



NumFOCUS Small Development Grants:

Help us Fund Advancements in the Open Data Stack



ABOUT NumFOCUS

The mission of NumFOCUS is to promote open practices in research, data, and scientific computing by serving as a fiscal sponsor for open source projects and organizing community-driven educational programs.

NumFOCUS is a 501(c)(3) public charity in the United States.

HOW WE SUPPORT OPEN SOURCE

NumFOCUS provides a stable, independent, and professional home for the open source projects powering contemporary scientific inquiry and business processes. We aim to ensure that funding and resources are available to sustain projects in the scientific data stack over the long haul.

“Pandas is used all over the world to help data scientists do their work. We’re all hungry for pandas2, a release that will vastly improve the experience of analyzing large datasets. The amount of time we have to wait to bring this release to the public is directly correlated to the amount of funds we have for developer time. In this way, NumFOCUS has directly fueled our progress. The public digital infrastructure is frighteningly fragile; NumFOCUS plays a huge part in ensuring that the projects we all depend on don’t pop out of existence entirely.”

—Jeff Reback, pandas Lead Developer



NumFOCUS SMALL DEVELOPMENT GRANTS PROGRAM

NumFOCUS awards small development grants to help our projects improve usability, grow their communities, and speed up the time to major releases. Eligibility is limited to our fiscally sponsored and affiliated projects, and calls for proposals are run 3 times per year.

SPONSORED PROJECTS:

- | | | | |
|---------------|-------------------|-----------------|----------|
| > Astropy | > IPython | > NumPy | > SciPy |
| > Blosc | > Julia | > Open Journals | > Shogun |
| > Bokeh | > JuMP | > pandas | > Stan |
| > Cantera | > Project Jupyter | > PyMC3 | > SunPy |
| > conda-forge | > Matplotlib | > PyTables | > SymPy |
| > Econ-ARK | > MathJax | > QuantEcon | > xarray |
| > FEniCS | > nteract | > rOpenSci | > yt |

AFFILIATED PROJECTS:

- | | | | |
|-----------|------------------|---------------|----------------|
| > ArviZ | > Dask | > Orange | > scikit-image |
| > Chainer | > Data Retriever | > pomegranate | > scikit-learn |
| > Conda | > DyND | > pvlib | > Spack |
| > CuPy | > Gensim | > Python(X,Y) | > Spyder |
| > Cython | > MDAnalysis | > QuTiP | > Statsmodels |
| > Dash | > Numba | > scikit-bio | > Theano |
| | | | > Yellowbrick |

In the first year of the program, NumFOCUS awarded \$13,000 in small development grants to our sponsored and affiliated projects. Last year we distributed \$60,000, and this year we have budgeted \$85,000. We have already distributed \$24,774 in 2019.

“We are really honored and grateful for having been chosen as one of the recipients of the NumFOCUS spring grant. It was what basically allowed us to keep developing the project during most of 2018.”










—Spyder, NumFOCUS Affiliated Project

“This is an excellent way for NumFOCUS to help the development of its projects — thanks for running it!”

—SunPy, NumFOCUS Sponsored Project

WHAT DO THE GRANTS FUND?

Past grants include:







PROJECT	PROPOSAL TITLE	AMOUNT
 ArviZ	Create educational material and give workshops related to exploratory analysis of Bayesian models with ArviZ (2019)	\$2,500
 Bokeh	Bokeh Docs Modernization (2018)	\$3,000
 Cantera	The 4th Annual Kinetics Code Conference (2019)	\$4,000
	Modernize, Reorganize, and Update Cantera's Documentation (2018)	\$3,000
	The 3rd Annual Kinetics Code Conference: Charting near- and long-term directions for Cantera software development (2018)	\$3,000
 conda-forge	conda-forge sprint at SciPy 2019 (2018)	\$3,000
 Gensim	Organize Gensim Documentation & Improve Discovery (2019)	\$5,000
	Modern user-friendly documentation (2018)	\$3,000
	FastText tutorials (2018)	\$3,000
 Julia	JuliaImages developer meeting (2019)	\$4,000
	(Sheehan Olver) BlockBandedMatrices.jl: add support for general array backends (GPU) (2018)	\$3,000
	JuliaImages developer meeting (2018)	\$3,000
 MDAnalysis	MDAnalysis tutorial and hackathon (2018)	\$2,500
	Widening platform availability for MDAnalysis: Full Python 3 Support (2017)	\$1,500
 NumPy	NumExpr-3.0 Beta (2017)	\$3,000
 Open Journals	Open Journals website update (2018)	\$2,800





