Fortune favours the bold

Chu-Cheng Hsieh
http://linkedin.com/in/chucheng

LinkedIn Data Scientist Salaries in Mountain View, CA
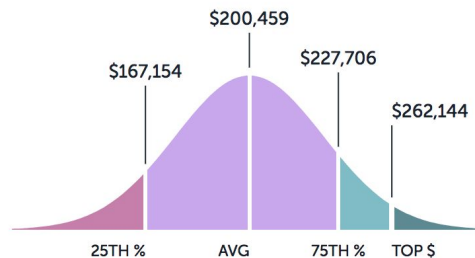
Data based on fewer than 20 profiles

How much does a LinkedIn Data Scientist in Mountain View, CA make?

A Data Scientist at LinkedIn in Mountain View, CA earns an average of $200,459, ranging from $167,154 at the 25th percentile to $227,706 at the 75th percentile, with top earners (the top 10%) earning more than $262,144. Compensation is derived from fewer than 20 profiles, including base salary, equity and bonus.

COMPENSATION | COMPANIES | LOCATIONS | TALENT | DIVERSITY | JOBS

Compensation

$200,459

$167,154

$227,706

$262,144

25TH % | AVG | 75TH % | TOP $

LinkedIn Salaries for Data Scientists in Mountain View, CA

TOTAL COMPENSATION

25TH PERCENTILE	$167,154
AVERAGE	$200,459
75TH PERCENTILE	$227,706
TOP EARNERS	$262,144

BREAK DOWN: AVERAGE

BASE	$150,408
EQUITY	$36,556
BONUS	$13,494
SIGNING BONUS	$20,367

Last updated: May 2nd, 2018.

#1 Another day in the Bay Area

#2 How to start a career

https://github.com/MrMimic/data-scientist-roadmap

Author: Swami Chandrasekaran

9. Data Munging — Denoising, Feature Extraction, Binning Sparse Values, Unbiased Estimators, Handling Missing Values, Data Scrubbing, Normalization, Dimensionality & Numerosity Reduction

8. Data Ingestion — Summary of Data Formats, Data Discovery, Data Sources & Acquisition, Data Integration, Data Fusion, Transformation & Enrichment, Data Survey

Sampling, Stratified Sampling, Principal Component Analysis, Google OpenRefine, How much Data?, Using ETL

5. Text Mining / NLP — Corpus, Named Entity Recognition, Text Analysis, UIMA, Term Document Matrix, Term Frequency & Weight, Support Vector Machines, Association Rules, Market Based Analysis, Feature Extraction, Using Mahout, Using Weka, Using NLTK

Clustering: Hierarchical Clustering, K-means Clustering
Regression: Perceptron, Linear Regression, Ranking, Logistic Regression
Classification: K-Nearest Neighbor, Naive Bayes Classifiers, Boosting, Decision Trees, Classification Rate

Neural Networks, Sentiment Analysis, Collaborative Filtering, Tagging, Vocabulary Mapping, Classify Text

Trees & Classification, Bias & Variance, Overfitting, Lift, Prediction, Classifier, Training & Test Data, Concepts, Inputs & Attributes, Unsupervised Learning, Supervised Learning, Categorical Var, Numerical Var, What is ML?

6. Visualization — Data Exploration in R (Hist, Boxplot etc), Uni, Bi & Multivariate Viz, ggplot2, Histogram & Pie (Uni), Tree & Tree Map, Scatter Plot (Bi), Line Charts (Bi), Spatial Charts, Survey Plot, Timeline, Decision Tree

Euclidean Distance, Least Fit, Causation, Pearson Coeff, Correlation

4. Machine Learning

10. Toolbox — MS Excel w/ Analysis ToolPak, Java, Python, R, R-Studio, Rattle, Weka, Knime, RapidMiner, Hadoop Dist of Choice, Spark, Storm, Flume, Scibe, Chukwa, Nutch, Talend, Scraperwiki, Webscraper, Flume, Sqoop, tm, RWeka, NLTK, RHIPE, D3.js, ggplot2, Shiny, Cassandra, MongoDB, IBM Languageware

1. Fundamentals — Matrices & Linear Algebra Fundamentals, Hash Functions, Binary Tree, O(n), Relational Algebra, DB Basics, Inner, Outer, Cross, Theta Join, CAP Theorem, Tabular Data, Data Frames & Series, Sharding, OLAP, Multidimensional Data Model, ETL, Reporting Vs BI Vs Analytics, JSON & XML, NoSQL, Regex, Vendor Landscape, Env Setup, Entropy

Prob Den Fn (PDF), ANOVA, Skewness, Continuos Distributions (Normal, Poisson, Gaussian), Cumul Dist Fn (CDF), Random Variables, Bayes Theorem, Probability Theory, Percentiles & Outliers, Histograms, Exploratory Data Analysis, Descriptive Statistics (mean, median, range, SD, Var), Pick a Dataset (UCI Repo), Install Pkgs

