

Transition to Tech

Academic to Data Scientist

Hercules Konstantopoulos
Data Scientist, Envizi

General Assembly
Sydney, 23 June 2016

Def: Alternative career path

University Faculty; Research Scientist; Lecturer

“The term ‘alternative career pathway’ in relation to leaving academic research is misleading. In reality, only a small number of PhD students will gain tenured jobs in academia. We need to change our thinking and consider staying in academia as the alternative pathway, and leaving as mainstream.”—Prof. Natalie Lundsteen, 2016

Fewer than 10% of PhD graduates remain in academia; best study conducted so far by UK Royal Society: 0.5% of PhDs become faculty, a further 3.5% remain in fields related to their doctoral research.

Def: Data Scientist

Too many definitions available

Type A: Provides in depth analyses that transcend the complexity of regular reporting and often exceed the boundaries of a company's standard practice.

Type B: Designs and refines advanced automation processes centred around machine learning and distributed computing to derive knowledge.

What sorts of industries hire data scientists, and for what types of analysis?

Financial sector

- Quantitative modelling, risk assessment (insurance, credit, loans)
- Investment: futures, retirement (modelling), high-frequency trading (algorithms)
- Compliance (anti-fraud, money-laundering)
 - Big players: Big Four banks (esp. CBA), Optiver, Australian Super, etc.

Property management

- Market valuation modelling, targeted marketing (inc. social media)
 - Big players: Domain Group

What sorts of industries hire data scientists, and for what types of analysis?

Listing continued...

Digital Marketing

- Market segmentation (social media or proprietary data)
- Modelling spending habits
- Click-stream analytics
 - Big players: Quantium, SC-Austereo

Technology

- Product or business analyst, data scientist
- IT/analytics consulting
 - Big players: Atlassian, Freelancer, IBM, SAS, SAP, Teradata, Accenture, KPMG, PWC, EY...

What sorts of industries hire data scientists, and for what types of analysis?

Listing continued...

Retail

- Marketing, targeted advertising
- Price/size/distribution optimisation
- Big players: supermarkets, distributors, hardware mega stores, shopping centres, any shop with a loyalty card

Energy

- Renewables
- Fossils
- Big players: EnerNOC, Schneider, Siemens, Ecova, GE, Carnegie Wave Energy

What sorts of industries hire data scientists, and for what types of analysis?

Listing continued...

Employment

- Recommendation engine for job-seekers
 - Big Players: Seek, LinkedIn, Glassdoor, Freelancer

Public sector

- Government collects a *lot* of data, only now starting to do something with it
- Focus on improving services, cutting costs, increasing efficiency, finding bad people
- Big Players: Defence, Customs/Immigration, Health, Police, Tax, ABS, DHS...

What sorts of industries hire data scientists, and for what types of analysis?

Listing continued...

Telecommunications

- Path to customer turnover (churn)
- Identifying areas to improve service
- Cyber-security
 - Big Players: Telstra, Vodafone, ISPs, Government, BT-Global

Misc service providers and “left-field”

- E.g., postal service, entertainment industry, sports/gambling
- Fashion, e.g., StitchFix, recommendation engine
- Journalism, science policy, outreach

What sorts of industries hire data scientists, and for what types of analysis?

In terms of job description:

- Data Scientist
- Data Analyst
- Business Analyst
- Insights Analyst
- Quantitative Analyst
- Business Development Manager
- Developer, Technologist, Software Engineer
- Solutions Architect
- Product manager
- User experience engineer
- Medical Physicist
- Science Policy Officer
- Journalist... so many more jobs.

Is this the career path for me?

An example of a data scientist's tasks.

- I work for a software company focussed on sustainability and energy management.
- Through my data analysis we come up with ideas on how to improve our products, how to better quantify environmental savings (i.e., GHG emissions), and how to convince clients to decrease their dependence on fossils by investing in renewables.
- On the side I do pro-bono consulting on data analysis and social media.

Is this the career path for me?

A day in the life:

- 80% of data science is cleaning data.
- As a general rule small companies do not know how to find the rotten apples in their databases, so that is a huge part of my job.
- As the only PhD in the company I am often called in to refine statistics, data presentation and visualisation practises, and the numeracy of the staff in general.
- Two days per week I work as a software developer, mostly focussed on designing, maintaining, and refining the SQL procedures that form the backbone of Envizi.

Is this the career path for me?

How do I find out?

- Take a couple of courses. Are they fun or are they painful?
- Go fetch some data from the real world (try OECD, EUstat, US Bureau of Labor Statistics), clean it up, get an interesting result, and write up a report.
- Show off your machine-learning skills through kaggle competitions.
- Get stuck in a real, messy dataset and see if you come out the other end with a smile on your face.

