

# Data Science your Job Search

TDP NY - September 2018

Amit Bhatia & Maryam Jahanshahi

TapRecruit

<https://tinyurl.com/tdp-workshop-2018>

**tap**Recruit.co  
Write every job perfectly

1 80% of candidates apply to the same 20% of roles.

Strategies to improve job app success.

Search, titles, companies.

2 In job search, only 20% of the work is the amazing resume.

Mechanisms to 'warm' job apps.

Recruiters, referrals, networks.

3 Resume tailoring creates impactful resumes.

Generating bespoke resumes for jobs.

Format, structure, the 4Ws.

# Skills typically associated with data-associated roles

Metric definition  
Building dashboards  
Descriptive statistics & forecasting  
Experiment design & analysis  
Data modeling  
Statistical modeling & inference  
Machine learning  
Algorithmic optimization  
Data engineering

Business

Modeling

Production

# Conventional data-associated roles

Business	Modeling	Production
Data Analyst	Data Scientist	Data Engineer
Business Analyst	Quantitative Analyst	ML Engineer Software Engineer
BI Analyst Product Analyst	Product Data Scientist Research Scientist	Infrastructure Engineer Data Platform Developer BI Developer

# Arbitrage opportunities based on search

Data Analyst

**Senior** Data Analyst

Data Analyst **New York**

**Technical** Data Analyst

Data **Associate**

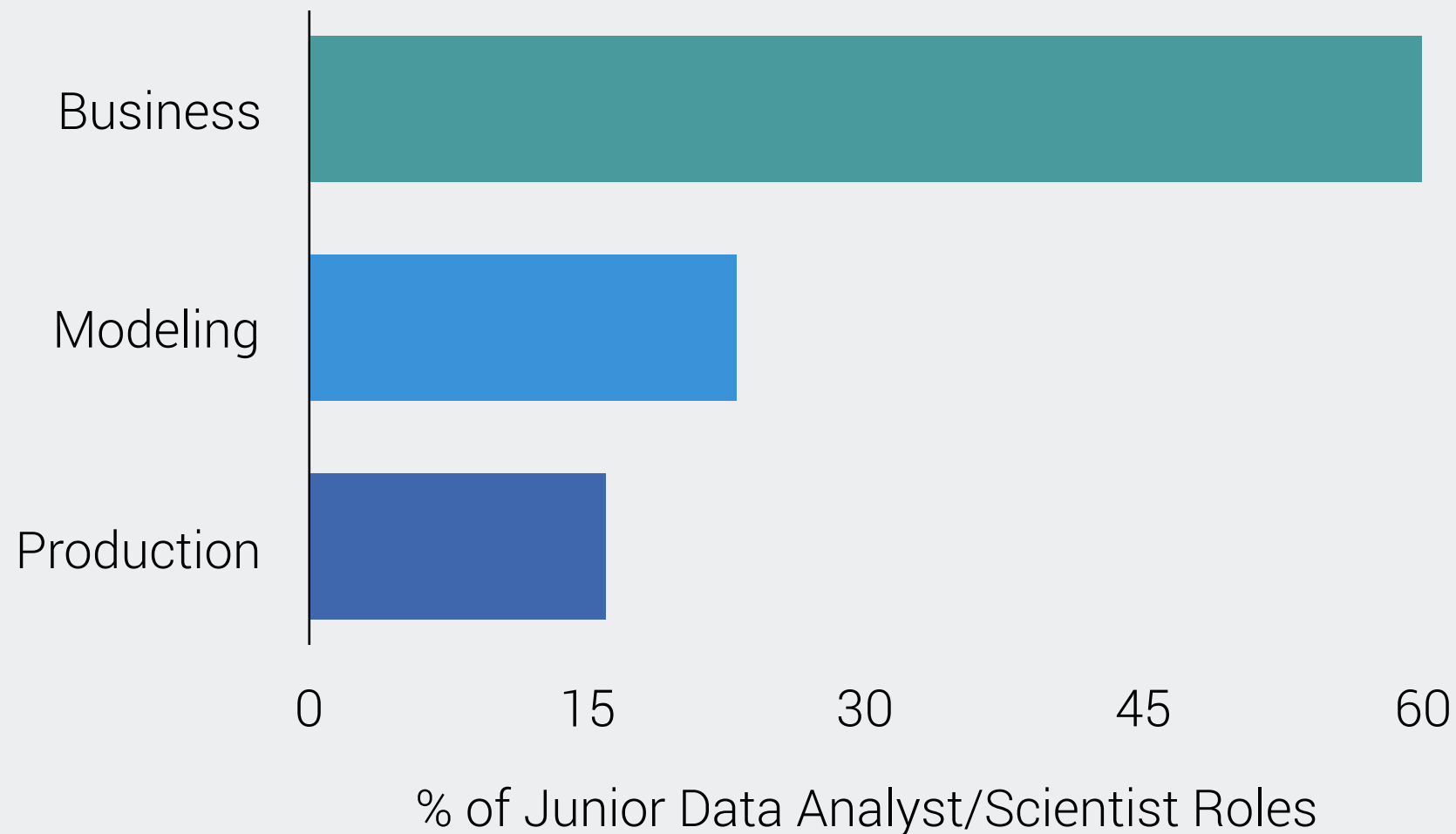
Fewer job seekers search for senior and when they find those roles, they apply at lower rates

Many companies add extra phrases to job titles which are typically penalized by search engines when candidates search for 'Data Analyst'.  
Fewer views = fewer applications!

