
Week 12

Foundations

Resumes

Source: SUSA

Overview

- Purpose
- What to Include and Formatting
- Do's, Don'ts, and Maybes
- Examples of Good Resumes
- Questions and Resume Work Time

Purpose

- Show yourself off and sell yourself in 1 page
- Secure an interview for a job/internship or a competitive club on campus
- **Don't focus on what you want, focus on what the employer wants**

Format

1. Header
2. Education
3. Experience/Extracurricular
4. Skills

Header

Name

Phone Number

Address

Email

Linkedin/Github/Personal Website if updated

Education

- High school education is okay if you're a freshman
 - DON'T put AP/SAT scores - no matter how proud you are of them
- Relevant coursework
 - Order from simplest to most complicated/specialized
- Full descriptive titles - not abbreviations

Experience

- Professional work experience
- Research
- Projects
- Volunteering
- For all of the above include company/location, role, dates, and bullet points of what you did

- Languages: Python, R, Java, ...
- Packages/Tools/Applications geared towards what you're applying for
- Soft skills **IF** you can back it up

General Tips

- Don't go above 1 page unless you have a really really good reason
- Space is limited, include only what you need
- Be clear, use spacing well
- Numbers and details are important
- Average time reviewing resumes - 16 seconds/computerized

General Tips

- Put name and contact info at very top
- Use a professional font and size (10-12pt)
- Organize into sections (with headers that stand out):
 - Header, Education, Experience, Skills, ...
- Aesthetically pleasing - templates available online

- Make descriptions specific:
 - Instead of saying “welded and soldered stuff”, say “Soldered ANSI Schedule 40 pipes using oxy-fuel welding and increased pipe yield strength by 93845 units.”
 - Adjectives are great
 - Quantifying what you did is also important
- Write out acronyms that aren’t obvious
- Proofread!
- Hobbies are ok

Don'ts

- Don't write an objective - unless it's very specific or unusual
- No paragraphs - succinct, descriptive phrases
- Don't list obvious skills - for others, list proficiency level
- For dates, only include month and year
- Don't include obvious skills (Microsoft Word/PPT, etc.)
- Don't lie!

Resume Examples:

EDUCATION

University of California, Berkeley | Berkeley, California | GPA: 3.84

B.A. in Mathematics and Computer Science

May 2020

Coursework Graduate Theoretical Statistics (Stat 210A), Graduate Computer Vision (CS 280), Machine Learning (CS 189), Convex Optimization (EE 127), Artificial Intelligence (CS 188), Efficient Algorithms (CS 170), Honors Complex Analysis (Math H185), Real Analysis (Math 104), Abstract Algebra (Math 113), Linear Algebra (Math 110)

PROFESSIONAL EXPERIENCE

Jane Street | New York

May 2018 – Aug 2018

Trading Intern

- Participated in the highly competitive trading internship program at Jane Street, focusing on financial models and market microstructure

LEADERSHIP AND EXTRACURRICULARS

Undergraduate Statistics Association | Berkeley, California

Sept 2016 – Present

Education Committee Director

- Leading a team to create a new course, hold public workshops, and teach a class of 50 students on [data science and machine learning](#)

Aug 2017 – May 2018

Project Manager

- Led a project of 11 members on utilizing statistics and machine learning to investigate the housing crisis in California
- Focused on creating data visualizations in R, time series predictions in Python, and interactive [web application](#) in JavaScript
- Selected to participate in the [Data for Good](#) competition against undergraduate/masters students, earned 2nd place and \$2800 dollars

Fall 2017

- Led project on creating time series and machine learning population models investigating the [demographic transition](#)

Fall 2016 - Spring 2017

- Worked with Environmental Progress to investigate shutting down the nuclear generator in Diablo Canyon

- Utilized R to parse millions of values and create time-series models resulting in predictions for CO₂ emissions in the future

Machine Learning at Berkeley | Berkeley, California

Jan 2018 – Present

Fall 2018

- Working with Zipline to conduct anomaly detection on drone flights using AWS S3, utilizing LSTMs and auto-encoder networks