 Orange Data Mining	Girls go Data Mining (2018)	\$3,000
	Text Analytics Introductory Course for Social Scientists (2017)	\$1,750
 Pandas	Improving and modernizing the introductory “Getting Started” pages of the pandas documentation (2019)	\$5,000
 Pomegranate	Improving Documentation, Examples, and Tutorials (2018)	\$3,000
	Adding compatibility with user-defined Python models (2018)	\$3,000
 PyTables	Better support for native HDF5 files (2018)	\$3,000
	h5py backend for PyTables (2017)	\$3,000
 SciPy	SciPy Development Documentation Overhaul (2019)	\$4,274
	Maturing a sparse array implementation for SciPy (2018)	\$3,000
	An Efficient, High-Level Implementation of Linear Programming (2018)	\$2,000
	American Meteorological Society Short Course on Open Source Radar Software (2017)	\$1,000
 Shogun	Fully integrate new parameter framework, unify API/interfaces, and release Shogun 7.0. (2018)	\$1,500
	Shogun website and UX redesign (2018)	\$2,500
 Spyder	Spyder 4: Making the Scientific Python Development Environment even better (2018)	\$3,000
 Statsmodels	Probability Plots and Generalized Additive Models (finish stalled pull requests) (2018)	\$3,000
 SunPy	Improving the Usability of SunPy’s Data Downloader (2018)	\$3,000
 SymPy	MatchPy C++ code generator for SymPy/symengine. (2018)	\$3,000
	SymPy 1.1 Release Support (2017)	\$3,000




GRANT TRACKING AND OUTCOMES

All recipients of NumFOCUS Small Development Grants are required to report back on what they accomplished with the funding and whether it was in line with their original proposal.

Sample Grant Outcomes:

PROJECT	PROPOSAL TITLE
 Cantera	<p>Modernize, Reorganize, and Update Cantera's Documentation (\$3k)</p> <p>"We achieved all of our goals for the small development grant. Our website is now live at https://cantera.org and the codebase has been split between API documentation and examples, tutorials, etc."</p>
 Gensim	<p>FastText tutorials (\$3k)</p> <p>"We achieved more than we proposed!</p> <ul style="list-style-type: none"> - update of FastText documentation & tutorials - also found and fixed several critical issues with FastText - FastText performance optimizations"
 Julia	<p>BlockBandedMatrices.jl: add support for general array backends (GPU) (\$3k)</p> <p>"We successfully implemented support for general array backends and the outputs are available to the public via the recently released BlockBandedMatrices.jl v3.0. We tested the results for SharedArrays and GPU Arrays for construction, matrix * vector and matrix * matrix products, showing the expected improvement in performance."</p>
 MDAnalysis	<p>MDAnalysis tutorial and hackathon (\$2.5k)</p> <p>"We held a 2 day workshop for 40 people, teaching the basics of our package and related tools, and some instruction on contributing to open source software. From the attendees it looks like 3 might become regular contributors to the package. We were able to provide travel support to 8 people, most from underrepresented minorities."</p>
 Orange Data Mining	<p>Girls go Data Mining (\$3k)</p> <p>"We held a two-day workshop for approximately 50 women of all ages and professions. In these two days we taught them the basics of data visualization, clustering and predictive modeling."</p>
 pomegranate	<p>Adding compatibility with user-defined Python models (\$3k)</p> <p>"We added in compatibility with user-defined Python models!</p> <p>We rewrote the internals of pomegranate to allow a fallback to calling Python methods if the distribution is not built-in. This allows people to write simple, pure Python distributions and models and drop them in as simply as they would a built-in distribution.</p> <p>An unintended consequence of this addition is it enables deep probabilistic models to be, fairly trivially, implemented. One can easily merge a deep network and a mixture model to yield deep mixture models, or deep networks and hidden Markov models to yield deep HMMs. Best of all, this functionality is not package dependent, and so users of Keras, Tensorflow, or PyTorch can do this all extremely easily.</p> <p>We then wrote two tutorial notebooks about the changes."</p>

