

Software Engineer
Data Engineer
Data Scientist



Who Am I

- 7 years in top 4 Bank, Finance Quantitative Analyst
- Lead a team in startup to build enterprise applications
- Computer Science Ph.D
- Senior SE in Capital One Shanghai Department (Acquired by Blackstone)
- Try best to provide you “DRIED FOODS”

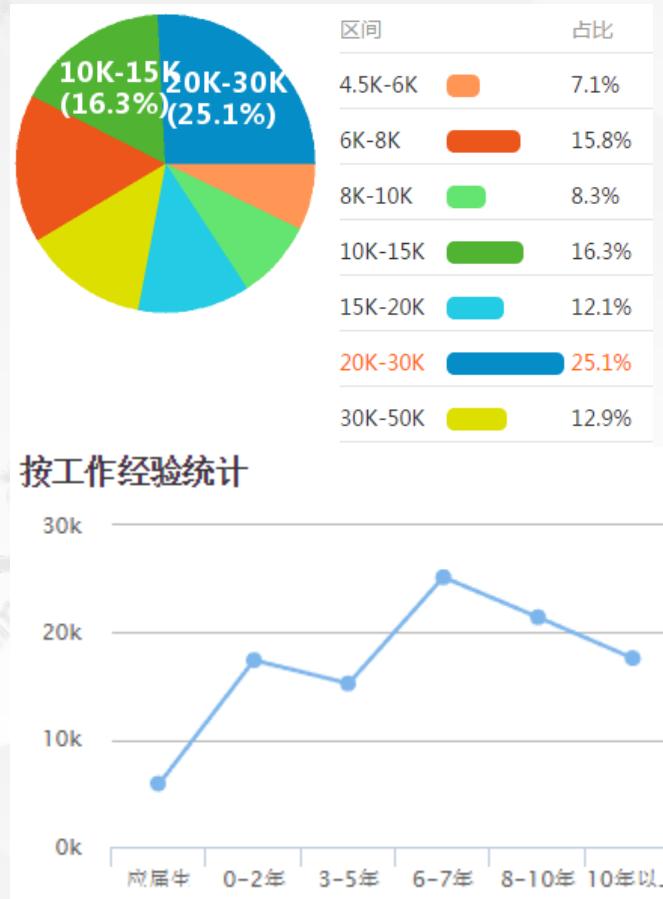
Why This Course

- Software Engineer
 - Data Engineer
 - Data Scientist
-
- Personal Interest
 - Cool & Hot
 - Job Openings
 - Great Compensation