How do I prepare for a career in data science?

Whatever skills you acquire during your PhD is what you have at the end.

- No time to do more than one PhD in one PhD's worth of time.
- Try to ensure it is enough for any good job, not too specialised.
- Advisors and managers need to know what those skills are.

Having a PhD demonstrates that you know how to learn and can master a difficult field quickly

How do I prepare for a career in data science?

Simple to acquire right skills during your PhD, with the right focus and guidance.

- Must program well in one common language.
- Good data wrangling skills, including SQL.
- Creativity & problem solving skills a given for PhDs, a top-shelf skill for companies.
- Communication skills very important - give public talks & media engagement.
- Type-B Data Science: machine learning.
- Hadoop is good but not essential.
- Those long lists of desirable experience on job ads... don't stress, that's a wish list.

How do I prepare for a career in data science?

But do not despair, additional training can happen after your PhD.

- Plenty of training providers; online:
 - Coursera, Udacity, edX
 - JHU Data Science Specialisation
 - Stanford Machine Learning
- ...and in person
 - Science to Data Science (S2DS)
 - Insight Fellowship program
 - General Assembly Data Science program

Where and how do I find jobs?

- Accept LinkedIn as your Lord and Saviour.
- Join industry Groups, follow influential people, accept invitations from recruiters.
- Go to every meet-up in your local area.
 - Sydney Python Meetup, Data Science Sydney, Melbourne Users of R etc.
- It is enough to meet a couple of people in each industry event... you'll get enough contacts quickly.
- Shamelessly exploit networks of former academics.

How do I write an industry CV?

- Crucial to understand that people do not read CVs before machines sort and grade them.
- Always insert strong verbs; machines need to be impressed, too.
- If you have leadership experience, use verbs that describe that.
- Be assertive and bold.
- Nobody cares about your publication list. Condense talks and papers into 'Communication Skills'.

Technology

Dr Hercules (Iraklís) Konstantopoulos

Data Scientist
Astrophysicist
iraklisk@me.com
+61 0468 785 620



Mad about new technology. Excellent with data. Will work for Petabytes.

Expert **data analyst** and data wrangler with ten years of experience in Data Science and Astrophysics. Fast learner, at home with various **programming** languages (Python, R, SQL) and tools (Pandas, Shiny, Tableau). **Management** experience includes spearheading analytics **R&D**; assembling a **database design** software team; mentoring students of all ranks; and leading **user experience** design. Author on 47 scientific publications, cited 877 times, H-index=19.

EXPERIENCE

*Data Scientist, Envizi (Sydney, Australia)
May 2015 onward*

Envizi is a global, cloud based software platform that distils complex sustainability and energy data into powerful insights. I lead R&D in the technology group and report directly to the CTO.

- * Advance Envizi to the cutting edge by introducing new analytical intelligence, refining the statistical treatment of our customers' data and introducing machine-learning algorithms.
- * Bolster customer retention and attract new clients by conducting tailored data analyses in collaboration with the National Accounts Manager.
- * Tailored and delivered custom reports to high-profile clients, including Fortune 500, four of the top ten Australian companies by revenue, and ASX Top Five companies.
- * Strengthened the fidelity of the database by pinpointing missing and misclassified data, in some cases increasing data completeness from 5% to 95%.

*John Stocker Research Fellow, Australian Astronomical Observatory (Sydney)
June 2012—May 2015*

A small number of John Stocker Fellowships are awarded annually across Australia by the Science and Industry Endowment Fund, covering all scientific disciplines.

- * Led an independent research program into galaxy groups and served as an architect of the SAMI Galaxy Survey, the largest multi-dimensional survey of galaxies to date.
- * Designed and coded samiDB, an innovative data archive using HDF5; assembled and managed a working group of 12 scientist-developers to advance and administer samiDB.
- * Led the user experience design of the SAMI Galaxy Survey visual data browser by charting the user pathway and designing the front-end user interface.
- * Created the Starfish Diagram, a new type of statistical data visualisation of individuals and their place within large datasets.

*Research Associate, The Pennsylvania State University (State College PA, USA)
May 2009—May 2012*

- * Led a research program into extra-galactic astrophysics and managed my supervisor's research group of 2 researchers and 4 students. All students succeeded in advancing to top-notch PhD programs, professorial and industry jobs.
- * Awarded a \$100k grant from NASA to analyse Hubble Space Telescope data, plus a NASA/Caltech Jet Propulsion Lab grant to advise on Herschel Space Telescope data.
- * Designed and implemented a data reduction pipeline to digitally process and analyse large sets of images from the Hubble Space Telescope.