There many data associate positions and these roles typically attract fewer applications than data analyst positions.

# Business-focussed roles comprise the biggest group of data science jobs

There are **three times** as many business-focussed data science roles as those focussed on modeling or production.



# Where are all the business-associated data roles?

Reporting

Reporting Analyst, Business Intelligence Analyst, Growth Analyst, Operations Research Analyst

Finance

Finance Analyst, Revenue Analyst

Sales

Sales Analyst, Sales Operations Analyst, Sales Engineer

Marketing

Advertising Analyst, Marketing Analyst, Marketing Operations Analyst, Web Analyst

Human Resources

HR Analyst, People Analyst, Workforce Analyst

Operations

Operations Analyst, Logistics Analyst, Supply Chain Analyst

1

80% of candidates  
apply to the same  
20% of roles.

Strategies to improve  
job app success.

Search, titles, companies.

2

In job search, only  
20% of the work is  
the amazing resume.

Mechanisms to  
'warm' job apps.

Recruiters, referrals, networks.

3

Resume tailoring  
creates impactful  
resumes.

Generating bespoke  
resumes for jobs.

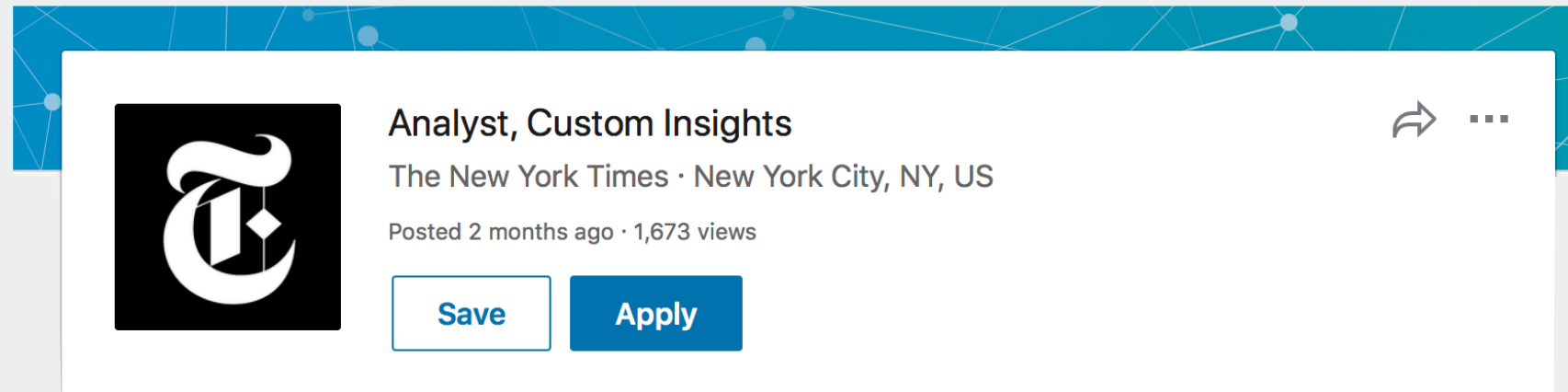
Format, structure, the 4Ws.



# There are different levels of warm applications



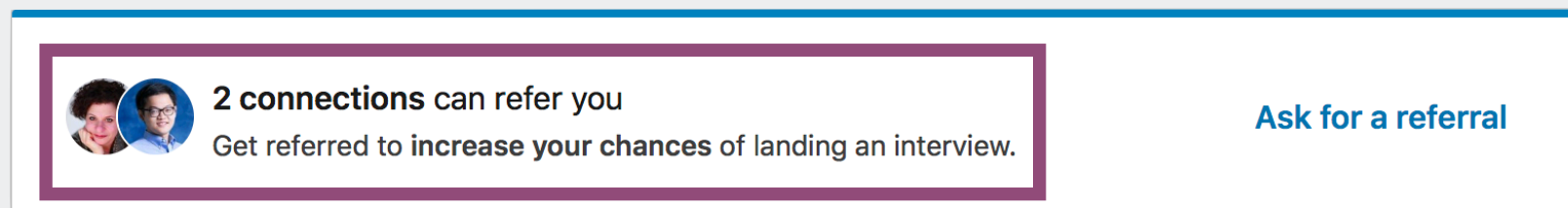
# LinkedIn can help you warm up your applications



The image shows a LinkedIn job post for 'Analyst, Custom Insights' at 'The New York Times' in 'New York City, NY, US'. The post was made '2 months ago' and has '1,673 views'. It includes a 'Save' button and an 'Apply' button. The New York Times logo is on the left, and a share icon and a three-dot menu are on the right.

