

SciPy user survey results

Author: Maja Gwózdź

Contents: Introduction | Graphical summary | Detailed summary | Conclusions

Introduction

This is an edited summary of the survey. Consult the comprehensive results [here](#). The SciPy user survey officially ran from September 18 to November 5 (2019) and was disseminated via Google Forms and advertised on the SciPy mailing list and website, Twitter, relevant departments of several universities, NumFOCUS, blogs, physics mailing lists, etc. In total, we received 185 responses. The questions and possible responses included in the survey were as follows (reproduced in the exact order of appearance):

1. What is your SciPy experience level? (multiple choice) [177 responses]
 - beginner
 - intermediate
 - advanced
 - expert
2. How often do you use SciPy? (multiple choice) [178 responses]
 - daily
 - weekly
 - monthly
 - yearly
 - only sporadically
3. Which parts of the documentation do you usually use? (checkboxes) [179 responses]
 - tutorials
 - developer's guide
 - API reference
 - unsure
4. How do you browse the documentation? (checkboxes) [179 responses]
 - online HTML documentation (non-mobile)
 - online HTML documentation (mobile)
 - PDF reference guide
 - docstring in terminal
 - help feature in an IDE
 - other

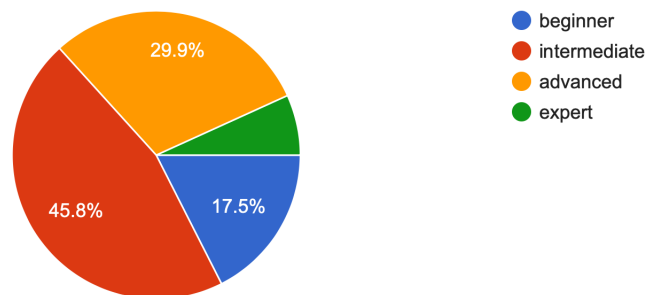
5. Are you satisfied with the current documentation? (linear scale) [180 responses]
 - scale from 1 (definitely not) to 5 (definitely yes)
6. Are you able to find the information you are looking for quickly? (multiple choice) [183 responses]
 - yes
 - no
7. Is the graphical layout clear? (linear scale) [179 responses]
 - scale from 1 (definitely not) to 5 (definitely yes)
8. Which of the following documentation features should be improved/added? (checkboxes) [179 responses]
 - better explanations
 - ease of navigation
 - graphical guides
 - more code samples
 - more references to external sources
 - more screenshots
 - more tutorials
9. Do you have any specific positive comments about the current documentation? (paragraph; long answer text) [62 responses]
10. Do you have any specific negative comments about the current documentation? (paragraph; long answer text) [57 responses]
11. Do you have any other ideas on how the documentation could be improved? (paragraph; long answer text) [57 responses]

Graphical summary

Below is a graphical summary of the closed-ended questions.

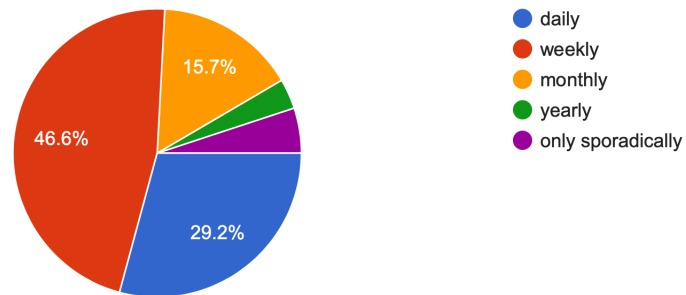
What is your SciPy experience level?

177 responses



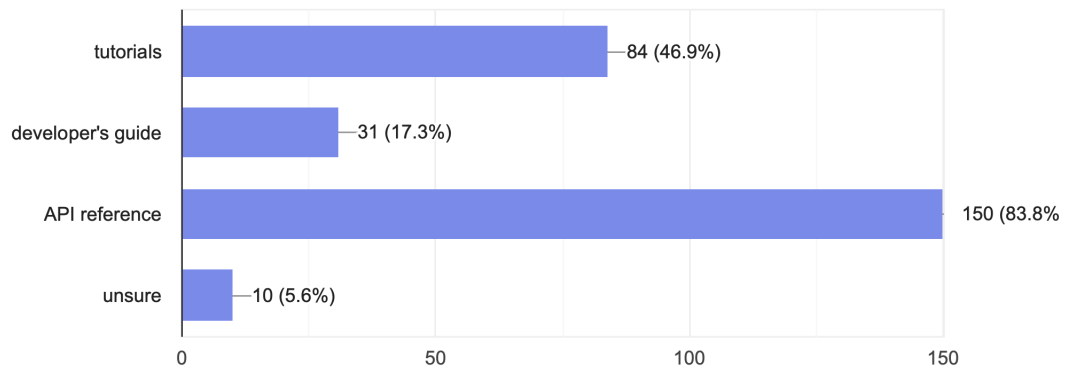
How often do you use SciPy?