Academic

CURRICULUM VITÆ

Iraklís S. Konstantopoulos

John Stocker Research Fellow
Australian Astronomical Observatory
PO Box 915, North Ryde NSW 1670
Australia

<http://irakliskonstantopoulos.com>
iraklis@aao.gov.au
ORCID: 0000-0003-2177-0146 | [ADS link](#)
[Bitbucket: iraklis_k](#)

Personal Information

Born 24 March 1983 in Ioánnina, Greece. Hellenic Nationality.

Academic Activity

- 2012–present John Stocker Postdoctoral Fellow, Australian Astronomical Observatory.
- 2010–2012 Research Associate, The Pennsylvania State University.
- 2009–2010 Postdoctoral Scholar, The Pennsylvania State University.
- 2009 Research Assistant, Gemini Observatory, Southern Operations Center.
- 2005–2009 Ph.D., University College London, UK: *Studies of the Formation and Evolution of Extragalactic Star Clusters*. Advisors: Linda Smith and Nate Bastian. Partly undertaken at ESO-Chile with Mark Gieles and Michael West.
- 2001–2005 M.Sci (Bachelor's and Master's degree in Astrophysics), University College London, UK. Master's thesis, supervised by Sarah Bridle: *Does the Universe Accelerate at the Same Rate in All Directions?*

Honours and Awards

- May 2014 *Science in the Shine Dome*, EMCR Travel Awardee representing the AAO.
- 2010–2012 Various NASA research grants of value ~\$100k.
- 2008 European Southern Observatory studentship.
- 2005 Perren Ph.D. studentship, University College London.
- 2005 Honours Degree in Astrophysics, University College London.

Team Membership

- 2014–present MaunaKea Spectroscopic Explorer (MSE; formerly ngCFHT): Science Team.
- 2013–present Rogier Windhorst's gravitational lensing team (100hr GTO allocation on JWST).
- 2013–present The GAMA and RESOLVE Galaxy Surveys: Full Member.
- 2012–present SAMI Galaxy Survey: 'Builder' member, Chair of Database Working Group.
- 2012–present Centre of Excellence for All-Sky Astrophysics (CAASTRO): Affiliate.
- 2010–present SHUCS Survey: 1/2 Founding Members, Lead Scientist for first two years.

Committees (selected)

- 2014–present Chair, Diversity Committee, Australian Astronomical Observatory.
- 2014–present Gender Action Committee, CAASTRO.
- 2013–present eScience Working Group, Australian Astronomy Decadal Plan.
- June 2013 Organising Committee, *Feeding, Feedback, Fireworks* conference.
- 2009–2012 Climate and Diversity Committee, Penn State Astronomy.
- 2009–2012 Green Team, Penn State Astronomy.

Personal Interests

Photography • Fencing, cycling, soccer, mountaineering • Travel, language.

Iraklis Konstantopoulos, CV—p.1

Technology

Academic

Dr Hercules (Iraklís) Konstantopoulos

Data Scientist
Astrophysicist

iraklisk@me.com
+61 0468 785 620



Mad about new technology. Excellent with data. Will work for Petabytes.

Expert **data analyst** and data wrangler with ten years of experience in Data Science and Astrophysics. Fast learner, at home with various **programming** languages (Python, R, SQL) and tools (Pandas, Shiny, Tableau). **Management** experience includes spearheading analytics **R&D**; assembling a **database design** software team; mentoring students of all ranks; and leading **user experience** design. Author on 47 scientific publications, cited 877 times, H-index=19.

CURRICULUM VITÆ Iraklís S. Konstantopoulos

John Stocker Research Fellow
Australian Astronomical Observatory
PO Box 915, North Ryde NSW 1670
Australia

<http://irakliskonstantopoulos.com>
iraklis@aao.gov.au
ORCID: 0000-0003-2177-0146 | [ADS link](#)
Bitbucket: [iraklis_k](#)

Personal Information

Born 24 March 1983 in Ioánnina, Greece. Hellenic Nationality.

Technology

EXPERIENCE

Data Scientist, Envizi (Sydney, Australia)

May 2015 onward

Envizi is a global, cloud based software platform that distils complex sustainability and energy data into powerful insights. I lead R&D in the technology group and report directly to the CTO.

- * Advance Envizi to the cutting edge by introducing new analytical intelligence, refining the statistical treatment of our customers' data and introducing machine-learning algorithms.
- * Bolster customer retention and attract new clients by conducting tailored data analyses in collaboration with the National Accounts Manager.
- * Tailored and delivered custom reports to high-profile clients, including Fortune 500, four of the top ten Australian companies by revenue, and ASX Top Five companies.
- * Strengthened the fidelity of the database by pinpointing missing and misclassified data, in some cases increasing data completeness from 5% to 95%.