**Analyst, Custom Insights**  
The New York Times · New York City, NY, US  
Posted 2 months ago · 1,673 views

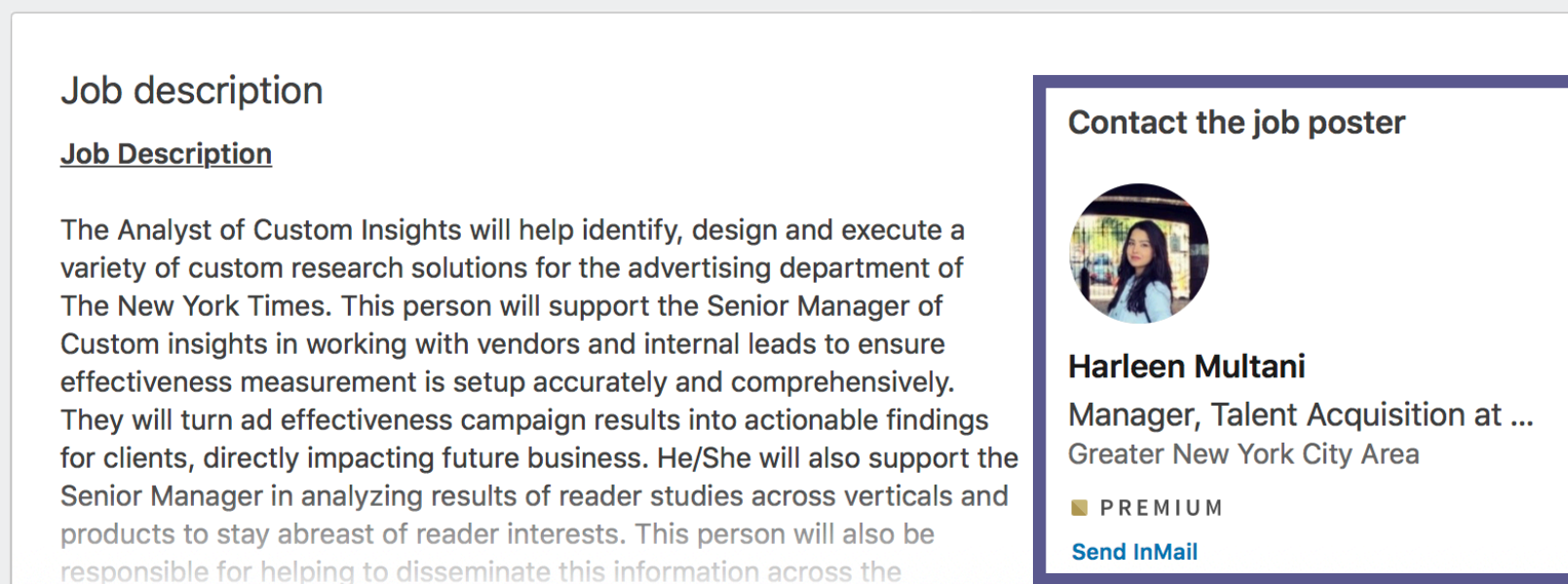
[Save](#) [Apply](#)



The image shows a referral section with two profile pictures and the text '2 connections can refer you'. Below this, it says 'Get referred to increase your chances of landing an interview.' and a blue button labeled 'Ask for a referral'.

**2 connections** can refer you  
Get referred to **increase your chances** of landing an interview.

[Ask for a referral](#)




The image shows two sections. The left section is titled 'Job description' and contains a paragraph about the role of an Analyst of Custom Insights at The New York Times. The right section is titled 'Contact the job poster' and features a profile for Harleen Multani, Manager of Talent Acquisition at The New York Times, with a 'Send InMail' button.

**Job description**  
Job Description

The Analyst of Custom Insights will help identify, design and execute a variety of custom research solutions for the advertising department of The New York Times. This person will support the Senior Manager of Custom insights in working with vendors and internal leads to ensure effectiveness measurement is setup accurately and comprehensively. They will turn ad effectiveness campaign results into actionable findings for clients, directly impacting future business. He/She will also support the Senior Manager in analyzing results of reader studies across verticals and products to stay abreast of reader interests. This person will also be responsible for helping to disseminate this information across the