178 responses



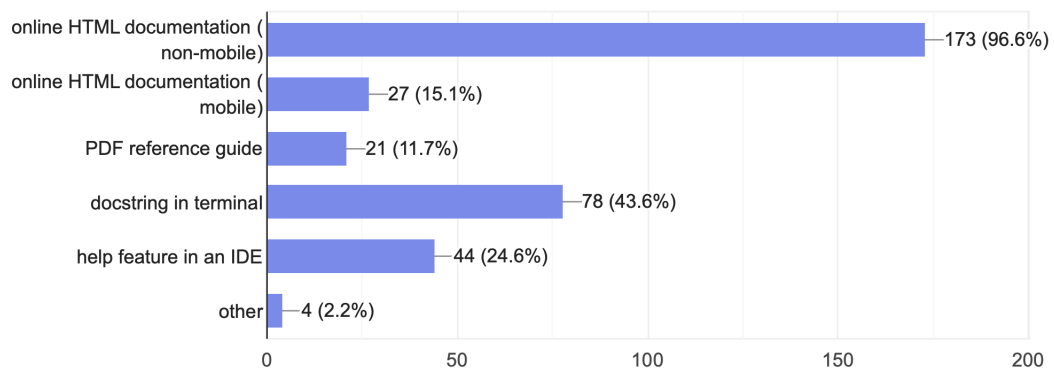
Which parts of the documentation do you usually use?

179 responses



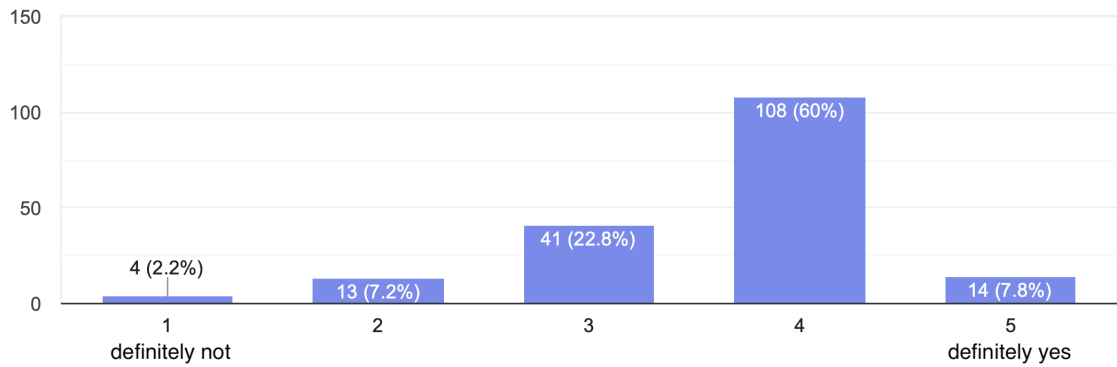
How do you browse the documentation?

179 responses



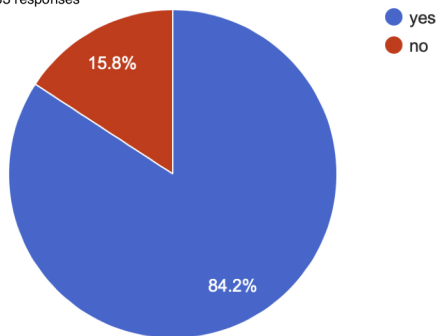
Are you satisfied with the current documentation?

180 responses



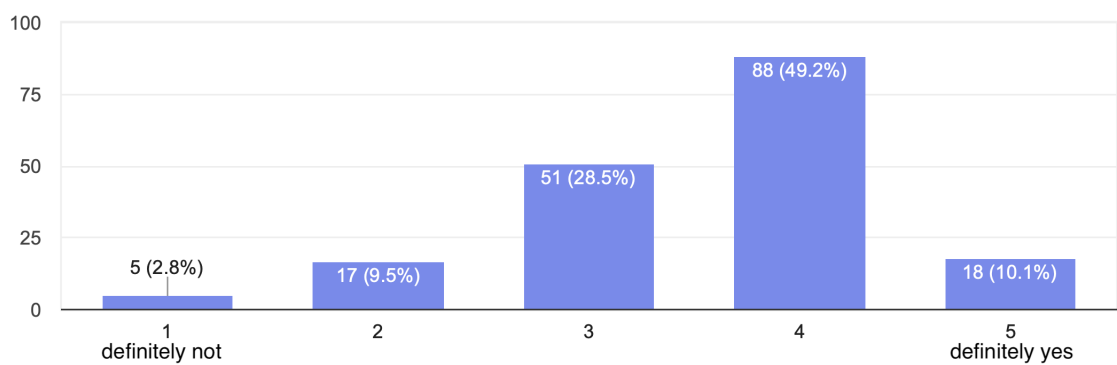
Are you able to find the information you are looking for quickly?

