Languages of Data Science

Python

R

SQL

Scala

java

c++

julia

js

php

go

ruby

vb

Choose depening on Company you work for, the role you have in a project and the age of existing application.

Roles in Datascience:Business Analyst

Database Engineer

Data Analyst

Data Engineeri

Data Scientist

Software Engineer

Statistician

Product Manager

Project Manager.

Introduction of Python

Open Source

IBM

WIki

Google

Yahoo

CERN

NASA

fb

Amazon

Insta

Reddit

Data Science : Pandas, Numpy, SciPy,Matplotlib

AI: TensorFlow, Pytorch, Keras. Scikit-learn

NLP: NLTK

Intro to R

Free Software

private, public and commercia uses

stats, mathemeticians, data miners use it

Uses array. Easy to convert math to code.

IBM, Google, FB, Microsoft, Bank of America, Ford, TechCrunch, Uber, Trulia

Largest repo for stat knowledge

Has >15k publicly released packages to conduct data analysis.

Integrates well with C++, java, C, .Net and Python.

matrix multiplication gives immediate results.

Stronger OOP.

Global Communities

useR

whyR

SatRdays

1. Ladies

r-projects.org

Intro to SQL

non procedural lang

simple and powerful

1974 at IBM

used for Structured Data

DBMS

uses Tables.

helps in getting a lot of job.

speeds workflow.

interpreter between you and database

It has specific standards so different sqls has similar syntax,

Many SQL Databases Available:

mysql

db2

postgres

sqlite

oracle database

maria

sql server

Other Data Science Langauages

scala - general purpose prg lang. functional programming, static type system. inter opearable with java. runs on JVM. Scalable language.

Apache Spark - provides api that makes parallel jobs easy to write

Optimized engine that supports computation graphs

Includes Shark(query Engine), MLlib, GraphX, and Spark Streaming.

Designed to be faster than Hadoop.

java - OOP tried and tested. Fast.

Weka(Data mining), Java-Ml(ml library) , Apache MLLib(Scalable ml), Deeplearning4

Hadoop - Data Managment for big data.

c++ --C with classes. Improves Speed, enables system programming, and gives you broader control over application.Develops programs that feed data ti custoners un real-tine.

TensorFlow built in c++ but can be used with python.

mongodb

Caffe with python and Matlab bindings.

julia --MIT. High performance numerical analysis and computational science. Seedy dev and fast programs. Compiled.

Calls C, Go, Java, MATLAB, R, Fortran, and Python libraries with refined parallelism.

built 2012.

juliaDB package for working with large persistent data sets.

js --

Tensorflow.js -- ML and Deepleaning in both node and browser.

brain.js, machinelearn.js.

R-js makes algebra possible in typescript.

php

go

ruby

vb