**Contact the job poster**

  
**Harleen Multani**  
Manager, Talent Acquisition at ...  
Greater New York City Area  
PREMIUM  
[Send InMail](#)



Help Robin identify the recruiter/hiring manager and refer you

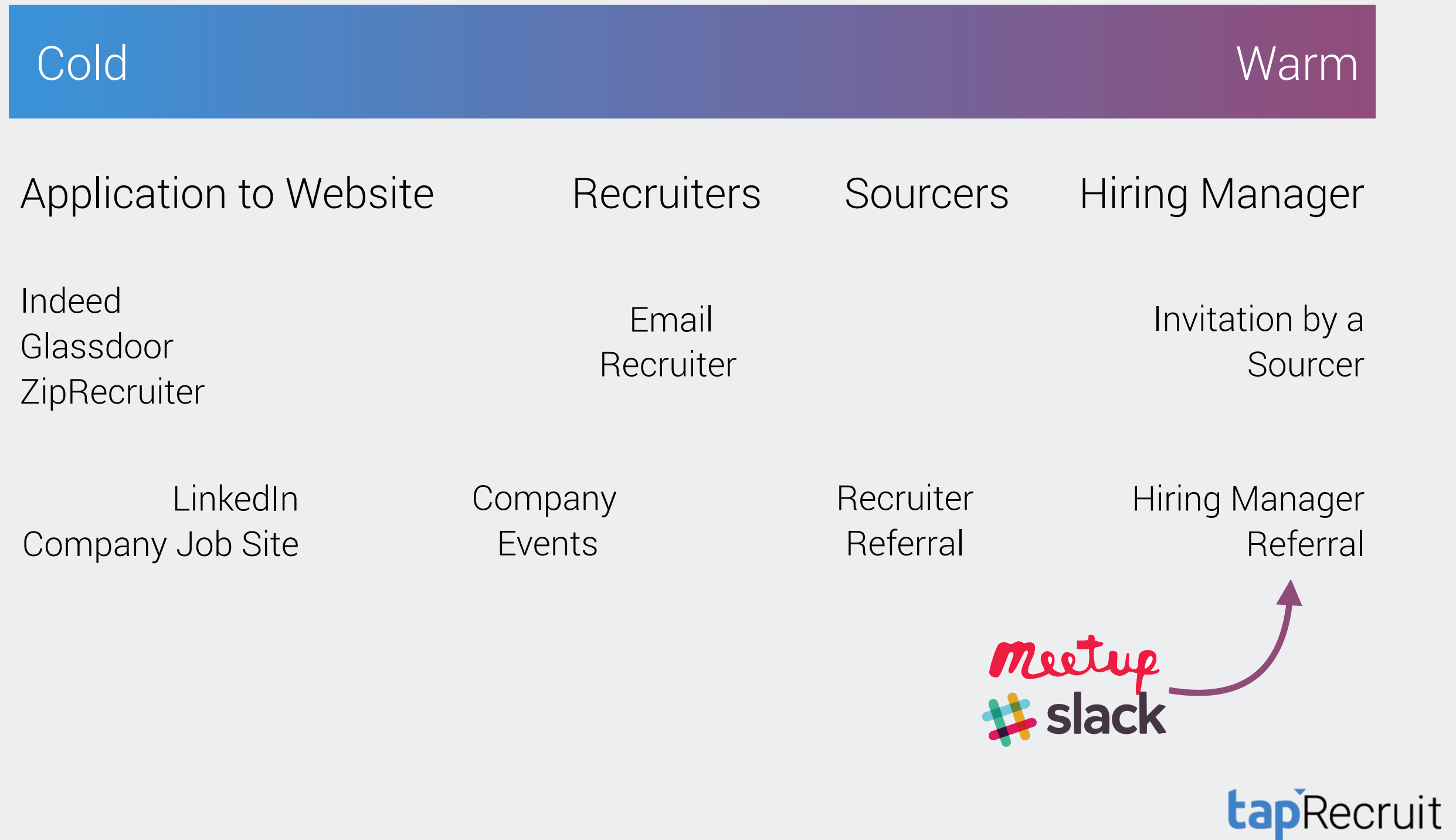
Contact the recruiter (or in some cases, the job poster is the hiring manager)



# Help Robin Help You!

- **Tailor the message:**
  - If Robin isn't your BFF, take time to personalize your message (their background or public work).
- **Be specific in your 'ask':**
  - What job are you interested in? Include a link. Include your CV. Make it super easy for Robin.
  - Maybe you are interested in understanding how data science works at a company like NYT? Tell Robin you'd like to set up an informational interview or chat. Make it super easy for Robin.

# Networking improves access to hiring managers



# NY-based data-focussed Meetup Groups

## General Groups

NYC Machine Learning  
Data Visualization Meetup  
Data Storytelling NYC

NY-HackR  
New York Python  
Looker NYC

## Affinity Group Focussed

❤️ Taiwanese Data Professionals :) ❤️

Women Who Code  
Women in Machine Learning and Data  
Science (WiMLDS)  
Women in Software Engineering  
Write/Speak/Code

R Ladies  
Py-Ladies

1

80% of candidates  
apply to the same  
20% of roles.

Strategies to improve  
job app success.

Search, titles, companies.

2

In job search, only  
20% of the work is  
the amazing resume.

Mechanisms to  
'warm' job apps.

Recruiters, referrals, networks.

3

Resume tailoring  
creates impactful  
resumes.