183 responses



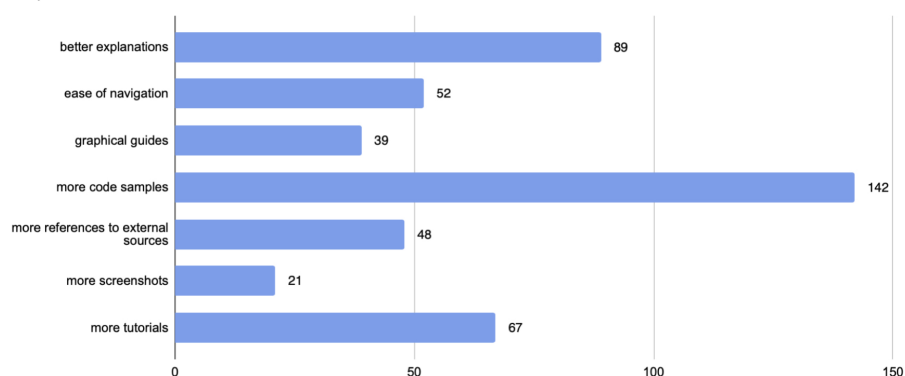
Is the graphical layout clear?

179 responses



Which of the following documentation features should be improved/added?

179 responses



Detailed summary

The present section includes a bullet-point summary of the three open-ended questions. They are addressed in their order of appearance in the survey. No duplicates have been included and only the relevant (subjectively speaking) responses are given. Some have been reproduced verbatim, while others have been edited and heavily truncated.

Positive comments

- overall, the documentation is:
 - (mostly) consistent
 - comprehensive
 - thorough
 - detailed
 - useful
 - clear
 - well written
 - easy to read
 - easy to navigate
- the external references are appreciated
- the quality of the documentation has recently improved
- users are able to find the desired information relatively quickly
- the references to source code implementation are appreciated
- the ‘See also’ section is helpful
- searching for particular functions yields satisfactory results
- parsing \LaTeX is very useful
- one respondent particularly likes the documentation for `scipy.optimize`

Negative comments

- the files are not well related (isolated features)
- some parts are not easy to follow (*e.g.*, distributions in `scipy.stats`)
- the back-and-forth navigation is difficult
- in some places, the links to ‘Contributing to SciPy’ are broken
- it would help if the documentation was more explicit about some of the assumptions used in the statistical functions
- many advanced features and functionalities are not easily discoverable
- it is sometimes difficult to discover that some important library functionality exists at all
- advanced features accessible through optional argument of many functions are either not well explained or lack adequate code samples for new users
- sometimes, functions and their arguments are described using the jargon of a specific field
- searching for terms does not work
- some functions that wrap low-level routines are sometimes difficult to understand and would benefit from additional description
- the examples are not basic
- some explanations are too vague or assume too much knowledge
- it is difficult to switch between SciPy versions in the API reference
- there are too few examples and images
- Google search results often yield confusing links to deprecated functions or old versions of SciPy
- old, outdated, or deprecated parts of the documentation ought to be marked explicitly (*e.g.*, in pink or red)
- the quality of the documentation is not uniform
- certain warnings about future deprecation (especially, regarding some algorithms) ought to be included in the documentation
- the navigation ought to be improved
- the visual representation of code does not consistently follow the convention of being shown in monospaced type
- the navigation between the tutorials and the API documentation is not straightforward
- some parameters are not properly explained (or are missing), especially in `scipy.stats` or `scipy.optimize`
- mobile experience is lacking
- it would be beneficial to have mathematical definitions included where necessary (*e.g.*, in `scipy.fftpack`)
- some algorithms are not clear and there do not exist specific references to them
- it is very difficult to find specific algorithms
- the limitations and/or accuracy of a given implementation are/is not easily obtainable

- the API reference describes features in an abstract way and does not have many examples
- the RV class documentation should be improved
- parameter descriptions sometimes fail to explain their purpose
- the page <http://docs.scipy.org/> is very confusing and should be updated to the latest version
- there is too little graphical representation
- sometimes, there is too much information given; it could be beneficial to split or hide optional arguments

Improvement ideas

- a documentation sprint or a buddy system (with a reviewer from a different scientific discipline)
- the search function ought to be improved
- interlinks between related functions would be welcome
- more examples (especially for beginners) and references to external documentation
- interactive examples
- a direct link from `scipy.org` to SciPy documentation
- more visualizations
- a better navigation system
- more cross-references
- integrating the existing GitHub issues with the current function would help (also, integration with Stack Overflow would be welcome)
- Jupyter notebooks with more elaborate explanations
- the documentation would benefit greatly if it were easier to edit
- including links from the API reference to the relevant source code would be appreciated
- the documentation should be more easily accessible to beginners
- improving the consistency and completeness for the documentation of similar objects
- it would be useful to have a critical comparison of different methods (where available)
- developer information should be clearly separated from user information
- adding description of the handling (or non-handling) of edge cases (*e.g.*, NaN inputs)
- including more explanations of the mathematical concepts

Conclusions

Overall, the results appear to be really useful and could serve as a helpful roadmap for future development and improvement of the documentation. The number of the responses, to my mind, was neither too low (hence, insignificant in a loose sense) nor too high (difficult to analyse).