Spring 2018

- Competed in the [Spotify Recommender System challenge](#), creating a recommendation system utilizing Siamese networks and LSTMs
- Worked with a dataset of 1 million playlists (2.2 million songs) to recommend songs to be added to the playlist

Principles and Techniques of Data Science | University of California, Berkeley

Aug 2018 – Present

Undergraduate Student Instructor

- uGSI for DS 100, teaching weekly classes and lab, hold office hours, write discussions/labs/homework for class of 800 students

Introduction to Machine Learning Staff | University of California, Berkeley

Jan 2018 – May 2018

Academic Intern

- Created lecture material for CS 189 Machine Learning at UC Berkeley, creating demos, interactive graphics, and IPython notebooks

PROJECTS [Github](#)

Optimization Research | Berkeley, California

April 2018 – Present

Advisors: Prof. Michael Jordan and Prof. Ben Recht

- Investigating gradient-free optimization algorithms with normalized step size approaches and Adagrad with Horia Mania

Tetris AI

July 2017 – Present

- Created game engines in Python and Java, AI's using Deep Q-Networks via Pytorch and a genetic algorithm with heuristics

Generative Models for Pose Transfer

Feb 2018 – May 2018

- Utilized pix2pix and OpenPose for [generatively transferring pose between persons](#), outperforming nearest neighbor approaches

PUBLICATIONS AND AWARDS

Datafest Competition 1st Place | University of California, Berkeley

Apr 2018

- Created a webapp and generalized linear models for the Datafest competition, earning 1st place and \$500 dollars

"Different Definitions of Conic Sections in Hyperbolic Geometry" with Dr. Jonathan Rosenberg

Mar 2016

Involve Research Journal, [arXiv](#)

- Research paper [published](#) in the Involve Research Journal comparing the various definitions of conic sections in hyperbolic geometry

SKILLS

Technical: Python, Java, Latex, R, Mathematica, Microsoft Excel

Resume Examples:



Education

University of California, Berkeley

B.A. Computer Science, Aug 2016 - May 2020, Technical GPA: 4.0

Relevant Coursework

- * CS 170: Efficient Algorithms and Intractable Problems, CS 188: Introduction to Artificial Intelligence (Fall 2018), EECS 127: Optimization Models in Engineering (Fall 2018), CS 161: Computer Security (Fall 2018), Data Science 100: Principles and Techniques of Data Science, Stat 198: Data Science with Kaggle
- Activities & Societies: Statistics Undergraduate Students Association (SUSA) Education Committee, Upsilon Pi Epsilon (Computer Science Honor Society), Computer Science Mentors, UC Jazz

Work Experience

Palo Alto Networks

Summer 2018

Software Engineering Intern

- Created a backend in Java Spring Boot for a ReactJS website for the customer support team to search for information about firewalls and licenses while on service calls
- Wrote microservices using Python Tornado with a Apache Cassandra/Solr backend, deployed using Docker, Kubernetes and Istio designed to handle 60 million update requests per day (70 per second) from PANW products
- Won first place in a security hackathon (Capture the Flag) by exploiting the most vulnerabilities in a website

Passage AI

Summer 2017

Software Developer Intern

- Wrote APIs using AWS Lambda to create microservices for IBM Watson services
- Created a version of NLU/P stack using IBM Watson services such as sentiment analysis and entity extraction
- Extended the platform to voice-enabled conversational interfaces such as Amazon Alexa

UC Berkeley Computer Science Department

Aug. 2018 - Present

Undergraduate Student Instructor, CS 170: Efficient Algorithms and Intractable Problems

- Taught a section of 20-30 students topics from the most widely taken upper division CS class at UC Berkeley

Projects

Megafind

October 2017

A webapp that allows professors to have a live lecture session with students that provides a real-time transcription of the lecture, with relevant terms hyperlinked to external resources, and the professor's presentation parsed and converted into notes. Built using Node.js, websockets, and Google Cloud Machine Learning API. Won 1st Place Overall and Best Education Hack at Cal Hacks 4.0, the largest collegiate hackathon in the world.