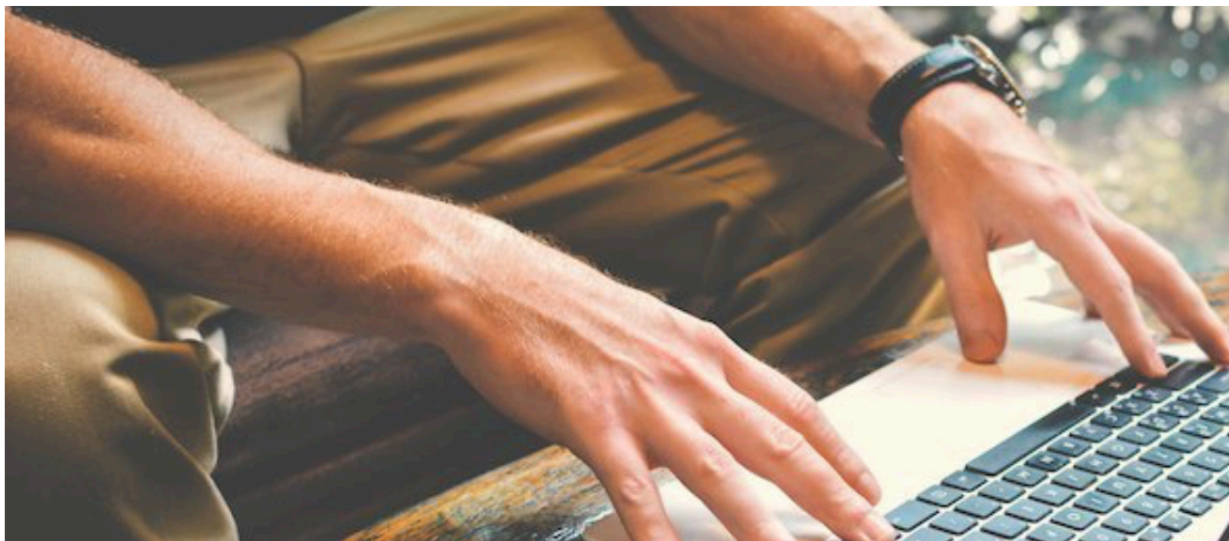
Generating bespoke  
resumes for jobs.

Format, structure, the 4Ws.

# The elephant in the room: ATS Filtering

## Beat the Robots: How to Get Your Resume Past the System & Into Human Hands

by Mark Slack and Erik Bowitz



<https://www.themuse.com/advice/beat-the-robots-how-to-get-your-resume-past-the-system-into-human-hands>

### Key Takeaways:

1. Use standard resume structure
  - Robots can't parse non-standard sections
2. Use exact keywords
  - Robots can't reconcile 'project manage' vs 'project managed'
3. Use (typically paid) text comparison tools to compare your resume vs the JD
  - Only Robots can beat other Robots



# Effective resumes make it easy to find relevant information quickly

## Sourabh Bajaj

<http://www.sourabhbajaj.com>

Email : [sourabh@sourabhbajaj.com](mailto:sourabh@sourabhbajaj.com)  
Mobile : +1-123-456-7890

### EDUCATION

- Georgia Institute of Technology** Atlanta, GA  
*Master of Science in Computer Science; GPA: 4.00* Aug. 2012 – Dec. 2013
- Birla Institute of Technology and Science** Pilani, India  
*Bachelor of Engineering in Electrical and Electronics; GPA: 3.66 (9.15/10.0)* Aug. 2008 – July. 2012

### EXPERIENCE

- Google** Mountain View, CA  
*Software Engineer* Oct 2016 - Present
  - Tensorflow:** TensorFlow is an open source software library for numerical computation using data flow graphs; primarily used for training deep learning models.
  - Apache Beam:** Apache Beam is a unified model for defining both batch and streaming data-parallel processing pipelines, as well as a set of language-specific SDKs for constructing pipelines and runners.
- Coursera** Mountain View, CA  
*Senior Software Engineer* Jan 2014 - Oct 2016
  - Notifications:** Service for sending email, push and in-app notifications. Involved in features such as delivery time optimization, tracking, queuing and A/B testing. Built an internal app to run batch campaigns for marketing etc.
  - Nostos:** Bulk data processing and injection service from Hadoop to Cassandra and provides a thin REST layer on top for serving offline computed data online.
- Georgia Institute of Technology** Atlanta, GA  
*Research and Teaching Assistant* Jan 2012 - Dec 2013
  - Research Assistant - Machine Learning:** Research on machine learning for portfolio hedging and replication algorithms. Modeling low-risk & continuous-return strategies. Developed the python library QSTK.
  - Teaching Assistant - Computational Investing:** The online course on Coursera, had more than 100,000 students enrolled. It was featured on the 11 Alive News and the Atlanta Journal Constitution. Involved in creating assignment, exams and conducting recitation sessions. Also taught the on-campus version of the course.

### PROJECTS

- QuantSoftware Toolkit:** Open source python library for financial data analysis and machine learning for finance.
- Github Visualization:** Data Visualization of Git Log data using D3 to analyze project trends over time.
- Recommendation System:** Music and Movie recommender systems using collaborative filtering on public datasets.
- Mac Setup:** Book that gives step by step instructions on setting up developer environment on Mac OS.