 PyTables	<p>Better support for native HDF5 files (\$3k)</p> <p>“The padding in native HDF5 files is totally honored in the new PyTables 3.5. Additionally, users can create HDF5 files with padding by using the NumPy capability of introduce padding in structured arrays. Also, a new release, with the respective wheels and latest HDF5 libraries, has been assembled.”</p>
 SciPy	<p>An Efficient, High-Level Implementation of Linear Programming (\$2k)</p> <p>“We implemented all of the proposed enhancements to <code>scipy.optimize.linprog</code>. Specifically, we added callback function support to the interior-point and revised simplex solvers, so that now users can run custom code each iteration regardless of the solution method. We created an option for users to provide an initial basic feasible solution when using the revised simplex method, essentially eliminating the need for the first phase of the simplex algorithm. We improved the clarity and presentation of the <code>linprog</code> documentation and carefully documented and organized the <code>linprog</code> test suite so that future contributors can easily add unit tests for the desired combinations of methods and options.</p> <p>Finally, we benchmarked the SciPy linear programming methods against one another and two other open source (GPL-licensed) solvers: CVXOPT and GLPK. <code>linprog</code>, the only Python-accessible, BSD/MIT-like licensed linear programming solver we know of, is shown to significantly outperform CVXOPT and even approach the effectiveness of GLPK. SciPy 1.3 will offer users a greatly enhanced linear programming suite for essentially unrestricted use on nearly any platform.”</p>
 Shogun	<p>Fully integrate new parameter framework, unify API/interfaces, and release Shogun 7.0. (\$1.5k)</p> <p>“On April 13-18, we hosted a developer code-sprint in Belgrade with four core developers in attendance. This was a highly productive meeting where the core developers were able to focus on pressing technical issues and design questions that are hard to ‘outsource’ to external contributors.</p> <p>We made major progress in integrating Shogun’s new parameter framework within the libraries’ low-level functionalities (clone, equals, generic put/get, serialization, etc). This further allowed for progress in deploying the new Shogun API in the examples/docs, in particular in the form of proof-of-concept implementations that can now be easily mimicked and further applied by external contributors such as Google Summer of Code students.</p> <p>Another long-term goal we finally achieved was to fully finalize the transition from a GPL license to a more liberal BSD-style license. Given that Shogun is almost two decades old, with contributors in the hundreds to be consulted regarding license issues, we are very pleased having managed this transition.</p> <p>Finally, we solved a number of technical issues, such as developing concepts for reducing the interface build complexity and various bugs that have been collected over the past months.</p> <p>The meeting resulted in major progress towards a major release (7.0) that we hope to publish soon.”</p>
 Spyder	<p>Spyder 4: Making the Scientific Python Development Environment even better (\$3k)</p> <p>“We accomplished a complete overhaul to the underlying infrastructure for completion, linting and introspection, allowing for improved performance and stability. We’re currently finishing the last pieces of its integration with Spyder.</p> <p>We also merged a figure explorer, allowing for easy viewing of all generated figures in a dedicated pane into the master and consider it production ready.”</p>