John Stocker Research Fellow, Australian Astronomical Observatory (Sydney)

June 2012—May 2015

A small number of John Stocker Fellowships are awarded annually across Australia by the Science and Industry Endowment Fund, covering all scientific disciplines.

- * Led an independent research program into galaxy groups and served as an architect of the SAMI Galaxy Survey, the largest multi-dimensional survey of galaxies to date.
- * Designed and coded samiDB, an innovative data archive using HDF5; assembled and managed a working group of 12 scientist-developers to advance and administer samiDB.
- * Led the user experience design of the SAMI Galaxy Survey visual data browser by charting the user pathway and designing the front-end user interface.
- * Created the Starfish Diagram, a new type of statistical data visualisation of individuals and their place within large datasets.

Academic

Academic Activity

2012–present	John Stocker Postdoctoral Fellow, Australian Astronomical Observatory.
2010–2012	Research Associate, The Pennsylvania State University.
2009–2010	Postdoctoral Scholar, The Pennsylvania State University.
2009	Research Assistant, Gemini Observatory, Southern Operations Center.
2005–2009	Ph.D., University College London, UK: <i>Studies of the Formation and Evolution of Extragalactic Star Clusters</i> . Advisors: Linda Smith and Nate Bastian. Partly undertaken at ESO-Chile with Mark Gieles and Michael West.

Honours and Awards

May 2014	<i>Science in the Shine Dome</i> , EMCR Travel Awardee representing the AAO.
2010–2012	Various NASA research grants of value ~\$100k.
2008	European Southern Observatory studentship.
2005	Perren Ph.D. studentship, University College London.