Hermes

January 2018 - Present

An interactive student experience for watching lecture videos. Allows professors to pass a URL to a webcast for a lecture and provides a rich interface for students to view the lecture, take notes, ask questions, and write code. Building for use by Berkeley CS students in the fall semester of CS 61A.

Uizard - Automating UI Code

Spring 2018

Through the Data Consulting Committee of SUSA, consulted for the startup Uizard to create a convolutional autoencoder using Keras and PyTorch.

Bearmax

November 2016

A chatbot that uses IBM Watson to help the user diagnose themselves with a sickness without having to go to a doctor's office. Won Best Beginner Hack and Best Health App at Cal Hacks 3.0

Card Cimator

September 2015

Developed a Chrome extension that auto-creates debate citations using the metadata on web pages, used by my high school debate team and several other teams in the Coast Forensic League

Success Predictor

August 2018

Created a webapp for Seabed.vc that scrapes LinkedIn for information on a company's founders and predicts the probability of success.

Awards

- 1st Place Overall, Cal Hacks 4.0
- Best Education Hack, Cal Hacks 4.0
- Winner, Cal DataFest 2018
- Best Beginner Hack, Cal Hacks 3.0
- Best Health App, Cal Hacks 3.0
- 1st Place, Palo Alto Networks Capture the Flag
- 3rd Place, CSUA Hackathon
- Louis Armstrong Outstanding Musician Award, Gunn High School
- Outstanding Achievement Award, CS Department, Gunn High School

Skills

Languages: Python, Java, SQL, Node.js, Web Development, C

Applications: Flask, AWS, Docker, ReactJS, Spring Boot, Heroku, scikit-learn, Selenium, Keras, Apache Cassandra, MongoDB

Resume Examples:

<div><div></div><div></div></div>	
ACADEMICS	
University of California, Berkeley Computer Science & Applied Math (3.8 GPA)	Expected 2020
Graduated East Brunswick High School, New Jersey	June 2016
Selected Participant for The New York Math Circle	June-July 2014
Worked on number theory applications under the instruction of Putnam Fellows at NYU's Courant Institute.	
WORK/RESEARCH EXPERIENCE	
Stanford Research Institute (SRI International)	June-Aug 2018
<i>Computer Vision Intern</i>	
<ul style="list-style-type: none">Developed novel training regime to eliminate multiplications in Deep Neural Networks during inference.Reduced storage by 90%, energy consumption by 80%, increased computation by 10, 95% retention in accuracy.Paper submitted to AAAI, patent filed for learned variable length precision systems used for efficient inference.	
Berkeley Deep Drive	Feb 2018- Present
<i>Undergraduate Researcher</i>	
<ul style="list-style-type: none">Investigating state-of-the-art technologies in computer vision & machine learning for automotive applications.Reinforcement Learning Based Channel Sparsification for Efficient Inference in ConvNets.Deep Reinforcement Learning based Optimization of Autonomous Vehicle traffic in multi-agent environments.	
Search for Extraterrestrial Intelligence (SETI) Research Center	Sept 2017-Feb 2018
<i>Machine Learning Researcher</i>	
<ul style="list-style-type: none">Researched deep learning methods for fast radio burst identification and radio frequency mitigation.Implemented localization networks (Faster R-CNN & Mask R-CNN) and denoising autoencoders.	
RISE Lab at University of California, Berkeley	Sept 2016-Sept 2017
<i>System Administrator + Undergraduate Researcher</i>	
<ul style="list-style-type: none">Configured continuous delivery pipeline and managed infrastructure using Ansible, Docker, and Jenkins.Conducted research to deploy low latency ML models using multi-armed bandit selection policy for Clipper.	
Commvault: Data and Information Management Software	June-Aug 2016
<i>Software Developer Intern</i>	
<ul style="list-style-type: none">Developed localization framework, form based authentication, and system for managing authorized keys.Improved accessibility of company's webpages and helped with maintenance in cloud migration.	
ACTIVITIES	
Machine Learning @ Berkeley - Project Manager	Sept 2018- Present
<ul style="list-style-type: none">Assisting Zipline in deploying and detecting anomalies for autonomous drones that deliver lifesaving supplies.	
Principles and Techniques of Data Science - Academic Intern	Jan- June 2018
<ul style="list-style-type: none">Assisted students in data collection, cleaning, visualization, and predictive modelling through lab and coursework.	
Statistics Undergraduate Student Association - Data Consultant	Sept 2017-May 2018
<ul style="list-style-type: none">Provided data consulting for a variety of startups using deep learning and modern statistical methods.	
Berkeley Quant Club - Project Manager	Sept 2016- May 2017
<ul style="list-style-type: none">Modeled stock market trends with machine learning and quantitative techniques in Python and Quantopian.	
PROJECTS	
Code2Pix - Data Consulting for Uizard	Jan-April 2018
<ul style="list-style-type: none">Uizard released Pix2Code, the first paper in the field of generating code from GUIs, our project was Code2Pix, a deep learning compiler, used to improve the accuracy of the original task in a generative-adversarial manner.	
Emotion Classification - Data Consulting for IPMD Inc.	Aug-Dec 2017
<ul style="list-style-type: none">Implemented novel ensemble convolutional network architecture for emotion classification with biomedical devices.	
Albumify - MHacks (Won Top 10)	June 2017
<ul style="list-style-type: none">Web application deployed on Google App Engine using style transfer to combine album art with profile pictures.	
SKILLS	
<ul style="list-style-type: none">Ansible, AWS, Docker, \LaTeX, Java, Javascript, Jenkins, Python, SQL, UnixData Analysis (Matplotlib, Numpy, Pandas), Deep Learning (PyTorch, Keras, Tensorflow)	