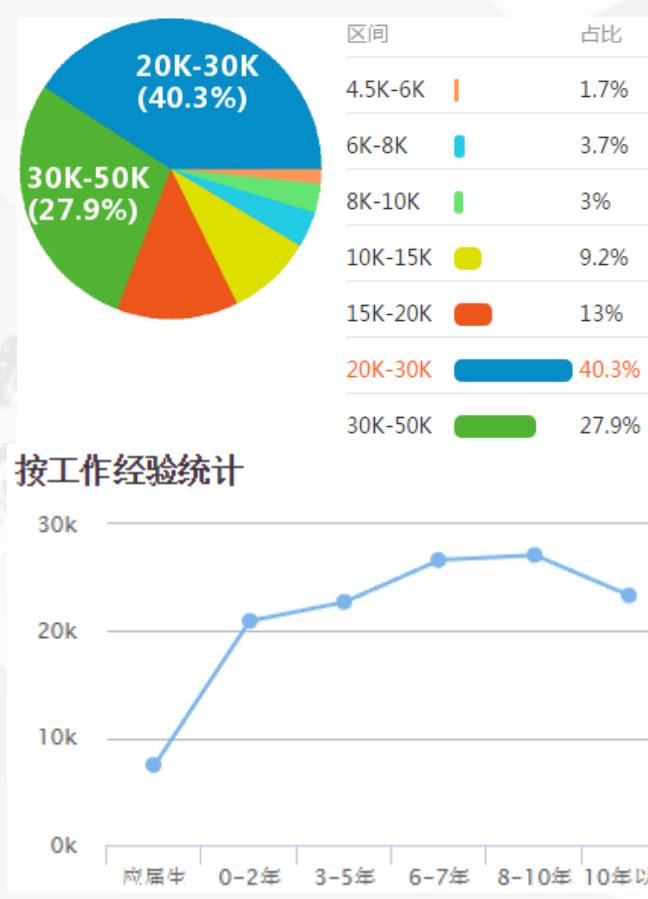
Career Path Insight

- 北京

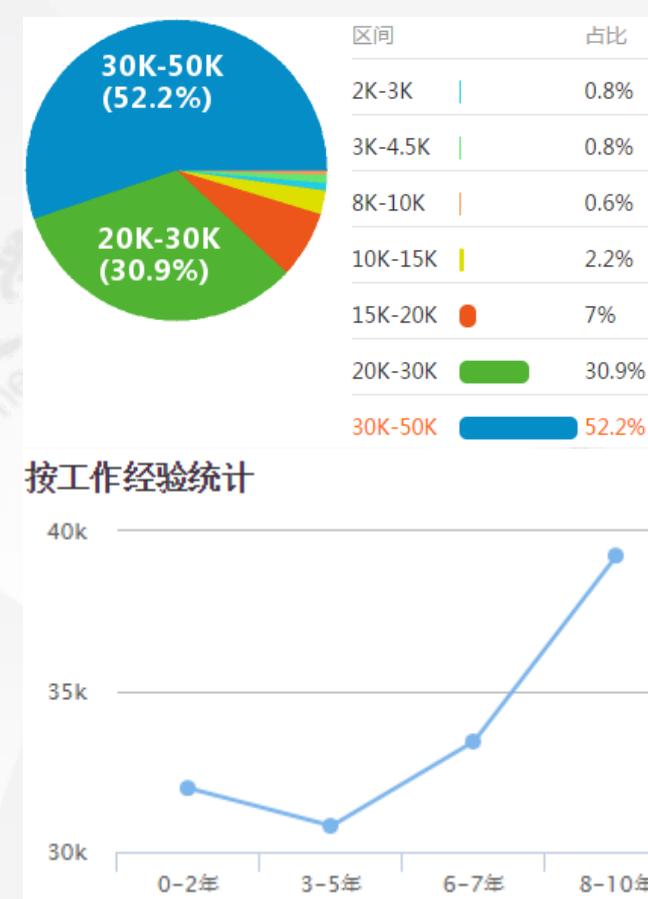
- Software Engineer (37843)



- Data Engineer (22200)

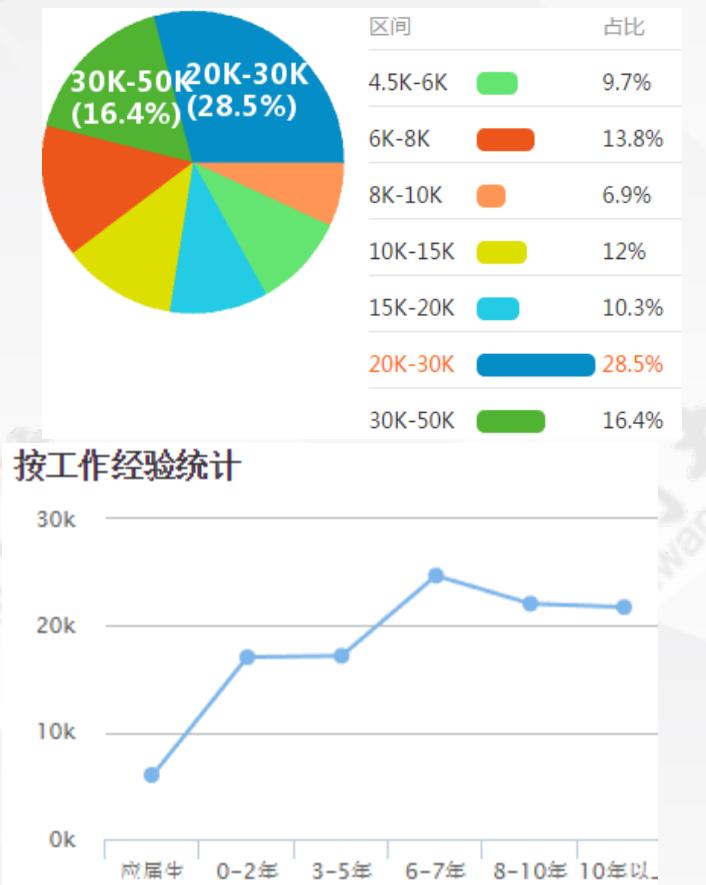


- Data Scientist (356)



Career Path Insight

- 上海
- Software Engineer (51625)



- Data Engineer (21050)



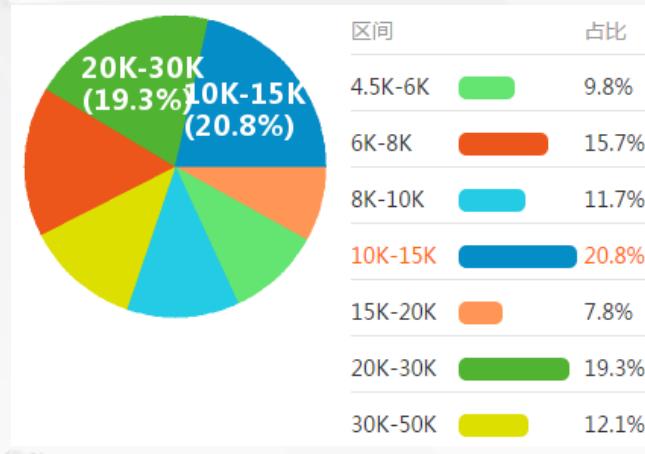
- Data Scientist (354)