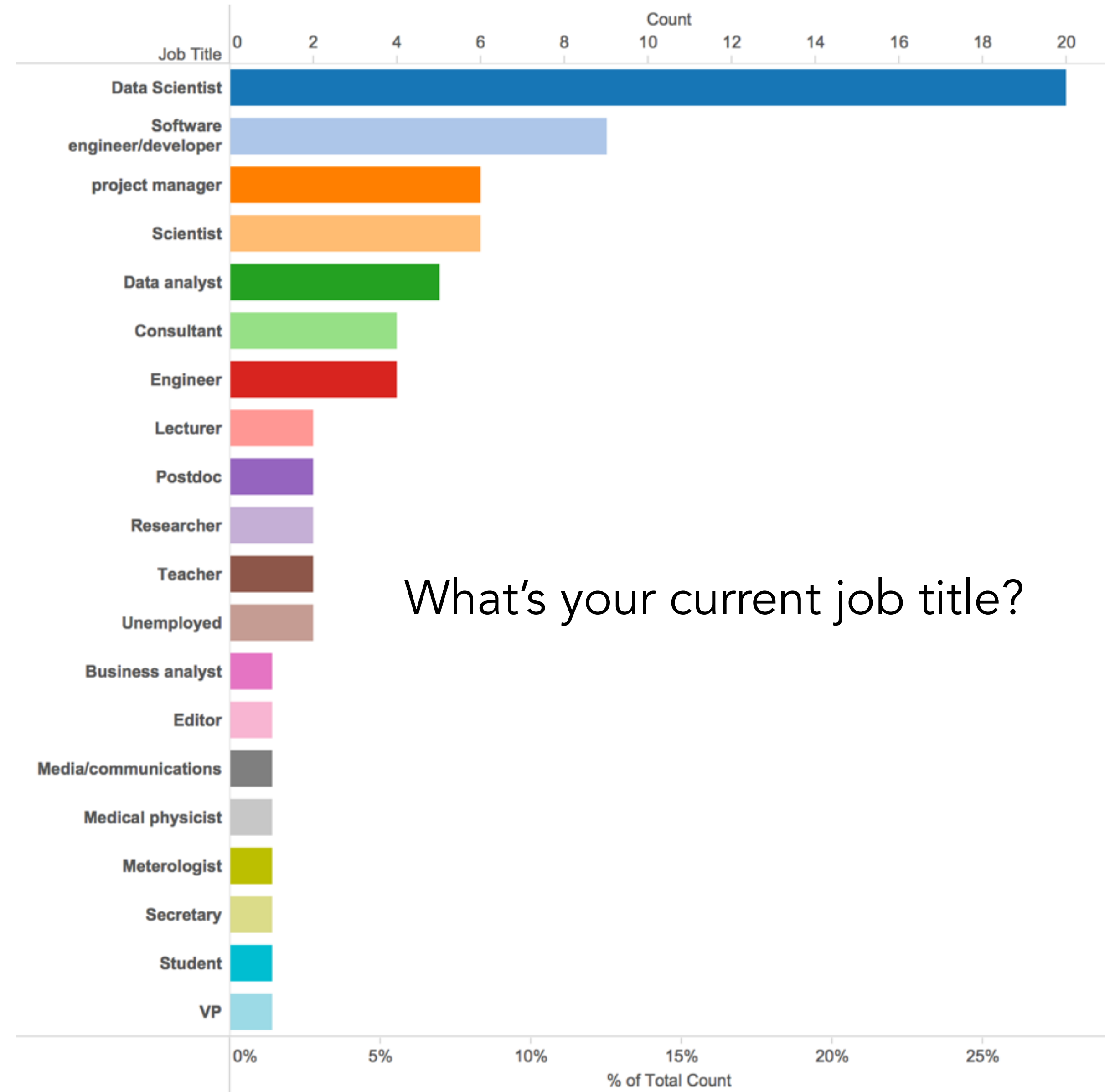
Team Membership

2014–present	MaunaKea Spectroscopic Explorer (MSE; formerly ngCFHT): Science Team.
2013–present	Rogier Windhorst's gravitational lensing team (100hr GTO allocation on JWST).
2013–present	The GAMA and RESOLVE Galaxy Surveys: Full Member.
2012–present	SAMI Galaxy Survey: 'Builder' member, Chair of Database Working Group.
2012–present	Centre of Excellence for All-Sky Astrophysics (CAASTRO): Affiliate.

