מכון טכנולוגי חולון
Holon Institute of Technology

ג"סמסטר סתיו תשפ
Fall semester 2023

Hirsch Index Prediction

☞ Introduction to Data Science

☞ Lecturers: Dr. Naama Kopelman,
Dr. Jonathan Schler

Link to Slideshow

Edo Koren

Abstract

The h-index is an author-level metric that measures both the productivity and citation impact of the publications, initially used for an individual scientist or scholar.

The h-index correlates with obvious success indicators such as winning the Nobel Prize, being accepted for research fellowships and holding positions at top universities.

Example

The h index is 3 because there are at least 3 publications with 3 citations

Publications of a scientist	Citations per publication
Publication No.1	**3**
Publication No.2	2
Publication No.3	**21**
Publication No.4	**7**

Abstract - cont.

While there's an algorithm for calculating the h-index – we assume that with a given information about scientists we could predict their h-index as well with the using of some basic machine learning algorithms.

There are dozens of websites providing mass information about thousands of scientist among thousands of institutes

ACADEMIA

LinkedIn

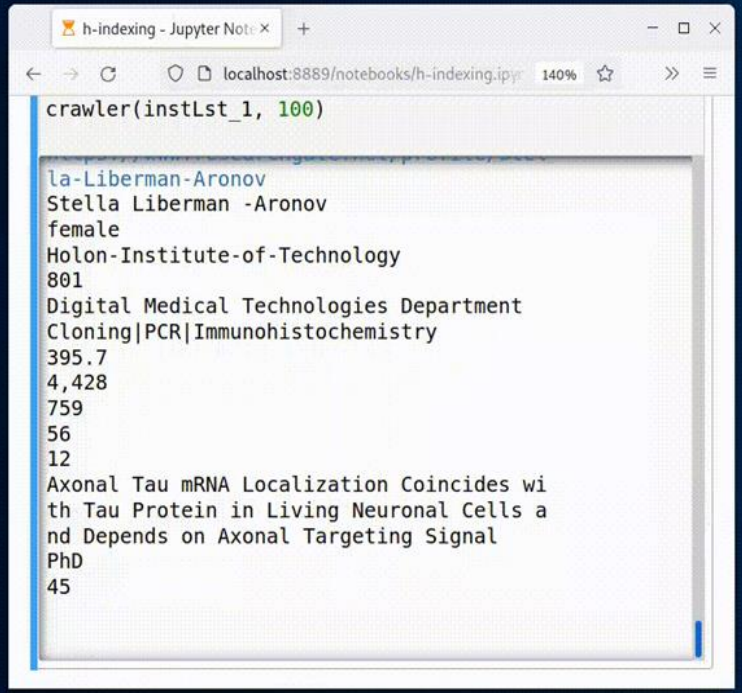
Google Scholar

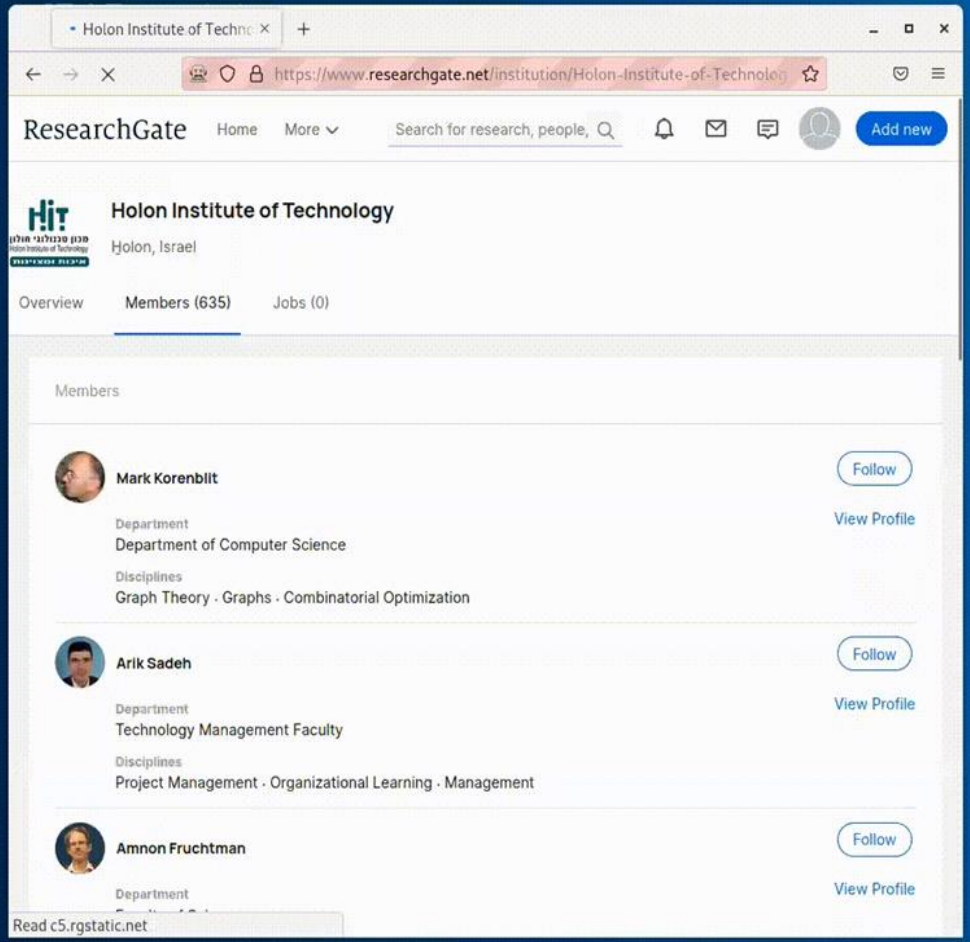
Scraping Data

Data Scraping

“Data scraping used for extracting data from websites. Web scraping software may directly access the World Wide Web using the Hypertext Transfer Protocol or a web browser.” ~wikipedia

We used both crawling and API. the crawling itself took about 3 days of gathering data with Selenium.

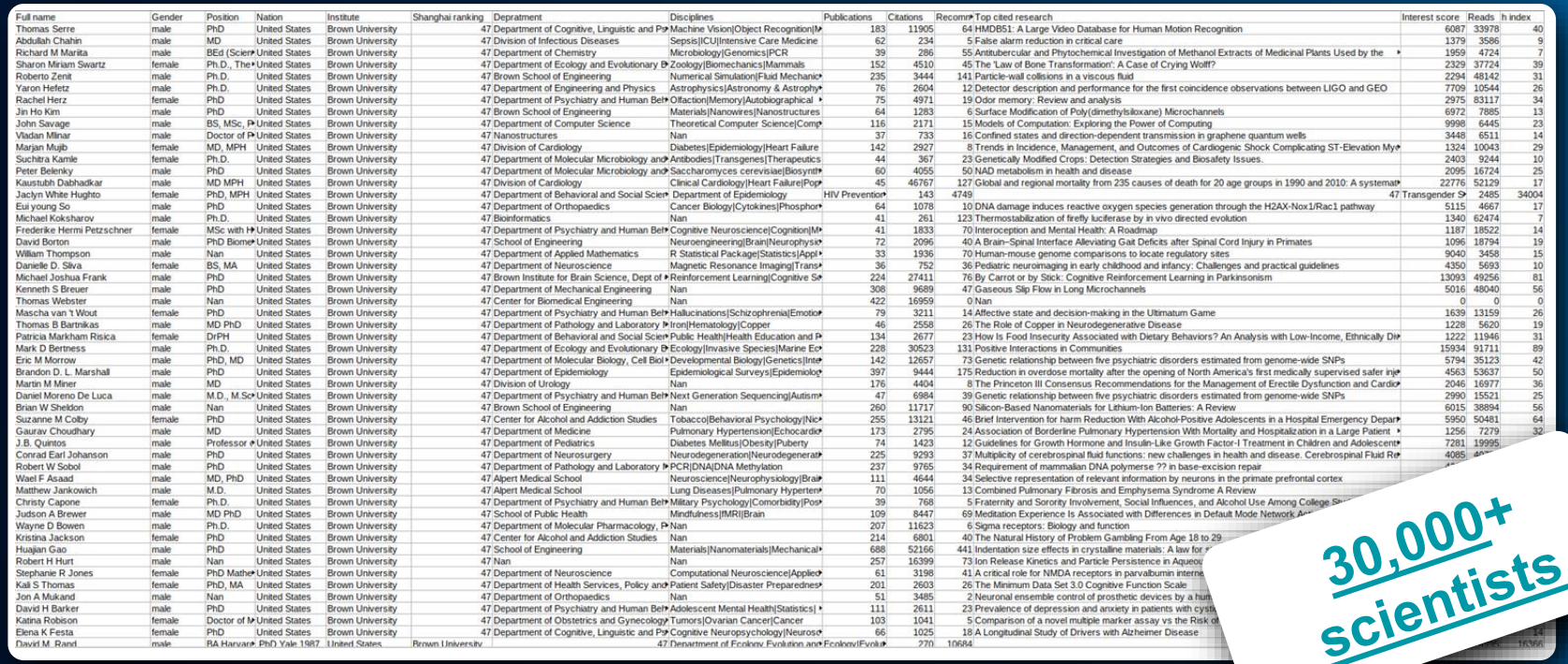

h-indexing - Jupyter Note
localhost:8889/notebooks/h-indexing.ipy
140%
crawler(instLst_1, 100)
la-Liberman-Aronov
Stella Liberman -Aronov
female
Holon-Institute-of-Technology
801
Digital Medical Technologies Department
Cloning|PCR|Immunohistochemistry
395.7
4,428
759
56
12
Axonal Tau mRNA Localization Coincides wi
th Tau Protein in Living Neuronal Cells a
nd Depends on Axonal Targeting Signal
PhD
45


Holon Institute of Techno
https://www.researchgate.net/institution/Holon-Institute-of-Technolog
ResearchGate
Home
More
Search for research, people,
Add new
Holon Institute of Technology
Holon, Israel
Overview
Members (635)
Jobs (0)
Members
Mark Korenblit
Follow
Department
Department of Computer Science
View Profile
Disciplines
Graph Theory · Graphs · Combinatorial Optimization
Arik Sadeh
Follow
Department
Technology Management Faculty
View Profile
Disciplines
Project Management · Organizational Learning · Management
Amnon Fruchtman
Follow
Department
View Profile
Read c5.rgstatic.net