Writing Resume Tools

Word

- built in templates
- templates uploaded to bCourses

LaTeX

Online Resume creators (watch out for pay to download sites)

Other Tools

Cover Letters

LinkedIn

Personal Website

Github

Review

Slides from data8.org

Definition of r

Correlation Coefficient (r) =

average of	product of	x in standard units	and	y in standard units
---------------	------------	---------------------------	-----	---------------------------

Measures how clustered the scatter is around a straight line

The Correlation Coefficient r

- Measures **linear** association
- Based on standard units
- $-1 \leq r \leq 1$
 - $r = 1$: scatter is perfect straight line sloping up
 - $r = -1$: scatter is perfect straight line sloping down
- $r = 0$: No linear association; *uncorrelated*

(Demo)

Regression to the Mean

A statement about x and y pairs

- Measured in *standard units*
- Describing the deviation of x from 0 (the average of x 's)
- And the deviation of y from 0 (the average of y 's)

On average, y deviates from 0 less than x deviates from 0

Regression
Line

$$y_{(\text{su})} = r \times x_{(\text{su})}$$

Correlation

Not true for all points — a statement about averages

Regression Line Equation

In original units, the regression line has this equation:

$$\frac{\text{estimate of } y - \text{average of } y}{\text{SD of } y} = r \times \frac{\text{the given } x - \text{average of } x}{\text{SD of } x}$$

