

DATA SCIENCE RESUME

USE OUR CHECKLIST TO FINE-TUNE
YOUR RESUME

EMMA

ding

EMMA ding



emmading.com



info@emmading.com



How to Craft a Great Data Science Resume

[The Goal](#)

[Understanding How to Use This Guide](#)

[General Resume Guidelines](#)

[Sections and Orders](#)

[1 Contact Information](#)

[2 Professional Experience](#)

[3 Technical Skills](#)

[4 Education](#)

[Experience Summary \(Optional\)](#)

[Projects \(Optional\)](#)

[Publications \(Optional\)](#)

[Pro Tips](#)

[Checklist](#)

[Action Items](#)

▼ The Goal



The goal of a data science resume is **to land interviews**.

To increase your chances of making it to the interview stage, your resume needs to stay focused on its goal. There are two things to keep in mind to ensure your resume is focused:

1. **Highlight data science-related skills and experience.**

- What recruiters and hiring managers care about are things related to the position to which you are applying.

2. **Avoid anything not relevant to data science.**

- Too much information will only distract recruiters and hiring managers.



So, the big rule of data science resumes is **Keep It Relevant!**



Understanding How to Use This Guide

This document contains suggestions and tips for crafting a great data science resume. However, there are a few things to keep in mind to make this guide work well for you.



The suggestions made here apply to **big and medium size companies and startups** that receive a **large volume of resumes** (hundreds or thousands of resumes per month) and need quick ways to **say No**.

These companies typically have an **automated resume screening** to select candidates. That means...

- They don't review every single resume manually.
- We need to beat automated screening by including certain phrases from job descriptions and repeating keywords.

The tips in this guide are designed to help you beat those automated resume screenings and also get by quick resume scans. But that also means...

- These suggestions may **not work for small startups** that are looking for candidates with specific skills.



General Resume Guidelines



Recruiters and hiring managers only spend 5 - 10 seconds on your resume! So keep it **relevant** and **concise**!

- **Keep it to one page. Always.**
 - Reduce font size and page margin if necessary, but always stay on one page. Recruiters and hiring managers will not read a second page!
- **Keep it to one column.**
 - It's easier to read and ensures that you can control what the recruiter sees first (you are going to put the most important information at the top).



+1 (425) [REDACTED] | Authorised to work in US | phd16shikhav@[REDACTED] | linkedin.com/in/[REDACTED]

EXPERIENCE SUMMARY

Ph.D. in Machine Learning with 2+ years of industry experience working in high-performance, worldwide scale projects on fraud detection, payments, warehouse management & promotion targeting. She is skilled in supervised & unsupervised machine learning algorithms, building end-to-end ML pipelines, applied statistics, generating insights from multiple data sources, metric-driven thinking, Python & SQL.

PROFESSIONAL EXPERIENCE

Principal Analyst

Toast (Payments Optimization, Restaurant POS)

- Led the design & development of a new optimization framework for a 700+ strength field reps org; responsible for reporting, root-cause analysis & 3 impact assessment of multiple levers on payment system margin
- Led deep dive analysis of downgraded & fraudulent transactions with \$14 mn volume, \$1mn potential savings in processing fees, directly influences staffing decision on fraud ops team

Nov 2022 - Current Seattle, WA

Data Scientist

PayPal (Fraud Prevention Team)

- Utilised Python to discover 21 new fraud patterns in velocity based rolling aggregations using EDA & fuzzy string matching-based feature generation. \$8mn reduction in quarterly fraud loss
- Specialised in identifying & developing of on-demand model explainability using interpretable ML methods to enable better fraud detection
- Implemented framework & system design for metric reporting & anomaly detection of 500+ active rules for 400+ merchants with \$10 mn daily TPV
- Formulated novel metrics for measuring reject rule performance in absence of ground truth using embeddings & clustering algorithms; consumed by 20+ data scientists for weekly rule tuning to fight adaptive fraud
- Developed & productionised flexible, parameterized code module using JS to enable testing of novel fraud rules, improving new feature release time by 3 weeks/qtr
- Led sprint planning & execution of revamping internal product documentation for a team of 5 data scientists, 50% reduction in new hire onboarding time

June 2021 - July 2022 Hyderabad, India

Analytic Researcher

Flipkart (Large e-comm, India)