Web crawling - CSV output

Full name	Gender	Position	Nation	Institute	Shanghai ranking	Depratment	Disciplines	Publications	Citations	Recom	Top cited research	Interest score	Reads	h index
Thomas Serre	male	PhD	United States	Brown University	47	Department of Cognitive, Linguistic and Psy	Machine Vision|Object Recognition|M	183	11905	64	HMDB51: A Large Video Database for Human Motion Recognition	6087	33978	40
Abdullah Chahin	male	MD	United States	Brown University	47	Division of Infectious Diseases	Sepsis|ICU|Intensive Care Medicine	62	234	5	False alarm reduction in critical care	1379	3586	9
Richard M Marika	male	BEd (Scien	United States	Brown University	47	Department of Chemistry	Microbiology|Genomics|PCR	39	286	55	Antitubercular and Phytochemical Investigation of Methanol Extracts of Medicinal Plants Used by the	1959	4724	7
Sharon Miriam Swartz	female	Ph.D., The	United States	Brown University	47	Department of Ecology and Evolutionary B	Zoology|Biomechanics|Mammals	152	4510	45	The 'Law of Bone Transformation': A Case of Crying Wolff?	2329	37724	39
Roberto Zenit	male	Ph.D.	United States	Brown University	47	Brown School of Engineering	Numerical Simulation|Fluid Mechanic	235	3444	141	Particle-wall collisions in a viscous fluid	2294	48142	31
Yaron Hefetz	male	Ph.D.	United States	Brown University	47	Department of Engineering and Physics	Astrophysics|Astronomy & Astrophy	76	2604	12	Detector description and performance for the first coincidence observations between LIGO and GEO	7709	10544	26
Rachel Herz	female	PhD	United States	Brown University	47	Department of Psychiatry and Human Beh	Olfaction|Memory|Autobiographical	75	4971	19	Odor memory: Review and analysis	2975	83117	34
Jin Ho Kim	male	PhD	United States	Brown University	47	Brown School of Engineering	Materials|Nanowires|Nanostructures	64	1283	6	Surface Modification of Poly(dimethylsiloxane) Microchannels	6972	7885	13
John Savage	male	BS, MSc, P	United States	Brown University	47	Department of Computer Science	Theoretical Computer Science|Comp	116	2171	15	Models of Computation: Exploring the Power of Computing	9998	6445	23
Vladan Milnar	male	Doctor of P	United States	Brown University	47	Nanostructures	Nan	37	733	16	Confined states and direction-dependent transmission in graphene quantum wells	3448	6511	14
Marjan Mujib	female	MD, MPH	United States	Brown University	47	Division of Cardiology	Diabetes|Epidemiology|Heart Failure	142	2927	8	Trends in Incidence, Management, and Outcomes of Cardiogenic Shock Complicating ST-Elevation Myo	1324	10043	29
Suchitra Kamle	female	Ph.D.	United States	Brown University	47	Department of Molecular Microbiology and	Antibodies|Transgenes|Therapeutics	44	367	23	Genetically Modified Crops: Detection Strategies and Biosafety Issues.	2403	9244	10
Peter Belenky	male	PhD	United States	Brown University	47	Department of Molecular Microbiology and	Saccharomyces cerevisiae|Biosynth	60	4055	50	NAD metabolism in health and disease	2095	16724	25
Kaustubh Dabhadkar	male	MD MPH	United States	Brown University	47	Division of Cardiology	Clinical Cardiology|Heart Failure|Pop	45	46767	127	Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: A systemat	22776	52129	17
Jaclyn White Hughto	female	PhD, MPH	United States	Brown University	47	Department of Behavioral and Social Scien	Department of Epidemiology	HIV Prevention	143	4749	47	Transgender St	2485	34004
Eui young So	male	PhD	United States	Brown University	47	Department of Orthopaedics	Cancer Biology|Cytokines|Phosphor	64	1078	10	DNA damage induces reactive oxygen species generation through the H2AX-Nox1/Rac1 pathway	5115	4667	17
Michael Koksharov	male	Ph.D.	United States	Brown University	47	Bioinformatics	Nan	41	261	123	Thermostabilization of firefly luciferase by in vivo directed evolution	1340	62474	7
Frederike Hermi Petzschner	female	MSc with H	United States	Brown University	47	Department of Psychiatry and Human Beh	Cognitive Neuroscience|Cognition|M	41	1833	70	Interoception and Mental Health: A Roadmap	1187	18522	14
David Borton	male	PhD Biome	United States	Brown University	47	School of Engineering	Neuroengineering|Brain|Neurophysi	72	2096	40	A Brain–Spinal Interface Alleviating Gait Deficits after Spinal Cord Injury in Primates	1096	18794	19
William Thompson	male	Nan	United States	Brown University	47	Department of Applied Mathematics	R Statistical Package|Statistics|Appl	33	1936	70	Human-mouse genome comparisons to locate regulatory sites	9040	3458	15
Danielle D. Silva	female	BS, MA	United States	Brown University	47	Department of Neuroscience	Magnetic Resonance Imaging|Trans	36	752	36	Pediatric neuroimaging in early childhood and infancy: Challenges and practical guidelines	4350	5693	10
Michael Joshua Frank	male	PhD	United States	Brown University	47	Brown Institute for Brain Science, Dept of	Reinforcement Learning|Cognitive Sc	224	27411	76	By Carrot or by Stick: Cognitive Reinforcement Learning in Parkinsonism	13093	49256	81
Kenneth S Breuer	male	PhD	United States	Brown University	47	Department of Mechanical Engineering	Nan	308	9689	47	Gaseous Slip Flow in Long Microchannels	5016	48040	56
Thomas Webster	male	Nan	United States	Brown University	47	Center for Biomedical Engineering	Nan	422	16959	0	Nan	0	0	0
Mascha van 't Wout	female	PhD	United States	Brown University	47	Department of Psychiatry and Human Beh	Hallucinations|Schizophrenia|Emotio	79	3211	14	Affective state and decision-making in the Ultimatum Game	1639	13159	26
Thomas B Bartnikas	male	MD PhD	United States	Brown University	47	Department of Pathology and Laboratory M	Iron|Hematology|Copper	46	2558	26	The Role of Copper in Neurodegenerative Disease	1228	5620	19
Patricia Markham Risica	female	DrPH	United States	Brown University	47	Department of Behavioral and Social Scien	Public Health|Health Education and P	134	2677	23	How Is Food Insecurity Associated with Dietary Behaviors? An Analysis with Low-Income, Ethnically Di	1222	11946	31
Mark D Bertness	male	Ph.D.	United States	Brown University	47	Department of Ecology and Evolutionary B	Ecology|Invasive Species|Marine Ec	228	30523	131	Positive Interactions in Communities	15934	91711	89
Eric M Morrow	male	PhD, MD	United States	Brown University	47	Department of Molecular Biology, Cell Biol	Developmental Biology|Genetics|Inte	142	12657	73	Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs	5794	35123	42
Brandon D. L. Marshall	male	PhD	United States	Brown University	47	Department of Epidemiology	Epidemiological Surveys|Epidemiolo	397	9444	175	Reduction in overdose mortality after the opening of North America's first medically supervised safer inje	4563	53637	50
Martin M Miner	male	MD	United States	Brown University	47	Division of Urology	Nan	176	4404	8	The Princeton III Consensus Recommendations for the Management of Erectile Dysfunction and Cardio	2046	16977	36
Daniel Moreno De Luca	male	M.D., M.Sc	United States	Brown University	47	Department of Psychiatry and Human Beh	Next Generation Sequencing|Autism	47	6984	39	Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs	2990	15521	25
Brian W Sheldon	male	Nan	United States	Brown University	47	Brown School of Engineering	Nan	260	11717	90	Silicon-Based Nanomaterials for Lithium-Ion Batteries: A Review	6015	38894	56
Suzanne M Colby	female	PhD	United States	Brown University	47	Center for Alcohol and Addiction Studies	Tobacco|Behavioral Psychology|Nic	255	13121	46	Brief Intervention for harm Reduction With Alcohol-Positive Adolescents in a Hospital Emergency Depar	5950	50481	64
Gaurav Choudhary	male	MD	United States	Brown University	47	Department of Medicine	Pulmonary Hypertension|Echocardio	173	2795	24	Association of Borderline Pulmonary Hypertension With Mortality and Hospitalization in a Large Patient	1256	7279	32
J.B. Quintos	male	Professor	United States	Brown University	47	Department of Pediatrics	Diabetes Mellitus|Obesity|Puberty	74	1423	12	Guidelines for Growth Hormone and Insulin-Like Growth Factor-I Treatment in Children and Adolescent	7281	19995	16
Conrad Earl Johanson	male	PhD	United States	Brown University	47	Department of Neurosurgery	Neurodegeneration|Neurodegenerati	225	9293	37	Multiplicity of cerebrospinal fluid functions: new challenges in health and disease. Cerebrospinal Fluid Re	4085	40756	53
Robert W Sobol	male	PhD	United States	Brown University	47	Department of Pathology and Laboratory M	PCR|DNA|DNA Methylation	237	9765	34	Requirement of mammalian DNA polymerse ?? in base-excision repair	4268	23084	52
Wael F Asaad	male	MD, PhD	United States	Brown University	47	Alpert Medical School	Neuroscience|Neurophysiology|Brai	111	4644	34	Selective representation of relevant information by neurons in the primate prefrontal cortex	2241	13008	25
Matthew Jankowich	male	M.D.	United States	Brown University	47	Alpert Medical School	Lung Diseases|Pulmonary Hyperten	70	1056	13	Combined Pulmonary Fibrosis and Emphysema Syndrome A Review	5109	3715	15
Christy Capone	female	Ph.D.	United States	Brown University	47	Department of Psychiatry and Human Beh	Military Psychology|Comorbidity|Pos	39	768	5	Fraternity and Sorority Involvement, Social Influences, and Alcohol Use Among College Students: A Pro	4186	6053	13
Judson A Brewer	male	MD PhD	United States	Brown University	47	School of Public Health	Mindfulness|fMRI|Brain	109	8447	69	Meditation Experience Is Associated with Differences in Default Mode Network Activity and Connectivity	4596	66486	42
Wayne D Bowen	male	Ph.D.	United States	Brown University	47	Department of Molecular Pharmacology, P	Nan	207	11623	6	Sigma receptors: Biology and function	5010	5485	61
Kristina Jackson	female	PhD	United States	Brown University	47	Center for Alcohol and Addiction Studies	Nan	214	6801	40	The Natural History of Problem Gambling From Age 18 to 29	3423	30253	50
Huajian Gao	male	PhD	United States	Brown University	47	School of Engineering	Materials|Nanomaterials|Mechanical	688	52166	441	Indentation size effects in crystalline materials: A law for strain gradient plasticity	27679	182728	112
Robert H Hurt	male	Nan	United States	Brown University	47	Nan	Nan	257	16399	73	Ion Release Kinetics and Particle Persistence in Aqueous Nano-Silver Colloids	8292	38946	66
Stephanie R Jones	female	PhD Mathe	United States	Brown University	47	Department of Neuroscience	Computational Neuroscience|Applied	61	3198	41	A critical role for NMDA receptors in parvalbumin interneurons for gamma rhythm induction and behavio	1608	8589	22
Kali S Thomas	female	PhD, MA	United States	Brown University	47	Department of Health Services, Policy and	Patient Safety|Disaster Preparednes	201	2603	26	The Minimum Data Set 3.0 Cognitive Function Scale	1134	15127	27
Jon A Mukand	male	Nan	United States	Brown University	47	Department of Orthopaedics	Nan	51	3485	2	Neuronal ensemble control of prosthetic devices by a human with tetraplegia	1725	6924	10
David H Barker	male	PhD	United States	Brown University	47	Department of Psychiatry and Human Beh	Adolescent Mental Health|Statistics|	111	2611	23	Prevalence of depression and anxiety in patients with cystic fibrosis and parent caregivers: results of T	1294	13445	25
Katina Robison	female	Doctor of M	United States	Brown University	47	Department of Obstetrics and Gynecology	Tumors|Ovarian Cancer|Cancer	103	1041	5	Comparison of a novel multiple marker assay vs the Risk of Malignancy Index for the prediction of epith	5085	2287	17
Elena K Festa	female	PhD	United States	Brown University	47	Department of Cognitive, Linguistic and Psy	Cognitive Neuropsychology|Neurosc	66	1025	18	A Longitudinal Study of Drivers with Alzheimer Disease	5236	5593	14
David M. Rand	male	BA Harvard	PhD Yale 1987	United States	Brown University	47	Department of Ecology Evolution and	Ecology|Evolu	270	10684	31	Evolution of ne	4995	16366