Committees (selected)

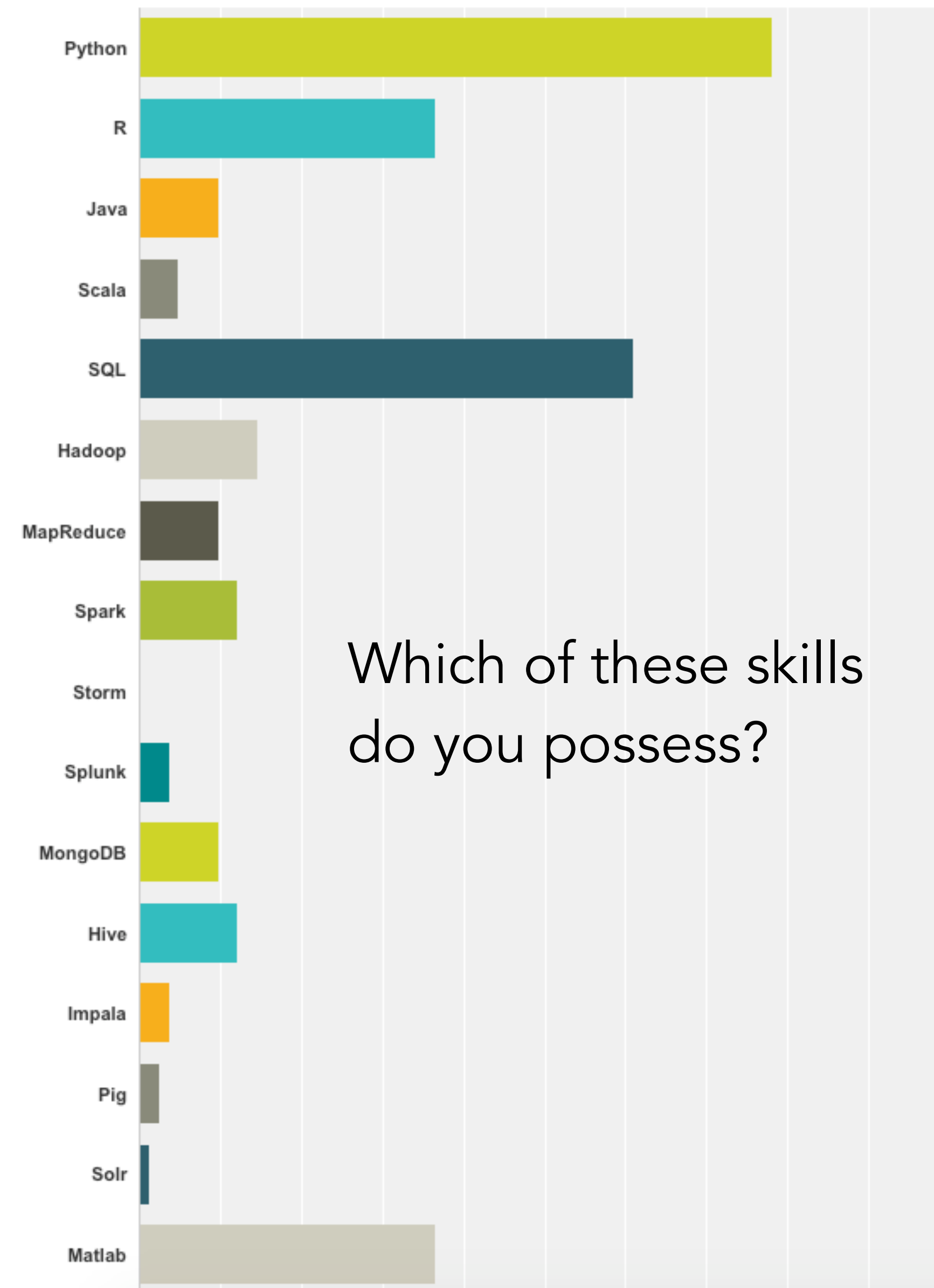
2014–present	Chair, Diversity Committee, Australian Astronomical Observatory.
2014–present	Gender Action Committee, CAASTRO.
2013–present	eScience Working Group, Australian Astronomy Decadal Plan.
June 2013	Organising Committee, <i>Feeding, Feedback, Fireworks</i> conference.
2009–2012	Climate and Diversity Committee, Penn State Astronomy.
2009–2012	Green Team, Penn State Astronomy.

