978    3.0        0.000581
173   Computer Science   1979    3.0        0.000581
174   Computer Science   1980    5.0        0.000968
175   Computer Science   1981    6.0        0.001162
176   Computer Science   1982    2.0        0.000387
177   Computer Science   1983   12.0        0.002324
178   Computer Science   1984    2.0        0.000387
179   Computer Science   1985   11.0        0.002131
180   Computer Science   1986   24.0        0.004648
181   Computer Science   1987   20.0        0.003874
182   Computer Science   1988   51.0        0.009878
183   Computer Science   1989   29.0        0.005617
184   Computer Science   1990   39.0        0.007554
185   Computer Science   1991   62.0        0.012009
186   Computer Science   1992   71.0        0.013752
187   Computer Science   1993   98.0        0.018981
188   Computer Science   1994  114.0        0.022080
189   Computer Science   1995  374.0        0.072439
190   Computer Science   1996  349.0        0.067596
191   Computer Science   1997  401.0        0.077668
192   Computer Science   1998  420.0        0.081348
193   Computer Science   1999  389.0        0.075344
194   Computer Science   2000  451.0        0.087352
195   Computer Science   2001  366.0        0.070889
196   Computer Science   2002  882.0        0.170831
197   Computer Science   2003  957.0        0.185357
2017
```

Computer Science

	domain	Year	Count	domain_Score
165	Computer Science	1963	1.0	0.000194
166	Computer Science	1972	5.0	0.000968
167	Computer Science	1973	1.0	0.000194
168	Computer Science	1974	12.0	0.002324
169	Computer Science	1975	1.0	0.000194
170	Computer Science	1976	1.0	0.000194
171	Computer Science	1977	1.0	0.000194
172	Computer Science	1978	3.0	0.000581
173	Computer Science	1979	3.0	0.000581
174	Computer Science	1980	5.0	0.000968
175	Computer Science	1981	6.0	0.001162
176	Computer Science	1982	2.0	0.000387
177	Computer Science	1983	12.0	0.002324
178	Computer Science	1984	2.0	0.000387
179	Computer Science	1985	11.0	0.002131
180	Computer Science	1986	24.0	0.004648
181	Computer Science	1987	20.0	0.003874
182	Computer Science	1988	51.0	0.009878
183	Computer Science	1989	29.0	0.005617
184	Computer Science	1990	39.0	0.007554
185	Computer Science	1991	62.0	0.012009
186	Computer Science	1992	71.0	0.013752
187	Computer Science	1993	98.0	0.018981
188	Computer Science	1994	114.0	0.022080
189	Computer Science	1995	374.0	0.072439
190	Computer Science	1996	349.0	0.067596
191	Computer Science	1997	401.0	0.077668
192	Computer Science	1998	420.0	0.081348
193	Computer Science	1999	389.0	0.075344
194	Computer Science	2000	451.0	0.087352
195	Computer Science	2001	366.0	0.070889
196	Computer Science	2002	882.0	0.170831
197	Computer Science	2003	957.0	0.185357

2017
Computer Science

	domain	Year	Count	domain_Score
165	Computer Science	1963	1.0	0.000194
166	Computer Science	1972	5.0	0.000968
167	Computer Science	1973	1.0	0.000194
168	Computer Science	1974	12.0	0.002324
169	Computer Science	1975	1.0	0.000194
170	Computer Science	1976	1.0	0.000194
171	Computer Science	1977	1.0	0.000194
172	Computer Science	1978	3.0	0.000581
173	Computer Science	1979	3.0	0.000581
174	Computer Science	1980	5.0	0.000968

```
175  Computer Science  1981    6.0        0.001162
176  Computer Science  1982    2.0        0.000387
177  Computer Science  1983   12.0        0.002324
178  Computer Science  1984    2.0        0.000387
179  Computer Science  1985   11.0        0.002131
180  Computer Science  1986   24.0        0.004648
181  Computer Science  1987   20.0        0.003874
182  Computer Science  1988   51.0        0.009878
183  Computer Science  1989   29.0        0.005617
184  Computer Science  1990   39.0        0.007554
185  Computer Science  1991   62.0        0.012009
186  Computer Science  1992   71.0        0.013752
187  Computer Science  1993   98.0        0.018981
188  Computer Science  1994  114.0        0.022080
189  Computer Science  1995  374.0        0.072439
190  Computer Science  1996  349.0        0.067596
191  Computer Science  1997  401.0        0.077668
192  Computer Science  1998  420.0        0.081348
193  Computer Science  1999  389.0        0.075344
194  Computer Science  2000  451.0        0.087352
195  Computer Science  2001  366.0        0.070889
196  Computer Science  2002  882.0        0.170831
197  Computer Science  2003  957.0        0.185357
2017
Computer Science
       domain  Year  Count  domain_Score
594  Software  1963    0.0      0.000000
595  Software  1972    0.0      0.000000
596  Software  1973    0.0      0.000000
597  Software  1974    5.0      0.004955
598  Software  1975    1.0      0.000991
599  Software  1976    0.0      0.000000
600  Software  1977    1.0      0.000991
601  Software  1978    4.0      0.003964
602  Software  1979    2.0      0.001982
603  Software  1980    1.0      0.000991
604  Software  1981    1.0      0.000991
605  Software  1982    1.0      0.000991
606  Software  1983    0.0      0.000000
607  Software  1984    4.0      0.003964
608  Software  1985    1.0      0.000991
609  Software  1986    0.0      0.000000
610  Software  1987    4.0      0.003964
611  Software  1988    4.0      0.003964
612  Software  1989    3.0      0.002973
613  Software  1990   10.0      0.009911
614  Software  1991   14.0      0.013875
615  Software  1992    8.0      0.007929
```

```
616   Software   1993   18.0        0.017839
617   Software   1994   23.0        0.022795
618   Software   1995   50.0        0.049554
619   Software   1996   74.0        0.073340
620   Software   1997   69.0        0.068385
621   Software   1998   90.0        0.089197
622   Software   1999   94.0        0.093162
623   Software   2000   81.0        0.080278
624   Software   2001   76.0        0.075322
625   Software   2002  179.0        0.177403
626   Software   2003  191.0        0.189296
2014
Software
               domain  Year  Count  domain_Score
165  Computer Science  1963    1.0      0.000194
166  Computer Science  1972    5.0      0.000968
167  Computer Science  1973    1.0      0.000194
168  Computer Science  1974   12.0      0.002324
169  Computer Science  1975    1.0      0.000194
170  Computer Science  1976    1.0      0.000194
171  Computer Science  1977    1.0      0.000194
172  Computer Science  1978    3.0      0.000581
173  Computer Science  1979    3.0      0.000581
174  Computer Science  1980    5.0      0.000968
175  Computer Science  1981    6.0      0.001162
176  Computer Science  1982    2.0      0.000387
177  Computer Science  1983   12.0      0.002324
178  Computer Science  1984    2.0      0.000387
179  Computer Science  1985   11.0      0.002131
180  Computer Science  1986   24.0      0.004648
181  Computer Science  1987   20.0      0.003874
182  Computer Science  1988   51.0      0.009878
183  Computer Science  1989   29.0      0.005617
184  Computer Science  1990   39.0      0.007554
185  Computer Science  1991   62.0      0.012009
186  Computer Science  1992   71.0      0.013752
187  Computer Science  1993   98.0      0.018981
188  Computer Science  1994  114.0      0.022080
189  Computer Science  1995  374.0      0.072439
190  Computer Science  1996  349.0      0.067596
191  Computer Science  1997  401.0      0.077668
192  Computer Science  1998  420.0      0.081348
193  Computer Science  1999  389.0      0.075344
194  Computer Science  2000  451.0      0.087352
195  Computer Science  2001  366.0      0.070889
196  Computer Science  2002  882.0      0.170831
197  Computer Science  2003  957.0      0.185357
2015
```