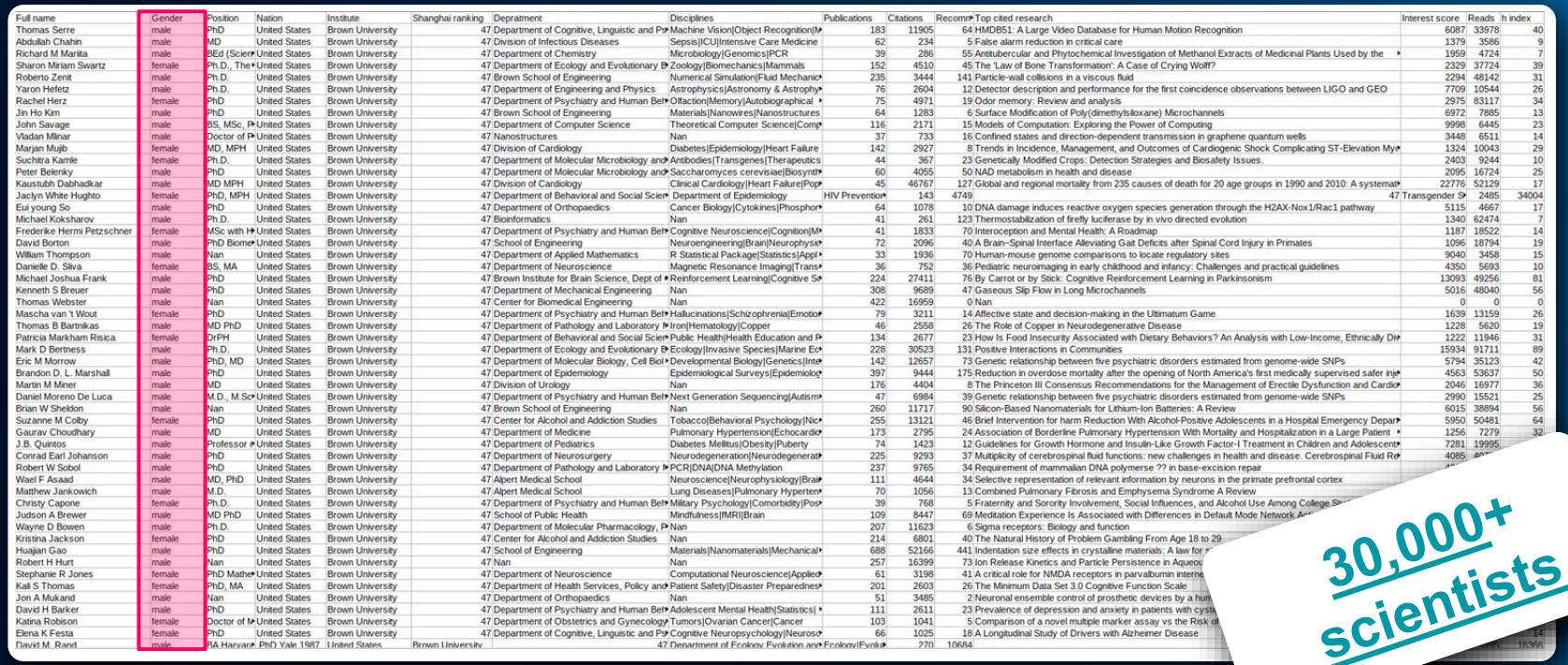
Web crawling - CSV output



Full name	Gender	Position	Nation	Institute	Shanghai ranking	Depratment	Disciplines	Publications	Citations	Recom	Top cited research	Interest score	Reads	h index
Thomas Serre	male	PhD	United States	Brown University	47	Department of Cognitive, Linguistic and Psy	Machine Vision|Object Recognition|M	183	11905	64	HMDB51: A Large Video Database for Human Motion Recognition	6087	33978	40
Abdullah Chahin	male	MD	United States	Brown University	47	Division of Infectious Diseases	Sepsis|ICU|Intensive Care Medicine	62	234	5	False alarm reduction in critical care	1379	3586	9
Richard M Marita	male	BEd (Scien	United States	Brown University	47	Department of Chemistry	Microbiology|Genomics|PCR	39	286	55	Antitubercular and Phytochemical Investigation of Methanol Extracts of Medicinal Plants Used by the	1959	4724	7
Sharon Miriam Swartz	female	Ph.D., The	United States	Brown University	47	Department of Ecology and Evolutionary B	Zoology|Biomechanics|Mammals	152	4510	45	The 'Law of Bone Transformation': A Case of Crying Wolff?	2329	37724	39
Roberto Zenit	male	Ph.D.	United States	Brown University	47	Brown School of Engineering	Numerical Simulation|Fluid Mechanic	235	3444	141	Particle-wall collisions in a viscous fluid	2294	48142	31
Yaron Hefetz	male	Ph.D.	United States	Brown University	47	Department of Engineering and Physics	Astrophysics|Astronomy & Astrophy	76	2604	12	Detector description and performance for the first coincidence observations between LIGO and GEO	7709	10544	26
Rachel Herz	female	PhD	United States	Brown University	47	Department of Psychiatry and Human Beh	Olfaction|Memory|Autobiographical	75	4971	19	Odor memory: Review and analysis	2975	83117	34
Jin Ho Kim	male	PhD	United States	Brown University	47	Brown School of Engineering	Materials|Nanowires|Nanostructures	64	1283	6	Surface Modification of Poly(dimethylsiloxane) Microchannels	6972	7885	13
John Savage	male	BS, MSc, P	United States	Brown University	47	Department of Computer Science	Theoretical Computer Science|Comp	116	2171	15	Models of Computation: Exploring the Power of Computing	9998	6445	23
Vladan Milnar	male	Doctor of P	United States	Brown University	47	Nanostructures	Nan	37	733	16	Confined states and direction-dependent transmission in graphene quantum wells	3448	6511	14
Marjan Mujib	female	MD, MPH	United States	Brown University	47	Division of Cardiology	Diabetes|Epidemiology|Heart Failure	142	2927	8	Trends in Incidence, Management, and Outcomes of Cardiogenic Shock Complicating ST-Elevation Myo	1324	10043	29
Suchitra Kamle	female	Ph.D.	United States	Brown University	47	Department of Molecular Microbiology and	Antibodies|Transgenes|Therapeutics	44	367	23	Genetically Modified Crops: Detection Strategies and Biosafety Issues.	2403	9244	10
Peter Belenky	male	PhD	United States	Brown University	47	Department of Molecular Microbiology and	Saccharomyces cerevisiae|Biosynth	60	4055	50	NAD metabolism in health and disease	2095	16724	25
Kaustubh Dabhadkar	male	MD MPH	United States	Brown University	47	Division of Cardiology	Clinical Cardiology|Heart Failure|Pop	45	46767	127	Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: A systemat	22776	52129	17
Jaclyn White Hughto	female	PhD, MPH	United States	Brown University	47	Department of Behavioral and Social Scien	Department of Epidemiology	HIV Prevention	143	4749	47	Transgender St	2485	34004
Eui young So	male	PhD	United States	Brown University	47	Department of Orthopaedics	Cancer Biology|Cytokines|Phosphor	64	1078	10	DNA damage induces reactive oxygen species generation through the H2AX-Nox1/Rac1 pathway	5115	4667	17
Michael Kokshar ov	male	Ph.D.	United States	Brown University	47	Bioinformatics	Nan	41	261	123	Thermostabilization of firefly luciferase by in vivo directed evolution	1340	62474	7
Frederike Hermi Petzschner	female	MSc with H	United States	Brown University	47	Department of Psychiatry and Human Beh	Cognitive Neuroscience|Cognition|M	41	1833	70	Interoception and Mental Health: A Roadmap	1187	18522	14
David Borton	male	PhD Biome	United States	Brown University	47	School of Engineering	Neuroengineering|Brain|Neurophysi	72	2096	40	A Brain–Spinal Interface Alleviating Gait Deficits after Spinal Cord Injury in Primates	1096	18794	19
William Thompson	male	Nan	United States	Brown University	47	Department of Applied Mathematics	R Statistical Package|Statistics|Appl	33	1936	70	Human-mouse genome comparisons to locate regulatory sites	9040	3458	15
Danielle D. Silva	female	BS, MA	United States	Brown University	47	Department of Neuroscience	Magnetic Resonance Imaging|Trans	36	752	36	Pediatric neuroimaging in early childhood and infancy: Challenges and practical guidelines	4350	5693	10
Michael Joshua Frank	male	PhD	United States	Brown University	47	Brown Institute for Brain Science, Dept of	Reinforcement Learning|Cognitive Sc	224	27411	76	By Carrot or by Stick: Cognitive Reinforcement Learning in Parkinsonism	13093	49256	81
Kenneth S Breuer	male	PhD	United States	Brown University	47	Department of Mechanical Engineering	Nan	308	9689	47	Gaseous Slip Flow in Long Microchannels	5016	48040	56
Thomas Webster	male	Nan	United States	Brown University	47	Center for Biomedical Engineering	Nan	422	16959	0	Nan	0	0	0
Mascha van 't Wout	female	PhD	United States	Brown University	47	Department of Psychiatry and Human Beh	Hallucinations|Schizophrenia|Emotio	79	3211	14	Affective state and decision-making in the Ultimatum Game	1639	13159	26
Thomas B Bartnikas	male	MD PhD	United States	Brown University	47	Department of Pathology and Laboratory M	Iron|Hematology|Copper	46	2558	26	The Role of Copper in Neurodegenerative Disease	1228	5620	19
Patricia Markham Risica	female	DrPH	United States	Brown University	47	Department of Behavioral and Social Scien	Public Health|Health Education and P	134	2677	23	How Is Food Insecurity Associated with Dietary Behaviors? An Analysis with Low-Income, Ethnically Di	1222	11946	31
Mark D Bertness	male	Ph.D.	United States	Brown University	47	Department of Ecology and Evolutionary B	Ecology|Invasive Species|Marine Ec	228	30523	131	Positive Interactions in Communities	15934	91711	89
Eric M Morrow	male	PhD, MD	United States	Brown University	47	Department of Molecular Biology, Cell Biol	Developmental Biology|Genetics|Inte	142	12657	73	Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs	5794	35123	42
Brandon D. L. Marshall	male	PhD	United States	Brown University	47	Department of Epidemiology	Epidemiological Surveys|Epidemiolo	397	9444	175	Reduction in overdose mortality after the opening of North America's first medically supervised safer inje	4563	53637	50
Martin M Miner	male	MD	United States	Brown University	47	Division of Urology	Nan	176	4404	8	The Princeton III Consensus Recommendations for the Management of Erectile Dysfunction and Cardio	2046	16977	36
Daniel Moreno De Luca	male	M.D., M.Sc	United States	Brown University	47	Department of Psychiatry and Human Beh	Next Generation Sequencing|Autism	47	6984	39	Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs	2990	15521	25
Brian W Sheldon	male	Nan	United States	Brown University	47	Brown School of Engineering	Nan	260	11717	90	Silicon-Based Nanomaterials for Lithium-Ion Batteries: A Review	6015	38894	56
Suzanne M Colby	female	PhD	United States	Brown University	47	Center for Alcohol and Addiction Studies	Tobacco|Behavioral Psychology|Nic	255	13121	46	Brief Intervention for harm Reduction With Alcohol-Positive Adolescents in a Hospital Emergency Depar	5950	50481	64
Gaurav Choudhary	male	MD	United States	Brown University	47	Department of Medicine	Pulmonary Hypertension|Echocardio	173	2795	24	Association of Borderline Pulmonary Hypertension With Mortality and Hospitalization in a Large Patient	1256	7279	32
J.B. Quintos	male	Professor	United States	Brown University	47	Department of Pediatrics	Diabetes Mellitus|Obesity|Puberty	74	1423	12	Guidelines for Growth Hormone and Insulin-Like Growth Factor-I Treatment in Children and Adolescent	7281	19995	
Conrad Earl Johanson	male	PhD	United States	Brown University	47	Department of Neurosurgery	Neurodegeneration|Neurodegenerati	225	9293	37	Multiplicity of cerebrospinal fluid functions: new challenges in health and disease. Cerebrospinal Fluid Re	4085		
Robert W Sobol	male	PhD	United States	Brown University	47	Department of Pathology and Laboratory M	PCR|DNA|DNA Methylation	237	9765	34	Requirement of mammalian DNA polymerse ?? in base-excision repair			
Wael F Asaad	male	MD, PhD	United States	Brown University	47	Alpert Medical School	Neuroscience|Neurophysiology|Brai	111	4644	34	Selective representation of relevant information by neurons in the primate prefrontal cortex			
Matthew Jankowich	male	M.D.	United States	Brown University	47	Alpert Medical School	Lung Diseases|Pulmonary Hyperten	70	1056	13	Combined Pulmonary Fibrosis and Emphysema Syndrome A Review			
Christy Capone	female	Ph.D.	United States	Brown University	47	Department of Psychiatry and Human Beh	Military Psychology|Comorbidity|Pos	39	768	5	Fraternity and Sorority Involvement, Social Influences, and Alcohol Use Among College St			
Judson A Brewer	male	MD PhD	United States	Brown University	47	School of Public Health	Mindfulness|fMRI|Brain	109	8447	69	Meditation Experience Is Associated with Differences in Default Mode Network Ac			
Wayne D Bowen	male	Ph.D.	United States	Brown University	47	Department of Molecular Pharmacology, P	Nan	207	11623	6	Sigma receptors: Biology and function			
Kristina Jackson	female	PhD	United States	Brown University	47	Center for Alcohol and Addiction Studies	Nan	214	6801	40	The Natural History of Problem Gambling From Age 18 to 29			
Huajian Gao	male	PhD	United States	Brown University	47	School of Engineering	Materials|Nanomaterials|Mechanical	688	52166	441	Indentation size effects in crystalline materials: A law for			
Robert H Hurt	male	Nan	United States	Brown University	47	Nan	Nan	257	16399	73	Ion Release Kinetics and Particle Persistence in Aqueou			
Stephanie R Jones	female	PhD Mathe	United States	Brown University	47	Department of Neuroscience	Computational Neuroscience|Applied	61	3198	41	A critical role for NMDA receptors in parvalbumin interne			
Kali S Thomas	female	PhD, MA	United States	Brown University	47	Department of Health Services, Policy and	Patient Safety|Disaster Preparednes	201	2603	26	The Minimum Data Set 3.0 Cognitive Function Scale			
Jon A Mukand	male	Nan	United States	Brown University	47	Department of Orthopaedics	Nan	51	3485	2	Neuronal ensemble control of prosthetic devices by a hum			
David H Barker	male	PhD	United States	Brown University	47	Department of Psychiatry and Human Beh	Adolescent Mental Health|Statistics|	111	2611	23	Prevalence of depression and anxiety in patients with cyst			
Katina Robison	female	Doctor of M	United States	Brown University	47	Department of Obstetrics and Gynecology	Tumors|Ovarian Cancer|Cancer	103	1041	5	Comparison of a novel multiple marker assay vs the Risk of			
Elena K Festa	female	PhD	United States	Brown University	47	Department of Cognitive, Linguistic and Psy	Cognitive Neuropsychology|Neurosc	66	1025	18	A Longitudinal Study of Drivers with Alzheimer Disease			
David M. Rand	male	BA Harvard	PhD Yale 1987	United States	Brown University	47	Department of Ecology, Evolution and	Ecology|Evolut	270	10684				16366

