

چرا پایتون

چرا پایتون برای آموزش علم داده انتخاب شده؟

به غیر از پایتون زبان‌های دیگری مانند R ، Scala ، Julia ، Matlab و SQL نیز برای علوم داده استفاده می‌شوند؛ اما پایتون به خصوص در صنعت از همه پرکاربردتر است (منظور از Matlab ، اسکرپت نوشته شده در متلب است).

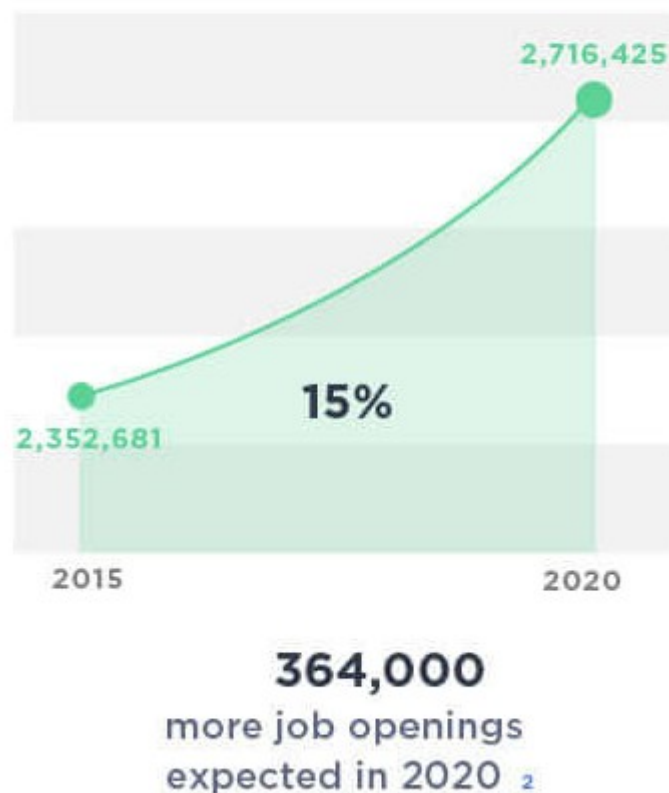
استفاده از پایتون دائما در حال افزایش است



آمار نشان می‌دهد ۶۵/۶٪ پژوهشگران داده از پایتون استفاده می‌کنند و امروزه به دلیل استفاده‌ی گسترده‌تر از پایتون در صنعت، تعداد کاربران پایتون از R پیشی گرفته است.

تقاضا برای پایتون در حال افزایش است

IT'S IN DEMAND



طبق آمار [forbes](#) تا سال ۲۰۲۰ بیش از ۳۶۴۰۰۰ شغل جدید برای برنامه نویسان پایتون ایجاد خواهد شد.

پایتون کتابخانه‌ها و ابزارهایی قوی و متعدد دارد

پایتون در تمامی حوزه‌های علم داده دارای کتابخانه‌های متعددی است که اکثر آنها توسط شرکت‌های بزرگ یا دانشگاه‌ها پشتیبانی و دائماً به‌روزرسانی می‌شوند. از کتابخانه‌های پایتون می‌توان به موارد زیر اشاره کرد.

- آنالیز و اکتشاف داده (*Data exploration & analysis*): SciPy , NumPy , Pandas
 - تصویرسازی و تجسم داده (*Data visualization*): Datashader , Seaborn , Matplotlib
 - یادگیری ماشین کلاسیک (*Classical machine learning*): StatsModels , Scikit-Learn
 - یادگیری عمیق (*Deep learning*): Caffe , Pytorch , TensorFlow , Keras
 - بیگ دیتا (*Big data*): h5py/pytables , Dask , HDFS , Apache Hadoop , Apache Spark
- پایتون در مواردی مانند پردازش زبان طبیعی و پردازش تصویر نیز کتابخانه‌هایی مانند `OpenCV/cv2` , `Spacy` , `nltk` , `scikit-image` دارد.

کدهای پایتون قابل فهم است

پایتون به نسبت زبان‌های دیگر نوشتار ساده‌تری دارد و به همین دلیل سریع قابل یادگیری است و این باعث می‌شود کار، به خصوص در پروژه‌های بزرگ، راحت‌تر شود. در واقع نوشتار پایتون به قدری ساده است که جمله‌ی زیر بسیار به شوخی گفته می‌شود:

پایتون، شبه کد قابل اجرا است

جامعه‌ی توسعه دهندگان بزرگ

فرض کنید در نوشتن کدی به مشکلی بر می‌خورید ؛ می‌توانید توضیح مشکل خود را در گوگل سرچ کنید و احتمالا از قبل این مشکل برای دیگران پیش‌آمده و راه حل آن در سایت‌هایی مانند *stackoverflow* موجود است. حتی اگر این مشکل از قبل برای کسی پیش نیامده باشد، تعداد زیادی متخصص در هر حوزه‌ی کار با پایتون در وبسایت‌های مختلف هستند تا به سوالات شما پاسخ دهند. این دقیقا فایده‌ی داشتن جامعه‌ی بزرگی از توسعه‌دهندگان است.

چند پارادایمی بودن

پایتون از جمله‌ی زبان‌هایی است که از اکثر پارادایم‌های معروف برنامه‌نویسی مانند `object-oriented programming` , `structured programming` و `procedural programming` پشتیبانی می‌کند.

پایتون در مقابل R

پایتون و R دو زبان پرتعداد در حوزه‌ی علم داده هستند. اشتراکات زیادی از جمله رایگان و متن باز بودن بین دو زبان وجود دارد. مستندات هر دو زبان با کمک کاربران نوشته شده است و هر دو اجتماعات فعالی در اینترنت دارند. اما هرکدام برتری‌هایی نسبت به دیگری دارند.

- پایتون کتابخانه‌های متعدد و قوی‌ای دارد که در بالا به آنها اشاره شده است.
- تقریبا هر ۱۸ ماه یک‌بار نسخه‌ی جدید و با ثباتی از پایتون عرضه می‌شود.
- پایتون دارای نوشتار ساده است.
- R توانایی کشیدن نمودارهای با کیفیت و متنوعی را دارد.
- R به سرعت در حال توسعه است و تقریبا برای همه‌ی تکنیک‌های آماری پکیج دارد.

اما تفاوت اصلی میان این دو، کاربرد بیشتر پایتون در صنعت است. به صورتی که اکثر شرکت‌ها برای بخش *production* از پایتون استفاده می‌کنند و R بیشتر به استفاده‌های آکادمیک محدود شده است. به این خاطر که هدف این دوره کاملا آکادمیک نبوده، در آن از پایتون استفاده شده است.