```
Computer Science
              domain  Year   Count  domain_Score
165  Computer Science  1963    1.0      0.000194
166  Computer Science  1972    5.0      0.000968
167  Computer Science  1973    1.0      0.000194
168  Computer Science  1974   12.0      0.002324
169  Computer Science  1975    1.0      0.000194
170  Computer Science  1976    1.0      0.000194
171  Computer Science  1977    1.0      0.000194
172  Computer Science  1978    3.0      0.000581
173  Computer Science  1979    3.0      0.000581
174  Computer Science  1980    5.0      0.000968
175  Computer Science  1981    6.0      0.001162
176  Computer Science  1982    2.0      0.000387
177  Computer Science  1983   12.0      0.002324
178  Computer Science  1984    2.0      0.000387
179  Computer Science  1985   11.0      0.002131
180  Computer Science  1986   24.0      0.004648
181  Computer Science  1987   20.0      0.003874
182  Computer Science  1988   51.0      0.009878
183  Computer Science  1989   29.0      0.005617
184  Computer Science  1990   39.0      0.007554
185  Computer Science  1991   62.0      0.012009
186  Computer Science  1992   71.0      0.013752
187  Computer Science  1993   98.0      0.018981
188  Computer Science  1994  114.0      0.022080
189  Computer Science  1995  374.0      0.072439
190  Computer Science  1996  349.0      0.067596
191  Computer Science  1997  401.0      0.077668
192  Computer Science  1998  420.0      0.081348
193  Computer Science  1999  389.0      0.075344
194  Computer Science  2000  451.0      0.087352
195  Computer Science  2001  366.0      0.070889
196  Computer Science  2002  882.0      0.170831
197  Computer Science  2003  957.0      0.185357
2014
Computer Science


In [65]: test_df

Out[65]:                                             Authors      index id  \
         0                 Ahmed Osman, Wojciech Samek  1802.00209v1
         1                 Ji Young Lee, Franck Dernoncourt  1603.03827v1
         2   Iulian Vlad Serban, Tim Klinger, Gerald Tesau...  1606.00776v2
         3   Sebastian Ruder, Joachim Bingel, Isabelle Aug...  1705.08142v2
         4   Iulian V. Serban, Chinnadhurai Sankar, Mathie...  1709.02349v2
         5   Kelvin Guu, Tatsunori B. Hashimoto, Yonatan O...  1709.08878v1
```

6	Iulian V. Serban, Chinnadhurai Sankar, Mathie...	1801.06700v1
7	Darko Brodic, Alessia Amelio, Zoran N. Milivo...	1609.06492v1
8	Mateusz Malinowski, Mario Fritz	1610.01076v1
9	Tony Beltramelli	1705.07962v2
10	Fred Richardson, Douglas Reynolds, Najim Dehak	1504.00923v1
11	Hieu Pham, Melody Y. Guan, Barret Zoph, Quoc ...	1802.03268v2
12	Brenden M. Lake, Tomer D. Ullman, Joshua B. T...	1604.00289v3
13	Hao Wang, Dit-Yan Yeung	1604.01662v2
14	Tejas D. Kulkarni, Karthik R. Narasimhan, Ard...	1604.06057v2
15	Deepak Pathak, Ross Girshick, Piotr Dollár, T...	1612.06370v2
16	Muhammad Ghifary, W. Bastiaan Kleijn, Mengjie...	1409.6041v1
17	Lionel Pigou, Aäron van den Oord, Sander Diel...	1506.01911v3
18	Rakesh Achanta, Trevor Hastie	1509.05962v2
19	Jeff Donahue, Philipp Krähenbühl, Trevor Darrell	1605.09782v7
20	Zachary C. Lipton	1606.03490v3
21	Sahil Garg, Irina Rish, Guillermo Cecchi, Aur...	1701.06106v2
22	Weifeng Ge, Yizhou Yu	1702.08690v2
23	Tanmay Gupta, Kevin Shih, Saurabh Singh, Dere...	1704.00260v2
24	Jan Hendrik Metzen, Mummadi Chaithanya Kumar,...	1704.05712v3
25	Quynh Nguyen, Matthias Hein	1704.08045v2
26	Chris Donahue, Zachary C. Lipton, Akshay Bals...	1705.07904v3
27	Mahesh Chandra Mukkamala, Matthias Hein	1706.05507v2
28	Chunyuan Li, Hao Liu, Changyou Chen, Yunchen ...	1709.01215v2
29	Mateusz Buda, Atsuto Maki, Maciej A. Mazurowski	1710.05381v1
..	...	...
70	Petr Baudi, Jan Pichl, Tomá Vyskoil, Jan ...	1603.06127v4
71	Jiatao Gu, Zhengdong Lu, Hang Li, Victor O. K...	1603.06393v3
72	Iulian Vlad Serban, Alberto García-Durán, Cag...	1603.06807v2
73	Chia-Wei Liu, Ryan Lowe, Iulian V. Serban, Mi...	1603.08023v2
74	Iulian Vlad Serban, Alessandro Sordoni, Ryan ...	1605.06069v3
75	Dirk Weissenborn	1606.03864v2
76	Marc Dymetman, Chunyang Xiao	1607.02467v2
77	Ondrej Bajgar, Rudolf Kadlec, Jan Kleindienst	1610.00956v1
78	James Bradbury, Stephen Merity, Caiming Xiong...	1611.01576v2
79	Jakob N. Foerster, Justin Gilmer, Jan Chorows...	1611.09434v2
80	Micha Daniluk, Tim Rocktäschel, Johannes Wel...	1702.04521v1
81	Zhouhan Lin, Minwei Feng, Cicero Nogueira dos...	1703.03130v1
82	Samuel Rönnqvist, Niko Schenk, Christian Chia...	1704.08092v1
83	Lara J. Martin, Prithviraj Ammanabrolu, Xinyu...	1706.01331v3
84	Tong Wang, Xingdi Yuan, Adam Trischler	1706.01450v1
85	Wei Wen, Yuxiong He, Samyam Rajbhandari, Minj...	1709.05027v7
86	Huda Hakami, Danushka Bollegala, Hayashi Kohei	1709.06673v2
87	Zhengdong Lu, Haotian Cui, Xianggen Liu, Yuku...	1709.08853v4
88	Bin Bi, Hao Ma	1709.10204v2
89	Mirco Ravanelli, Philemon Brakel, Maurizio Om...	1710.00641v1
90	Baolin Peng, Xiujun Li, Jianfeng Gao, Jingjin...	1801.06176v1
91	Andrew L. Maas, Peng Qi, Ziang Xie, Awni Y. H...	1406.7806v2
92	William Chan, Ian Lane	1504.01482v1

```
93                       David Krueger, Roland Memisevic   1511.08400v7
94      Noam Shazeer, Azalia Mirhoseini, Krzysztof Ma...   1701.06538v1
95        Yacine Jernite, Samuel R. Bowman, David Sontag   1705.00557v1
96                          Zhengyang Wang, Shuiwang Ji    1705.06824v1
97      Kyunghyun Cho, Bart van Merrienboer, Caglar G...    1406.1078v3
98              Zhiyuan Tang, Dong Wang, Zhiyong Zhang     1505.04630v5
99          Haim Sak, Andrew Senior, Françoise Beaufays    1402.1128v1

                                             paperTitle  Year  \
0    Dual Recurrent Attention Units for Visual Ques...  2018
1    Sequential Short-Text Classification with Recu...  2016
2    Multiresolution Recurrent Neural Networks: An ...  2016
3    Learning what to share between loosely related...  2017
4              A Deep Reinforcement Learning Chatbot    2017
5          Generating Sentences by Editing Prototypes    2017
6    A Deep Reinforcement Learning Chatbot (Short V...  2018
7    Document Image Coding and Clustering for Scrip...  2016
8    Tutorial on Answering Questions about Images w...  2016
9    pix2code: Generating Code from a Graphical Use...  2017
10   A Unified Deep Neural Network for Speaker and ...  2015
11   Efficient Neural Architecture Search via Param...  2018
12   Building Machines That Learn and Think Like Pe...  2016
13          Towards Bayesian Deep Learning: A Survey    2016
14   Hierarchical Deep Reinforcement Learning: Inte...  2016
15          Learning Features by Watching Objects Move   2016
16   Domain Adaptive Neural Networks for Object Rec...  2014
17   Beyond Temporal Pooling: Recurrence and Tempor...  2015
18          Telugu OCR Framework using Deep Learning    2015
19                      Adversarial Feature Learning    2016
20              The Mythos of Model Interpretability    2016
21   Neurogenesis-Inspired Dictionary Learning: Onl...  2017
22   Borrowing Treasures from the Wealthy: Deep Tra...  2017
23   Aligned Image-Word Representations Improve Ind...  2017
24   Universal Adversarial Perturbations Against Se...  2017
25   The loss surface of deep and wide neural networks  2017
26   Semantically Decomposing the Latent Spaces of ...  2017
27   Variants of RMSProp and Adagrad with Logarithm...  2017
28   ALICE: Towards Understanding Adversarial Learn...  2017
29   A systematic study of the class imbalance prob...  2017
..                                                 ...   ...
70   Sentence Pair Scoring: Towards Unified Framewo...  2016
71   Incorporating Copying Mechanism in Sequence-to...  2016
72   Generating Factoid Questions With Recurrent Ne...  2016
73   How NOT To Evaluate Your Dialogue System: An E...  2016
74   A Hierarchical Latent Variable Encoder-Decoder...  2016
75   Neural Associative Memory for Dual-Sequence Mo...  2016
76   Log-Linear RNNs: Towards Recurrent Neural Netw...  2016
77   Embracing data abundance: BookTest Dataset for...  2016
```