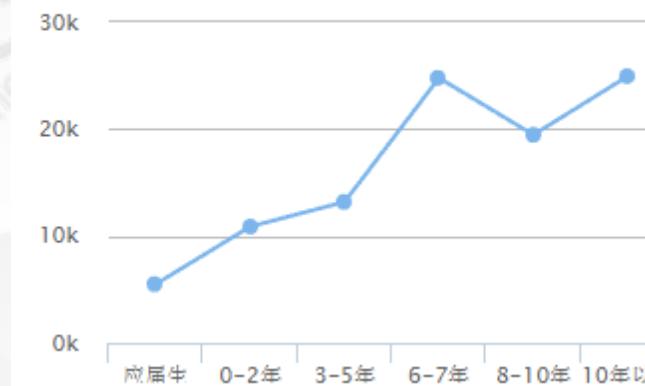
Career Path Insight

- 深圳

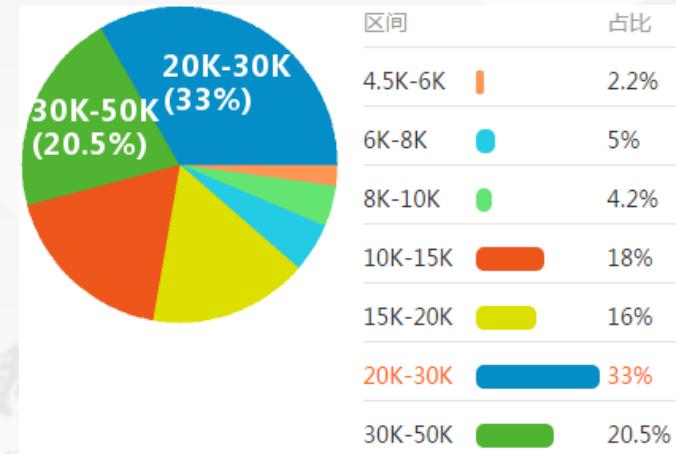
- Software Engineer (64865)



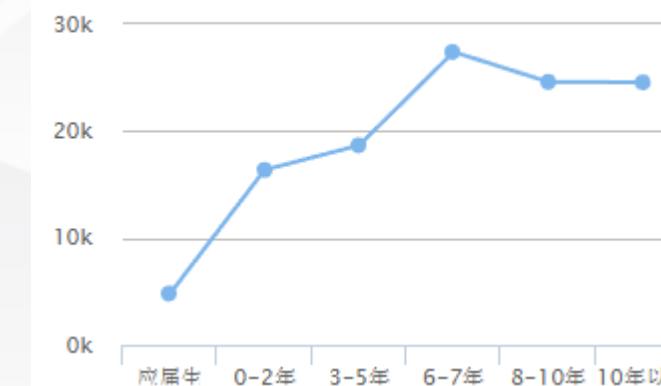
按工作经验统计



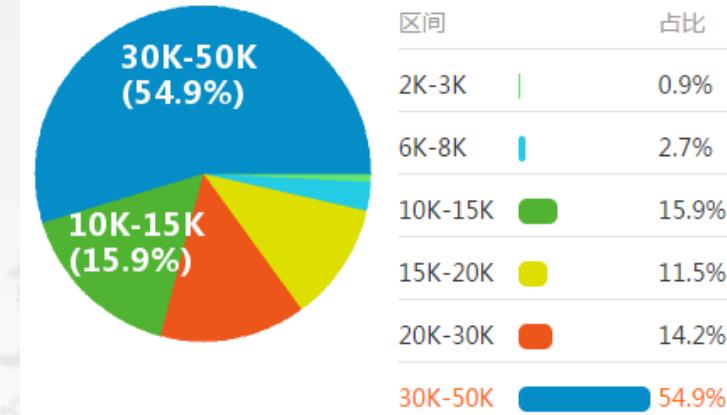
- Data Engineer (9030)



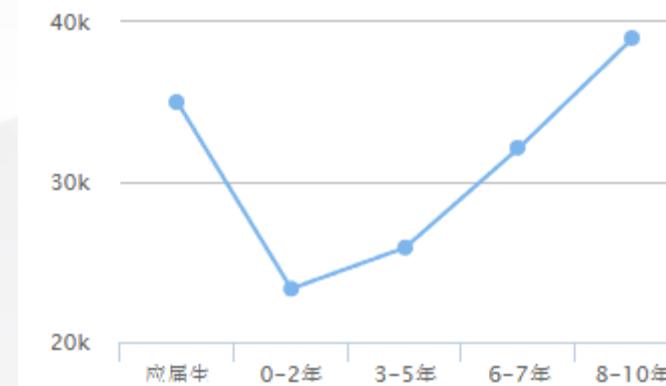
按工作经验统计



- Data Scientist (113)