<https://github.com/sb2nov/resume>

## Alexander Eakins

777 Springboard St. #144  
Los Angeles, CA 90001  
(111) 111-1111  
[email@youreemail.com](mailto:email@youreemail.com)

### EDUCATION

#### Harvard University — Bachelor of Science, Data Science

AUGUST 2011 - MAY 2015

#### Springboard — Data Science Certificate

MAY 2017

### EXPERIENCE

#### Random Corp, 141 Random Circle Suite #991 Data Science Intern

MAY 2017 - AUGUST 2018

- Report actionable, statistical, and analytical insights to executives for effective strategic positioning in market place
- Shadow data scientists and assist in developing algorithms for predictive modeling
- Analyze and process sophisticated data sets using SAS, MySQL, and Excel
- Write python scripts to automate everyday tasks such as organizing desktop files, automatically redefining strings in MySQL database to promote consistency, and creating file and folder structures for new projects

### SKILLS & KNOWLEDGE

#### Languages

- R
- MySQL
- Python
- PHP
- C

#### Technical Skills

- Data mining/processing
- Statistical analysis

#### Soft Skills

- Business acumen
- Oral/written comm
- Strategic planning

### AWARDS

Walnut Festival International  
Fingerstyle Guitar  
Championship | 1st Place

Writers of the Future Contest  
(\$750 prize) | 2nd Place

<https://www.springboard.com/blog/data-science-resume/>



# Ensure hiring managers understand your educational experiences

## EDUCATION

### Teachers College, Columbia University

New York, USA

- M.S. Applied Statistics, GPA: 3.74/4.00
- Coursework: Computational Statistics, Multivariate Analysis, Linear Models and Experimental Design, Data Mining.
- Work: Teaching Assistant for Social Network Analysis, Quantitative Methods Program.

May 2018

### National Chiao Tung University (NCTU)

Hsinchu, Taiwan

- M.A. Communication Studies GPA: 4.00/4.00 (EP Credential) December 2012
- Concentration: Network Science, Data Visualization and Multimedia Analysis.

### National Cheng Kung University (NCKU)

Tainan, Taiwan

- B.A. Taiwanese Literature; B.A. Economics, GPA: 3.57/4.00 (EP) July 2010

# Jian-Yang (Jimmy) Wu

jimmy@gmail.com ♦ 123 456 789 ♦ GitHub: hotdogornotahotdog

## Experience

- Include title, company, location, dates.
- Maximum of 3 bullets per role.

## Education

- Include degree, location, dates, GPA/rankings, thesis title, teaching experience, relevant coursework.
- Maximum of 3 bullets per degree unless your primary context for data science projects are from your degree

## Skills

- 3-4 lines maximum organized by thematic.

## Other

- Awards / Leadership / Volunteer Experiences / Hobbies

# Jimmy (Jian-Yang) Wu

jimmy@gmail.com ♦ 123 456 789 ♦ GitHub: hotdogornotahotdog

## Experience

- Include title, company, location, dates.
- Maximum of 3 bullets per role.



## Education

- Include degree, location, dates, GPA/rankings, thesis title, teaching experience, relevant coursework.
- Maximum of 3 bullets per degree unless your primary context for data science projects are from your degree

## Skills

- 3-4 lines maximum organized by thematic.

## Other

- Awards / Leadership / Volunteer Experiences / Hobbies

# How candidates typically describe data science projects

## **Data science projects via R:**

**Winter Olympics:** Used tidyverse, ggplot2, and plotly to clean and visualize the data, including medal counts, host country advantage, and country success

**Fire in NYC and FDNY Response:** Used dplyr, leaflet, and ggmap to map the relationship between fire incidences and firehouses, and to calculate and visualize the geo distance

**Kickstarter:** Used tm, wordcloud, and quanteda to do text-mining for Kickstarter's projects

**U.S. Senator on Twitter:** Used ggnetwork, ggnet, igraph, and visNetwork to visualize the network between Senators on Twitter

**Yelp Las Vegas:** Used R Shiny to create an interactive map for selecting restaurants in Las Vegas

## **Problems:**

1. Emphasis on tools rather than analysis
2. Hard to gauge size and relevance of project

# How candidates typically describe data science projects

## **Data science projects via R:**

**Winter Olympics:** Used tidyverse, ggplot2, and plotly to clean and visualize the data, including medal counts, host country advantage, and country success

**Fire in NYC and FDNY Response:** Used dplyr, leaflet, and ggmap to map the relationship between fire incidences and firehouses, and to calculate and visualize the geo distance

**Kickstarter:** Used tm, wordcloud, and quanteda to do text-mining for Kickstarter's projects

**U.S. Senator on Twitter:** Used ggnetwork, ggnet, igraph, and visNetwork to visualize the network between Senators on Twitter

**Yelp Las Vegas:** Used R Shiny to create an interactive map for selecting restaurants in Las Vegas

## **Problems:**

1. Emphasis on tools rather than analysis
2. Hard to gauge size and relevance of project
3. Lacks a clear narrative of why you did the analysis and why the metrics that you chose worked

# How candidates typically describe data science projects

## **Data science projects via R:**

**Winter Olympics:** Used tidyverse, ggplot2, and plotly to clean and visualize the data, including medal counts, host country advantage, and country success

**Fire in NYC and FDNY Response:** Used dplyr, leaflet, and ggmap to map the relationship between fire incidences and firehouses, and to calculate and visualize the geo distance

**Kickstarter:** Used tm, wordcloud, and quanteda to do text-mining for Kickstarter's projects

**U.S. Senator on Twitter:** Used ggnetwork, ggnet, igraph, and visNetwork to visualize the network between Senators on Twitter

**Yelp Las Vegas:** Used R Shiny to create an interactive map for selecting restaurants in Las Vegas

## **Problems:**

1. Emphasis on tools rather than analysis
2. Hard to gauge size and relevance of project
3. Lacks a clear narrative of why you did the analysis and why the metrics that you chose worked
4. Lack of prioritization

**Kickstarter:** Used tm, wordcloud, and quanteda to do text-mining for Kickstarter's projects

Where?