Central Limit Theorem, Monte Carlo Method, Hypothesis Testing, p-Value, Chi² Test, Estimation, Confid Int (CI), MLE, Kernel Density Estimate, Regression, Covariance, Tableau, IBM ManyEyes, InfoVis, D3.js

Data Frames, Reading CSV Data, Reading Raw Data, Subsetting Data, Manipulate Data Frames, Functions, Variables, Expressions, R Basics, R Setup R Studio

Lists, Factors, Arrays, Matrices, Vectors, Factor Analysis, Rapid Miner, IBM SPSS

Job & Task Tracker, MR Programming, Sqoop: Loading Data in HDFS, Flume, Scibe: For Unstruct Data, SQL with Pig, DWH with Hive, Scribe, Chukwa For Weblog, Using Mahout, Name & Data Nodes, Setup Hadoop (IBM / Cloudera / HortonWorks), Data Replication Principles, HDFS, Hadoop Components, Map Reduce Fundamentals

Zookeeper Avro, Storm: Hadoop Realtime, Rhadoop, RHIPE, rmr, Cassandra, MongoDB, Neo4j

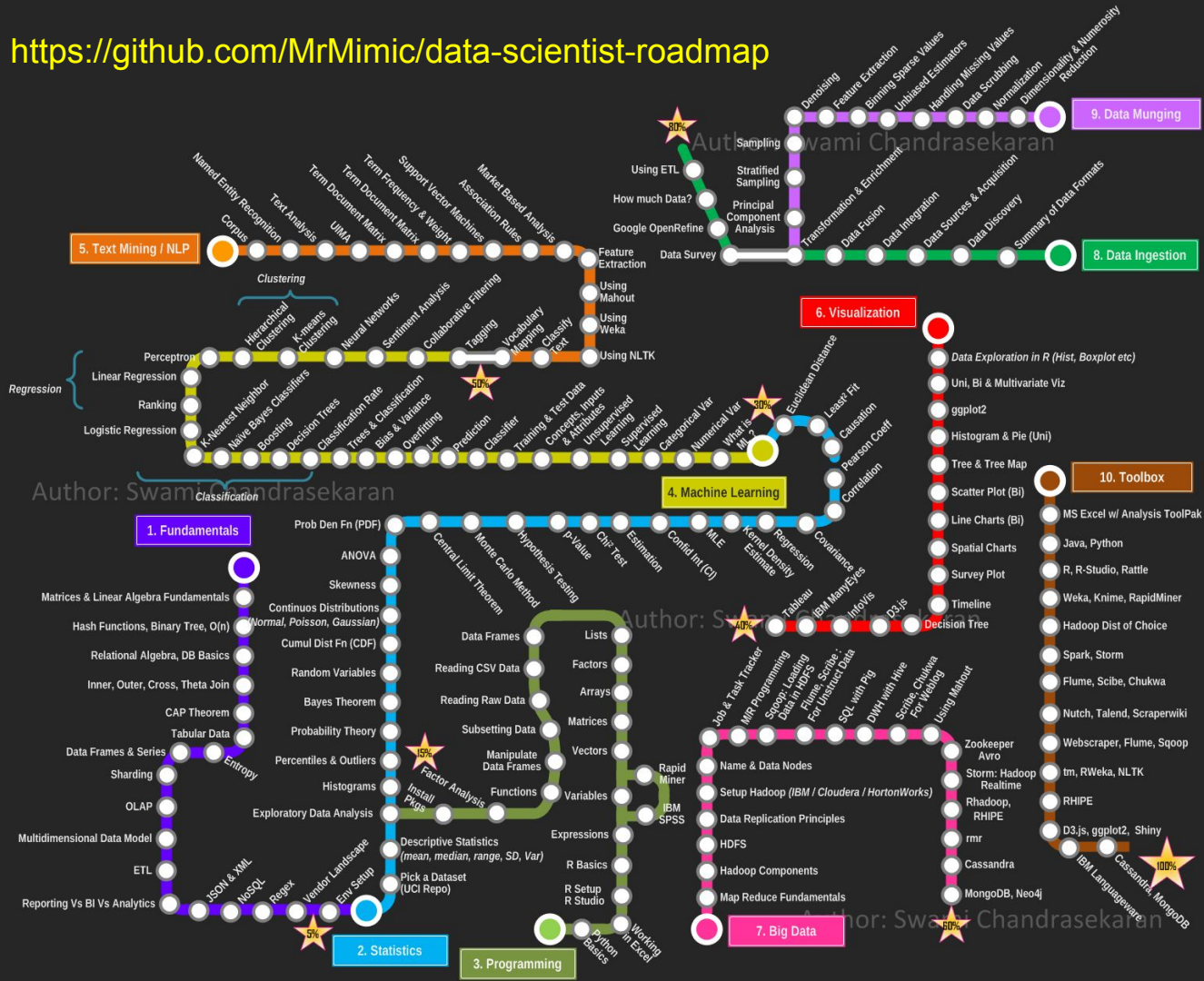**2. Statistics**

3. Programming
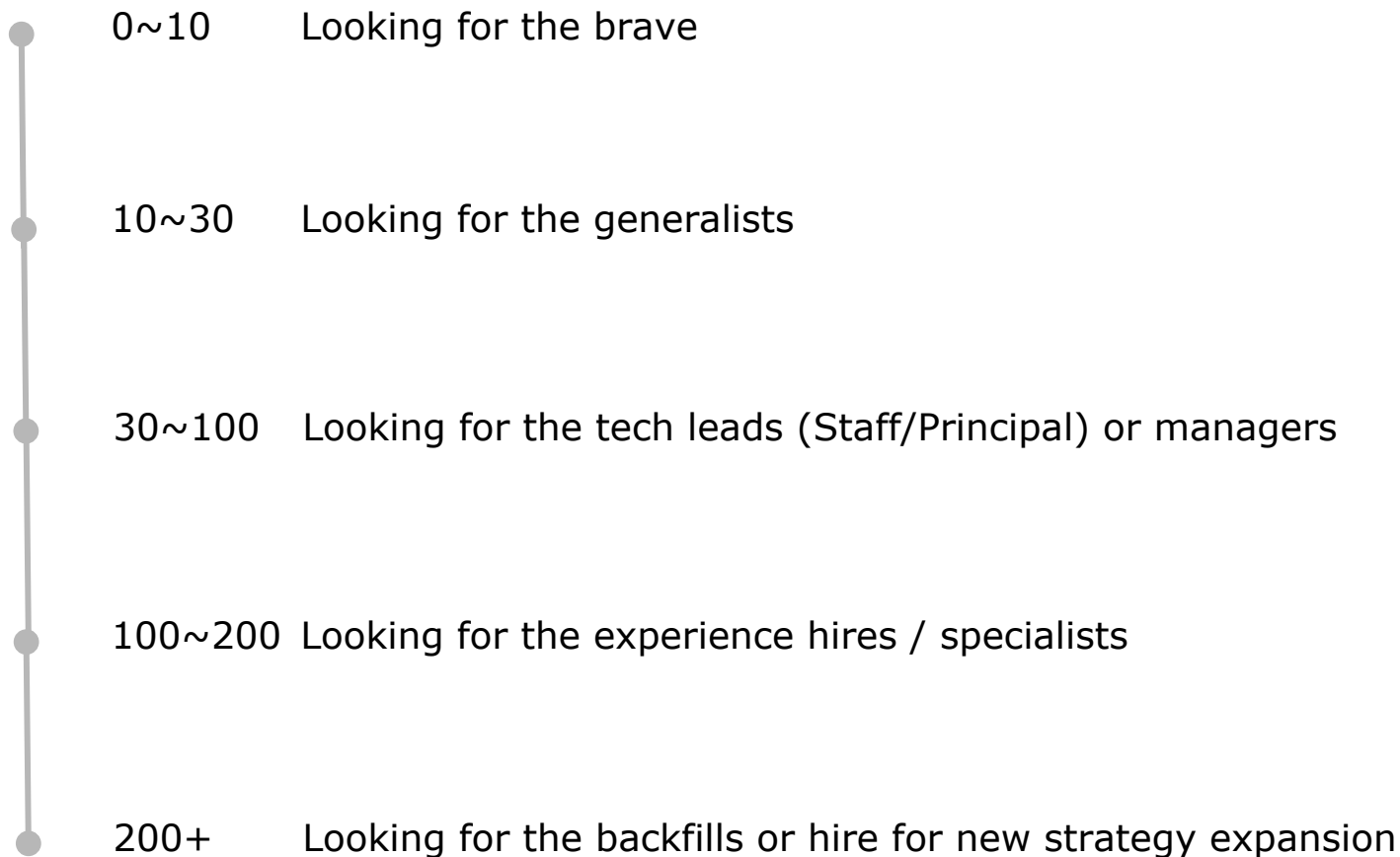
7. Big Data

Python Basics, Working in Excel

8

Find a role/team that allow you
to practice **specific skills**

Specific skills are skills that
you cannot be trained for.
If society can train you, it can train