- Devised product-affinity based item slotting plan using association rule mining, demonstrating a 27% reduction in order picking time for one of India's largest e-commerce warehouse for 130K SKUs across 700 categories
- Developed full factorial experiment design & quantified tradeoff metrics for layout/item placement & routing

Oct 2018 - Mar 2019 Bengaluru, India

BUSINESS ANALYST

ZS Associates

- Developed HCP-level regression models with variable transformations to devise a sales force sizing allocation & targeting plan for \$30mn US client for promoting its medical device to market worth \$700mn
- Created demand estimation models, sales force structure & sizing for Fortune 500 \$2.5bn car rental service for 10 corporate markets worth \$2mn; resulting in annual savings of \$0.4mn

Oct 2015 - May 2016 New Delhi, India

EDUCATION

Ph.D | Machine Learning & Data Science

Indian Institute of Management, Ahmedabad

2016 - 2021 Ahmedabad, India

Bachelor of Technology | Electronics & Communication

National Institute of Technology, Jalandhar

2011 - 2015 Jalandhar, India

DISSERTATION WORK

Streaming Data Analytics for Intelligent Systems

- Developed novel prediction algorithms for detecting data distribution shifts in sensor data streams with demonstrated applications in transportation & human activity recognition domains with ~ 30% savings in labeling costs
- Published research in Springer publication, presented at ACM and Grace Hopper conferences

Aug 2018 - Mar 2019 Ahmedabad, India

TECHNICAL SKILLS

Methods: Regression, Classification, Decision Trees, Random Forests, XGBoost, SVM, K-NN, Clustering, Neural Networks, Explainable ML, A/B Testing, Exploratory Data Analysis, Data Visualization, Concept Drift Detection Tools & Languages: Python (NumPy, Pandas, scikit-learn, matplotlib, seaborn, ELI5), SQL, Looker, TensorFlow



Data Scientist • Artificial Intelligence • Client Management

Data Scientist with 6 years of experience building Machine Learning and Deep Learning applications, optimizing solutions to enhance model performance, and generating insights. Proficient in Predictive Modelling, building Neural Networks, Data Processing as well as scripting languages including Python & C++. Highly experienced in compiling metrics & generate insights to provide strategic analysis & actionable recommendations and in Client Relations & Management.

AREAS OF EXPERTISE

- Machine Learning & Deep Learning
- Computer Vision (CNN) and NLP
- Algorithm Optimization & Multi-Factor Model
- Feature Engineering ***
- Data Analysis and Visualization
- Statistical Analysis & Data Modelling
- Quantitative & Qualitative Research

TECHNICAL SKILLS

- Programming: Python (Scikit-learn, TensorFlow, Keras, PyTorch, OpenCV, C++, C, SQL, MySQL, Linux)
- Tools: AWS(S3), Git, GitHub, C-Lion, IntelliJ, PyCharm
- Tableau, SQL, R, SPSS
- Spreadsheet Modelling
- Predictive Modelling
- Business Process Modelling
- MS Visual Studio/MS Office

EDUCATION

- Rutgers Business School, The State University Of New Jersey - Dec 2022 Master of IT & Analytics - GPA: 3.8/4
- Adjunct Professor (2022) - Formulate and execute interactive lectures on Production & Operation Management
- Vice President - Student association 2021-22, Rutgers Business School

RELEVANT PROJECTS

- Deep Fake Faces | Python | TensorFlow | May 2022
 - Developed deep fake faces using GANS and implemented W-GANS to improve model efficacy by 40%.
 - Suppressed the shortcomings of vanilla GANS such as vanishing gradients and lowering the potency of discriminator against the generator.
- SMS Spam Detection | Python | Scikit Learn | Dec 2021
 - Performed data pre-processing such as handling missing data, duplicate values, and unwanted outliers on dataset with 55k rows.
 - Trained and evaluated performance of 50 machine learning models like Naive Bayes, KNN, Random Forest, Logistic Regression, SVM and Neural Nets to detect Spam SMS.
 - Implemented semantic relationships between words using TF-IDF and Word2Vec.

▼ 📈 Sections and Orders



How do you decide what to put on your resume?