按工作经验统计



Career Path Insight

- 全国软件工程师平均工资 : **¥ 10,160**/月, 取自 398,949 份样本
- 全国数据工程师平均工资 : **¥ 18,040**/月, 取自 85,685 份样本
- 全国数据科学家平均工资 : **¥ 29,420**/月, 取自 1079 份样本

Career Path Insight

- B

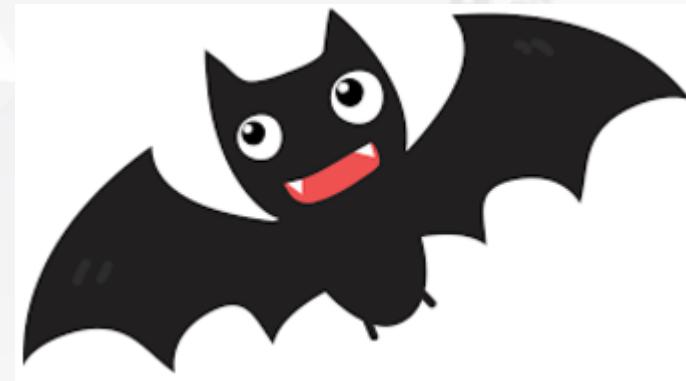
level	最高	中间值	最低	股票
T1,	140000	106000	70000	
T2	180000	136000	90000	
T3	260000	194000	130000	
T4	360000	266000	180000	
T5	450000	331000	230000	30W-40W
T6	600000	426000	290000	35W-55万
T7	700000	572000	390000	45W-70万
T8	1000000	736000	500000	70W-100万
T9	1300000	980000	660000	90W-120W

- A

级别	薪资	股票 (4年拿完)
P5	15W-25W	无
P6	20W-35W	无
P7	30W-50W	2400股
P8	45W-80W	6400股
P9	80W-100W	16000股
P10		

- T

级别		基本定义	薪酬范围
T1 应届生	1.1	低端岗位	10k+
	1.2	本科生	
	1.3	研究生	
T2	2.1	博士生	10K , 14w
	2.2	工程师	15K , 21w
	2.3	工程师	20K , 28w
T3 高级	3.1	资深工程师	23K以上 , 32w以上
	3.2	技术专家	50-70w
	3.3	总监级	60w、70w
T4 专家		总监以上，基本没有	大于100万，另外100万股票/期权



Career Path Insight

- 美国



BUREAU OF LABOR STATISTICS

	High Wage	Median Wage	Projected Job Openings	Growth%	National wide # of Jobs
Software Engineer	\$138,800	\$100,690	186,600	17%	1,114,000
System Analyst	\$129,980	\$85,800	118,600	21%	567,800
Network Engineer	\$115,300	\$77,810	30,200	8%	382,600
Electrical Design Engineer	\$148,000	\$89,630		4%	315,900
Data Scientist	\$148,000	\$110,620	2,700	11%	25600
Data Analyst	\$124,000	\$95,900	140,000	19%	495,500
Finance Analyst	\$160,760	95,320	32,300	12%	268,360

软件工程师

猿之路

- 1. 学习基本语法
- 2. 读别人的代码
- 3. 修改别人的代码
- 4. 开始写自己的小程序
- 5. 参与一个开源项目

基础课程

- 1. 数据结构和算法 
- 2. 数据库 
- 3. 系统设计
- 4. 操作系统
- 5. 计算机网络
- 6. 编译原理
- 7. 分布式系统 
- 8. 图像处理
- 9. 安全/加密解密
- 10. 人工智能 
-

面试

- 数据结构 (70-80%)
- 系统设计
- 简历 (课程, 项目, 经验, 等等) (10%)
- 计算机语言 (不限, Java, Python, C++, 等)
- 行为面试 (10%)
- 项目相关 : 操作系统, 网络, 等

面试

- 各种数据结构
- 图论 : BFS / DFS
- 递归 (permutation, combination,... 2^n , m^n , $n!$...etc. like NP problems)
- 动态规划
- 系统设计 (面向对象, 分布式系统, 网络爬虫, Hadoop, 社交网络, 推荐系统, 等)
- 基础计算机知识
- 数学, 离散数学, 统计, 概率

面试

- 算法是否过关，是否能写出递归和动规
- Coding是否过关，是否能在IDE/白板中写出Bug Free
- Design是否过关，是否能给出Tradeoff
- 项目经历整理，能够流利说出架构，难点，自己的贡献
- 加分项: GitHub, Blog, 参与Open source