```
78                  Quasi-Recurrent Neural Networks  2016
79  Input Switched Affine Networks: An RNN Archite...  2016
80  Frustratingly Short Attention Spans in Neural ...  2017
81     A Structured Self-attentive Sentence Embedding  2017
82  A Recurrent Neural Model with Attention for th...  2017
83  Event Representations for Automated Story Gene...  2017
84  A Joint Model for Question Answering and Quest...  2017
85  Learning Intrinsic Sparse Structures within Lo...  2017
86  Why PairDiff works? -- A Mathematical Analysis...  2017
87  Object-oriented Neural Programming (OONP) for ...  2017
88  A Neural Comprehensive Ranker (NCR) for Open-D...  2017
89  Improving speech recognition by revising gated...  2017
90  Integrating planning for task-completion dialo...  2018
91  Building DNN Acoustic Models for Large Vocabul...  2014
92  Deep Recurrent Neural Networks for Acoustic Mo...  2015
93        Regularizing RNNs by Stabilizing Activations  2015
94  Outrageously Large Neural Networks: The Sparse...  2017
95  Discourse-Based Objectives for Fast Unsupervis...  2017
96  Learning Convolutional Text Representations fo...  2017
97  Learning Phrase Representations using RNN Enco...  2014
98  Recurrent Neural Network Training with Dark Kn...  2015
99  Long Short-Term Memory Based Recurrent Neural ...  2014


                                  Abstract  citation_count  \
0   We propose an architecture for VQA which utili...               1
1   Recent approaches based on artificial neural n...             105
2   We introduce the multiresolution recurrent neu...              59
3   Multi-task learning is motivated by the observ...               9
4   We present MILABOT: a deep reinforcement learn...               0
5   We propose a new generative model of sentences...              36
6   We present MILABOT: a deep reinforcement learn...               2
7   The paper introduces a new method for discrimi...               4
8   Together with the development of more accurate...               1
9   Transforming a graphical user interface screen...              26
10  Learned feature representations and sub-phonem...               0
11  We propose Efficient Neural Architecture Searc...              55
12  Recent progress in artificial intelligence (AI...             346
13  While perception tasks such as visual object r...              23
14  Learning goal-directed behavior in environment...             215
15  This paper presents a novel yet intuitive appr...              95
16  We propose a simple neural network model to de...              48
17  Recent studies have demonstrated the power of ...              56
18  In this paper, we address the task of Optical ...              10
19  The ability of the Generative Adversarial Netw...             357
20  Supervised machine learning models boast remar...             346
21  In this paper, we focus on online representati...               2
22  Deep neural networks require a large amount of...              34
23  An important goal of computer vision is to bui...               7
```

```
24  While deep learning is remarkably successful o...        14
25  While the optimization problem behind deep neu...        70
26  We propose a new algorithm for training genera...        25
27  Adaptive gradient methods have become recently...        13
28  We investigate the non-identifiability issues ...         0
29  In this study, we systematically investigate t...        55
..  ...                                                     ...
70  We review the task of Sentence Pair Scoring, p...        19
71  We address an important problem in sequence-to...         0
72  Over the past decade, large-scale supervised l...         0
73  We investigate evaluation metrics for dialogue...         0
74  Sequential data often possesses a hierarchical...         0
75  Many important NLP problems can be posed as du...         0
76  We introduce LL-RNNs (Log-Linear RNNs), an ext...         0
77  There is a practically unlimited amount of nat...         0
78  Recurrent neural networks are a powerful tool ...         0
79  There exist many problem domains where the int...         0
80  Neural language models predict the next token ...         0
81  This paper proposes a new model for extracting...         0
82  We introduce an attention-based Bi-LSTM for Ch...         0
83  Automated story generation is the problem of a...         0
84  We propose a generative machine comprehension ...         0
85  Model compression is significant for the wide ...         0
86  Representing the semantic relations that exist...         0
87  We propose Object-oriented Neural Programming ...         0
88  This paper proposes a novel neural machine rea...         0
89  Speech recognition is largely taking advantage...         0
90  Training a task-completion dialogue agent with...         0
91  Deep neural networks (DNNs) are now a central ...         0
92  We present a novel deep Recurrent Neural Netwo...         0
93  We stabilize the activations of Recurrent Neur...         0
94  The capacity of a neural network to absorb inf...         0
95  This work presents a novel objective function ...         0
96  Visual question answering is a recently propos...         0
97  In this paper, we propose a novel neural netwo...         0
98  Recurrent neural networks (RNNs), particularly...         0
99  Long Short-Term Memory (LSTM) is a recurrent n...         0

             domain  domain_score
0   Computer Science             0
1   Computer Science             0
2   Computer Science             0
3   Computer Science             0
4   Computer Science             0
5   Computer Science             0
6   Computer Science             0
7           Software             0
8   Computer Science             0
```

```
9                 Software               0
10        Computer Science               0
11                Software               0
12        Computer Science               0
13        Computer Science               0
14             Neuroscience              0
15        Computer Science               0
16        Computer Science               0
17        Computer Science               0
18        Computer Science               0
19                Political               0
20                 History               0
21        Computer Science               0
22        Computer Science               0
23        Computer Science               0
24                Political               0
25            Earth Science              0
26                Political               0
27            Earth Science              0
28        Computer Science               0
29        Computer Science               0
..                     ...             ...
70        Computer Science               0
71        Computer Science               0
72        Computer Science               0
73        Computer Science               0
74                Software               0
75        Computer Science               0
76        Computer Science               0
77        Computer Science               0
78      Biomedical Research              0
79        Computer Science               0
80        Computer Science               0
81                Business               0
82                Political               0
83        Computer Science               0
84        Computer Science               0
85        Computer Science               0
86        Computer Science               0
87        Computer Science               0
88        Computer Science               0
89                  Health               0
90                Business               0
91        Computer Science               0
92        Computer Science               0
93                Software               0
94        Computer Science               0
95        Computer Science               0
```

```
96      Computer Science            0
97               Software           0
98      Computer Science            0
99      Computer Science            0

[100 rows x 8 columns]
```

In [101]:
```python
# paper influence by giving wieghts to each parameter:
# 1) author popularity : 0.2,
# 2) domain-populaarity : 0.2, more popular domain, less influence paper is
# 3) citations_count : 0.8
metric_paper_list = []
domain_paper_list = []
for index, rows in train_df.iterrows():
    authors = rows["Authors"]
    citations = rows["citation_count"]
    domain = rows["domain"]
    year = rows['Year']
    authors_list = authors.split(',')
    count = 1
    curr_sum = 0
    for author in authors_list:
        auth_score = df_n.loc[df_n['author_name'] == author, 'SVS - index'].iloc[0]
        curr_sum = curr_sum + auth_score
        curr_sum = curr_sum / count
        count = count + 1
    domain_list = df_new[df_new['domain'] == domain]
    domain_score = 0
    if not domain_list.empty:
        domain_score = domain_list.loc[domain_list['Year'] == year, 'domain_Score'].
    paper_score = 0.5*curr_sum + 0.8*citations - domain_score + 2 * citations/(2018

    metric_paper_list.append(paper_score)
    domain_paper_list.append(domain_score)
train_df['reach_score'] = metric_paper_list
train_df['domain_score'] = domain_paper_list
```

```
---------------------------------------------------------------------------

AttributeError                            Traceback (most recent call last)

<ipython-input-101-87815812b2a2> in <module>()
    10     domain = rows["domain"]
    11     year = rows['Year']
---> 12     authors_list = authors.split(',')
    13     count = 1
    14     curr_sum = 0
```

```
AttributeError: 'numpy.float64' object has no attribute 'split'