Where did you get the data from?  
Scraping? API? Database?

How?

How big was the dataset?  
148,000 Kickstarter projects

Why?

Why did you do the analysis?  
What was the goal/hypothesis of the project?  
Hypothesized that readability of project descriptions  
could predict project success

What?

What type of analysis did you do?  
Was it descriptive or predictive?  
What skills/packages/libraries did you use?  
Natural Language Processing techniques such  
as TFIDF (via quanteda, tm)

**Kickstarter:** Used tm, wordcloud, and quanteda to do text-mining for Kickstarter's projects



Analyzed 148,000 Kickstarter projects using Natural Language Processing techniques such as TFIDF (via quanteda, tm) to generate a model to predict project success from readability.

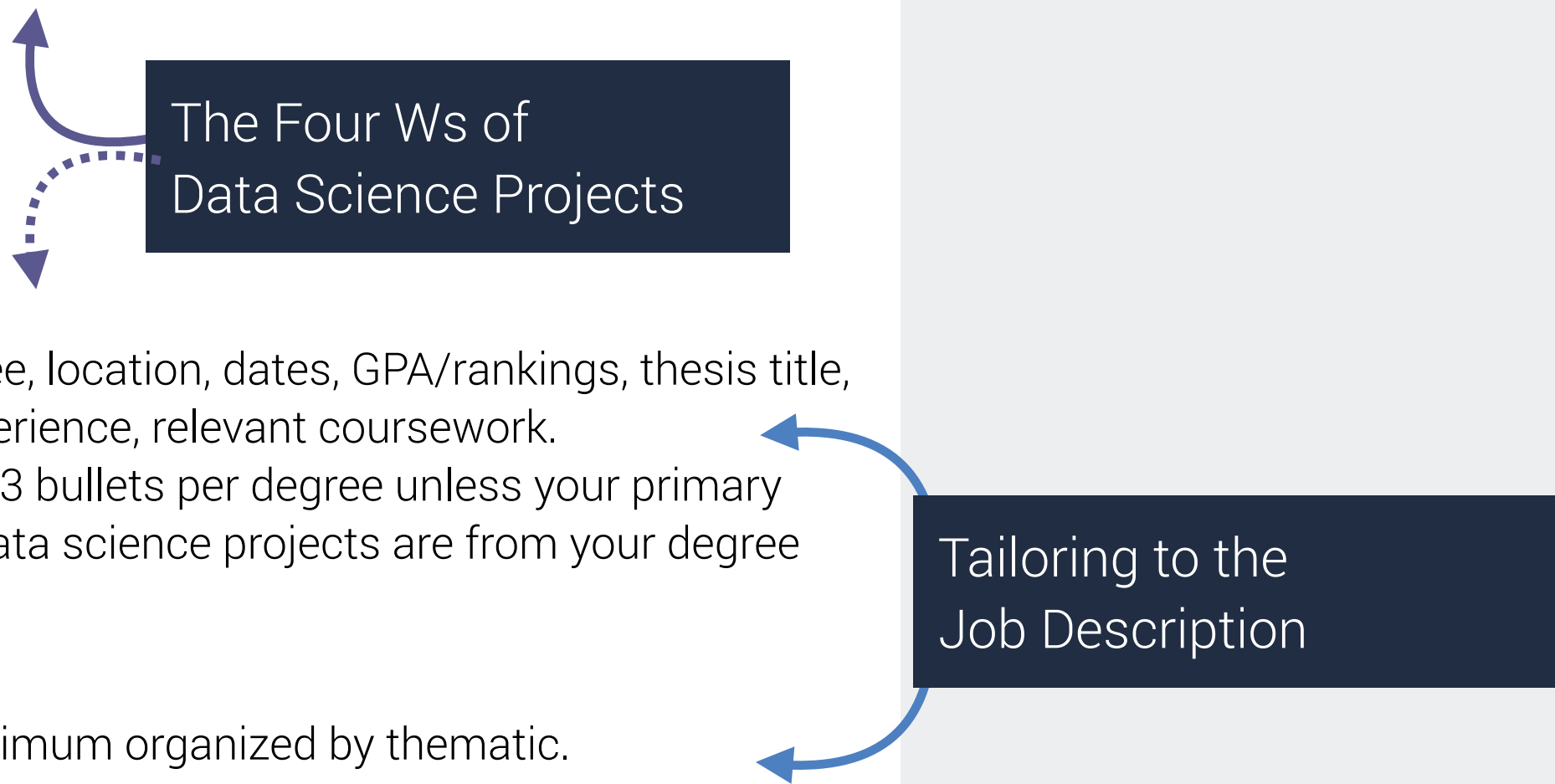


# Jimmy (Jian-Yang) Wu

jimmy@gmail.com ♦ 123 456 789 ♦ GitHub: hotdogornotahotdog

## Experience

- Include title, company, location, dates.
- Maximum of 3 bullets per role.