Full Stack Web Development

- HTML / DHTML / CSS
- Java Script, Ajax, JQuery, AngularJS, BackboneJS
- APIs: Java API, Web API, MVC, Entity Framework, Spring, Hibernate, Ruby and Rails, Python, node.js, PHP
- Database: SQL Server, Oracle, MySQL, MongoDB, NoSQL
- Web App Servers: Tomcat, Jetty, JBoss, WebLogic, WebSphere, IIS
- For examples:
 - {Client Side Technologies} + {AngularJS} + {ASP.NET Web API, ADO.NET, Entity Frameworks, LINQ} + {SQL Server}
 - {Client Side Technologies} + {AngularJS} + {Java APIs, Ruby and Rails} + {Oracle,JDBC}
 - {Client Side Technologies} + {BackboneJS} + {Java APIs, Ruby and Rails} + {Oracle,JDBC}
 - {Client Side Technologies} + {AngularJS} + {Node.js} + {MongoDB}

大数据

- Big data is like teenage sex: everyone talks about it, nobody really knows how to do it

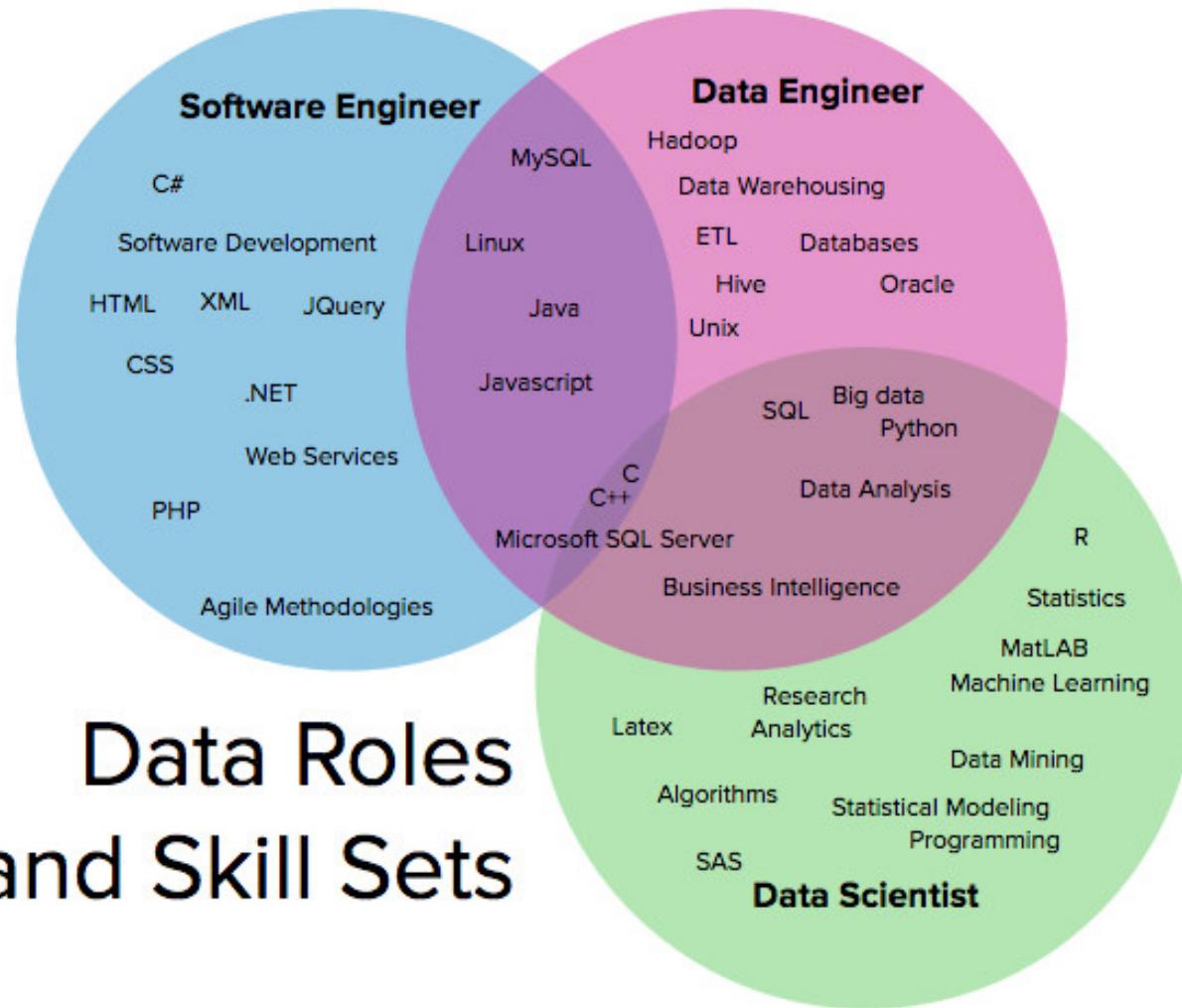
数据工程师 / 数据科学家

	Data Engineer	Data Scientist
对CS的要求	较强	一般
对统计和ML的要求	一般	较强
对具体行业的理解	一般	较深入
工资水平	~\$110k/year	>\$120k/year
Position高低	中	中高
岗位数量	大量	中等