✓ Required

- Contact Information
- Professional Experience
- Technical Skills
- Education

💡 Optional

- Experience Summary
- Projects
- Publications
- Certificates

🚫 Avoid

- Hobbies
- Volunteer Experience
- Interests



Remove unnecessary sections! Too many sections make it hard to find useful information.

Let's now look more closely at each section. Below you will find tips and examples for each section of a data science resume.

▼ 1 Contact Information

FULL NAME

Tel: | Email: name@gmail.com | LinkedIn: | Address

This section always goes at the very top of your resume. It should have at the least:

- Full Name
- Telephone
- Email
- LinkedIn Profile
- Address

▼ Some other things to keep in mind:

- You can also include your GitHub profile.
 - Pin the important repos on your homepage (you can pin up to 6).
 - Don't include a link to your GitHub if it's empty.
- You do not need to specify an exact address.
 - Keeping it at the city level is good enough.

▼ 2 Professional Experience

Data Scientist – Foundation & Infrastructure, Airbnb | San Francisco, CA Mar. 2019 – present

- Architect, build, and launch efficient and reliable data models, schemas and pipelines to enable analysis and modeling;
- Build tools for auditing, error logging and validating data tables;
- Analyze data to design business metrics and size the opportunity for new business and product initiatives;
- Define logging needs in partnership with engineering;



If you have 2+ years of experience, this section should **always go at the top of your resume**, under Contact Information, as this is the section recruiters and hiring managers **care most about**.

This section contains positions you held at different companies. For each position make sure you include:

- Position
- Company
- Location
- Time Duration

Do not include non-relevant information. For example you do not need to say Full-Time Data Scientist. You can just say Data Scientist to be more concise.



If a company you worked with is not well-known, it's helpful to have one sentence high-level explaining what the organization does and what is it amazing to work there.

▼ Describing Responsibilities and Achievements

Under each position, you will describe your responsibilities and achievements. Here are some tips for doing that effectively.

▼ Use strong verbs.

- Strong verbs are those that highlight your role and the results.
- Use verbs like "**led, drove, created, optimized, and improved**" rather than verbs like "**processed, coded, and completed**".

185+ Action Verbs That Will Majorly Impress Hiring Managers

Responsible for ... Experience in... Most resume bullet points start with the same tired old words hiring managers have read over and over-to the point where they've lost a lot of their meaning. They also don't do much to show off

📘 <https://www.themuse.com/advice/185-powerful-verbs-that-will-make-your-resume-awesome>



▼ Prioritize.

- For each position, list your experiences and achievements in order of **importance** rather than chronologically.

▼ Be concise.

- No more than 5 bullet points



- Create and maintain databases (Presto, Snowflake, Redshift) and ETL pipelines written in SQL and Python to track and store product information.
- Curated data marts/models using dbt to stitch together experimental data with revenue/sales data across disparate sources (e.g., Amplitude, Salesforce).
- Create and maintain critical dashboards to prepare and evaluate advertising and developer product graduation readiness; present key analytical findings that can improve advertising products at Pinterest.
- Designed, implemented, and launched 15 dashboards in SQL and Tableau to measure the success of products at Pinterest.
- Influenced product roadmaps by identifying opportunities and quantifying the impact of product launches through A/B experimentation and measurement.
- Applied software engineering best practices to analytics code by developing, documenting, and maintaining a reusable programming code library using Bazel, Docker, Git, Phabricator, and related tools. Tested software for bugs and operating speed, fixing bugs and documenting processes to increase performance and efficiency by 38%.
- Designed the architecture for comparing data between similar products in Alpha / Beta across unified metrics such as adoption, CSAT, and revenue; enabled statistical comparisons (A/B tests) of over 100 product feature launches.
- Strengthened graduation recommendation by direct customer feedback for 80% of graduating alphas/betas, ensuring 80% or more of product launches are high quality.

- Limit each bullet point to 2 lines (physical lines, not sentences).



- Built, deployed, monitored and documented the employee resignation prediction model for Hong Kong and India which generates resignation probabilities and outputs top 3 resignation reasons for each predicted employee for management to take intervention actions upon to improve employee retention, socialized with senior management and stakeholders as end-to-end product owner for their feedback, sign-offs and support to roll out across markets
- Analyzed company-wise internal mobility (IM) data to refine IM eligibility definition, generate insights on ineligible IMs, simulated multiple population bases as the denominator for stakeholders to decide IM rate evaluation metric, overlay with promotion data to get cross markets/department breakdowns and yearly trend which is used to debunk the myth that IM slows down promotion which helps cultivate encouraging culture of IM within company

▼ Be impact-oriented.