Amazing!! We asked former astronomers ten questions. What follows next will blow your mind!! Click here to find out more!!



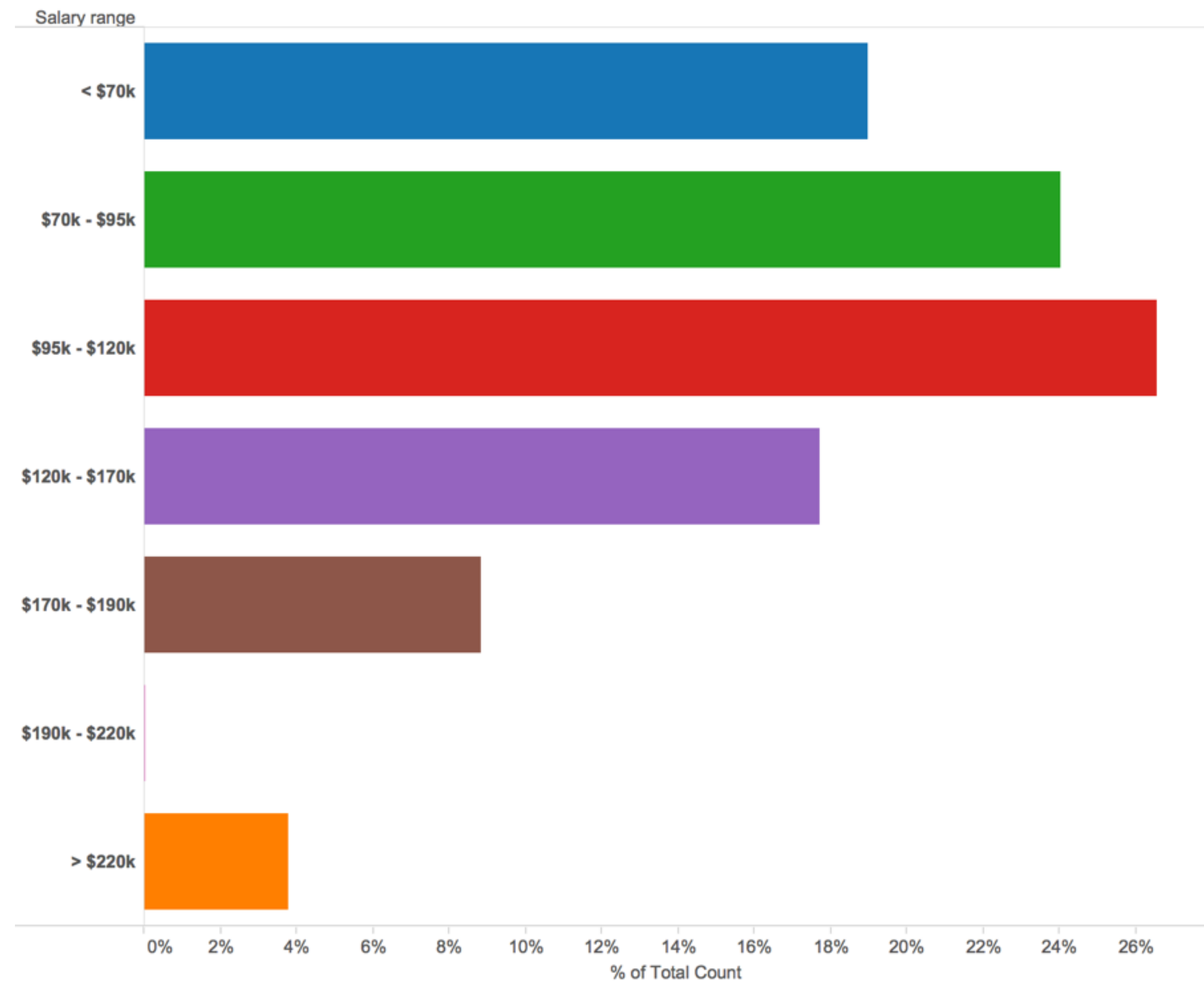
GA ↗ Academic to Data Scientist

Amazing!! We asked former astronomers ten questions. What follows next will blow your mind!! Click here to find out more!!



Amazing!! We asked former astronomers ten questions. What follows next will blow your mind!! Click here to find out more!!

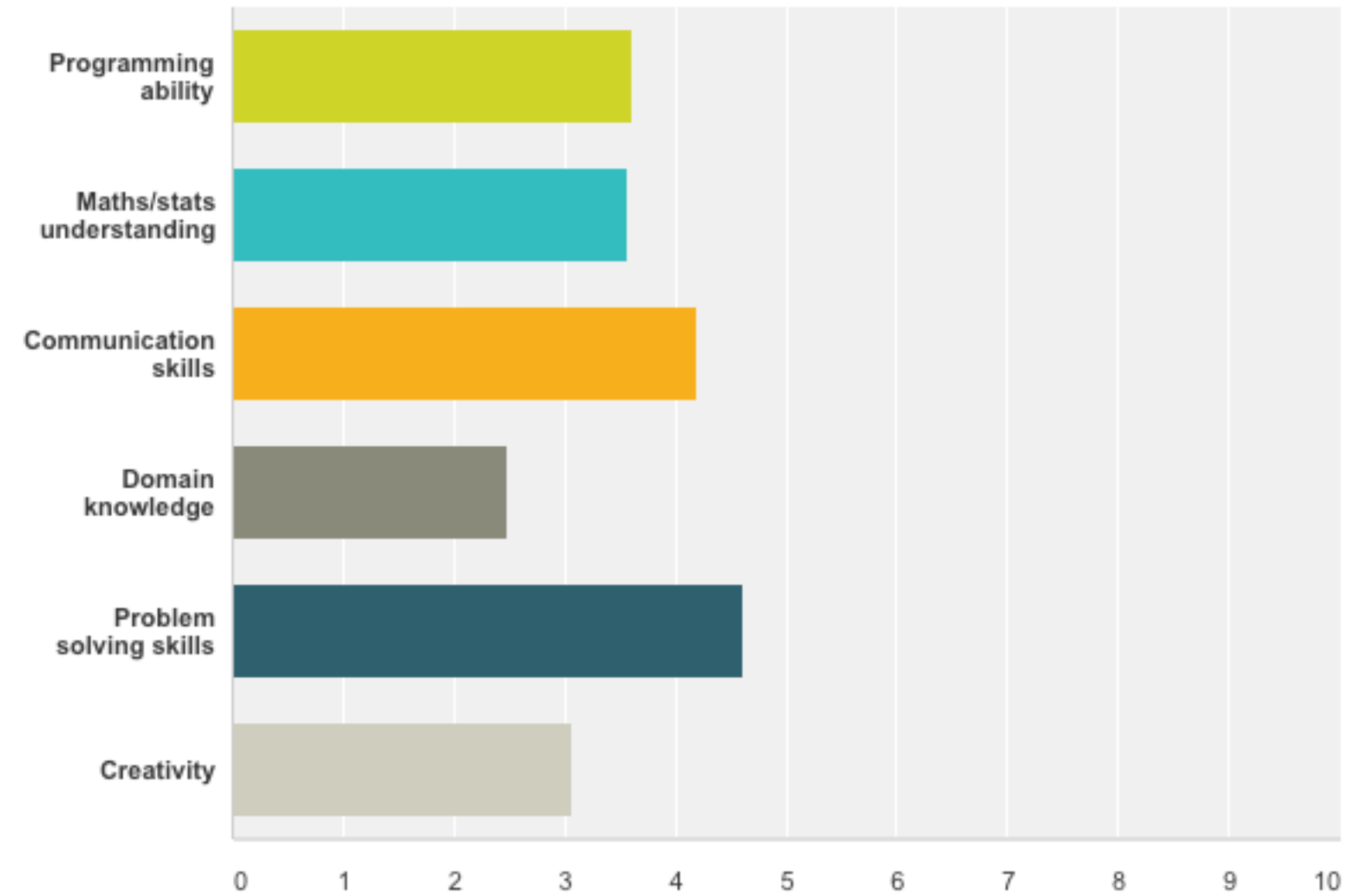
What is your current salary range (in AUD)?