 Statsmodels	<p>Probability Plots and Generalized Additive Models (finish stalled pull requests) (\$3k)</p> <p>“The main outcome of the project is the merging for Generalized Additive Models (GAM) and penalized regression splines into statsmodels. The supported features now are able to match the reference package in R.”</p>
 SunPy	<p>Improving the Usability of SunPy's Data Downloader (\$3k)</p> <p>“The first, and most complex, phase of implementing registration for search parameters was completed. This involved implementing a metaclass to enable custom tab completion of class attributes.</p> <p>The next phase of having the data sources register their search attributes is underway, as is the third component of re-implementing the file downloader to provide better feedback and error recovery to the user.”</p>
 SymPy	<p>MatchPy C++ code generator for SymPy/symengine (\$3k)</p> <p>“The code generator of MatchPy has been ported into C++. The abstract syntax tree expressions of MatchPy rely on SymEngine (C++ library). Associative/commutative pattern matching expressions defined in MatchPy can now be exported into C++ decision trees, performing the same matching logic.”</p>

The above work was funded with \$100K over 3 years.

Imagine what we could do with \$1M!



NumFOCUS SEEKS CORPORATE SPONSORSHIPS TO FULLY FUND OUR SMALL DEVELOPMENT GRANTS






Help NumFOCUS:

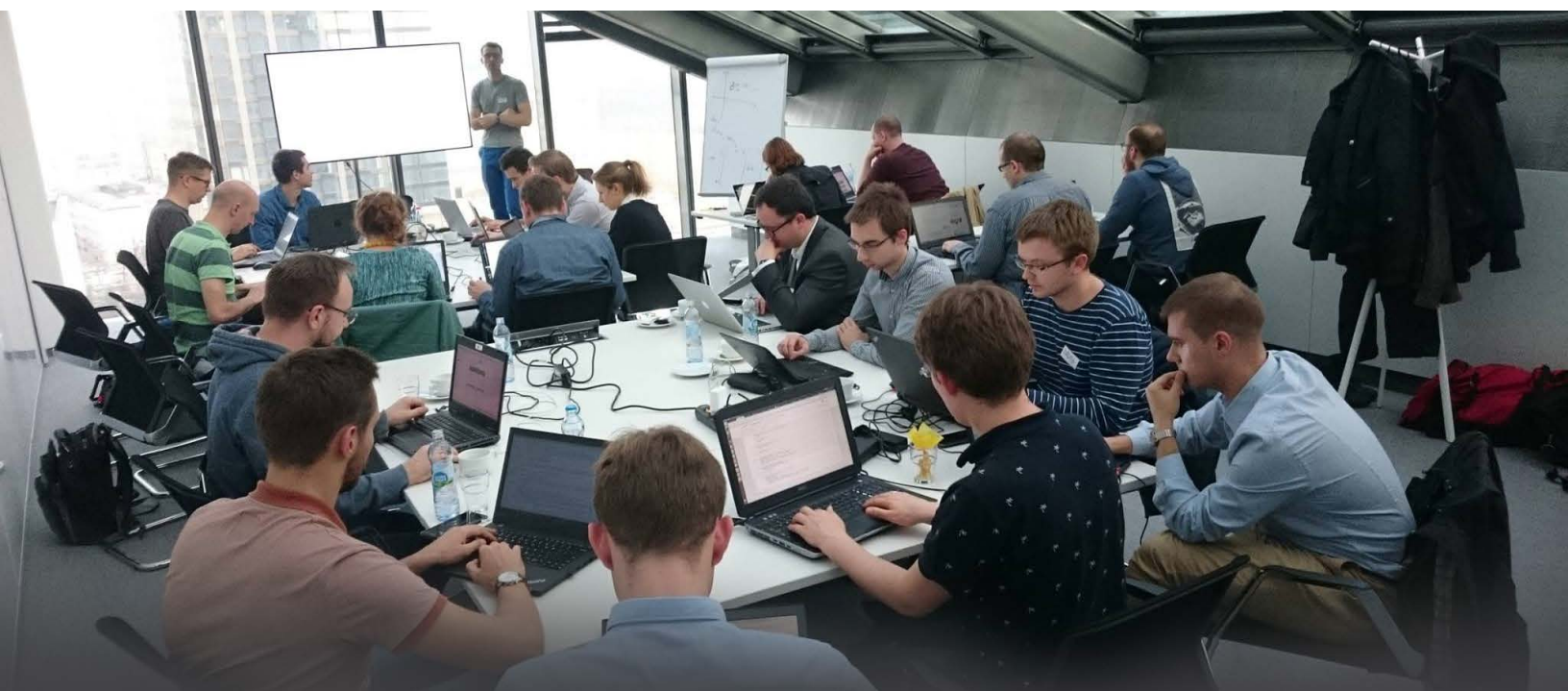
- > Increase the amount of funding per proposal
- > Fund more proposals
- > Accelerate the development of more user-friendly, efficient, and extensible open source scientific software