In [102]: train_df

Out[102]:        Abstract  Authors  Year  index id  paperTitle  citation_count  domain  \
        0           2443     2128  1998         0       8355               3       5
        1           2854     5070  2003         1       9844               9      17
        2           2443     6113  2002         2       2261               0       5
        3           1896     1383  2002         3       9457               1       8
        4            998     4312  2003         4       2882               0       5
        5           2470     3937  2003         5       3641               0      17
        6           3624     5487  1999         6       4249               2       3
        7           2443      913  1995         7       5946               0       5
        8           2079     3868  2002         8       9143               0       5
        9           1036     8274  2002         9       3765               0       5
        10          2443     8585  1998        10       4710               2      18
        11          1686     6707  1999        11       7690               2       5
        12          2443     8089  1998        12       9439               0       5
        13           499     4169  2003        13       6596               1       5
        14          3210     7959  2003        14       8483               0       5
        15          4834     3197  1998        15       1413               0       3
        16          4795     1435  2002        16       9164               0       5
        17          2443     4760  1993        17       4495               8       3
        18          2198      992  2003        18       6235               2      19
        19          2443      419  1998        19       6382              14      17
        20          2443     8703  1998        20       1868               0       5
        21          2443     6553  2002        21       7902               0       9
        22          4210      646  2002        22       5098               0      18
        23          2443     3802  1984        23       8821               1      10
        24          2361     6536  2002        24       1564               0       5
        25          1281     4323  2002        25       8110               2       8
        26          4936     4891  2000        26       9485               1       5
        27          2443     6223  2001        27       6971               3       3
        28          2443     2511  1997        28       3266               0      18
        29          2443     6113  1999        29       3961               0      17
        ...          ...      ...   ...       ...        ...             ...     ...
        9970        5674     8902  2000      9970       3390               0       3
        9971        2443     1227  1999      9971       8782               1       5
        9972        2443     1464  1994      9972       2599               0       5
        9973        4702     6565  2002      9973        579              27       5
        9974        3642     2459  2003      9974       3075               0       3
        9975         862     4442  2001      9975       4037               0       5
        9976        1198      750  2003      9976       2964               3       3
        9977        1442     6518  1998      9977        194               3       5
        9978        2443     6037  1997      9978       6775               0       5
```

```
9979    2443    1465    2002    9979     483       6       5
9980    2443     671    1994    9980    4420       0       5
9981    2443     866    2001    9981    1873       0       5
9982    1322    8219    1999    9982    8815       1      10
9983    3974    8420    2002    9983    2412       4       5
9984    2443    5517    2001    9984    9281       2       3
9985    4812    1599    2002    9985    9251       1       5
9986     799    6557    2000    9986    3917       0       3
9987     848    9356    1998    9987    4364       0       3
9988    2443    3090    1979    9988    7966       0       7
9989     584     578    2000    9989    1324       7      18
9990    1470    8358    2002    9990    9758       3       5
9991    4788    6529    2001    9991    5594       0       5
9992    2443    8297    2002    9992    8350       0       0
9993    2443    2804    2002    9993    7078       0      11
9994    2004    1207    2001    9994    7863       2       3
9995    2443    5827    2002    9995    4388       0       5
9996    2443    4478    1997    9996    7122       0       5
9997    2443    7302    2003    9997    2107       0       5
9998    2443     996    2001    9998    9884       0       5
9999    2093    1390    2000    9999    8598       1      17


        reach_score   domain_score
0        1.255956        0.081348
1        3.809434        0.230189
2       -0.512493        0.170831
3        0.000904        0.166365
4       -0.556072        0.185357
5       -0.690566        0.230189
6        0.748332        0.083889
7       -0.217316        0.072439
8       -0.512493        0.170831
9       -0.512493        0.170831
10       0.732408        0.089197
11       0.773969        0.075344
12      -0.244044        0.081348
13      -0.056072        0.185357
14      -0.556072        0.185357
15      -0.308866        0.102955
16      -0.512493        0.170831
17       3.959962        0.013346
18       0.162338        0.279221
19       6.796226        0.067925
20      -0.244044        0.081348
21      -0.838509        0.279503
22      -0.532210        0.177403
23       0.489583        0.003472
24      -0.512493        0.170831
```

```
25        0.500904      0.166365
26        0.237943      0.087352
27        1.219733      0.093422
28       -0.205154      0.068385
29       -0.294340      0.098113
...            ...           ...
9970     -0.300286      0.100095
9971      0.273969      0.075344
9972     -0.066241      0.022080
9973     12.987507      0.170831
9974     -0.743565      0.247855
9975     -0.212667      0.070889
9976      0.756435      0.247855
9977      1.255956      0.081348
9978     -0.233004      0.077668
9979      2.487507      0.170831
9980     -0.066241      0.022080
9981     -0.212667      0.070889
9982      0.322917      0.059028
9983      1.487507      0.170831
9984      0.719733      0.093422
9985     -0.012493      0.170831
9986     -0.300286      0.100095
9987     -0.308866      0.102955
9988     -0.130435      0.043478
9989      3.259167      0.080278
9990      0.987507      0.170831
9991     -0.212667      0.070889
9992     -0.383721      0.127907
9993     -0.887324      0.295775
9994      0.719733      0.093422
9995     -0.512493      0.170831
9996     -0.233004      0.077668
9997     -0.556072      0.185357
9998     -0.212667      0.070889
9999      0.183019      0.105660

[10000 rows x 9 columns]
```

10   Time Analysis of Paper citation count

```
In [69]: # Graph 1: Graph which will tell the domain and the total number of papers in the dom

In [70]: df_domain_count = paper_domain_dict

In [71]: df_domain_count

Out[71]: {'Computer Science': 5163,
          'Political': 265,
```

```
       'Geometry': 553,
       'Business': 1049,
       'Software': 1009,
       'Statistics': 154,
       'Health': 161,
       'History': 288,
       'Art': 86,
       'Earth Science': 570,
       'Literature': 142,
       'Biomedical Research': 221,
       'Microprocessor': 162,
       'Maths': 80,
       'N/A': 18,
       'Economics': 23,
       'Physics': 24,
       'Biology': 15,
       'Chemistry': 4,
       'Neuroscience': 13}
```

```python
In [72]: df1 = pd.DataFrame(list(paper_domain_dict.items()), columns=['Domain', 'Count'])
```

```python
In [73]: df1 = df1.sort_values(by='Count', ascending=False)
         df1 = df1.head(10)
```

```python
In [74]: import matplotlib.pyplot as plt_1

         x = df1['Domain']
         y = df1['Count']
         plt_1.bar(x,y, width = 0.8, color = ['red', 'green'])
         plt_1.xlabel('Domain')
         plt_1.ylabel('Count')
         plt_1.xticks(rotation=90);
```
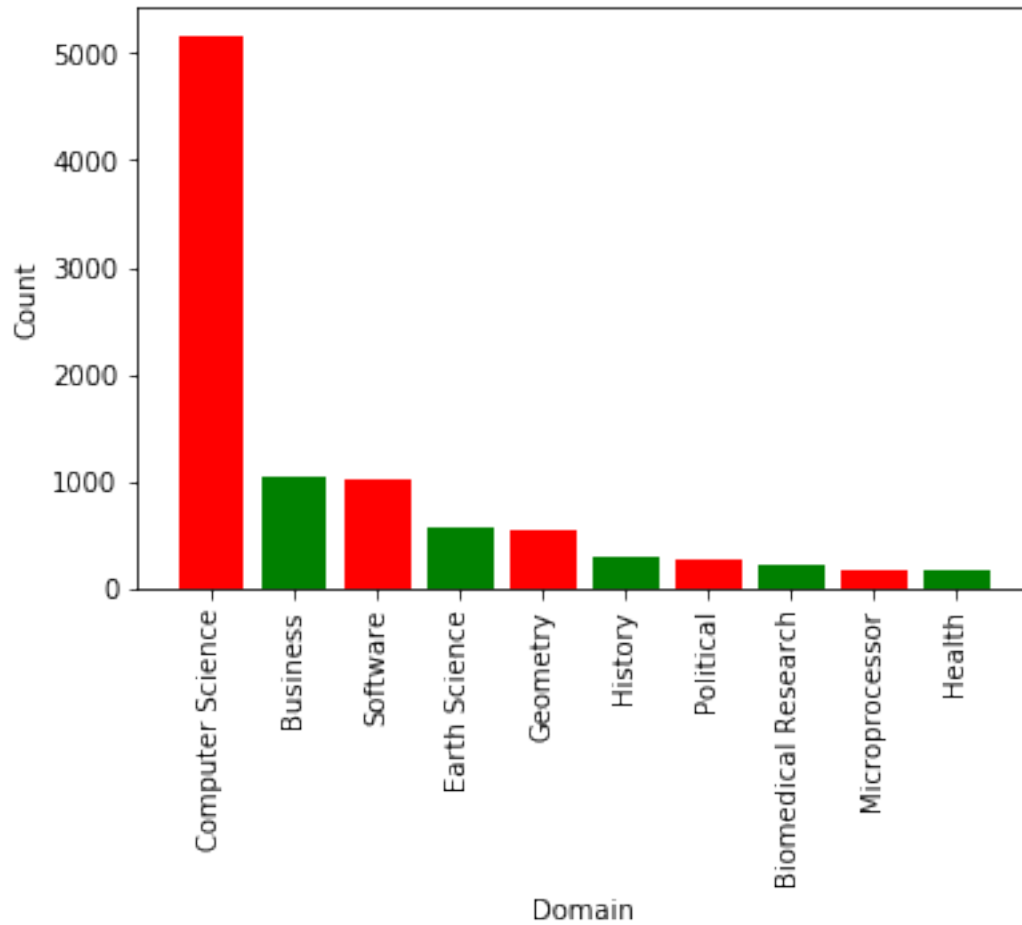
```
In [75]: mydf = df_n.head(10)