The Four Ws of  
Data Science Projects

## Education

- Include degree, location, dates, GPA/rankings, thesis title, teaching experience, relevant coursework.
- Maximum of 3 bullets per degree unless your primary context for data science projects are from your degree

Tailoring to the  
Job Description

## Skills

- 3-4 lines maximum organized by thematic.

## Other

- Awards / Leadership / Volunteer Experiences / Hobbies

# Relevant coursework should match job description

## **Fordham University**

New York, NY

*Master of Arts in Applied Statistics, GPA: 3.54*

*2016-2018*

Relevant Coursework: Multivariate Statistical Analysis, Linear Models, Experimental Design and Analysis.

## **New York University**

New York, NY

*MA in Applied Quantitative Research (GPA: 3.71)*

*Expected 2019*

- Thesis: The Impact of the Motherhood Penalty on Single Mothers' Earnings.
- Relevant Coursework: Applied Quantitative Analysis, Data Analysis and Visualization, Design of Social Research, Social Stratification.

# Skills should match job descriptions

**Data Analysis:** Excel, R (tidyverse and base)

**Data Visualization:** Tableau, PowerBI, ggplot2, plotly

**Proficient:** Python, Java, SQL, Caffe, Git

**Familiar with:** Theano, MongoDB, Hadoop, Redshift

**Data Science:** Python, Pandas, Numpy, Seaborn

**Databases:** MySQL, Postgres

Order by the importance  
and relevance to the job  
description

**Data:** 4+ years of data experience with a focus on Machine Learning (including Classification, Regression, Feature Engineering) and Statistics (Regression, Confidence intervals, Bayesian and Monte Carlo Methods).

**Programming:** 3+ years Python (scikit-learn, NumPy, SciPy, Pandas) and Git.

**Communication:** 3+ years experience presenting, public speaking, writing (articles, papers, blogposts) and communicating results.

# Jimmy (Jian-Yang) Wu

jimmy@gmail.com ♦ 123 456 789 ♦ GitHub: hotdogornotahotdog

- Points you want to highlight - 4 bullets in 2 columns

## Experience

- Include title, company, location, dates.
- Maximum of 3 bullets per role.

## The Four Ws of Data Science Projects

## Education

- Include degree, location, dates, GPA/rankings, thesis title, teaching experience, relevant coursework.
- Maximum of 3 bullets per degree unless your primary context for data science projects are from your degree

## Skills

- 3-4 lines maximum organized by thematic.

## Other

- Awards / Leadership / Volunteer Experiences / Hobbies

## Tailoring to the Job Description

# Tailoring your resume header to the job description

## **Procurement Analyst**

**HelloFresh - New York City**

### **You are / have ...**

- Bachelor's degree in Business, Information Systems, Computer Science, Statistics, Economics, or Maths
- Have 0-2 years of experience
- Excellent Excel and analytical skills to create various databases and spreadsheets
- Preferably have experience with Tableau, SQL -- understanding of basic statistical analysis
- Detail-oriented, inherent drive to seek new information and explore topics in-depth
- Solid research background and robust project management skills
- Strong written and verbal communication skills, proven track record understanding demand forecasting, and forecasting models

Key skills

Experiences to emphasize

# Tailoring your resume header to the job description

## **Procurement Analyst**

**HelloFresh - New York City**

### **You are / have ...**

- Bachelor's degree in Business, Information Systems, Computer Science, Statistics, Economics, or Maths
- Have 0-2 years of experience
- Excellent Excel and analytical skills to create various

Jimmy (Jian-Yang) Wu

jimmy@gmail.com ♦ 123 456 789

- Economics Masters candidate in with over 2 years experience in using Excel and other data tools for forecast modeling.
- Prior experience in SQL, Tableau and other data visualization systems.
- Experienced communicator with experience teaching undergraduate courses in statistics.
- Recipient of multiple awards / prizes.
- Avid user of on-demand Food Tech.

# Tailoring your resume header to the job description

## **Senior Business Analyst**

**TVision - New York City**

### **Requirements**

- BS/MS in a quantitative discipline
- 2+ years of experience in consulting, research or similar analytical field
- Familiarity with media research and/or passion for the media industry
- Strong presentation, communication and project management skills
- Experience in (or interest in learning) statistical analysis using one or more of the following (R, Python, SAS, SPSS, STATA, Matlab, Tableau, Spotfire, Qlik, Alteryx)
- Advanced fluency in Excel and PowerPoint
- Intermediate fluency in SQL
- Strong attention to detail, effective organization and task management skills

Thank you!