estimated y in standard units

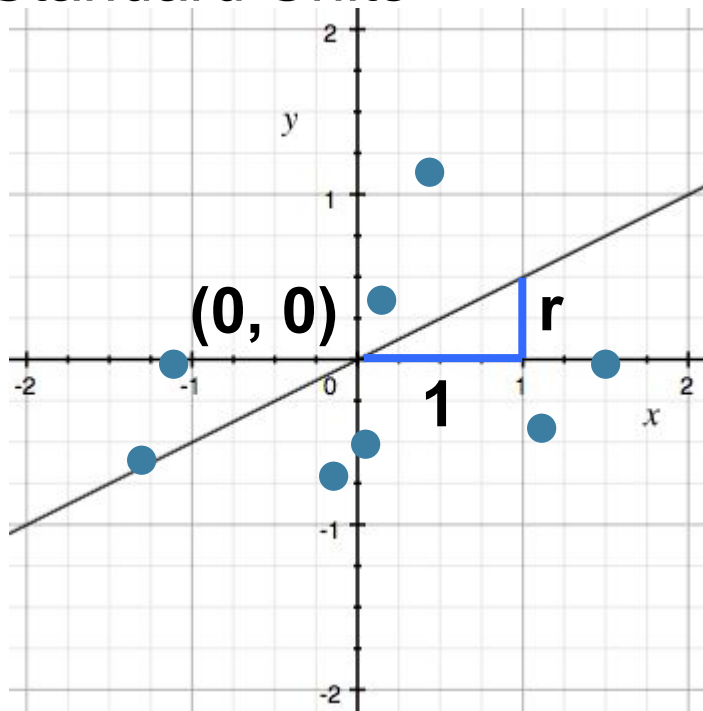
x in standard units

Lines can be expressed by *slope* & *intercept*

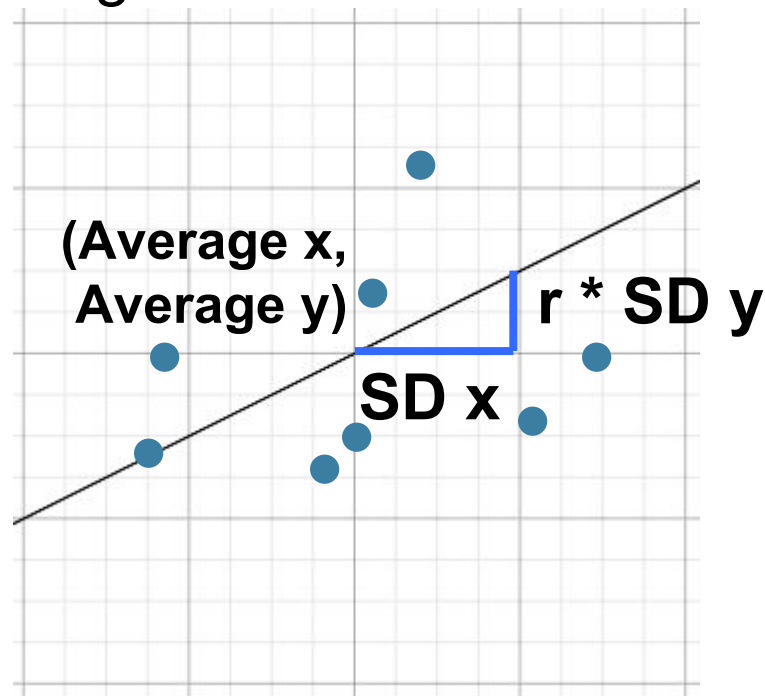
$$y = \text{slope} \times x + \text{intercept}$$

Regression Line

Standard Units



Original Units



Slope and Intercept

estimate of y = slope * x + intercept

$$\text{slope of the regression line} = r \cdot \frac{\text{SD of } y}{\text{SD of } x}$$

$$\text{intercept of the regression line} = \text{average of } y - \text{slope} \cdot \text{average of } x$$

(Demo)

How Big are Most of the Values?

No matter what the shape of the distribution,
the bulk of the data are in the range “average \pm a few SDs”

Chebyshev's Inequality

No matter what the shape of the distribution,
the proportion of values in the range “average $\pm z$ SDs” is
at least $1 - 1/z^2$

Chebyshev's Bounds

Range	Proportion
average \pm 2 SDs	at least $1 - 1/4$ (75%)
average \pm 3 SDs	at least $1 - 1/9$ (88.888...%)
average \pm 4 SDs	at least $1 - 1/16$ (93.75%)
average \pm 5 SDs	at least $1 - 1/25$ (96%)

No matter what the distribution looks like

Standard Units

- How many SDs above average?
 - **$z = (\text{value} - \text{average})/\text{SD}$**
 - Negative z : value below average
 - Positive z : value above average
 - $z = 0$: value equal to average
 - When values are in standard units: average = 0, SD = 1
 - Chebyshev: At least 96% of the values of z are between -5 and 5
- (Demo)
-

How Big are Most of the Values?