Data Engineer和Data Scientist的市场需求与需求对比

协同合作

Data Roles and Skill Sets



数据工程师

Google Data Engineer



Strategic Customer Engineer, Big Data/Data Engineering, Google Cloud

Google

Minimum qualifications:

- Bachelors degree in Computer Science or Math (or related technical major) or equivalent practical experience.
 - 5 years of experience in leading or working on complex technical projects with proven track record of customer successes. Demonstrated ability to work under tight deadlines and prioritize several complex projects simultaneously.
 - Experience in writing software in one or more languages such as Java, Go, Python, PHP or similar.
 - Working knowledge of any of the following: NoSQL databases, distributed data processing frameworks (Hadoop), Linux, Firewalls, Proxy Servers, Load Balancing, VPN, VPC, relational database systems, distributed compute environments, application development, big data analytics, fundamentals of cloud computing including virtualization, multi tenant cloud infrastructures, storage systems and content delivery networks.

Preferred qualifications:

- MS degree or PhD in Computer Science or Math and/or a technically-oriented MBA degree.
 - 9 years of relevant experience in software development, professional services, solution engineering, technical consulting; with expertise in architecting and rolling out new technology and solution initiatives.
 - Experience with data processing software (such as Hadoop, Spark, Pig, Hive) and open source software (such as Cassandra, MongoDB, RabbitMQ, NGINX, Redis, Elasticsearch) along with data processing algorithms (MapReduce, Flume).
 - Hands-on experience in big data, information retrieval, data mining or machine learning as well as experiences in building multi-tier high availability applications with modern web technologies. NoSQL, MongoDB, Hadoop, Spark, SparkML, Tensorflow
 - Strong customer-facing communication skills. Proven success in and genuine enthusiasm for working directly with customer.
 - Willing to travel 20% (domestic and international).

Data Engineer Positions

data engineer jobs

1 - 30 of 43,936 positions

1 2 3 4 5 > >>

Senior Data Engineer



JOB SUMMARY Provides technical expertise to design and develop softw...

H&R Block Manageme... ↗ Kansas City, MO

3 weeks ago

Data Engineer



Komodo Health is transforming healthcare predictive analytics with its pro...

Komodo Health

↗ San Francisco, CA

3 weeks ago

Data Engineer



Are you compelled to measure the value of engineering and marketing? A...

Zillow.com

↗ Seattle, WA

17 hours ago

Big Data Engineer



Prudential's Customer Office Data Technology team is seeking an experie...

Prudential

↗ Newark, NJ

9 hours ago

Big Data Engineer - Hadoop



Position Summary: Position Summary: As a Big Data Engineer, you will pr...

Requirements

- * Bachelor's degree in computer science or equivalent experience
- * Two years experience with Hadoop (includes Hive and Pig)
- * Two years working in a Linux environment
- * Two years of programming experience (Java or Python preferred)
- * Two years of experience with SQL and relational databases
- * Demonstrated ability to performance-tune MapReduce jobs
- * Strong analytical and research skills
- * Demonstrated ability to work independently as well as with a team
- * Ability to troubleshoot problems and quickly resolve issues
- * Strong communication skills