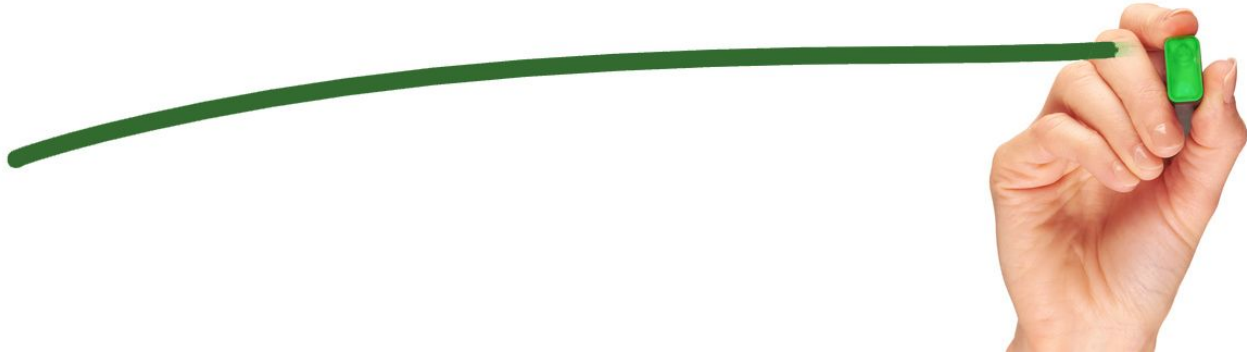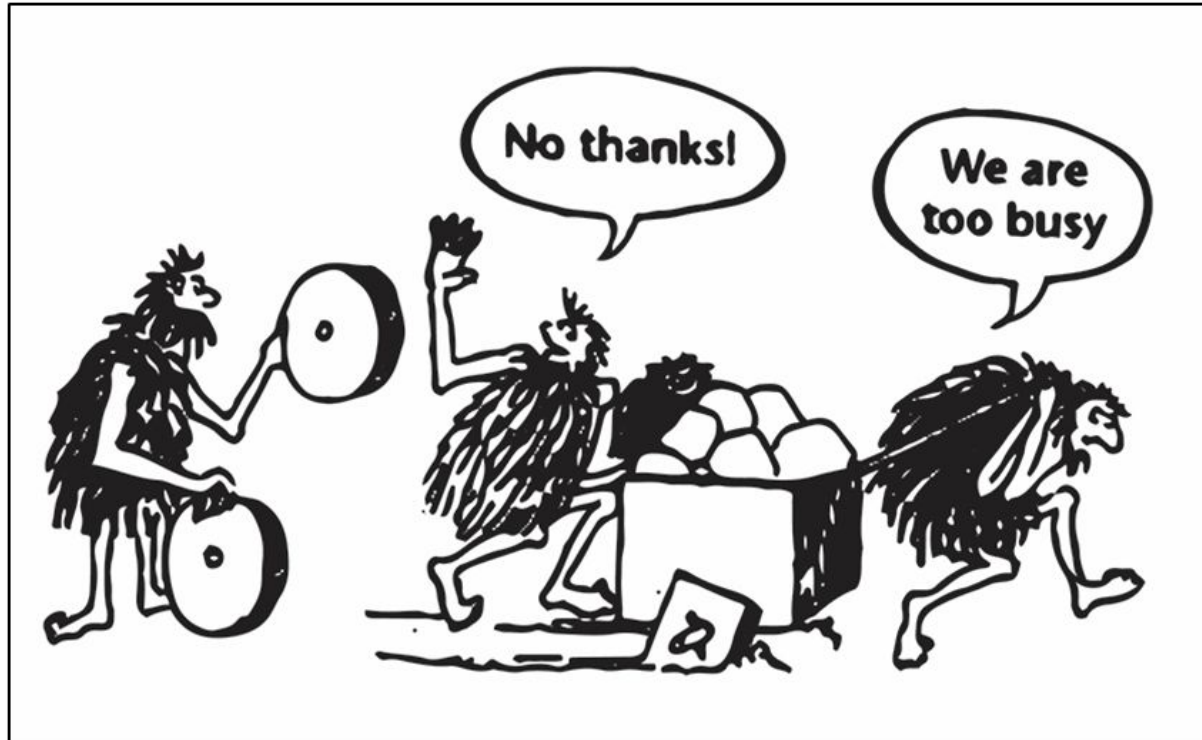someone else, and replace you.

0~10　　Looking for the brave

10~30　　Looking for the generalists

30~100　　Looking for the tech leads (Staff/Principal) or managers

100~200　Looking for the experience hires / specialists

200+　　Looking for the backfills or hire for new strategy expansion

#1 Busy is a decision

#2

MACRO PATIENCE
MICRO SPEED

#3 Thinking in bets

FOR MANY PEOPLE

TWENTY YEARS OF EXPERIENCE

IS REALLY

ONE YEAR OF EXPERIENCE
REPEATED TWENTY TIMES

—ANDY HARGADON

Book recommendations: