

TOP PYTHON / SCALA / R LIBRARIES IN DATA SCIENCE

PYTHON			SCALA			R		
	COMMITTS	CONTRIBUTORS		COMMITTS	CONTRIBUTORS		COMMITTS	CONTRIBUTORS
TensorFlow	34 689	1 505	DEEPLARNING4J	22 938	223	H2O	22 947	104
Orange3	22 811	1 099	Spark MLlib & ML	22 073	1 241	mlr	4 052	59
PYTORCH	11 763	668	PredictionIO	4 383	125	dmlc XGBoost	3 296	286
Keras	4 591	688	bigDL	2 358	50	caret	1 741	65
XGBoost	3 296	286	summingbird	1 780	33	LightGBM	1 125	82
CatBoost	1 746	63	DeepLearning.scala	1 874	15	Prophet	284	34
LightGBM	1 125	82	Spark + H2O SPARKLING WATER	1 416	33	ggplot2	4 070	152
dist-keras	1 125	5	Smile	1 195	30	plotly	3 185	24
elephas	170	13	Conjecture	155	8	ggvis	2 159	21
spark-deep-learning	68	11	VEGAS	215	16	DT DataTables	1 921	22
eli5	929	6	Breeze-viz	29	3	rCharts	638	11
matplotlib	25 984	727	akka	22 298	552	Corrplot	335	10
Bokeh	17 066	298	kafka	5 145	442	lattice	781	3
plotly	2 910	47	Breeze	3 431	89	dplyr	4 990	151
Seaborn	2 093	85	Epic	1 790	15	data.table	3 416	50
pydot	169	12	Puck	536	1	lubridate	1 501	51
SciPy	19 255	613	Saddle	184	10	jsonlite	914	12
NumPy	18 162	654	Scalab	44	1	knitr	5 528	105
Natural Language Toolkit	13 053	238	Slick	2 037	107	markdown	2 373	64
StatsModels Statistics in Python	10 244	153	ADAM	1 775	64	slidify	302	6
spaCy	8 655	226						
gensim	3 610	275						
pandas	17 243	1 179						
Scrapy	6 643	285						

LEGEND: ● Machine Learning ● Visualization ● Mathematics & Engineering ● Data Manipulation & Analysis ● Reproducible Research