- The first bullet point in each experience should be the **most impactful**.
- The impact should be measured by **business metrics**.
 - Led sprint planning and execution of revamping internal product documentation for a team of **4 data scientists**, **50% reduction** in new hire onboarding time.
 - Created and implemented framework and system design for metric reporting and anomaly detection of **500+ active rules for 400+ merchants** with **\$100M** daily Total Payment Volume (TPV).
 - Part of a 3-data-scientist team that owns feature engineering for a fraud detection system. Added **300 new features**, resulting in a reducing the false positive rate from **20% to 15%**, while keeping the false negative the same.

▼ Don't tell. Show.

Use numbers and metrics. Your experience and skills are more reliable and trustworthy when you can show quantitative impact.

- *significantly increased conversion rate* → *doubled conversion rate from 0.3% to 0.6% translating to \$2M increase in annual revenue*
- *processed survey data* → *processed 1TB of unstructured survey text data*
- *greatly improved model performance* → *improved accuracy by 20% and reduced training time by 30% using XGBoost model*
- *strong coding skills* → *contributed 2000 lines to Apache Airflow using Python*

▼ Avoid wordiness.

- Aim to be as concise as possible in your word choice.
 - “in order to” → “to”
 - “is able to” → “can”
- Example of a concise description

- **Predicted Human Behavior using TensorFlow 2 and Neural Network (Python):** Used Tensor Flow 2.0 performed to predict human behavior on Audiobook dataset whether customer will buy again or not. Used ADAM optimizer with parse_categorical_cross_entropy as we are dealing with categorical data. With accuracy as metrics I have predicted with 89% accuracy whether the customer will buy again or not.



Using ADAM optimizer with categorical data to predict customers' re-purchasing decision with 89% accuracy.

- **Reverse Engineered Facebook's Pricing Model for Viacom (Python, KNIME):** Conducted EDA on 15 months demographic data (till April 2019) to obtain insights for CPM (Cost per mile). Created different pricing models using Linear Regression, Random Forest and Gradient Boosting Tree. Gradient Boosting method achieved 96.7% accuracy over actual model after validating with a new dataset. Predicted 10 cents closer to Facebook's actual price which it charges to Viacom. Recommended Viacom to focus more on Female in age range 25-30 in USA to garner more views with less expense.



Created different pricing models, such as Linear Regression, Random Forest and Gradient Boosting Tree, and achieved 96.7% accuracy by cross validation.

▼ Don't get too technical.

- Only include the information relevant to the job to which you are applying. Too much technical jargon makes your resume difficult to understand and unnecessarily wordy.
- **Predicted Human Behavior using TensorFlow 2 and Neural Network (Python):** Used Tensor Flow 2.0 performed to predict human behavior on Audiobook dataset whether customer will buy again or not. Used ADAM optimizer with parse_categorical_cross_entropy as we are dealing with categorical data. With accuracy as metrics I have predicted with 89% accuracy whether the customer will buy again or not.
- Reduce "used ADAM optimizer with parse_categorical_cross_entropy" to simply "used ADAM optimizer".

▼ 3 Technical Skills

TECHNICAL CAPABILITIES

- Programming language: Python (NumPy, Pandas, ScikitLearn), R
- Statistics: time series forecasting, ARIMA
- Machine learning: Classification, Regression, Clustering, PCA, CNN, RNN
- Database: SQL
- Data Visualization: Tableau, Power BI, QlikView, Cognos

This section is important, but it can quickly become enormous and unorganized. Here are some tips for keeping this section neat and easy to understand.

- **✓ Organize.** All the relevant skills you have can become quite a list. To make things easy to digest, place them in categories.
- **✓ Remove common skills** such as Git, Notebook, MS Word/Excel.
- **✓ Don't include skills you are not good at.** Highlight the ones that are most relevant to the job. It can actually weaken your resume if you claim to be an expert in too many things.