No matter what the shape of the distribution,
the bulk of the data are in the range “average \pm a few SDs”

If a histogram is bell-shaped, then

- Almost all of the data are in the range
“average \pm 3 SDs”

Bounds and Normal Approximations

Percent in Range	All Distributions	Normal Distribution
average \pm 1 SD	at least 0%	about 68%
average \pm 2 SDs	at least 75%	about 95%
average \pm 3 SDs	at least 88.888...%	about 99.73%

Spring 2017 Practice Final

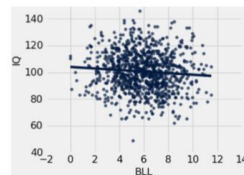
3. (25 points) Regression

The `lead` table (left) contains one row per child in a study of 1000 children's Blood Lead Levels (BLL) measured in micrograms per deciliter and their intelligence quotients (IQ). Assume that the data were collected by sampling children at random from a very large population. Summary statistics (middle) and a scatter diagram (right) are shown below. All BLLs are measured to one decimal place, and all IQ scores are integers.

BLL	IQ
7.9	90
6.2	78
3.2	110
4.1	128
7.3	88

(995 rows omitted)

Expression	Value
<code>np.average(lead.column('BLL'))</code>	6
<code>np.std(lead.column('BLL'))</code>	2
<code>np.average(lead.column('IQ'))</code>	100
<code>np.std(lead.column('IQ'))</code>	15
<code>correlation(lead, 'BLL', 'IQ')</code>	-0.1



- (a) (2 pt) What is the value of `correlation(lead, 'IQ', 'BLL')`?

Hint: The `correlation` function appears on your midterm study guide.

- (b) (3 pt) What is the estimated average IQ of a child with a BLL that is 2 standard deviations above the mean BLL? Use the regression line to find this estimate, assuming BLL and IQ are linearly related.

- (c) (4 pt) Write the equation of the regression line through this sample for the IQ y in terms of the BLL x .

$y =$

Spring 2017 Practice Final Answer

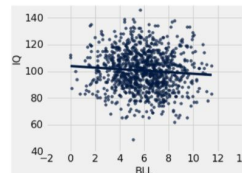
3. (25 points) Regression

The `lead` table (left) contains one row per child in a study of 1000 children's Blood Lead Levels (BLL) measured in micrograms per deciliter and their intelligence quotients (IQ). Assume that the data were collected by sampling children at random from a very large population. Summary statistics (middle) and a scatter diagram (right) are shown below. All BLLs are measured to one decimal place, and all IQ scores are integers.

BLL	IQ
7.9	90
6.2	78
3.2	110
4.1	128
7.3	88

(995 rows omitted)

Expression	Value
<code>np.average(lead.column('BLL'))</code>	6
<code>np.std(lead.column('BLL'))</code>	2
<code>np.average(lead.column('IQ'))</code>	100
<code>np.std(lead.column('IQ'))</code>	15
<code>correlation(lead, 'BLL', 'IQ')</code>	-0.1



- (a) (2 pt) What is the value of `correlation(lead, 'IQ', 'BLL')`?

Hint: The `correlation` function appears on your midterm study guide.

-0.1

- (b) (3 pt) What is the estimated average IQ of a child with a BLL that is 2 standard deviations above the mean BLL? Use the regression line to find this estimate, assuming BLL and IQ are linearly related.

$$2 * -0.1 * 15 + 100 = 97$$

- (c) (4 pt) Write the equation of the regression line through this sample for the IQ y in terms of the BLL x .

$$y = -0.1 \cdot \frac{15}{2} \cdot x + 100 - 6 \cdot (-0.1 \cdot \frac{15}{2}) = -0.75x + 104.5$$