30,000+ scientists

Web crawling - CSV output

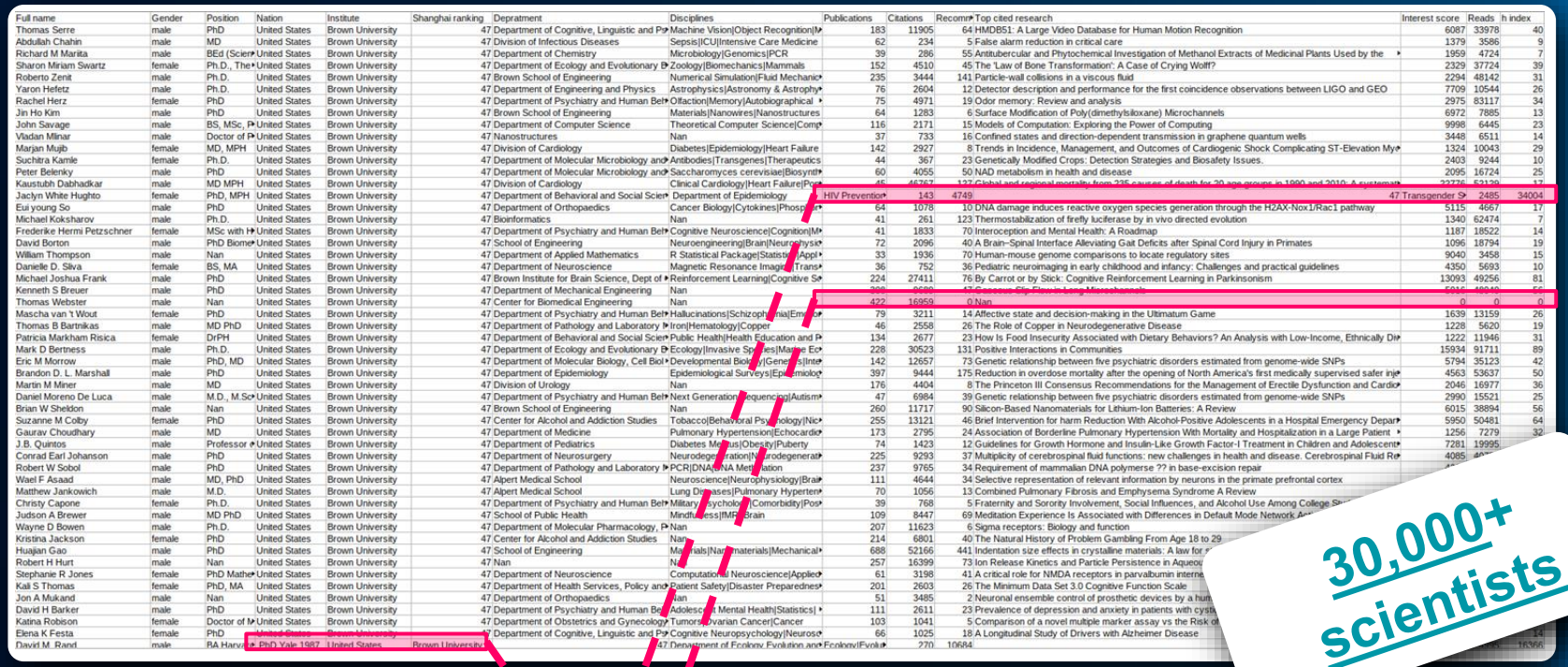


Full name	Gender	Position	Nation	Institute	Shanghai ranking	Depratment	Disciplines	Publications	Citations	Recom	Top cited research	Interest score	Reads	h index
Thomas Serre	male	PhD	United States	Brown University	47	Department of Cognitive, Linguistic and Psy	Machine Vision|Object Recognition|M	183	11905	64	HMDB51: A Large Video Database for Human Motion Recognition	6087	33978	40
Abdullah Chahin	male	MD	United States	Brown University	47	Division of Infectious Diseases	Sepsis|ICU|Intensive Care Medicine	62	234	5	False alarm reduction in critical care	1379	3586	9
Richard M Marita	male	BEd (Scien	United States	Brown University	47	Department of Chemistry	Microbiology|Genomics|PCR	39	286	55	Antitubercular and Phytochemical Investigation of Methanol Extracts of Medicinal Plants Used by the	1959	4724	7
Sharon Miriam Swartz	female	Ph.D., The	United States	Brown University	47	Department of Ecology and Evolutionary Bi	Zoology|Biomechanics|Mammals	152	4510	45	The 'Law of Bone Transformation': A Case of Crying Wolff?	2329	37724	39
Roberto Zenit	male	Ph.D.	United States	Brown University	47	Brown School of Engineering	Numerical Simulation|Fluid Mechanic	235	3444	141	Particle-wall collisions in a viscous fluid	2294	48142	31
Yaron Hefetz	male	Ph.D.	United States	Brown University	47	Department of Engineering and Physics	Astrophysics|Astronomy & Astrophy	76	2604	12	Detector description and performance for the first coincidence observations between LIGO and GEO	7709	10544	26
Rachel Herz	female	PhD	United States	Brown University	47	Department of Psychiatry and Human Beh	Olfaction|Memory|Autobiographical	75	4971	19	Odor memory: Review and analysis	2975	83117	34
Jin Ho Kim	male	PhD	United States	Brown University	47	Brown School of Engineering	Materials|Nanowires|Nanostructures	64	1283	6	Surface Modification of Poly(dimethylsiloxane) Microchannels	6972	7885	13
John Savage	male	BS, MSc, P	United States	Brown University	47	Department of Computer Science	Theoretical Computer Science|Comp	116	2171	15	Models of Computation: Exploring the Power of Computing	9998	6445	23
Vladan Milnar	male	Doctor of P	United States	Brown University	47	Nanostructures	Nan	37	733	16	Confined states and direction-dependent transmission in graphene quantum wells	3448	6511	14
Marjan Mujib	female	MD, MPH	United States	Brown University	47	Division of Cardiology	Diabetes|Epidemiology|Heart Failure	142	2927	8	Trends in Incidence, Management, and Outcomes of Cardiogenic Shock Complicating ST-Elevation Myo	1324	10043	29
Suchitra Kamle	female	Ph.D.	United States	Brown University	47	Department of Molecular Microbiology and	Antibodies|Transgenes|Therapeutics	44	367	23	Genetically Modified Crops: Detection Strategies and Biosafety Issues.	2403	9244	10
Peter Belenky	male	PhD	United States	Brown University	47	Department of Molecular Microbiology and	Saccharomyces cerevisiae|Biosynth	60	4055	50	NAD metabolism in health and disease	2095	16724	25
Kaustubh Dabhadkar	male	MD MPH	United States	Brown University	47	Division of Cardiology	Clinical Cardiology|Heart Failure|Pop	45	46767	127	Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: A systemat	22776	52129	17
Jaclyn White Hughto	female	PhD, MPH	United States	Brown University	47	Department of Behavioral and Social Scien	Department of Epidemiology	HIV Prevention	143	4749	47	Transgender St	2485	34004
Eui young So	male	PhD	United States	Brown University	47	Department of Orthopaedics	Cancer Biology|Cytokines|Phosphor	64	1078	10	DNA damage induces reactive oxygen species generation through the H2AX-Nox1/Rac1 pathway	5115	4667	17
Michael Koksharov	male	Ph.D.	United States	Brown University	47	Bioinformatics	Nan	41	261	123	Thermostabilization of firefly luciferase by in vivo directed evolution	1340	62474	7
Frederike Hermi Petzschner	female	MSc with H	United States	Brown University	47	Department of Psychiatry and Human Beh	Cognitive Neuroscience|Cognition|M	41	1833	70	Interoception and Mental Health: A Roadmap	1187	18522	14
David Borton	male	PhD Biome	United States	Brown University	47	School of Engineering	Neuroengineering|Brain|Neurophysi	72	2096	40	A Brain–Spinal Interface Alleviating Gait Deficits after Spinal Cord Injury in Primates	1096	18794	19
William Thompson	male	Nan	United States	Brown University	47	Department of Applied Mathematics	R Statistical Package|Statistics|Appl	33	1936	70	Human-mouse genome comparisons to locate regulatory sites	9040	3458	15
Danielle D. Silva	female	BS, MA	United States	Brown University	47	Department of Neuroscience	Magnetic Resonance Imaging|Trans	36	752	36	Pediatric neuroimaging in early childhood and infancy: Challenges and practical guidelines	4350	5693	10
Michael Joshua Frank	male	PhD	United States	Brown University	47	Brown Institute for Brain Science, Dept of	Reinforcement Learning|Cognitive Sc	224	27411	76	By Carrot or by Stick: Cognitive Reinforcement Learning in Parkinsonism	13093	49256	81
Kenneth S Breuer	male	PhD	United States	Brown University	47	Department of Mechanical Engineering	Nan	308	9689	47	Gaseous Slip Flow in Long Microchannels	5016	48040	56
Thomas Webster	male	Nan	United States	Brown University	47	Center for Biomedical Engineering	Nan	422	16959	0	Nan	0	0	0
Mascha van 't Wout	female	PhD	United States	Brown University	47	Department of Psychiatry and Human Beh	Hallucinations|Schizophrenia|Emotio	79	3211	14	Affective state and decision-making in the Ultimatum Game	1639	13159	26
Thomas B Bartnikas	male	MD PhD	United States	Brown University	47	Department of Pathology and Laboratory M	Iron|Hematology|Copper	46	2558	26	The Role of Copper in Neurodegenerative Disease	1228	5620	19
Patricia Markham Risica	female	DrPH	United States	Brown University	47	Department of Behavioral and Social Scien	Public Health|Health Education and P	134	2677	23	How Is Food Insecurity Associated with Dietary Behaviors? An Analysis with Low-Income, Ethnically Di	1222	11946	31
Mark D Bertness	male	Ph.D.	United States	Brown University	47	Department of Ecology and Evolutionary Bi	Ecology|Invasive Species|Marine Ec	228	30523	131	Positive Interactions in Communities	15934	91711	89
Eric M Morrow	male	PhD, MD	United States	Brown University	47	Department of Molecular Biology, Cell Biol	Developmental Biology|Genetics|Inte	142	12657	73	Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs	5794	35123	42
Brandon D. L. Marshall	male	PhD	United States	Brown University	47	Department of Epidemiology	Epidemiological Surveys|Epidemiolo	397	9444	175	Reduction in overdose mortality after the opening of North America's first medically supervised safer inje	4563	53637	50
Martin M Miner	male	MD	United States	Brown University	47	Division of Urology	Nan	176	4404	8	The Princeton III Consensus Recommendations for the Management of Erectile Dysfunction and Cardio	2046	16977	36
Daniel Moreno De Luca	male	M.D., M.Sc	United States	Brown University	47	Department of Psychiatry and Human Beh	Next Generation Sequencing|Autism	47	6984	39	Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs	2990	15521	25
Brian W Sheldon	male	Nan	United States	Brown University	47	Brown School of Engineering	Nan	260	11717	90	Silicon-Based Nanomaterials for Lithium-Ion Batteries: A Review	6015	38894	56
Suzanne M Colby	female	PhD	United States	Brown University	47	Center for Alcohol and Addiction Studies	Tobacco|Behavioral Psychology|Nic	255	13121	46	Brief Intervention for harm Reduction With Alcohol-Positive Adolescents in a Hospital Emergency Depar	5950	50481	64
Gaurav Choudhary	male	MD	United States	Brown University	47	Department of Medicine	Pulmonary Hypertension|Echocardio	173	2795	24	Association of Borderline Pulmonary Hypertension With Mortality and Hospitalization in a Large Patient	1256	7279	32
J.B. Quintos	male	Professor	United States	Brown University	47	Department of Pediatrics	Diabetes Mellitus|Obesity|Puberty	74	1423	12	Guidelines for Growth Hormone and Insulin-Like Growth Factor-I Treatment in Children and Adolescent	7281	19995	
Conrad Earl Johanson	male	PhD	United States	Brown University	47	Department of Neurosurgery	Neurodegeneration|Neurodegenerati	225	9293	37	Multiplicity of cerebrospinal fluid functions: new challenges in health and disease. Cerebrospinal Fluid Re	4085		
Robert W Sobol	male	PhD	United States	Brown University	47	Department of Pathology and Laboratory M	PCR|DNA|DNA Methylation	237	9765	34	Requirement of mammalian DNA polymerse ?? in base-excision repair			
Wael F Asaad	male	MD, PhD	United States	Brown University	47	Alpert Medical School	Neuroscience|Neurophysiology|Brai	111	4644	34	Selective representation of relevant information by neurons in the primate prefrontal cortex			
Matthew Jankowich	male	M.D.	United States	Brown University	47	Alpert Medical School	Lung Diseases|Pulmonary Hyperten	70	1056	13	Combined Pulmonary Fibrosis and Emphysema Syndrome A Review			
Christy Capone	female	Ph.D.	United States	Brown University	47	Department of Psychiatry and Human Beh	Military Psychology|Comorbidity|Pos	39	768	5	Fraternity and Sorority Involvement, Social Influences, and Alcohol Use Among College St			
Judson A Brewer	male	MD PhD	United States	Brown University	47	School of Public Health	Mindfulness|fMRI|Brain	109	8447	69	Meditation Experience Is Associated with Differences in Default Mode Network Ac			
Wayne D Bowen	male	Ph.D.	United States	Brown University	47	Department of Molecular Pharmacology, P	Nan	207	11623	6	Sigma receptors: Biology and function			
Kristina Jackson	female	PhD	United States	Brown University	47	Center for Alcohol and Addiction Studies	Nan	214	6801	40	The Natural History of Problem Gambling From Age 18 to 29			
Huajian Gao	male	PhD	United States	Brown University	47	School of Engineering	Materials|Nanomaterials|Mechanical	688	52166	441	Indentation size effects in crystalline materials: A law for			
Robert H Hurt	male	Nan	United States	Brown University	47	Nan	Nan	257	16399	73	Ion Release Kinetics and Particle Persistence in Aqueou			
Stephanie R Jones	female	PhD Mathe	United States	Brown University	47	Department of Neuroscience	Computational Neuroscience|Applied	61	3198	41	A critical role for NMDA receptors in parvalbumin interne			
Kali S Thomas	female	PhD, MA	United States	Brown University	47	Department of Health Services, Policy and	Patient Safety|Disaster Preparednes	201	2603	26	The Minimum Data Set 3.0 Cognitive Function Scale			
Jon A Mukand	male	Nan	United States	Brown University	47	Department of Orthopaedics	Nan	51	3485	2	Neuronal ensemble control of prosthetic devices by a hum			
David H Barker	male	PhD	United States	Brown University	47	Department of Psychiatry and Human Beh	Adolescent Mental Health|Statistics|	111	2611	23	Prevalence of depression and anxiety in patients with cyst			
Katina Robison	female	Doctor of M	United States	Brown University	47	Department of Obstetrics and Gynecology	Tumors|Ovarian Cancer|Cancer	103	1041	5	Comparison of a novel multiple marker assay vs the Risk of			
Elena K Festa	female	PhD	United States	Brown University	47	Department of Cognitive, Linguistic and Psy	Cognitive Neuropsychology|Neurosc	66	1025	18	A Longitudinal Study of Drivers with Alzheimer Disease			
David M. Rand	male	BA Harvard	PhD Yale 1987	United States	Brown University	47	Department of Ecology Evolution and	Ecology|Evolut	270	10684				16366

30,000+ scientists

☞ Female / Male achieved with an API requests and later with names-gender library

Web crawling - CSV output



Full name	Gender	Position	Nation	Institute	Shanghai ranking	Depratment	Disciplines	Publications	Citations	Recom	Top cited research	Interest score	Reads	h index
Thomas Serre	male	PhD	United States	Brown University	47	Department of Cognitive, Linguistic and Psy	Machine Vision|Object Recognition|M	183	11905	64	HMDB51: A Large Video Database for Human Motion Recognition	6087	33978	40
Abdullah Chahin	male	MD	United States	Brown University	47	Division of Infectious Diseases	Sepsis|ICU|Intensive Care Medicine	62	234	5	False alarm reduction in critical care	1379	3586	9
Richard M Marita	male	BEd (Scien	United States	Brown University	47	Department of Chemistry	Microbiology|Genomics|PCR	39	286	55	Antitubercular and Phytochemical Investigation of Methanol Extracts of Medicinal Plants Used by the	1959	4724	7
Sharon Miriam Swartz	female	Ph.D., The	United States	Brown University	47	Department of Ecology and Evolutionary B	Zoology|Biomechanics|Mammals	152	4510	45	The 'Law of Bone Transformation': A Case of Crying Wolff?	2329	37724	39
Roberto Zenit	male	Ph.D.	United States	Brown University	47	Brown School of Engineering	Numerical Simulation|Fluid Mechanic	235	3444	141	Particle-wall collisions in a viscous fluid	2294	48142	31
Yaron Hefetz	male	Ph.D.	United States	Brown University	47	Department of Engineering and Physics	Astrophysics|Astronomy & Astrophy	76	2604	12	Detector description and performance for the first coincidence observations between LIGO and GEO	7709	10544	26
Rachel Herz	female	PhD	United States	Brown University	47	Department of Psychiatry and Human Beh	Olfaction|Memory|Autobiographical	75	4971	19	Odor memory: Review and analysis	2975	83117	34
Jin Ho Kim	male	PhD	United States	Brown University	47	Brown School of Engineering	Materials|Nanowires|Nanostructures	64	1283	6	Surface Modification of Poly(dimethylsiloxane) Microchannels	6972	7885	13
John Savage	male	BS, MSc, P	United States	Brown University	47	Department of Computer Science	Theoretical Computer Science|Comp	116	2171	15	Models of Computation: Exploring the Power of Computing	9998	6445	23
Vladan Milnar	male	Doctor of P	United States	Brown University	47	Nanostructures	Nan	37	733	16	Confined states and direction-dependent transmission in graphene quantum wells	3448	6511	14
Marjan Mujib	female	MD, MPH	United States	Brown University	47	Division of Cardiology	Diabetes|Epidemiology|Heart Failure	142	2927	8	Trends in Incidence, Management, and Outcomes of Cardiogenic Shock Complicating ST-Elevation Myo	1324	10043	29
Suchitra Kamle	female	Ph.D.	United States	Brown University	47	Department of Molecular Microbiology and	Antibodies|Transgenes|Therapeutics	44	367	23	Genetically Modified Crops: Detection Strategies and Biosafety Issues.	2403	9244	10
Peter Belenky	male	PhD	United States	Brown University	47	Department of Molecular Microbiology and	Saccharomyces cerevisiae|Biosynth	60	4055	50	NAD metabolism in health and disease	2095	16724	25
Kaustubh Dabhadkar	male	MD MPH	United States	Brown University	47	Division of Cardiology	Clinical Cardiology|Heart Failure|Pos	45	46767	137	Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: A systemati	23776	52139	17
Jaclyn White Hughto	female	PhD, MPH	United States	Brown University	47	Department of Behavioral and Social Scien	Department of Epidemiology	HIV Prevention	143	4749	47	Transgender St	2485	34004
Eui young So	male	PhD	United States	Brown University	47	Department of Orthopaedics	Cancer Biology|Cytokines|Phosph	64	1078	10	DNA damage induces reactive oxygen species generation through the H2AX-Nox1/Rac1 pathway	5115	4667	17
Michael Koksharov	male	Ph.D.	United States	Brown University	47	Bioinformatics	Nan	41	261	123	Thermostabilization of firefly luciferase by in vivo directed evolution	1340	62474	7
Frederike Hermi Petzschner	female	MSc with H	United States	Brown University	47	Department of Psychiatry and Human Beh	Cognitive Neuroscience|Cognition|M	41	1833	70	Interoception and Mental Health: A Roadmap	1187	18522	14
David Borton	male	PhD Biome	United States	Brown University	47	School of Engineering	Neuroengineering|Brain|Neurophysi	72	2096	40	A Brain–Spinal Interface Alleviating Gait Deficits after Spinal Cord Injury in Primates	1096	18794	19
William Thompson	male	Nan	United States	Brown University	47	Department of Applied Mathematics	R Statistical Package|Statistical Appl	33	1936	70	Human-mouse genome comparisons to locate regulatory sites	9040	3458	15
Danielle D. Silva	female	BS, MA	United States	Brown University	47	Department of Neuroscience	Magnetic Resonance Imaging|Trans	36	752	36	Pediatric neuroimaging in early childhood and infancy: Challenges and practical guidelines	4350	5693	10
Michael Joshua Frank	male	PhD	United States	Brown University	47	Brown Institute for Brain Science, Dept of	Reinforcement Learning|Cognitive Sc	224	27411	76	By Carrot or by Stick: Cognitive Reinforcement Learning in Parkinsonism	13093	49256	81
Kenneth S Breuer	male	PhD	United States	Brown University	47	Department of Mechanical Engineering	Nan	208	8689	47	Crossover Slip Flow in Long Microchannels	5816	10949	56
Thomas Webster	male	Nan	United States	Brown University	47	Center for Biomedical Engineering	Nan	422	16959	0	Nan	0	0	0
Mascha van 't Wout	male	PhD	United States	Brown University	47	Department of Psychiatry and Human Beh	Hallucinations|Schizophrenia|Emo	79	3211	14	Affective state and decision-making in the Ultimatum Game	1639	13159	26
Thomas B Bartnikas	male	MD PhD	United States	Brown University	47	Department of Pathology and Laboratory M	Iron|Hematology|Copper	46	2558	26	The Role of Copper in Neurodegenerative Disease	1228	5620	19
Patricia Markham Risica	female	DrPH	United States	Brown University	47	Department of Behavioral and Social Scien	Public Health|Health Education and P	134	2677	23	How Is Food Insecurity Associated with Dietary Behaviors? An Analysis with Low-Income, Ethnically Div	1222	11946	31
Mark D Bertness	male	Ph.D.	United States	Brown University	47	Department of Ecology and Evolutionary B	Ecology|Invasive Species|Marine Ec	228	30523	131	Positive Interactions in Communities	15934	91711	89
Eric M Morrow	male	PhD, MD	United States	Brown University	47	Department of Molecular Biology, Cell Biol	Developmental Biology|Genetics|Inte	142	12657	73	Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs	5794	35123	42
Brandon D. L. Marshall	male	PhD	United States	Brown University	47	Department of Epidemiology	Epidemiological Surveys|Epidemiolo	397	9444	175	Reduction in overdose mortality after the opening of North America's first medically supervised safer inje	4563	53637	50
Martin M Miner	male	MD	United States	Brown University	47	Division of Urology	Nan	176	4404	8	The Princeton III Consensus Recommendations for the Management of Erectile Dysfunction and Cardio	2046	16977	36
Daniel Moreno De Luca	male	M.D., M.Sc	United States	Brown University	47	Department of Psychiatry and Human Beh	Next Generation Sequencing|Autism	47	6984	39	Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs	2990	15521	25
Brian W Sheldon	male	Nan	United States	Brown University	47	Brown School of Engineering	Nan	260	11717	90	Silicon-Based Nanomaterials for Lithium-Ion Batteries: A Review	6015	38894	56
Suzanne M Colby	female	PhD	United States	Brown University	47	Center for Alcohol and Addiction Studies	Tobacco|Behavioral Psychology|Nic	255	13121	46	Brief Intervention for harm Reduction With Alcohol-Positive Adolescents in a Hospital Emergency Depar	5950	50481	64
Gaurav Choudhary	male	MD	United States	Brown University	47	Department of Medicine	Pulmonary Hypertension|Echocardio	173	2795	24	Association of Borderline Pulmonary Hypertension With Mortality and Hospitalization in a Large Patient	1256	7279	32
J.B. Quintos	male	Professor	United States	Brown University	47	Department of Pediatrics	Diabetes Mellitus|Obesity|Puberty	74	1423	12	Guidelines for Growth Hormone and Insulin-Like Growth Factor-I Treatment in Children and Adolescent	7281	19995	
Conrad Earl Johanson	male	PhD	United States	Brown University	47	Department of Neurosurgery	Neurodegeneration|Neurodegenerati	225	9293	37	Multiplicity of cerebrospinal fluid functions: new challenges in health and disease. Cerebrospinal Fluid Re	4085		
Robert W Sobol	male	PhD	United States	Brown University	47	Department of Pathology and Laboratory M	PCR|DNA|DNA Mei ation	237	9765	34	Requirement of mammalian DNA polymerse ?? in base-excision repair			
Wael F Asaad	male	MD, PhD	United States	Brown University	47	Alpert Medical School	Neuroscience|Neurophysiology|Brai	111	4644	34	Selective representation of relevant information by neurons in the primate prefrontal cortex			
Matthew Jankowich	male	M.D.	United States	Brown University	47	Alpert Medical School	Lung Diseases|Pulmonary Hyperten	70	1056	13	Combined Pulmonary Fibrosis and Emphysema Syndrome A Review			
Christy Capone	female	Ph.D.	United States	Brown University	47	Department of Psychiatry and Human Beh	Military Psychology|Comorbidity|Pos	39	768	5	Fraternity and Sorority Involvement, Social Influences, and Alcohol Use Among College St			
Judson A Brewer	male	MD PhD	United States	Brown University	47	School of Public Health	Mindfulness|fMRI|Brain	109	8447	69	Meditation Experience Is Associated with Differences in Default Mode Network Ac			
Wayne D Bowen	male	Ph.D.	United States	Brown University	47	Department of Molecular Pharmacology, P	Nan	207	11623	6	Sigma receptors: Biology and function			
Kristina Jackson	female	PhD	United States	Brown University	47	Center for Alcohol and Addiction Studies	Nan	214	6801	40	The Natural History of Problem Gambling From Age 18 to 29			
Huajian Gao	male	PhD	United States	Brown University	47	School of Engineering	Materials|Nanomaterials|Mechanical	688	52166	441	Indentation size effects in crystalline materials: A law for			
Robert H Hurt	male	Nan	United States	Brown University	47	Nan	Na	257	16399	73	Ion Release Kinetics and Particle Persistence in Aqueou			
Stephanie R Jones	female	PhD Mathe	United States	Brown University	47	Department of Neuroscience	Computational Neuroscience|Applied	61	3198	41	A critical role for NMDA receptors in parvalbumin interne			
Kali S Thomas	female	PhD, MA	United States	Brown University	47	Department of Health Services, Policy and	Patient Safety|Disaster Preparednes	201	2603	26	The Minimum Data Set 3.0 Cognitive Function Scale			
Jon A Mukand	male	Nan	United States	Brown University	47	Department of Orthopaedics	Nan	51	3485	2	Neuronal ensemble control of prosthetic devices by a hum			
David H Barker	male	PhD	United States	Brown University	47	Department of Psychiatry and Human Beh	Adolescent Mental Health|Statistics|	111	2611	23	Prevalence of depression and anxiety in patients with cyst			
Katina Robison	female	Doctor of M	United States	Brown University	47	Department of Obstetrics and Gynecology	Tumor|Ovarian Cancer|Cancer	103	1041	5	Comparison of a novel multiple marker assay vs the Risk of			
Elena K Festa	female	PhD	United States	Brown University	47	Department of Cognitive, Linguistic and Psy	Cognitive Neuropsychology|Neurosc	66	1025	18	A Longitudinal Study of Drivers with Alzheimer Disease			
David M. Rand	male	BA Harvd	PhD Yale 1987	United States	Brown University	47	Department of Ecology Evolution and	Ecology|Evolu	270	10684				16366