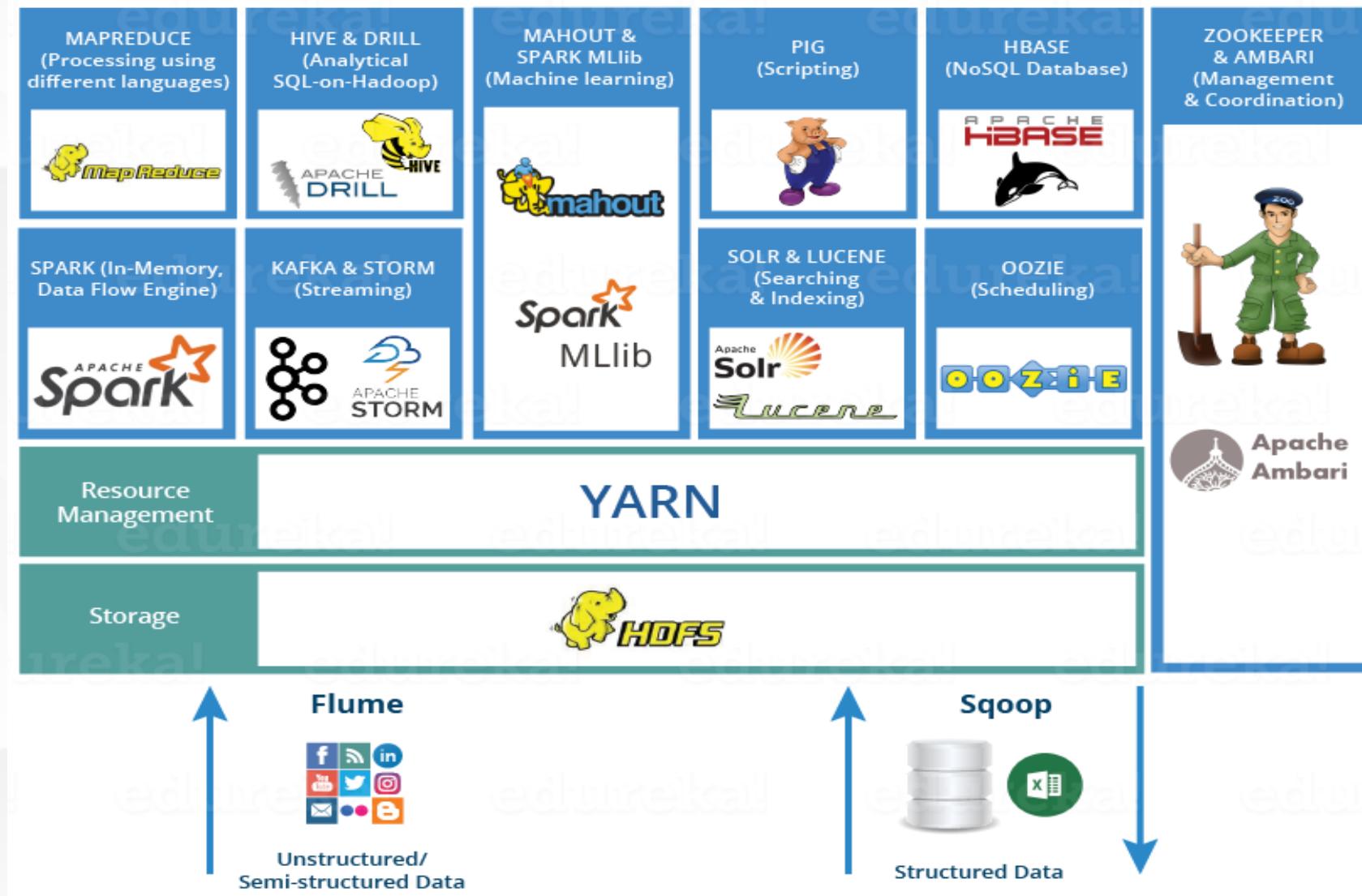
Required:

- * Programming language -- Java, Python, SQL
- * AWS - RDS, EC2, RedShift, S3 and others. AWS cloud Data migration, AWS Security
- * Database Oracle, complex SQL queries, performance tuning concepts, Backups, Recovery, DR, BCP
- * NoSQL -- HBase, MongoDB, Cassandra, Riak,
- * ETL Tools Data Stage, Informatica
- * Code/Build/Deployment -- git, svn, maven, sbt, jenkins, bamboo

Desired:

- * AWS Associate / Solution Architect certified
- * Batch processing -- Hadoop MapReduce, Cascading/Scaling, Apache Spark, AWS EMR
- * Stream processing -- Spark streaming, Apache Storm, Flink

Big Data Ecosystem



面试

- 数据结构和算法
- 编程语言
- 大数据通用处理平台
- 分布式存储与计算
- 并发程序设计
- 使用 Hadoop, Spark, Kafka, Hive 等工具
- 数据分析/数据仓库(SQL类)
- 开发ETL/数据流水线 (data pipelines)
- 数据可视化

数据科学家

Google Data Engineer



Data Scientist, Engineering

Google

Qualifications

Minimum qualifications:

- MS degree in a quantitative discipline (e.g., statistics, operations research, bioinformatics, economics, computational biology, computer science, mathematics, physics, electrical engineering, industrial engineering)
- 2 years of relevant work experience in data analysis or related field. (e.g., as a statistician / data scientist / computational biologist / bioinformatician).
- Experience with statistical software (e.g., R, Julia, MATLAB, pandas) and database languages (e.g., SQL).

Preferred qualifications:

- PhD degree in a quantitative discipline as listed in Minimum Qualifications.
- 4 years of relevant work experience (e.g., as a statistician / computational biologist / bioinformatician / data scientist), including deep expertise and experience with statistical data analysis such as linear models, multivariate analysis, stochastic models, sampling methods. Analytical engagements outside class work while at school can be included.
- Applied experience with machine learning on large datasets.
- Experience articulating business questions and using mathematical techniques to arrive at an answer using available data. Experience translating analysis results into business recommendations.
- Demonstrated skills in selecting the right statistical tools given a data analysis problem. Demonstrated effective written and verbal communication skills.
- Demonstrated leadership and self-direction. Demonstrated willingness to both teach others and learn new techniques.

