

stars 1.1k pypi v1.8.2 downloads 735k downloads 49k docs passing
Azure Pipelines succeeded

Scanpy – Single-Cell Analysis in Python

Scanpy is a scalable toolkit for analyzing single-cell gene expression data built jointly with [anndata](#). It includes preprocessing, visualization, clustering, trajectory inference and differential expression testing. The Python-based implementation efficiently deals with datasets of more than one million cells.

- Discuss usage on [Discourse](#) and development on [GitHub](#).
- Get started by browsing [tutorials](#), [usage principles](#) or the main [API](#).
- Follow changes in the [release notes](#).
- Find tools that harmonize well with anndata & Scanpy via the [external API](#) and the [ecosystem page](#).
- Consider citing [Genome Biology \(2018\)](#) along with original [references](#).

News

scVelo on the cover of Nature Biotechnology 2020-12-01

Scanpy's counterpart for RNA velocity, [scVelo](#), made it on the cover of [Nature Biotechnology](#) [[tweet](#)].

Scanpy selected among 20 papers for 20 years of Genome Biology 2020-08-01

[Genome Biology: Celebrating 20 Years of Genome Biology](#) selected the initial Scanpy paper for the year 2018 among 20 papers for 20 years [[tweet](#)].

COVID-19 datasets distributed as [h5ad](#) 2020-04-01

Key Contributors

[anndata graph](#) | [scanpy graph](#) | ✨ = maintainer

- [Isaac Virshup](#): lead developer since 2019 ✨
- [Gökçen Eraslan](#): developer, diverse contributions ✨
- [Sergei Rybakov](#): developer, diverse contributions ✨
- [Fidel Ramirez](#): developer, plotting ✨
- [Giovanni Palla](#): developer, spatial data
- [Malte Luecken](#): developer, community & forum
- [Philipp Angerer](#): developer, software quality, initial anndata conception ✨
- [Alex Wolf](#): lead developer 2016-2019, initial anndata & scanpy conception
- [Fabian Theis](#) & lab: enabling guidance, support and environment

In a joint initiative, the Wellcome Sanger Institute, the Human Cell Atlas, and the CZI distribute datasets related to COVID-19 via anndata's `h5ad` files: covid19cellatlas.org. It wasn't anticipated that the [initial idea](#) of sharing and backing an on-disk representation of `AnnData` would become so widely adopted. Curious? Read up more on the [format](#).

Latest additions

Version 1.8

1.8.2 2021-11-3

Docs

- Update conda installation instructions [PR 1974](#) L Heumos

Bug fixes

- Fix plotting after `scanpy.tl.filter_rank_genes_groups()` [PR 1942](#) S Rybakov
- Fix `use_raw=None` using `anndata.AnnData.var_names` if `anndata.AnnData.raw` is present in `scanpy.tl.score_genes()` [PR 1999](#) M Klein
- Fix compatibility with UMAP 0.5.2 [PR 2028](#) L Mcinnes
- Fixed non-determinism in `scanpy.pl.paga()` node positions [PR 1922](#) I Virshup

Ecosystem

- Added PASTE (a tool to align and integrate spatial transcriptomics data) to scanpy ecosystem.

1.8.1 2021-07-07

Bug fixes

- Fixed reproducibility of `scanpy.tl.score_genes()`. Calculation and output is now float64 type. [PR 1890](#) I Kucinski
- Workarounds for some changes/ bugs in pandas 1.3 [PR 1918](#) I Virshup
- Fixed bug where `sc.pl.paga_compare` could mislabel nodes on the paga graph [PR 1898](#) I Virshup
- Fixed handling of `use_raw` with `scanpy.tl.rank_genes_groups()` [PR 1934](#) I Virshup

1.8.0 2021-06-28

Metrics module

- Added `scanpy.metrics` module!

- Added `scanpy.metrics.gearys_c()` for spatial autocorrelation [PR 915](#) | Virshup
- Added `scanpy.metrics.morans_i()` for global spatial autocorrelation [PR 1740](#) | Virshup, G Palla
- Added `scanpy.metrics.confusion_matrix()` for comparing labellings [PR 915](#) | Virshup

Features

- Added `layer` and `copy` kwargs to `normalize_total()` [PR 1667](#) | Virshup
- Added `vcenter` and `norm` arguments to the plotting functions [PR 1551](#) | G Eraslan
- Standardized and expanded available arguments to the `sc.pl.rank_genes_groups*` family of functions. [PR 1529](#) | F Ramirez | Virshup - See examples sections of `rank_genes_groups_dotplot()` and `rank_genes_groups_matrixplot()` for demonstrations.
- `scanpy.tl.tsne()` now supports the metric argument and records the passed parameters [PR 1854](#) | Virshup

Ecosystem

- Added [Cubé](#) to ecosystem page [PR 1878](#) | C Lambden
- Added `triku` a feature selection method to the ecosystem page [PR 1722](#) | AM Ascensión
- Added `dorothea` and `progeny` to the ecosystem page [PR 1767](#) | P Badia-i-Mompel

Documentation

- Added [Community](#) page to docs [PR 1856](#) | Virshup
- Added rendered examples to many plotting functions [issue 1664](#) | A Schaar | L Zappia | bio-la | L Hetzel | L Dony | M Buttner | K Hrovatin | F Ramirez | Virshup | LouisK92 | mayarali
- Integrated [DocSearch](#), a find-as-you-type documentation index search. [PR 1754](#) | P Angerer
- ◦ Reorganized reference docs [PR 1753](#) | Virshup
- Clarified docs issues for `neighbors()`, `diffmap()`, `calculate_qc_metrics()` [PR 1680](#) | G Palla
- Fixed typos in grouped plot doc-strings [PR 1877](#) | C Rands
- Extended examples for differential expression plotting. [PR 1529](#) | F Ramirez - See `rank_genes_groups_dotplot()` or `rank_genes_groups_matrixplot()` for examples.

Bug fixes

- Fix `scanpy.pl.paga_path()` `TypeError` with recent versions of anndata [PR 1047](#) | P Angerer
- Fix detection of whether IPython is running [PR 1844](#) | Virshup
- Fixed reproducibility of `scanpy.tl.diffmap()` (added `random_state`) [PR 1858](#) | Kucinski
- Fixed errors and warnings from embedding plots with small numbers of categories after `sns.set_palette` was called [PR 1886](#) | Virshup
- Fixed handling of `gene_symbols` argument in a number of `sc.pl.rank_genes_groups*` functions [PR 1529](#) | F Ramirez | Virshup
- Fixed handling of `use_raw` for `sc.tl.rank_genes_groups` when no `.raw` is present [PR 1895](#) | Virshup

- `scanpy.pl.rank_genes_groups_violin()` now works for `raw=False` [PR 1669](#) M van den Beek

Development processes

- Switched to [flit](#) for building and deploying the package, a simple tool with an easy to understand command line interface and metadata [PR 1527](#) P Angerer
- Use [pre-commit](#) for style checks [PR 1684](#) [PR 1848](#) L Heumos I Virshup

Deprecations

- Dropped support for Python 3.6. [More details here.](#) [PR 1897](#) I Virshup
- Deprecated `layers` and `layers_norm` kwargs to `normalize_total()` [PR 1667](#) I Virshup
- Deprecated `MulticoreTSNE` backend for `scanpy.tl.tsne()` [PR 1854](#) I Virshup