Amazing!! We asked former astronomers ten questions. What follows next will blow your mind!! Click here to find out more!!

Please rank the following in order of importance for your current role/field

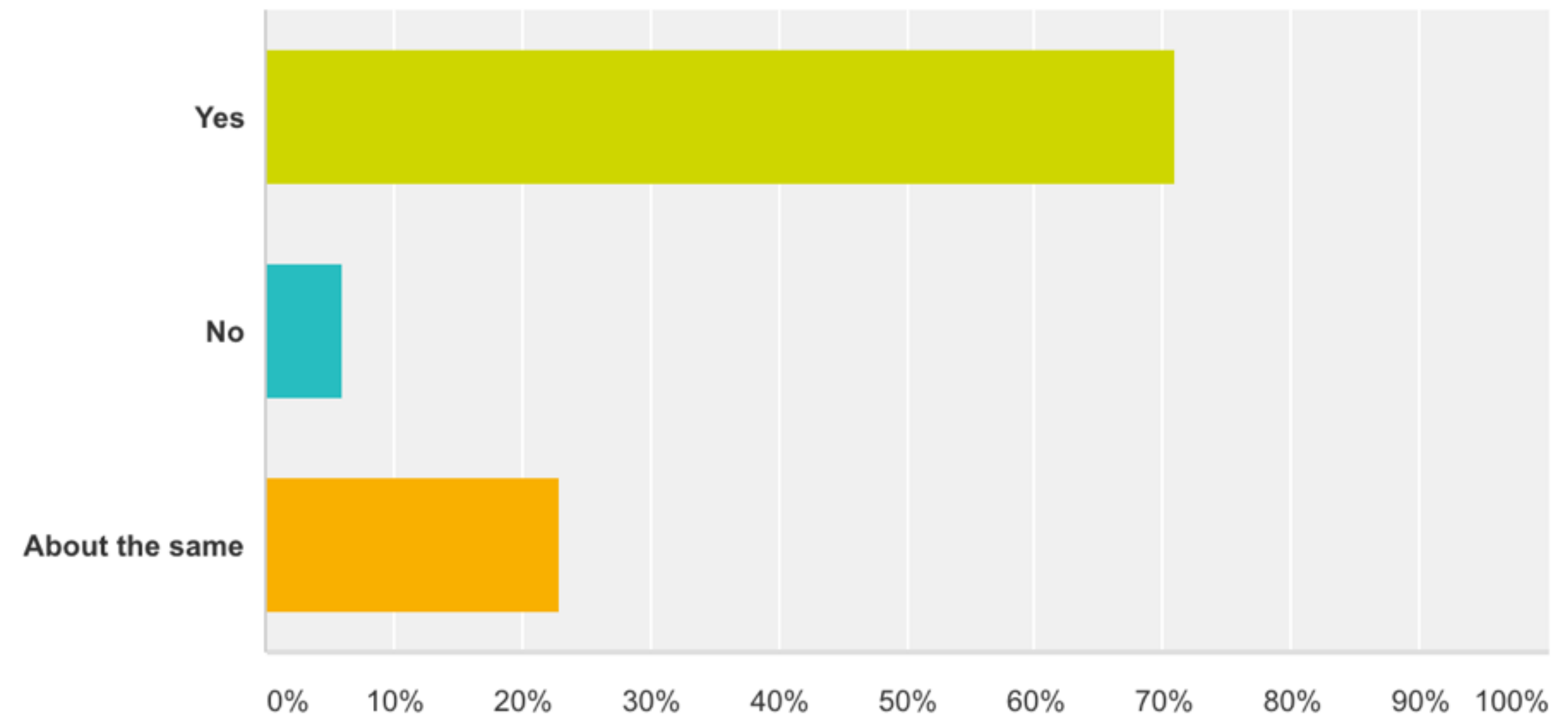
Answered: 83 Skipped: 0



Amazing!! We asked former astronomers ten questions. What follows next will blow your mind!! Click here to find out more!!

Would you describe yourself as happier since you left astronomy?

Answered: 83 Skipped: 0



With thanks to
Sean Farrell and Amanda Manypenny
for additional material.