30,000+ scientists

☞ Flaws

Data Cleansing

Data Cleansing

“the process of detecting and correcting corrupt or inaccurate records from a database and refers to identifying incomplete, incorrect, inaccurate or irrelevant parts of the data and then replacing, modifying, or deleting the dirty or coarse data.” ~wikipedia

We had several kinds of flaws. First, we simply scrolled down the data-frame and categorizing the most obvious.

Later, we used advanced techniques such as boxplots, searching for outliers.

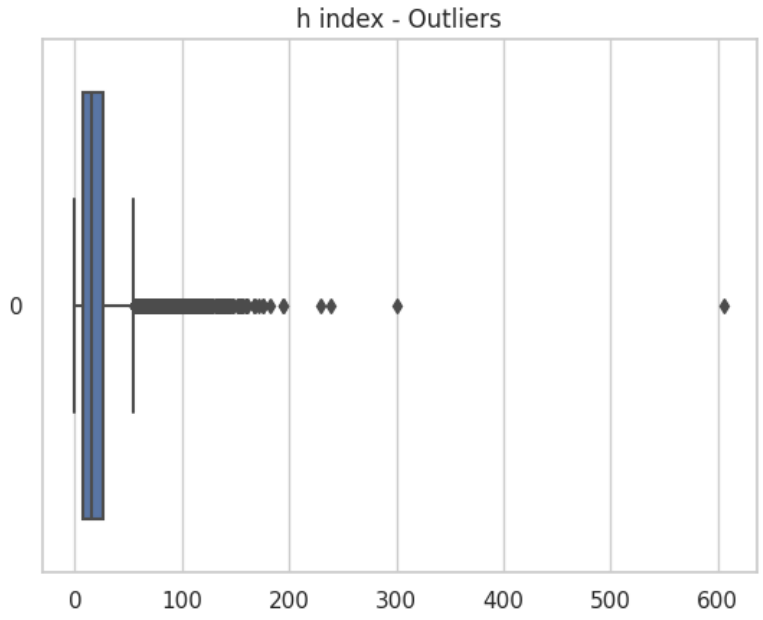
Data Cleansing cont.

Obvious data we chosen to remove:

- Rows with more than the expected columns
- Rows with strings within numeric cells
  - E.g a word instead of digits within *citations*
- Non scientist
  - 0 publications, 0 citations, 0 reads...
- Couldn't determine gender
  - Names that wer'ent featured in the API
- Duplicates

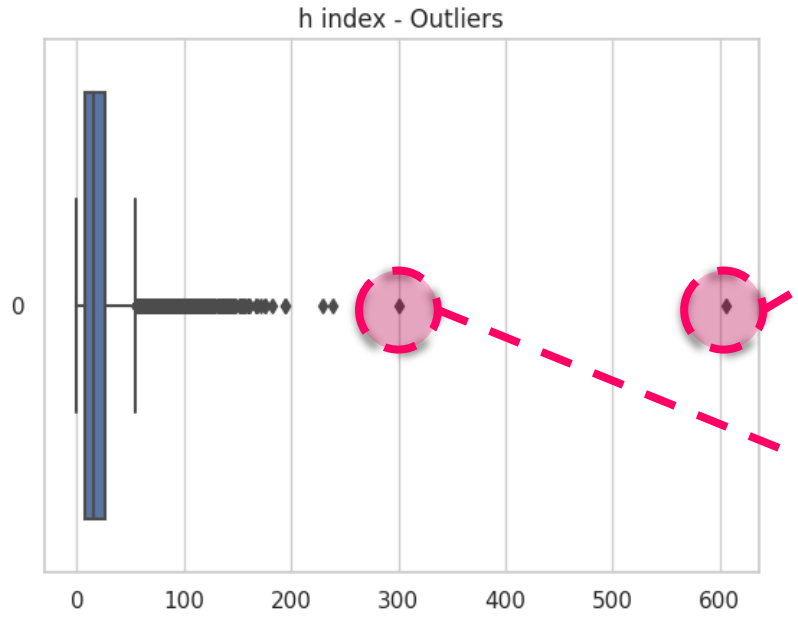
Corrupts and flaws

Data Cleansing cont.



Outliers

Data Cleansing cont.



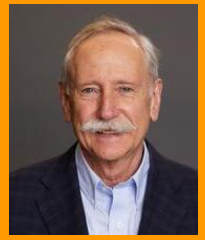
Outliers

Dr. Cynthia Edenburg

האוניברסיטה הפתוחה
THE OPEN UNIVERSITY OF ISRAEL
الجامعة المفتوحة

Actual h index: 3

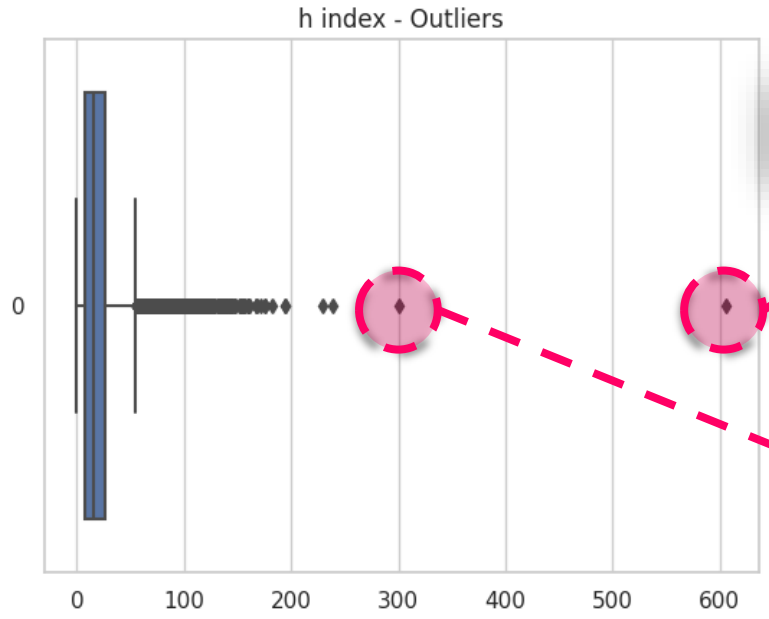
Prof. Walter C Willett



DIVISION OF NUTRITION
HARVARD MEDICAL SCHOOL

Actual h index: 300

Data Cleansing cont.

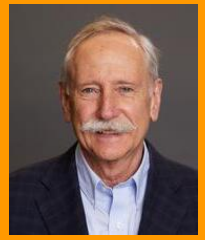


Outliers

```
df.drop(df[df["h index"]>600].index, inplace = True)
```

Actual h index: 3

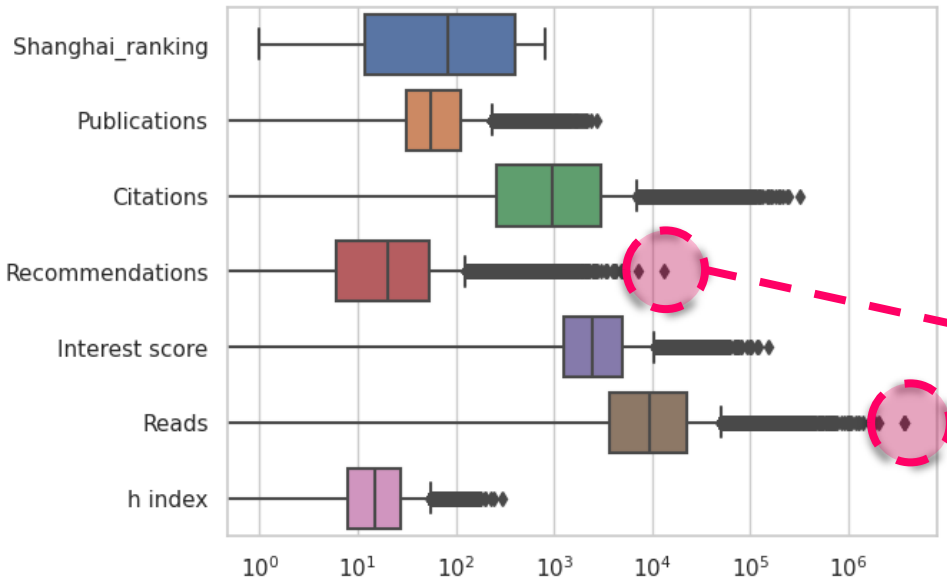
Prof. Walter C Willett



DIVISION OF NUTRITION
HARVARD MEDICAL SCHOOL

Actual h index: 300

Data Cleansing cont.



Outliers

☞ Not outliers.
Checked manually

☞ X axis was scaled logarithmically
for better understanding

Exploratory Data Analysis

EDA

“An approach of analyzing data sets to summarize their main characteristics, often using statistical graphics and other data visualization methods.” ~wikipedia

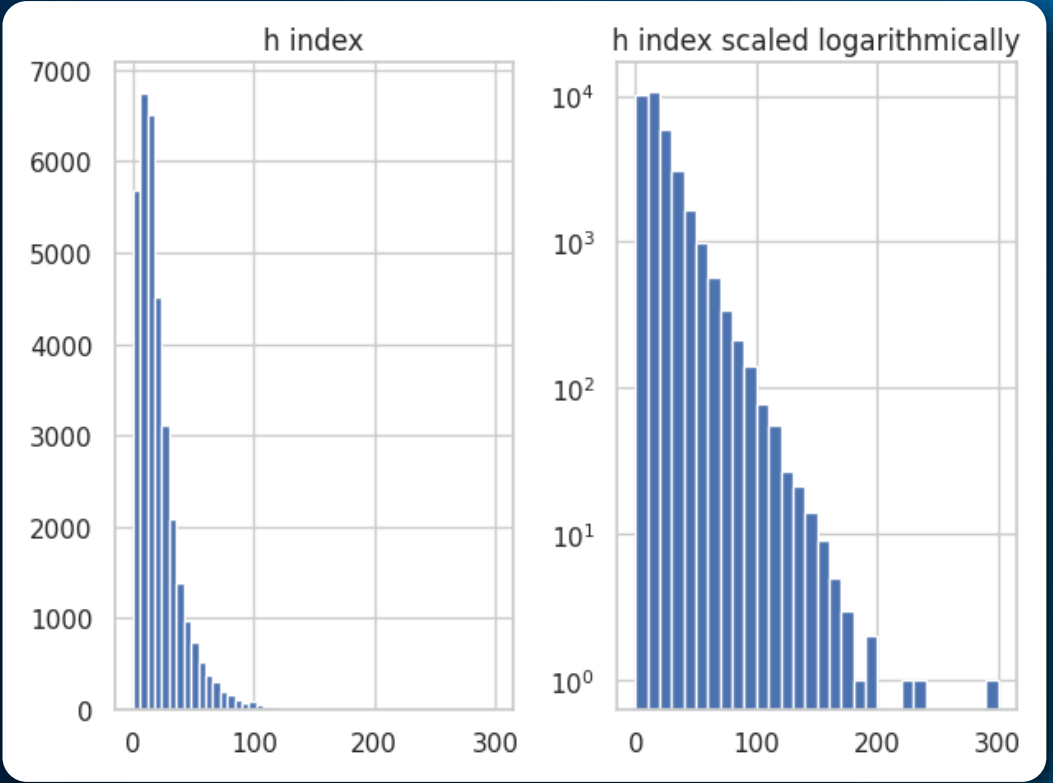
In this section we will discuss the relations between the features we collected and their impact on the h index score.

EDA cont.

After scaling the h index logarithmically, the decrease appeared to be linear.

We drew from this that the higher the h index is there are exponentially less people with that score.

Bars plot



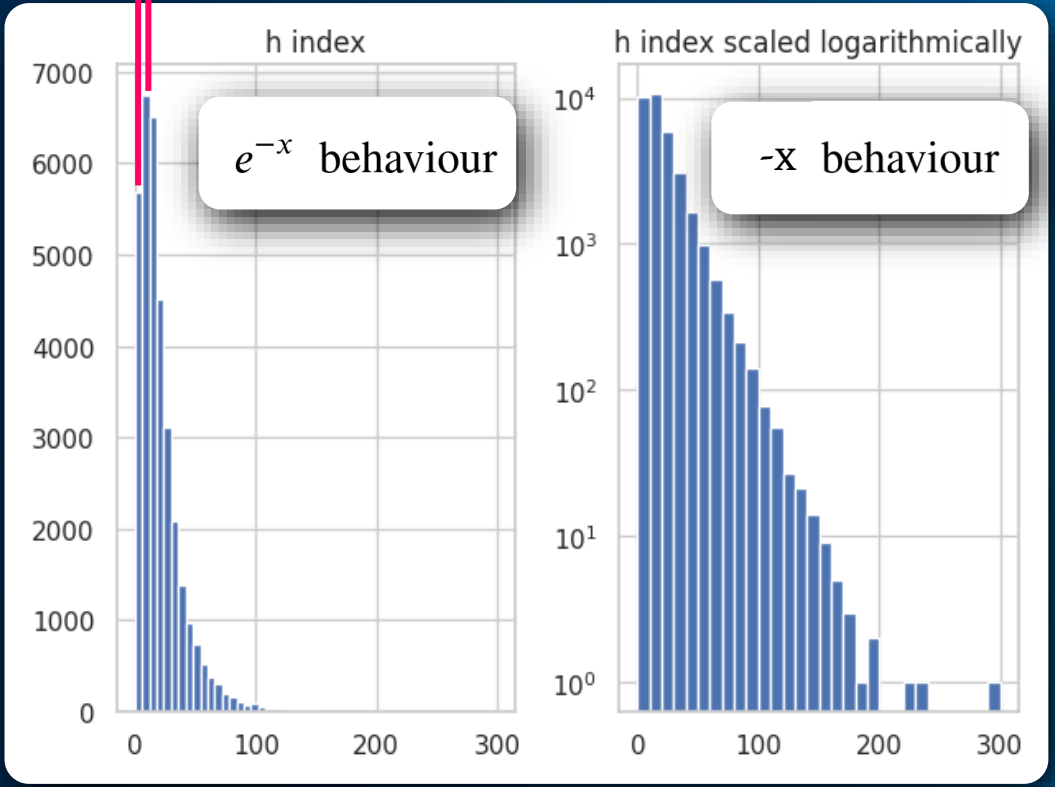
☞ Y axis was scaled logarithmically

EDA cont.

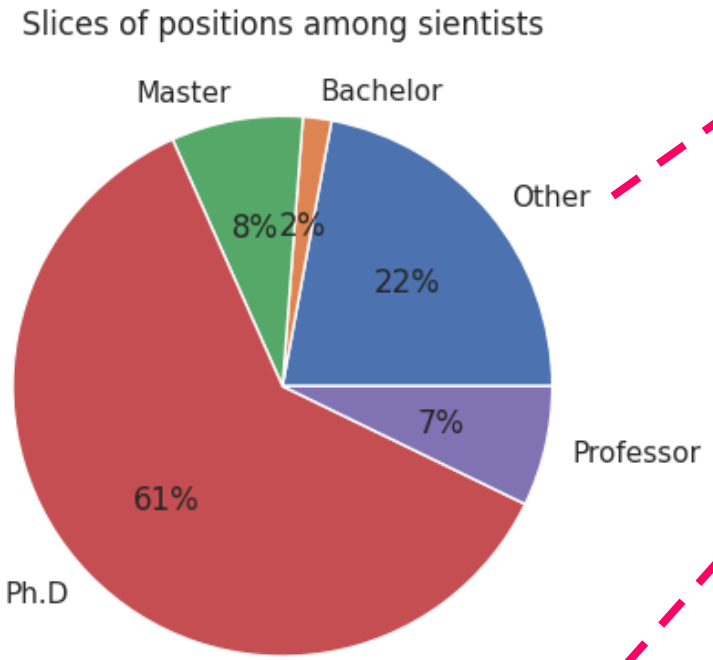
After scaling the h index logarithmically, the decrease appeared to be linear.

We drew from this that the higher the h index is there are exponentially less people with that score.

☞ Missing data because we had collected top 1000 h index scientists of each institute + many among the low h index were removed in the cleansing part.



EDA cont.



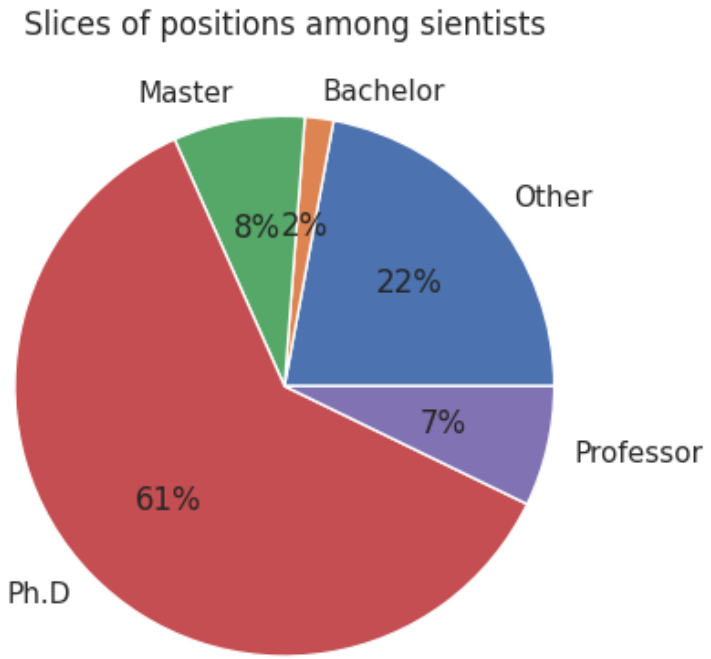
We added *Position-rank* categorical column ; professors received 4, Ph.D 3, Masters 2, Bachelors 1 and else 0.

☞ Weren't declaring their position or undergraduates.

```
def position_to_rank(position):
    try:
        position = position.replace(",", "").replace(".", "").replace("_", "").replace("-", "")
        position = position.lower()
    except:
        return 0

    if position.find("prof") != -1:
        return 4
    elif position.find("ph") != -1 or position.find("doc") != -1 or position.find("ds") != -1:
        return 3
    elif position.find("md") != -1 or position.find("ma") != -1 or position.find("ms") != -1:
        return 2
    elif position.find("ba") != -1 or position.find("bs") != -1 or position.find("be") != -1:
        return 1
    else:
        return 0
```

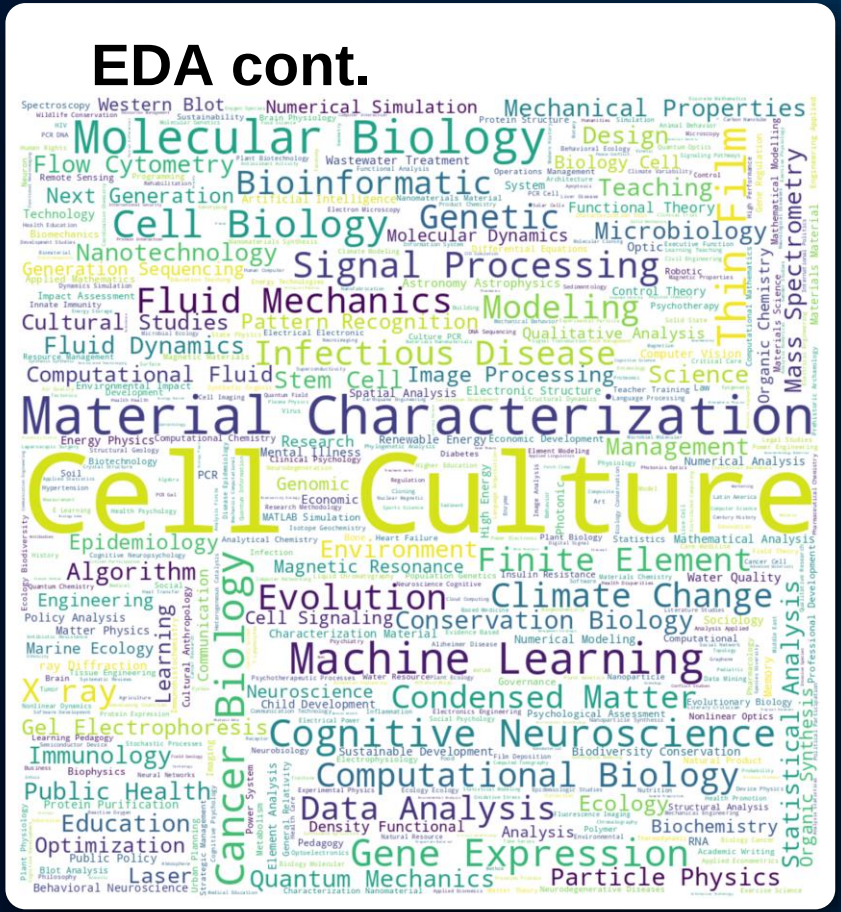
EDA cont.



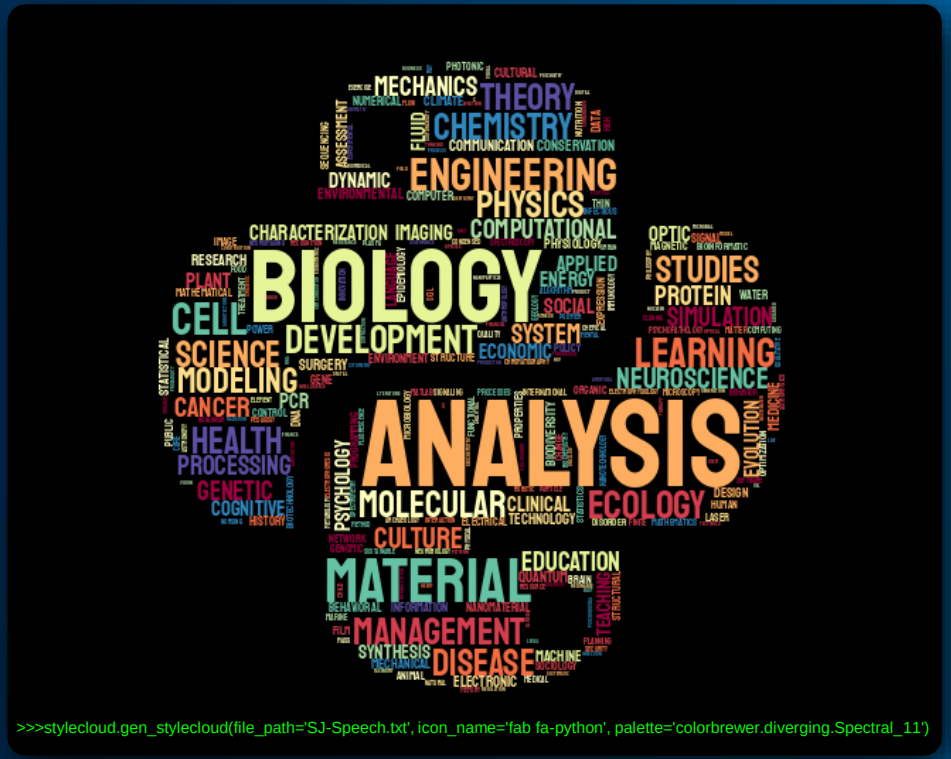
Assumptions

- Naturally there are much less professors than Ph.Ds.
- Professors may retired very often due to their age, or may be less ambitious than young Ph.Ds.
- Professors might be much more busy in lecturing, pedagogy and other activities rather than in researches.
- People with Bachelor degreed represent only 2 percents of researchers. It does make sense, because naturally the may continue to their Master degree or leave the academic environment, rather than being involved in researches.
- Most Master programs require conducting researches and submitting a thesis at the end of the program.

EDA cont.



Words cloud - Disciplines




```
>>>stylecloud.gen_stylecloud(file_path='SJ-Speech.txt', icon_name='fab fa-python', palette='colorbrewer.diverging.Spectral_11')
```

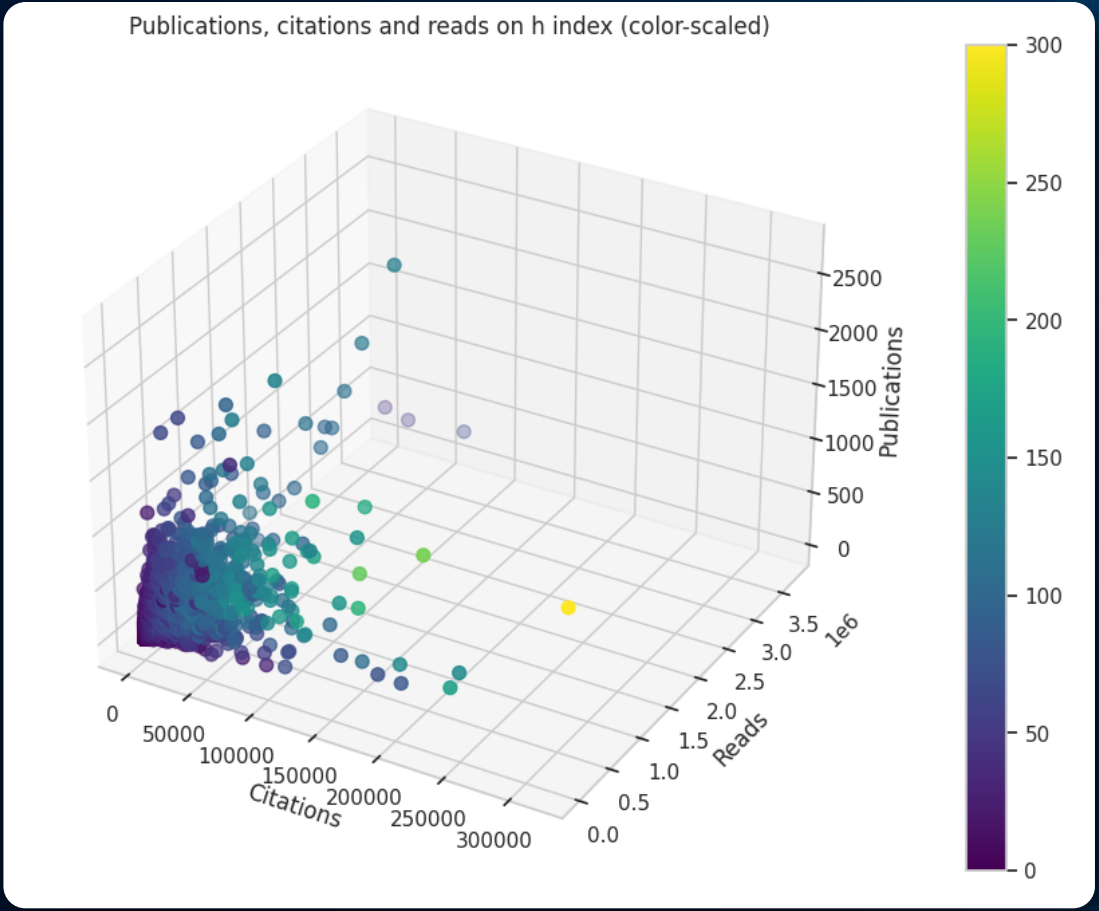
EDA cont.

Scatter plots

Publications and citation are the main and only factors to calculate the h index and indeed there is a strong correlation between them.

```
df.corr()["h index"].sort_values()
```

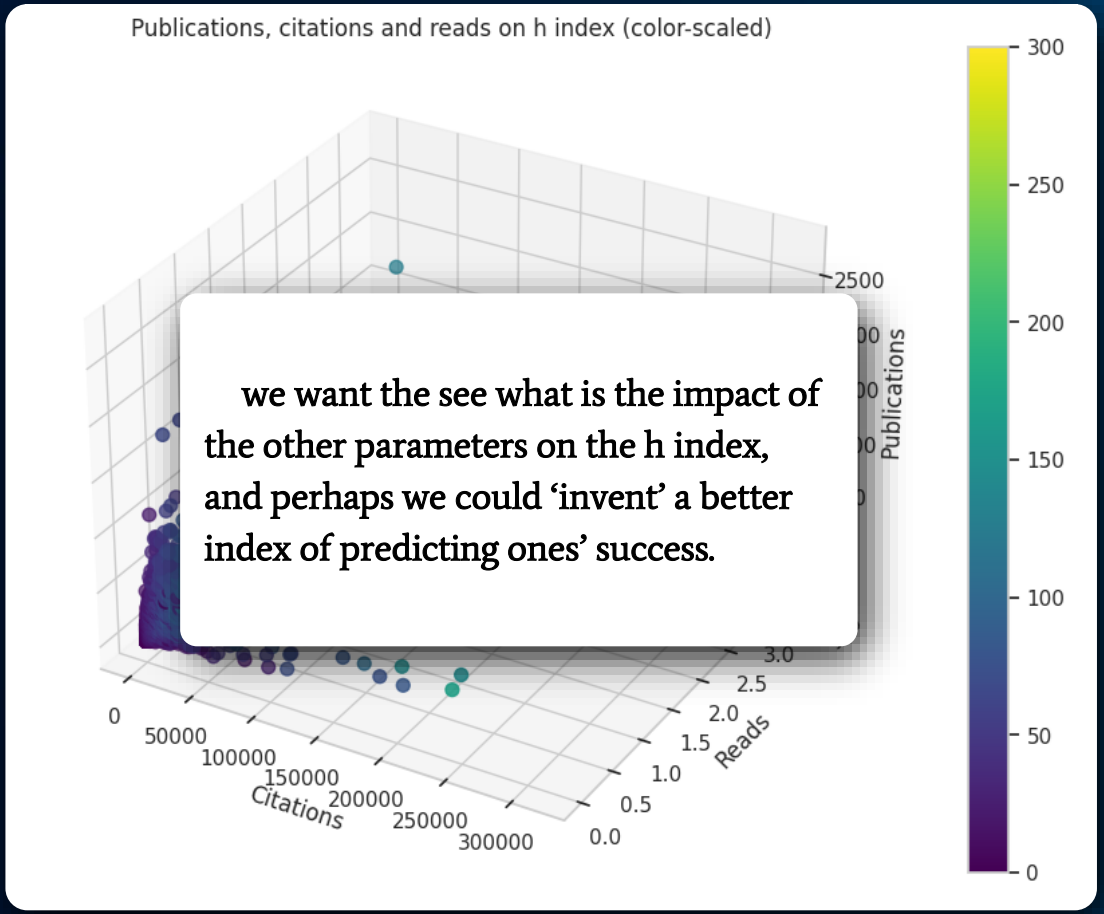
```
Shanghai_ranking   -0.301562
PISA rank          -0.289788
Position rank       0.053868
HDI score           0.173123
GDP (Millions $)    0.245993
Recommendations     0.296905
Reads               0.382567
Interest_score      0.643960
Citations           0.791692
Publications        0.798109
h index             1.000000
Name: h index, dtype: float64
```



Scatter plots

We added the '*Reads*' factor and plot in 4D.

- some scientists have very high h index although they published 'few' papers. That is probably because there are very few papers that reached widespread exposure.
- Number of reads and h-index has moderate correlation as well.



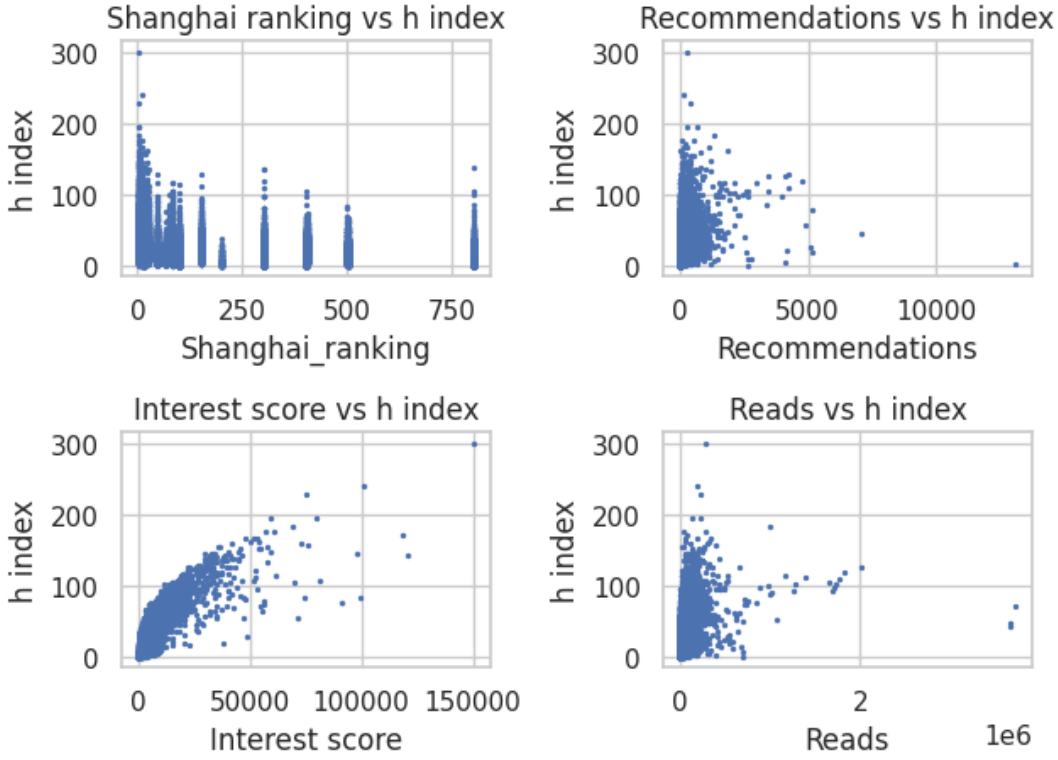
we want the see what is the impact of the other parameters on the h index, and perhaps we could 'invent' a better index of predicting ones' success.

Scatter plots

We added the '*Reads*' factor and plot in 4D.

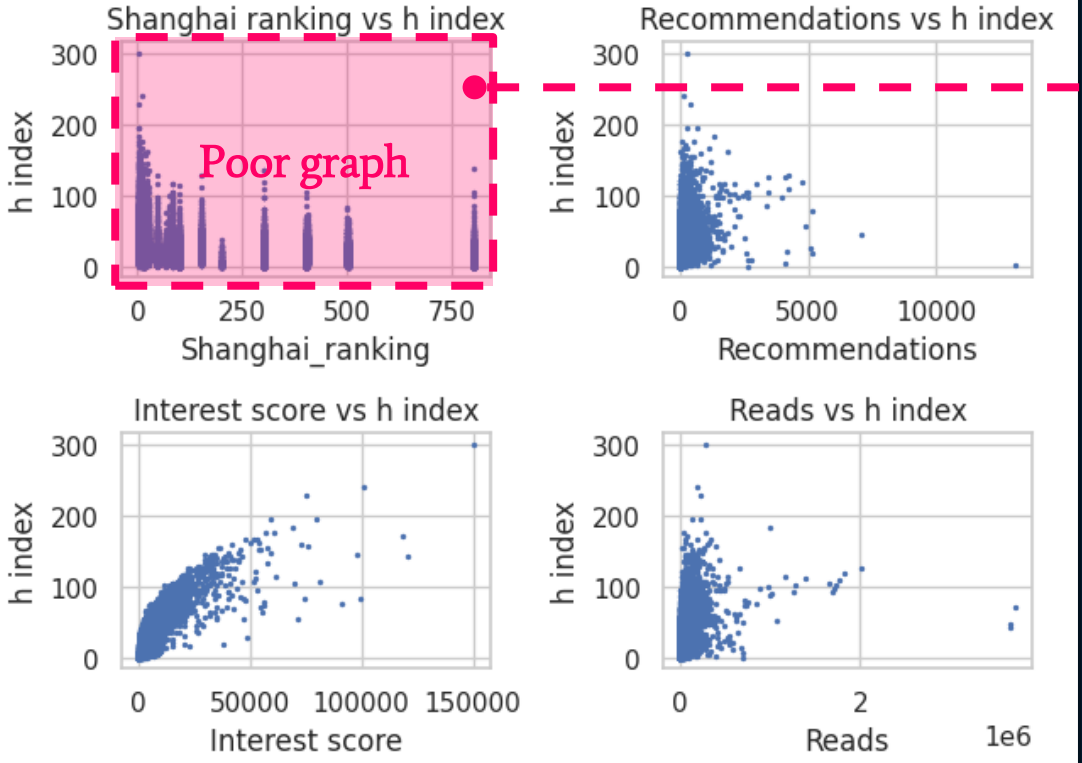
- some scientists have very high h index although they published 'few' papers. That is probably because there are very few papers that reached widespread exposure.
- Number of reads and h-index has moderate correlation as well.

EDA cont.



Scatter plots

EDA cont.

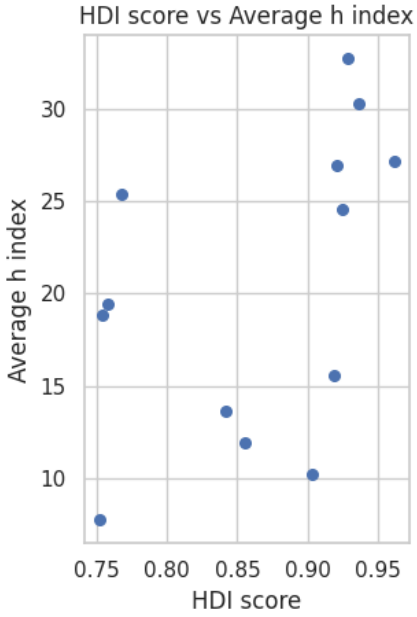


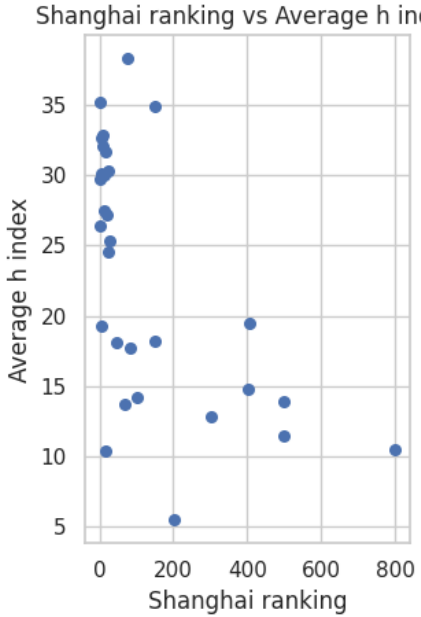
Scatter plots

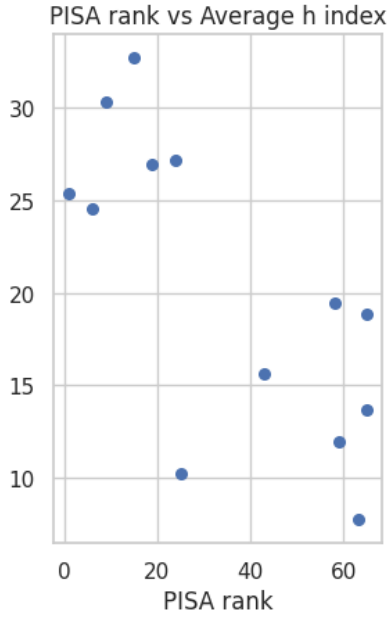
Instead, we'll calculate the average h index per Shanghai ranking, and late the average h index for each nation.

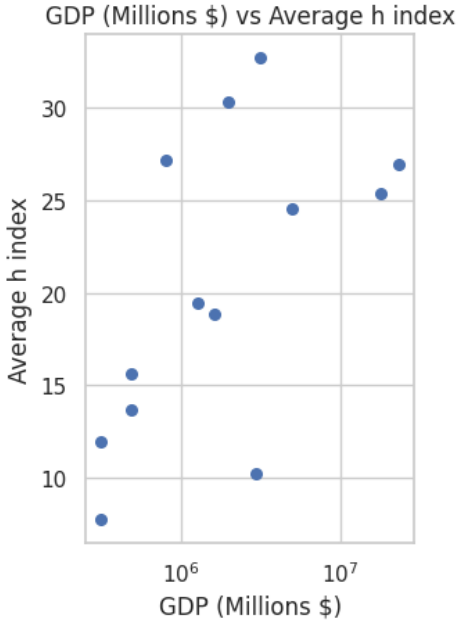
We'll add three more columns:

- GDP of a nation in M$
- HDI score of a nation
- PISA rank of a nation







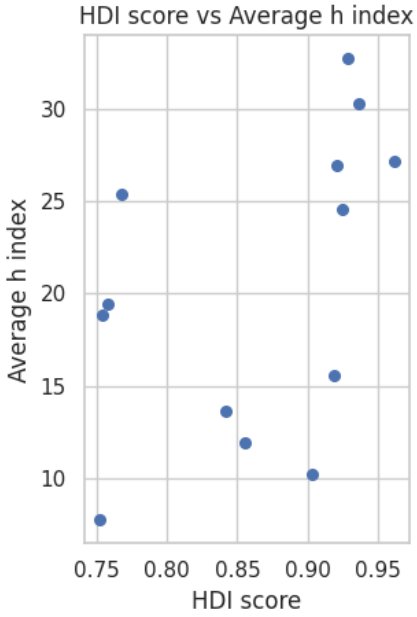


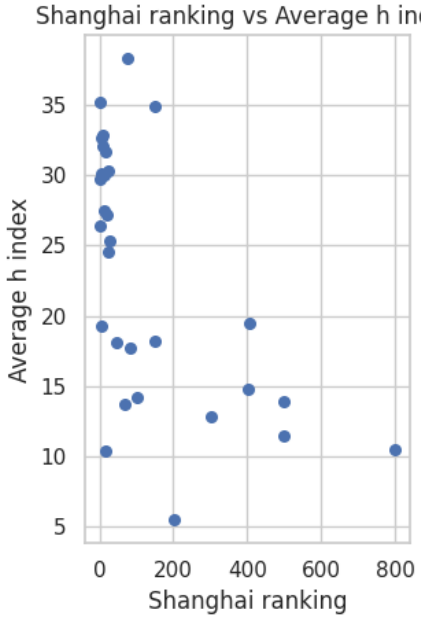
R-squared = 0.213
Pearson correlation coefficient = 0.462

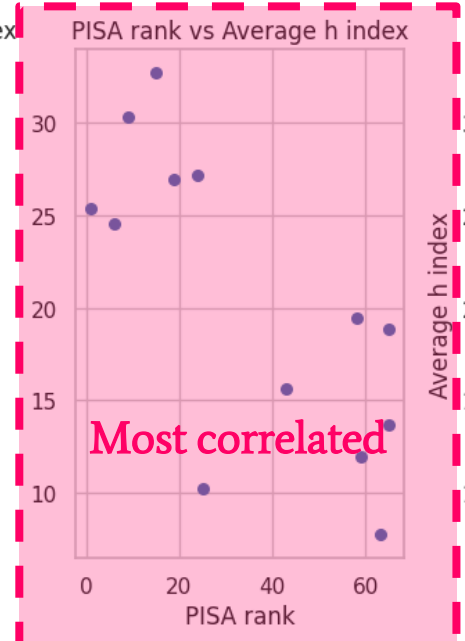
R-squared = 0.343
Pearson correlation coefficient = -0.586

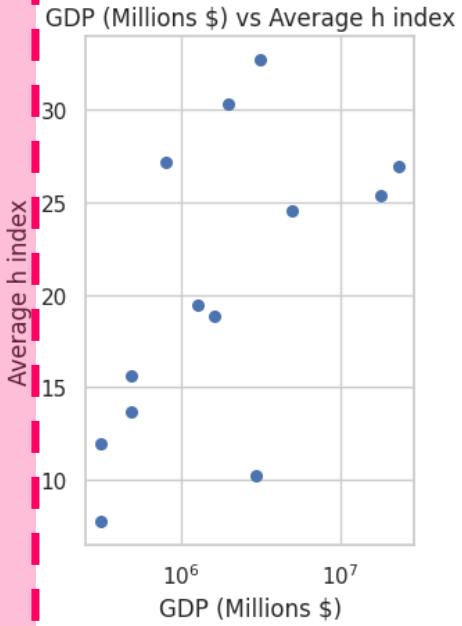
R-squared = 0.524
Pearson correlation coefficient = -0.724

R-squared = 0.157
Pearson correlation coefficient = 0.396









R-squared = 0.213
Pearson correlation coefficient = 0.462

R-squared = 0.343
Pearson correlation coefficient = -0.586

☞ It doesn't matter much where you study. Prestige has a low impact.