         x = mydf['author_name']
         y = mydf['citation_count']
         plt_1.bar(x,y, width = 0.8, color = ['blue', 'orange'])
         plt_1.xlabel('Authors')
         plt_1.ylabel('Citation Count')
         plt_1.xticks(rotation=90);
```
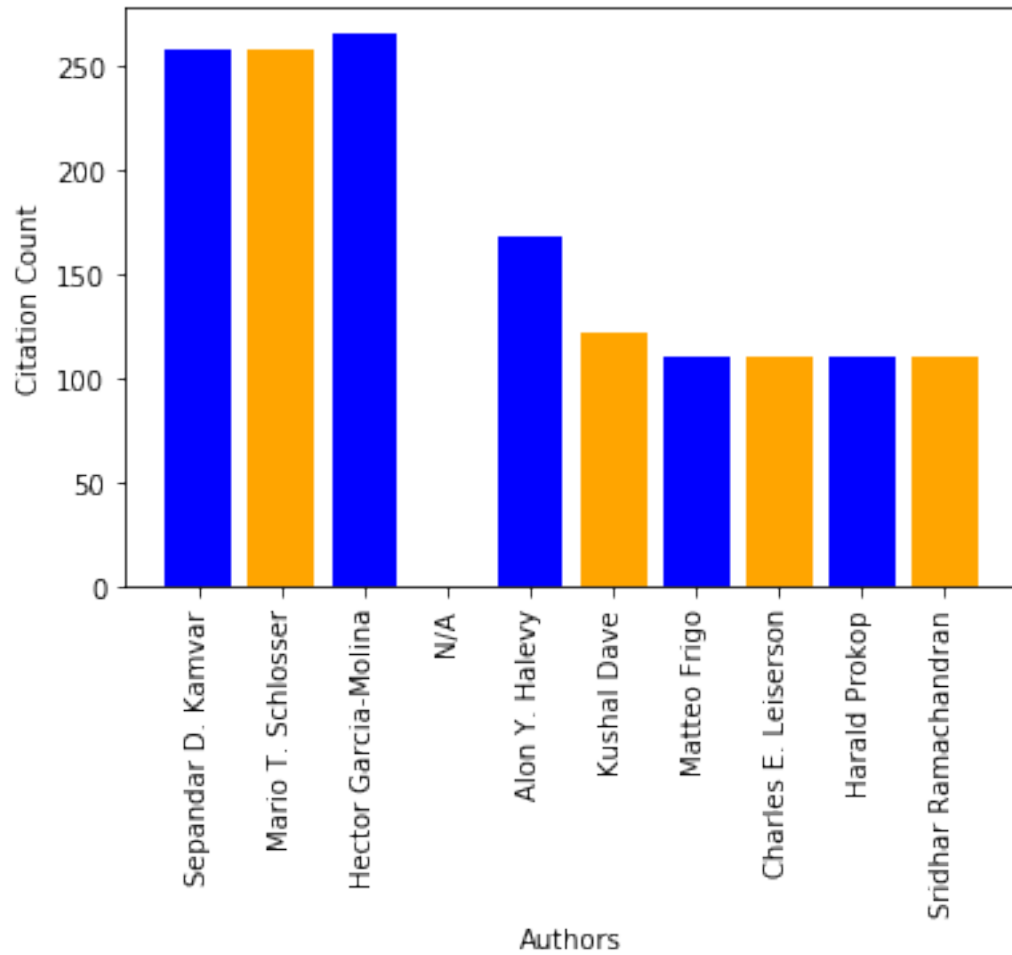
In [76]: mydf = df_n.head(10)

```
x = mydf['author_name']
y = mydf['SVS - index']
plt_1.bar(x,y, width = 0.8, color = ['magenta', 'cyan'])
plt_1.xlabel('Authors')
plt_1.ylabel('SVS-Index')
plt_1.xticks(rotation=90);
```
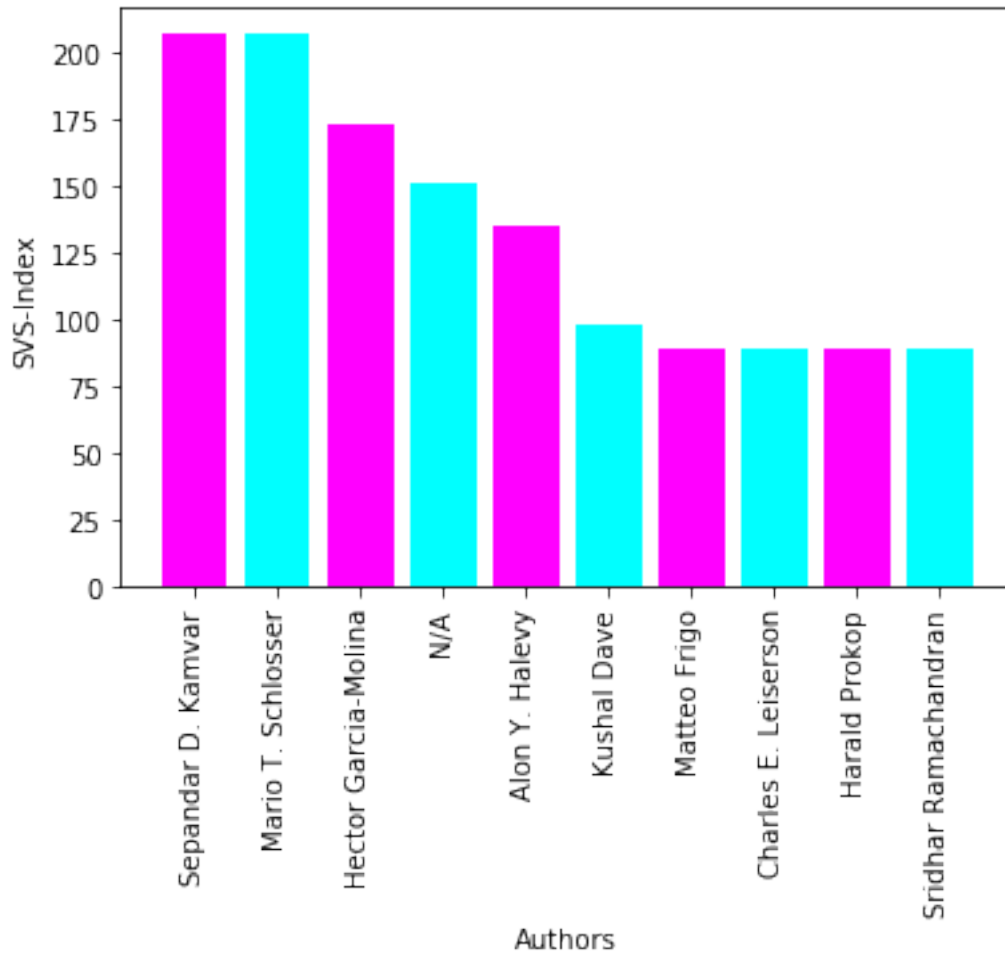
```
In [77]: import warnings
         warnings.filterwarnings("ignore")
         temp_df = df_new.ix[(df_new['domain'] == 'Computer Science')]
         temp_df1 = df_new.ix[(df_new['domain'] == 'Geometry')]
         temp_df2 = df_new.ix[(df_new['domain'] == 'Business')]