BENEFITS	PLATINUM	GOLD	SILVER	SILVER	EMERGING LEADER
	\$100,000	\$50,000	\$25,000	\$10,000	Under \$10,000
Subscription to NumFOCUS project release updates and news	✓	✓	✓	✓	✓
Use of the NumFOCUS Partner Badge on your website	✓	✓	✓	✓	✓
Logo placement on numfocus.org	✓	✓	✓	✓	✓
Company/Organization profile on NumFOCUS.org	✓	✓	✓		
Presentation at your headquarters, or to the team of your choice, by a NumFOCUS project core developer	✓	✓	✓		
Seat on the NumFOCUS Advisory Council	✓	✓			
Logo used on all promotional material and media channels	✓				
Naming of a sponsored educational workshop series	✓				
Invitation to attend a project developer summit partner day	✓				
Sponsor benefits at PyData events of your choice	Platinum at one event	Gold at one event	Silver at one event	No PyData benefit	No PyData benefit

We now receive more high-quality proposals than our current budget can accommodate. With additional funding, NumFOCUS will be able to offer more and larger grants to our projects.

Your generous corporate sponsorship will help fund proposals like these, which were unable to be awarded due to 2019 budget constraints:

PROJECT	PROPOSAL TITLE	AMOUNT REQUESTED
 SymPy	Code Generation in SymPy and SymEngine for Rubi module	\$3,000
 Bokeh	Migrate Bokeh Documentation to ReadTheDocs	\$4,000
 Orange	Make Your Own Data Science Textbook We will design an interactive web page where users would pick data science topics and the desired difficulty level and obtain an assembled textbook in PDF.	\$5,000
 QuTiP	Modernise distribution, testing and installation QuTiP deployment is currently suffering from increasing issues arising in setup, platform-specific deployment, broken gcc compilation, and lacks novel interactive ways to reach out to the community.	\$4,000
 Astropy	Developing Spectroscopic Reduction Tools The goal of this proposal is to fund an interested and available developer to jump-start the development of the specreduce Astropy affiliated package. The functionality this package will provide is of broad use to the astronomical community and is our most requested enhancement.	\$5,000



 Shogun	Improving Shogun's Observable framework To improve and extend Shogun's facilities to emit parameter changes of models at runtime. This enables improved logging, (scientific) debugging, and user transparency of the underlying C++ implementation.	\$5,000
 FEniCS Project	Writing a plug-in extension framework for ffc FEniCS' just-in-time (JIT) form compiler "ffc" turns DSL into C-code. To increase flexibility, we want to add a user-extensible plugin system.	\$4,000
 Blosc	Support for Split Frames Support for split frames would allow better integration with VCS (Version Control Systems, e.g. git) and remote storage (e.g. S3). It would allow improved modification time of data on-disk and better disk utilization (via e.g. git).	\$5,000
 Data Retriever	An Efficient, High-Level Implementation of Linear Programming Create a web interface to simplify the addition of new data packages to the Data Retriever. The web interface will automatically detect the properties of data sources provided as urls, provide an interface for entering metadata, and automatically submit the new data package to the central repository as a pull request.	\$5,000





NUMFOCUS

OPEN CODE = BETTER SCIENCE

If you're interested in learning more about where your contributions can make a difference in the NumFOCUS ecosystem, how NumFOCUS will acknowledge you as a Small Development Grant sponsor, or simply how to make a donation, please get in touch!

GET IN TOUCH

P.O. Box 90596 • Austin, TX 78709
info@numfocus.org
+1 (512) 831-2870

[LEARN MORE](#)