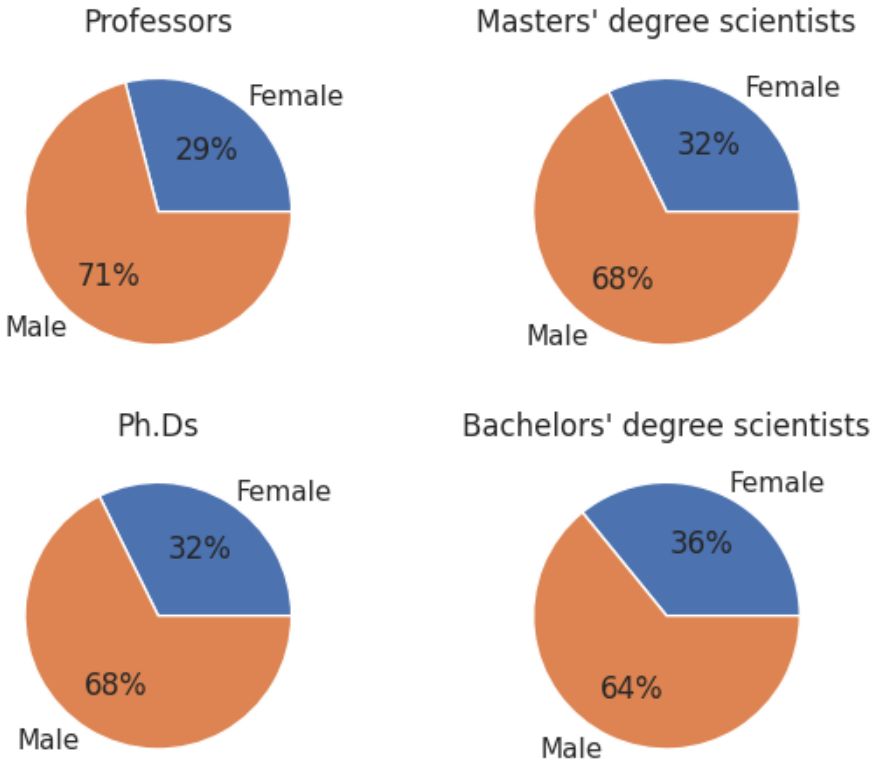
R-squared = 0.524
Pearson correlation coefficient = -0.724

☞ Primary education has a relatively strong correlation to academic success.

R-squared = 0.157
Pearson correlation coefficient = 0.396

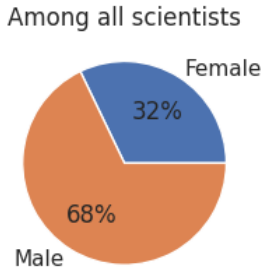
☞ Cash doesn't produce good scientists.

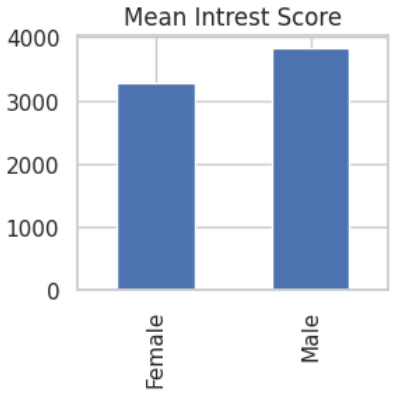
EDA cont.

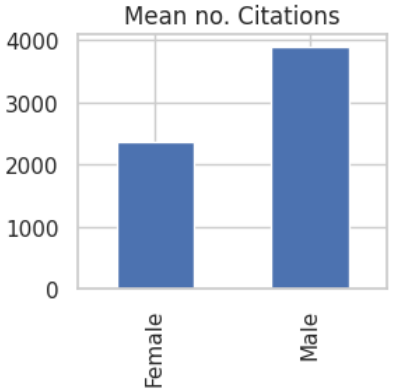


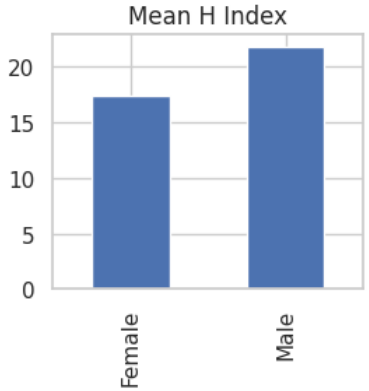
Gender biases

EDA cont.









Gender biases

Unfortunately, we witness that women are still underrepresented in many scientific fields in academia, especially in higher positions.

This is due to factors such as implicit bias, gender stereotypes, and lack of support.

Academic institutions and the scientific community need to work together to promote diversity, provide equal opportunities, and encourage young women to pursue science careers.

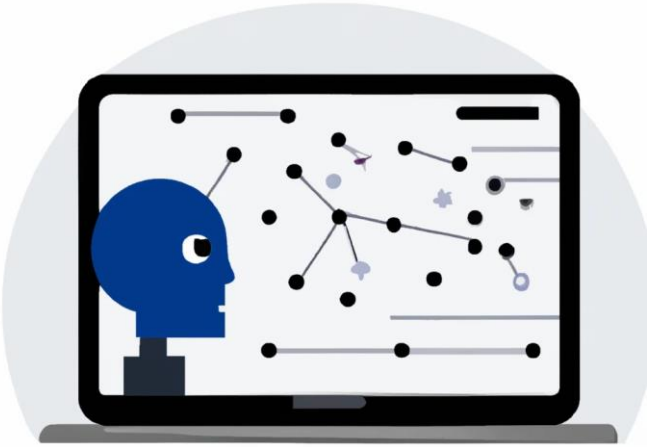
Machine learning models

Machine learning models

“A field of inquiry devoted to understanding and building methods that "learn" – that is, methods that leverage data to improve performance on some set of tasks.” ~wikipedia

We removed all the columns which are irrelevant for the model:

- Gender -> convert to Female to 0 and Male to 1
- remove 'Full name'
- remove 'Position' (there’s ‘*position rank*’)
- remove 'Nation'
- remove 'Disciplines'
- remove 'Department'
- remove 'Institute' (only in linear regression)
- remove 'Top cited research'



Machine learning cont.

Models we were tested two models:

- Linear regression
- Random Forest

Other models s.a KNN or K-means are less relevant because that the predict of a h index is a regression problem and not a categorical one.

For the models we used the sklearn library

Linear Regression

- Linear Regression is a statistical method that studies the relationship between different variables on data.
- Since h index is given based on the number of publications and how many times each was cited and it being on a continuous scale it's appropriate to use linear regression for our problem.
- We split the dataset to train and test sets, removed all non continuous/ numeric features and scaled it using Min-Max scaler.
- Our model yielded R2 score of over 78 percents

Random Forest

- Decision trees are a type of supervised machine learning algorithm used for classification and regression analysis. They use a tree-like model of decisions and their possible consequences to make predictions or classifications.
- Random Forest is a machine learning algorithm that combines multiple decision trees to make better predictions.
- In a decision tree, the data is split based on a single feature at each node. However, a Random Forest uses multiple decision trees and randomly selects a subset of features for each tree. By doing so, it reduces overfitting and increases the diversity of the trees in the forest, which leads to better generalization performance.
- When we performed this algorithm on our dataset with 10 estimators (trees in forest) which led to very poor fitting with accuracy score of 17 percents .
- After removing some features and we witnessed a drastic increase in the performance of our model , resulting in accuracy score of over 88 percents.
- Finally, we did a grid search (`GridSearchCV`) with different parameters to find an optimal combination resulting in over 93 percents fitting , 5.45 SSE score and R2 score of 98 percents

PCA - Principal component analysis

"*Technique for analyzing large datasets containing a high number of dimensions/features per observation, increasing the interpretability of data while preserving the maximum amount of information.*" ~wikipedia

We will try the use of PCA in order to increase the accuracy score of the linear regression from 78%.

Improvements attempts

PCA cont.

We reduced the components to 4 and a receive negative R-squared of -401 which is clearly not improving the linear regression model.

Improvements attempts

```
from sklearn.decomposition import PCA
pca = PCA(n_components=4)
X_train_pca = pca.fit_transform(X_train_numeric)
X_test_pca = pca.fit_transform(X_test_numeric)
X_pca_train_scaled = scaler.fit_transform(X_train_pca)
X_pca_test_scaled  = scaler.transform(X_test_pca)
X_pca_train_scaled
```

```
array([[1.22468017e-01, 2.70964071e-02, 4.61078565e-01, 3.05014256e-01],
       [9.99966551e-01, 2.92236972e-03, 4.63185396e-01, 3.48512198e-01],
       [1.14924484e-01, 8.91882107e-03, 4.71238739e-01, 3.76149163e-01],
       ...,
       [7.56769308e-03, 6.65245864e-03, 4.67337072e-01, 3.41693847e-01],
       [9.99966467e-01, 9.94378470e-04, 4.58407391e-01, 3.05618076e-01],
       [9.99966586e-01, 3.50022784e-03, 4.67312446e-01, 3.84817307e-01]])
```

```
clf_pca = sklearn.linear_model.LinearRegression()
clf_pca.fit(X_pca_train_scaled,y_trainl)
```

```
LinearRegression()
```

In a Jupyter environment, please rerun this cell to show the HTML representation or trust the n
On GitHub, the HTML representation is unable to render, please try loading this page with nbvie

```
y_pca_predict = clf_pca.predict(pca.fit_transform(X_test_scaled))
print(f"SSE = {mean_squared_error(y_testl , np.floor(y_pca_predict))}  " )
print(f"R2 = {r2_score(y_testl , np.floor(y_pca_predict))}")
```

```
SSE = 143937.80963112324
R2 = -401.21479530891526
```

Reducing columns

We will try another attempt, manually removing columns which has lower correlation the the h index:

- GDP
- HDI score
- Shanghai_ranking
- Reads
- PISA rank

Improvements attempts

```
y_predictl_new = clf_new.predict(X_test_scaled.drop(['GDP (Millions $)',
                                                     'HDI score',
                                                     'Shanghai_ranking',
                                                     'Reads',
                                                     'PISA rank'],axis=1))

print(f"SSE = {mean_squared_error(y_testl , np.floor(y_predictl_new))}  " )
print(f"R2 = {r2_score(y_testl , np.floor(y_predictl_new))}")
```

```
SSE = 84.91804056458308
R2 = 0.7627079890179692
```

Reducing columns

We will try another attempt, manually removing columns which has lower correlation the the h index:

- GDP
- HDI score
- Shanghai_ranking
- Reads
- PISA rank

Improvements attempts

```
y_predictl_new = clf_new.predict(X_test_scaled.drop(['GDP (Millions $)',
                                                     'HDI score',
                                                     'Shanghai_ranking',
                                                     'Reads',
                                                     'PISA rank'],axis=1))

print(f"SSE = {mean_squared_error(y_testl , np.floor(y_predictl_new))}  " )
print(f"R2 = {r2_score(y_testl , np.floor(y_predictl_new))}")
```

```
SSE = 84.91804056458308
R2 = 0.7627079890179692
```

☞ It worsen our accuracy score from 0.78 to 0.76. Not an improvement.

Demonstration

We will send our great lecturer Dr. Jonathan Schler into each of the model to see how his h index would be predicted on each.

We made sure Dr. Schler isn't in the training set. We create a data frame of one row and concatenate him to the training set with the index -1.

Dr. Jonathan Schler

Actual h index: 20

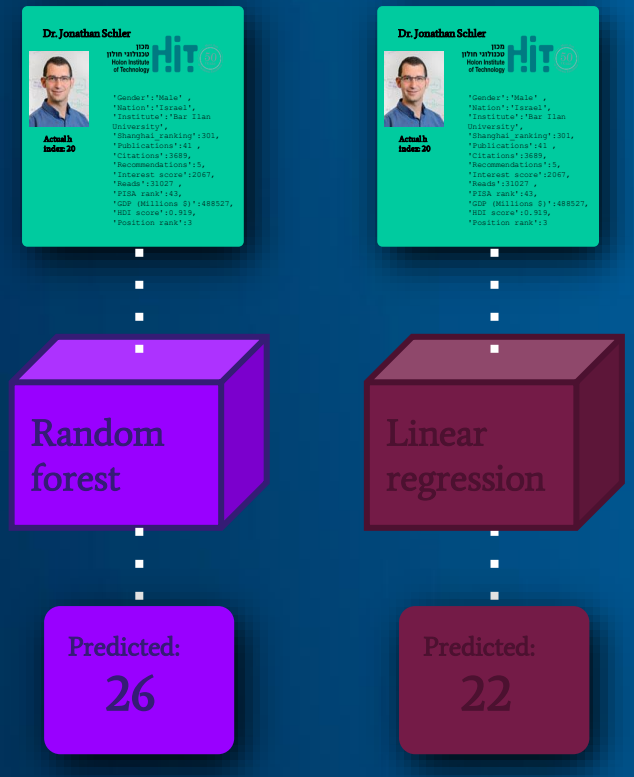
מכון טכנולוגי חולון
Holon Institute of Technology

```
'Gender':'Male' ,
'Nation':'Israel',
'Institute':'Bar Ilan
University',
'Shanghai_ranking':301,
'Publications':41 ,
'Citations':3689,
'Recommendations':5,
'Interest score':2067,
'Reads':31027 ,
'PISA rank':43,
'GDP (Millions $)':488527,
'HDI score':0.919,
'Position rank':3
```

Demonstration

```
print("Dr. Schler's actual h index: " ,yl.iloc[-1])
print("predicted h index (Linear Regression): ",np.round(y_predictl[-1]))
print("predicted h index (Random Forest) ",y_predict[-1])
```

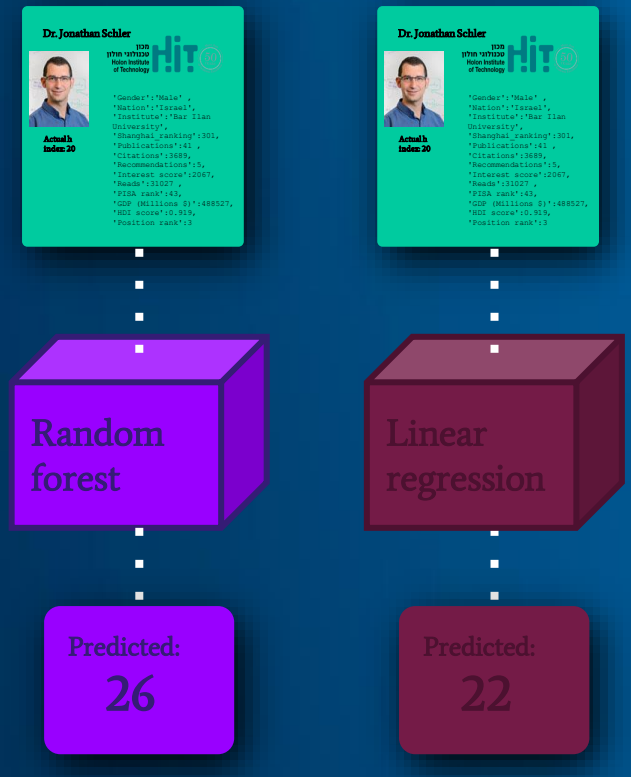
```
Dr. Schler's actual h index:  20
predicted h index (Linear Regression):  22.0
predicted h index (Random Forest)  26
```



Demonstration

Although the linear regression has an accuracy score of 78%, the predicted results, 22, is much more close to its actual score.

Whilst the random forest shows an accuracy of 98% which is almost perfect, but the score, 26, is far from its actual h index



Demonstration

Although the linear regression has an accuracy score of 78%, the predicted results, 22, is much more close to its actual score.

Whilst the random forest shows an accuracy of 98% which is almost perfect, but the score, 26, is far from its actual h index

- The only reasonable explanation for the random forest high accuracy with poor result is because of overfitting.
- Yet, it is important to note that it's not a good practice to assume it from one example.

Conclusions

Conclusions

(*heuristic approach*)

Whilst there are some features which are moderately correlated to the h index, a linear regression model, even with a medium accuracy, appear to be the optimal option available that doesn't uses a neural network.

We could predict a close h index of a scientists to his actual, relying of his other features.

Additionally, we saw that there are gaps and biases among genders, and nations that suffer from poor education to nations with advanced education system, in producing successful scientist.