In [78]: temp_df.plot(x='Year', y=['domain', 'Count'], figsize=(10,5), grid=True, color="green"
         temp_df1.plot(x='Year', y=['domain', 'Count'], figsize=(10,5), grid=True, color="red"
         temp_df2.plot(x='Year', y=['domain', 'Count'], figsize=(10,5), grid=True, color="magel

Out[78]: <matplotlib.axes._subplots.AxesSubplot at 0x1a9287da58>
```
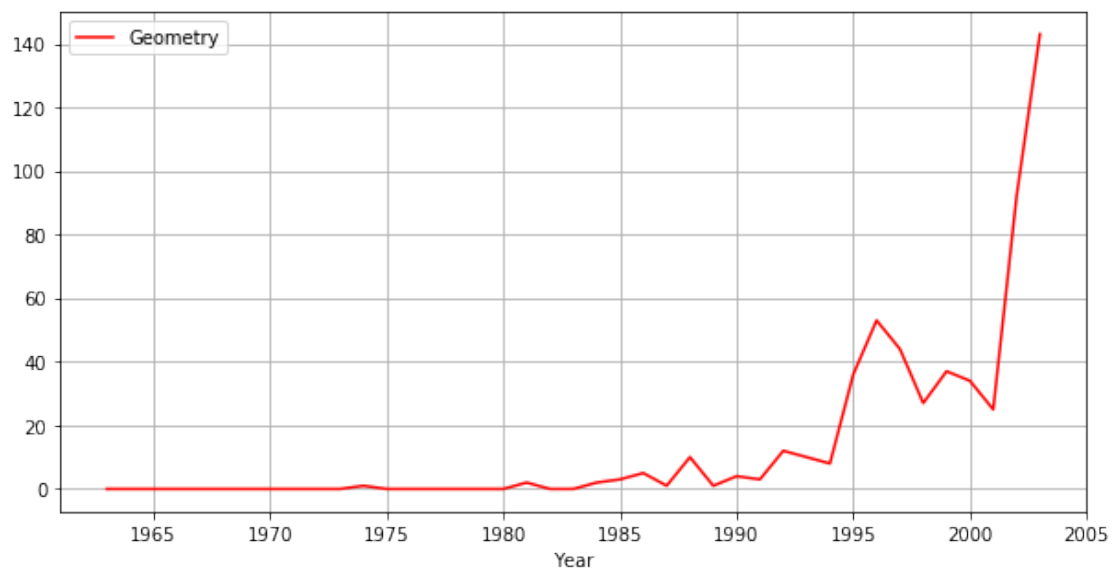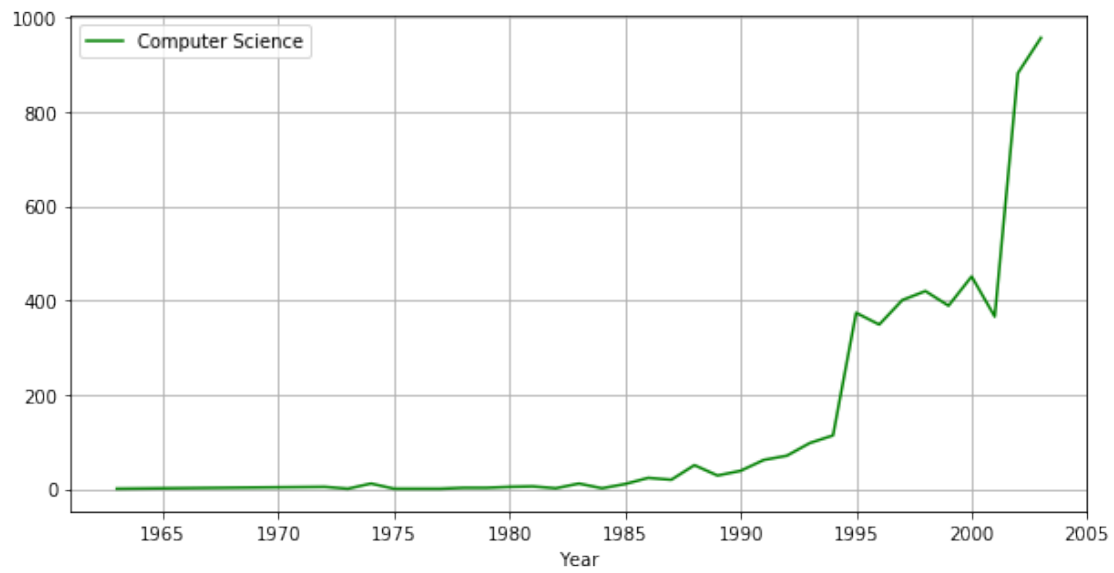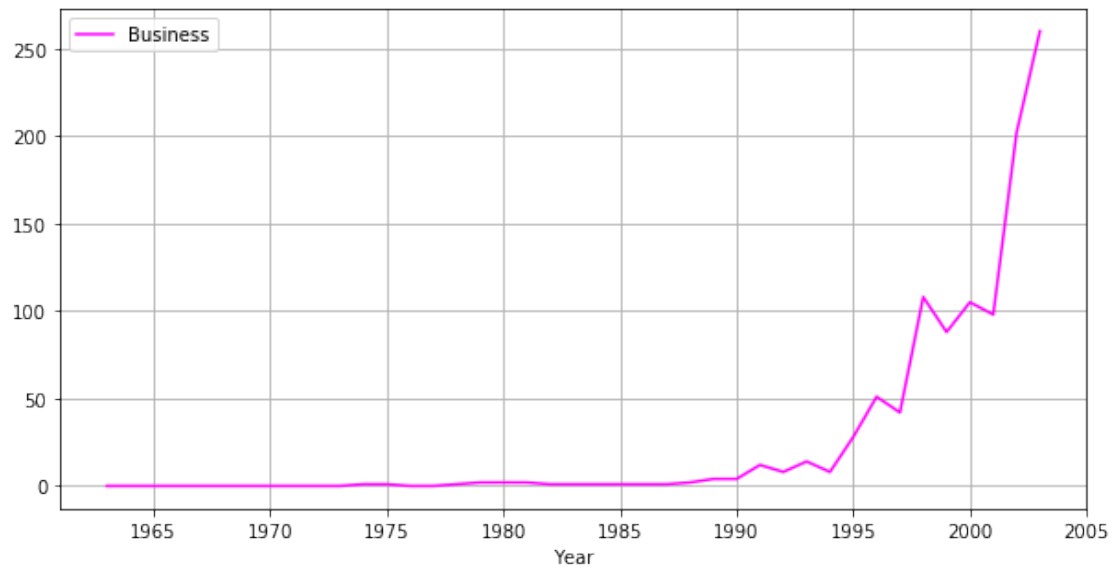
```
In [79]: x = temp_df['Year']
         y1 = temp_df['Count']
         y2 = temp_df1['Count']
         plt.plot(x, y1, x, y2)
         plt.legend(['Computer Science', 'Geometry'])
```

Out[79]: <matplotlib.legend.Legend at 0x1a94c75470>
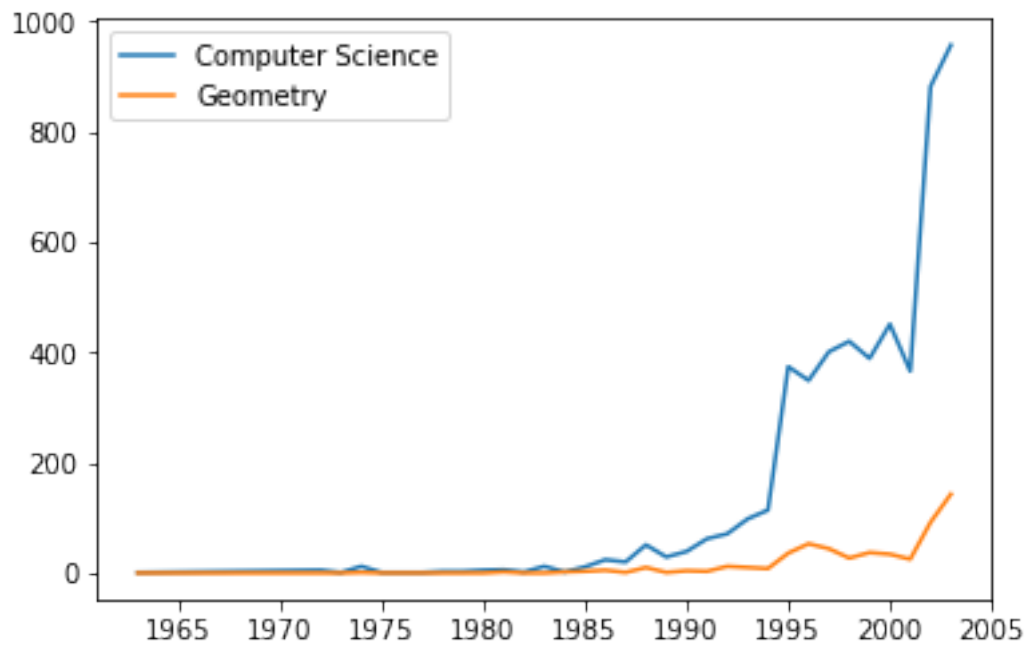
11   Baseline Model

