scanpy.tl.score_genes

scanpy.tl.score_genes(*adata, gene_list, ctrl_size=50, gene_pool=None, n_bins=25, score_name='score', random_state=0, copy=False, use_raw=None*)

Score a set of genes [Satija15].

The score is the average expression of a set of genes subtracted with the average expression of a reference set of genes. The reference set is randomly sampled from the `gene_pool` for each binned expression value.

This reproduces the approach in Seurat [Satija15] and has been implemented for Scanpy by Davide Cittaro.

Parameters:	**adata** : `AnnData`
	The annotated data matrix.
	gene_list : `Sequence` [`str`]
	The list of gene names used for score calculation.
	ctrl_size : `int` (default: `50`)
	Number of reference genes to be sampled from each bin. If `len(gene_list)` is not too low, you can set `ctrl_size=len(gene_list)`.
	gene_pool : `Optional` [`Sequence` [`str`]] (default: `None`)
	Genes for sampling the reference set. Default is all genes.
	n_bins : `int` (default: `25`)
	Number of expression level bins for sampling.
	score_name : `str` (default: `'score'`)
	Name of the field to be added in `.obs` .
	random_state : `Union` [`None` , `int` , `RandomState`] (default: `0`)
	The random seed for sampling.
	copy : `bool` (default: `False`)
	Copy `adata` or modify it inplace.

use_raw : `Optional` [`bool`] (default: `None`)

> Whether to use `raw` attribute of `adata` . Defaults to `True` if `.raw` is present.
>
> *Changed in version 1.4.5:* Default value changed from `False` to `None` .

Return type:	`Optional` [`AnnData`]
Returns:	: Depending on `copy` , returns or updates `adata` with an additional field `score_name` .

Examples

See this notebook.