面试

- 数学, 统计
- SQL
- 数据建模
- 机器学习
 - Linear Regression, Logistics Regression, Clustering, Decision Tree, Time Series, Random Number, Monte-Carlo, Bayesian, Bayes, SVM, KNN, etc.
- 算法
- 商业智能的 dashboards
- 会读paper, 会写research proposal
- R / Python / MATLAB / SAS
- 数据可视化

求职经验

- 学历 < 学校 < 实习 < 工作经验
- 简历，技能匹配
- 专业知识
- 沟通，文化

谢谢

beijing@dataapplab.com

WHICH PROGRAMMING LANGUAGE

SHOULD I LEARN FIRST?

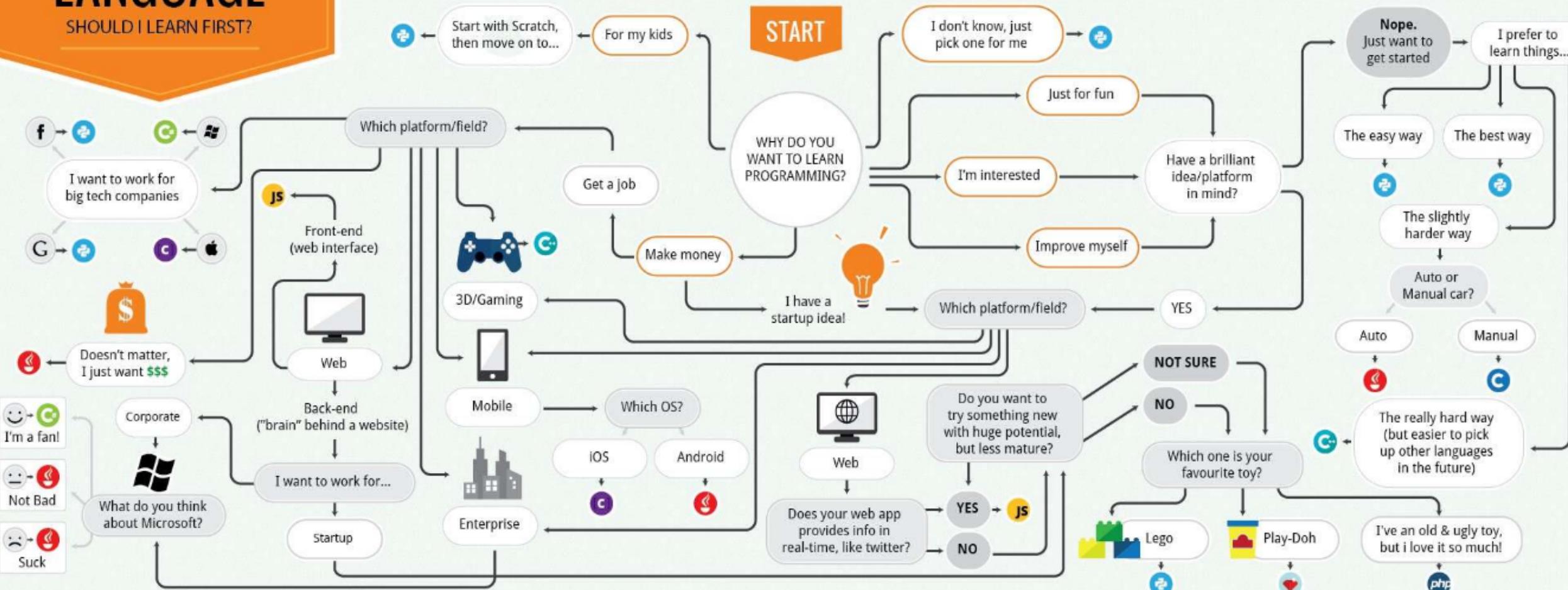


WHAT IS PROGRAMMING?

Writing very specific instructions to a very dumb, yet obedient machine.



LANGUAGES



The Lord Of The Rings

Python	Java	C	C++	JavaScript	C#	Ruby	PHP	Objective-C
POPULARITY	USED TO BUILD	POPULARITY	USED TO BUILD	POPULARITY	USED TO BUILD	POPULARITY	USED TO BUILD	POPULARITY
★★★★★	YouTube, Instagram, Spotify	★★★★★	Gmail, Minecraft, Most Android Apps, Enterprise applications	★★★★★	Operating systems, hardware	★★★★★	Enterprise and Windows applications	★★★★★
Avg. Salary \$107,000	🌐💻	Avg. Salary \$102,000	🌐📱💻	Avg. Salary \$102,000	📱💻	Avg. Salary \$104,000	🌐📱💻	Avg. Salary \$107,000