```
In [80]: from sklearn.linear_model import LinearRegression
         from sklearn.model_selection import train_test_split
         from sklearn.metrics import mean_squared_error
         from sklearn import preprocessing
```

```
In [81]: features = ['Abstract', 'paperTitle', 'domain', 'Authors', 'Year', 'citation_count']
```

```
In [82]: cat_cols = ["Authors", "paperTitle", "domain", "Abstract"]

         for col in cat_cols:
             print(col)
             lbl = preprocessing.LabelEncoder()
             lbl.fit(list(train_df[col].values.astype('str')) + list(test_df[col].values.astype
             train_df[col] = lbl.transform(list(train_df[col].values.astype('str')))
             test_df[col] = lbl.transform(list(test_df[col].values.astype('str')))
```

```
Authors
paperTitle
domain
Abstract
```

```
In [83]: X = train_df[features].values
         Y = train_df['reach_score'].values
```

```
In [84]: X_train, X_test, y_train, y_test = train_test_split(X, Y, test_size = 0.25)
```

```
In [85]: linear_regressor = LinearRegression()
         linear_regressor.fit(X_train, y_train)
         print (linear_regressor.coef_)
         y_test_pred = linear_regressor.predict(X_test)
```

```
[ 1.73513301e-06 -1.68650447e-06 -1.39947169e-03  1.92758015e-07
 -3.41163194e-02  4.99987591e-01]
```

```
In [86]: root_mean_square_error = np.sqrt(mean_squared_error(y_test,y_test_pred))
         root_mean_square_error
```

```
Out[86]: 0.1487473001505127
```

```
In [87]: linear_regressor.fit(X, Y)
```

```
Out[87]: LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None,
                  normalize=False)
```

```
In [88]: test_predictions = linear_regressor.predict(test_df[features])
```

```
In [89]: test_predictions

Out[89]: array([-4.90709040e-01,  5.15712090e+01,  2.85758772e+01,  3.53533670e+00,
               -9.52073560e-01,  1.70436162e+01,  1.39938554e-02,  1.05606505e+00,
               -4.33496667e-01,  1.20145368e+01, -8.90039644e-01,  2.64934983e+01,
                1.72094487e+02,  1.05694914e+01,  1.06567918e+02,  4.65778534e+01,
                2.31481609e+01,  2.71120452e+01,  4.09729420e+00,  1.77580333e+02,
                1.72077046e+02,  3.28200334e-02,  1.60394929e+01,  2.53858319e+00,
                6.01747180e+00,  3.40369808e+01,  1.15200396e+01,  5.52395990e+00,
               -9.53562560e-01,  2.65442686e+01,  4.03252516e+00,  7.50384915e+00,
                1.11026516e-02,  2.20632747e+02,  1.45736406e+01, -9.77209498e-01,
                2.56820787e+00,  1.35759567e+01, -9.22191255e-01, -9.25123513e-01,
                3.56599560e+00,  3.65715786e+01,  1.56398664e+00,  9.03817406e+00,
                4.54023145e+00,  6.03561447e+00,  6.02986853e+00,  1.35240870e+01,
                1.55439673e+01, -9.73314145e-01, -9.62452658e-01,  1.31128005e+02,
                2.53421839e+00, -9.46750106e-01,  5.74833079e-01,  7.12415285e+00,
               -9.23126359e-01,  8.61289015e+00,  1.40666206e+01,  3.56347605e+00,
                6.60098306e+00, -8.84313542e-01,  1.02112273e+02,  2.15637716e+02,
                1.68127732e+02,  2.46046927e+01,  3.56037075e+01,  1.63623921e+02,
                6.46036645e+01,  1.36085378e+01,  8.57094371e+00, -9.26106337e-01,
               -9.29170954e-01, -9.25199236e-01, -9.40776603e-01, -9.32579743e-01,
               -9.26653995e-01, -9.25740105e-01, -9.29409111e-01, -9.27892100e-01,
               -9.63328357e-01, -9.51138172e-01, -9.69917193e-01, -9.65615320e-01,
               -9.52106131e-01, -9.65239798e-01, -9.71668403e-01, -9.61651518e-01,
               -9.53584248e-01, -9.69045153e-01, -9.91752014e-01, -8.60368380e-01,
               -8.88352597e-01, -9.13808873e-01, -9.65819236e-01, -9.57820895e-01,
               -9.60964901e-01, -8.82251091e-01, -8.99908798e-01, -8.63573201e-01])

In [90]: test_df

Out[90]:       Authors     index id  paperTitle  Year  Abstract  citation_count  domain  \
         0           0  1802.00209v1        3328  2018      5457               1       5
         1          43  1603.03827v1        8035  2016      2744             105       5
         2          37  1606.00776v2        6071  2016      5207              59       5
         3          77  1705.08142v2        5391  2017      2425               9       5
         4          33  1709.02349v2         192  2017      5238               0       5
         5          49  1709.08878v1        4358  2017      5421              36       5
         6          34  1801.06700v1         193  2018      5237               2       5
         7          13  1609.06492v1        3297  2016      3667               4      18
         8          59  1610.01076v1        9376  2016      4775               1       5
         9          83  1705.07962v2        9930  2017      4802              26      18
         10         23  1504.00923v1         758  2015      2189               0       5
         11         27  1802.03268v2        3496  2018      5392              55      18
         12          9  1604.00289v3        2173  2016      2748             346       5
         13         24  1604.01662v2        9234  2016      5653              23       5
         14         80  1604.06057v2        4562  2016      2190             215      15
         15         15  1612.06370v2        5353  2016      4465              95       5
         16         63   1409.6041v1        3306  2014      5445              48       5
```

17	55	1506.01911v3	2056	2015	2757	56	5
18	70	1509.05962v2	8656	2015	1861	10	5
19	42	1605.09782v7	1248	2016	3237	357	17
20	93	1606.03490v3	8877	2016	3022	346	10
21	73	1701.06106v2	6155	2017	1898	2	5
22	88	1702.08690v2	2126	2017	936	34	5
23	79	1704.00260v2	1298	2017	528	7	5
24	40	1704.05712v3	9454	2017	5644	14	17
25	69	1704.08045v2	9067	2017	5655	70	6
26	11	1705.07904v3	8013	2017	5415	25	17
27	56	1706.05507v2	9622	2017	410	13	6
28	12	1709.01215v2	1107	2017	5221	0	5
29	58	1710.05381v1	1058	2017	2028	55	5
..	...	...	...	...	...	...	...
70	68	1603.06127v4	8027	2016	5512	19	5
71	45	1603.06393v3	4866	2016	4898	0	5
72	35	1603.06807v2	4354	2016	2595	0	5
73	10	1603.08023v2	4639	2016	5212	0	5
74	36	1605.06069v3	348	2016	2877	0	18
75	16	1606.03864v2	6141	2016	2270	0	5
76	57	1607.02467v2	5510	2016	5188	0	5
77	65	1610.00956v1	3632	2016	3947	0	5
78	39	1611.01576v2	7364	2016	2785	0	2
79	38	1611.09434v2	4969	2016	3937	0	5
80	60	1702.04521v1	4277	2017	2460	0	5
81	97	1703.03130v1	702	2017	4604	0	3
82	75	1704.08092v1	619	2017	5203	0	17
83	51	1706.01331v3	3818	2017	668	0	5
84	82	1706.01450v1	375	2017	5403	0	5
85	87	1709.05027v7	5354	2017	2357	0	5
86	29	1709.06673v2	9846	2017	2805	0	5
87	94	1709.08853v4	6313	2017	5394	0	5
88	8	1709.10204v2	475	2017	4607	0	5
89	61	1710.00641v1	4839	2017	2981	0	9
90	6	1801.06176v1	5020	2018	4796	0	3
91	3	1406.7806v2	2168	2014	935	0	5
92	89	1504.01482v1	2971	2015	5304	0	5
93	14	1511.08400v7	7577	2015	5549	0	18
94	64	1701.06538v1	6730	2017	3316	0	5
95	90	1705.00557v1	3227	2017	4730	0	5
96	95	1705.06824v1	5351	2017	4887	0	5
97	50	1406.1078v3	5357	2014	1976	0	18
98	96	1505.04630v5	7520	2015	2784	0	5
99	26	1402.1128v1	5525	2014	2231	0	5

	domain_score
0	0
1	0

2	0
3	0
4	0
5	0
6	0
7	0
8	0
9	0
10	0
11	0
12	0
13	0
14	0
15	0
16	0
17	0
18	0
19	0
20	0
21	0
22	0
23	0
24	0
25	0
26	0
27	0
28	0
29	0
..	...
70	0
71	0
72	0
73	0
74	0
75	0
76	0
77	0
78	0
79	0
80	0
81	0
82	0
83	0
84	0
85	0
86	0
87	0
88	0

```
89          0
90          0
91          0
92          0
93          0
94          0
95          0
96          0
97          0
98          0
99          0

[100 rows x 8 columns]
```

In [91]: *#Make a csv of final predicted results*
         submission = pd.DataFrame(
             {'key': test_df.paperTitle, 'reach_score': test_predictions},
             columns = ['key', 'reach_score'])
         submission.to_csv('submission.csv', index = False)

In [92]: import eli5
         from eli5.sklearn import PermutationImportance

         perm = PermutationImportance(linear_regressor, random_state=1).fit(X, Y)
         eli5.show_weights(perm, feature_names = features)

Out[92]: <IPython.core.display.HTML object>

In [93]: *#Random Forest Model*
         from sklearn.ensemble import RandomForestRegressor
         random_forest = RandomForestRegressor(n_estimators=24, max_features=None, max_depth=2
                                 min_samples_leaf=3, random_state=0)
         random_forest.fit(X_train, y_train)
         y_test_pred_1 = random_forest.predict(X_test)

In [94]: root_mean_square_error_rf = np.sqrt(mean_squared_error(y_test,y_test_pred_1))
         root_mean_square_error_rf

Out[94]: 0.19336991124693018

In [95]: random_forest.fit(X, Y)
         test_predictions_rf = random_forest.predict(test_df[features])

         #Make a csv of final predicted results
         submission = pd.DataFrame(
             {'key': test_df.paperTitle, 'Reach-Score': test_predictions_rf},
             columns = ['key', 'Reach-Score'])
         submission.to_csv('submission_rf1.csv', index = False)

In [96]: test_predictions_rf

```
Out[96]: array([-5.60720511e-02,  5.87618513e+01,  2.98986549e+01,  3.94337167e+00,
                -5.56072051e-01,  1.77673386e+01,  4.43927949e-01,  1.37394336e+00,
                -5.60720511e-02,  1.25936302e+01, -5.56072051e-01,  2.63652758e+01,
                 8.16738815e+01,  1.09390684e+01,  8.16738815e+01,  4.85614339e+01,
                 2.32885750e+01,  2.66315704e+01,  4.41926153e+00,  8.16738815e+01,
                 8.16738815e+01,  4.43927949e-01,  1.66053888e+01,  2.93552643e+00,
                 6.42414756e+00,  3.54302853e+01,  1.19995908e+01,  5.92564588e+00,
                -5.56072051e-01,  2.64941142e+01,  4.44433064e+00,  7.93089754e+00,
                 4.43927949e-01,  8.16738815e+01,  1.48935535e+01, -6.33967821e-01,
                 2.94209714e+00,  1.39944756e+01, -5.56072051e-01, -5.56072051e-01,
                 3.92598246e+00,  3.61730630e+01,  1.93403247e+00,  9.45983062e+00,
                 4.90887933e+00,  6.43916189e+00,  6.43883364e+00,  1.40084040e+01,
                 1.59786832e+01, -5.56072051e-01, -5.56072051e-01,  8.16738815e+01,
                 2.92372531e+00, -5.67888999e-01,  9.60203330e-01,  7.41802698e+00,
                -5.56072051e-01,  8.91561000e+00,  1.47362965e+01,  3.92598246e+00,
                 6.94463110e+00, -5.56072051e-01,  8.16738815e+01,  8.16738815e+01,
                 8.16738815e+01,  2.54429094e+01,  3.59877965e+01,  8.16738815e+01,
                 8.16738815e+01,  1.40204216e+01,  8.93295392e+00, -5.56072051e-01,
                -5.56072051e-01, -5.56072051e-01, -5.67888999e-01, -5.56072051e-01,
                -5.56072051e-01, -5.56072051e-01, -3.97741327e-01, -5.56072051e-01,
                -5.56072051e-01, -7.43565300e-01, -7.26893939e-01, -5.56072051e-01,
                -5.56072051e-01, -5.56072051e-01, -5.56072051e-01, -5.56072051e-01,
                -5.56072051e-01, -6.05418873e-01, -7.43565300e-01, -5.56072051e-01,
                -5.56072051e-01, -5.67888999e-01, -5.56072051e-01, -5.56072051e-01,
                -5.56072051e-01, -5.67888999e-01, -5.56072051e-01, -5.56072051e-01])
```

```python
In [97]: import eli5
         from eli5.sklearn import PermutationImportance

         perm = PermutationImportance(random_forest, random_state=1).fit(X, Y)
         eli5.show_weights(perm, feature_names = features)
```

```
Out[97]: <IPython.core.display.HTML object>
```

12 Advanced Model

```python
In [98]: test_X = test_df[features]
         import lightgbm as lgb
         #Train and run the lgb model
         def run_lgb(train_X, train_y, val_X, val_y, test_X):
             params = {
                 "objective" : "regression", "metric" : "rmse", "num_leaves" : 30,
                 "min_child_samples" : 100, "learning_rate" : 0.1, "bagging_fraction" : 0.7,
                 "feature_fraction" : 0.5, "bagging_frequency" : 5, "bagging_seed" : 2018,
                 "verbosity" : -1
             }

             lgtrain = lgb.Dataset(train_X, label=train_y)
```

```
        lgval = lgb.Dataset(val_X, label=val_y)
        model = lgb.train(params, lgtrain, 2000, valid_sets=[lgval], early_stopping_rounds
        
        pred_test_y = model.predict(test_X, num_iteration=model.best_iteration)
        pred_val_y = model.predict(val_X, num_iteration=model.best_iteration)
        return pred_test_y, model, pred_val_y

    X_train, X_test, y_train, y_test
    pred_test, model, pred_val = run_lgb(X_train, y_train, X_test, y_test, test_X)
```

```
Training until validation scores don't improve for 100 rounds.
[100]       valid_0's rmse: 1.40951
Early stopping, best iteration is:
[89]        valid_0's rmse: 1.40822
```

```
In [99]: #Make a csv of final predicted results
        submission = pd.DataFrame(
            {'key': test_df.paperTitle, 'Reach-Score': pred_test},
            columns = ['key', 'Reach-Score'])
        submission.to_csv('advanced_lgb.csv', index = False)
```

```
In [103]: test_df = test_df.sort_values(by='citation_count', ascending=False)
          test_df.head(10)
```

Out[103]:

	Authors	index id	paperTitle	Year	Abstract	citation_count	domain \
33	54	1502.08029v5	3013	2015	2750	443	5
63	31	1507.04808v3	2169	2015	5226	433	5
19	42	1605.09782v7	1248	2016	3237	357	17
20	93	1606.03490v3	8877	2016	3022	346	10
12	9	1604.00289v3	2173	2016	2748	346	5
64	17	1508.04395v2	3658	2015	2274	338	3
67	81	1509.06664v4	7476	2015	5650	329	5
51	30	1511.05234v2	1825	2015	4902	264	5
14	80	1604.06057v2	4562	2016	2190	215	15
62	71	1506.08909v3	8959	2015	4346	206	5

	domain_score
33	0
63	0
19	0
20	0
12	0
64	0
67	0
51	0
14	0
62	0