▼ 4 Education

Massachusetts Institute of Technology (MIT)
Master of Engineering in Computer Science
Bachelor of Science in Computer Science

Cambridge, MA
September 2020
June 2019

- Concentration: Artificial Intelligence
- Master's Thesis: "De-identification of free-text clinical notes using NLP"
- Relevant Coursework: Natural Language Processing, Artificial Intelligence, Autonomy Decision Making, Computer Vision, Machine Learning, Computational Statistics, Algorithm

For the education section, there are some things you should always include and others that are optional.

▼ **Must Have:**

- School
- Degree
- Time of Graduation

▼ **Optional**

- Relevant coursework
- GPA



If you have 2+ years of experience, you don't need to include the optional information., Leave more room on your resume for Professional Experience instead.

▼ Experience Summary (Optional)

SUMMARY

-
- Dedicated Ph.D. with 5 years of time series analysis and statistical analysis with satellite data
 - 4+ years of industry working experience as a data scientist in machine learning modeling
 - Hands-on coding experience in Python and SQL with documentation for reproducibility
 - Strong interpersonal communication and reporting skills as a team contributor

This section is an optional summary that goes before the professional experience. Here are some additional examples of what it can look like.

SUMMARY

Expert in the field of data analytics using the latest technologies in data science and machine learning. Worked for a Fortune 100 company before joining Meta (formerly Facebook) as a data scientist for the recruiting product org. Foundation in mathematics, moved into the data world after completing a master's in management information systems (MIS).

EXPERIENCE SUMMARY

Ph.D. in Machine Learning with 2+ years of industry experience working in high-performance, worldwide scale projects on fraud detection, payments, warehouse management & promotion targeting. She is skilled in supervised & unsupervised machine learning algorithms, building end-to-end ML pipelines, applied statistics, generating insights from multiple data sources, metric-driven thinking, Python & SQL.



The experience summary is not as important as the professional experience section, so only put this if you have room you need to fill.

▼ Projects (Optional)

This optional section can help you fill out your resume when you don't have enough work experience or your work experience is not relevant to Data Science. However, there are a few things to keep in mind.

- **Only include the most impactful projects.**
- **Avoid cookie-cutter projects** (popular projects, e.g. titanic, sentiment analysis of tweets, stock trading, chatbots, etc).
 - These type of projects don't help demonstrate your creativity or technical depth.
- **Project names should make sense to non-technical people.**



- **Adding links helps!**
 - Follow [this link](#) to create your own website to host your projects.

Movie Night 🎬 | Python (Surprise, SciPy)

USE CASE: HELP A GROUP OF USERS WITH VARYING TASTES FIND THE BEST MOVIE FOR THEIR MOVIE NIGHT

Berkeley, CA

May - Jun. 2020

- Built collaborative filtering recommender for individual users using SVD++ on MovieLens data, outperforming baseline model (RMSE = .87 and 1.4).
- Reviewed 15+ papers on "cold start"; solved it by asking new users to rate 20 movies with highest harmonic means of rating entropy and log frequency.
- Extended recommender system to user group by aggregating individual recommendations ("average-without-misery") or merging user profiles.

Reopen Navigator (Dash App) 🎬 | Python (LightGBM), Dash, Heroku

USE CASE: GAUGE COVID-19 INFECTION RISK AND OUTCOME SEVERITY BASED ON USER SOCIAL BUBBLE, STATE, AND CONDITIONS

San Francisco, CA

Jun. - Aug. 2020

- Scraped 20.6 million COVID-19 case and viral test data from state reports; reviewed 50+ medical journals to estimate average COVID-19 infectivity.
- Trained and deployed LightGBM classifier (F1 score = .84) to predict new user's treatment outcome based on state, demographics, and conditions.

Data Mining, Database, & NLP Projects 🎬 | Spark, SQL, Python (Scikit-Learn)

COMPILATION: THE DATA INCUBATOR PROJECTS INVOLVING WEB SCRAPING, DISTRIBUTED COMPUTING, AND LARGE-SCALE NETWORKS

San Francisco, CA

Jun. - Aug. 2020

- Web-scraped 100,000 photo captions in New York Social Diaries and applied NLP and network analysis to constructing social graphs of NYC elites.
- Processed 100 GBs of Stack Overflow posts with Apache Spark and used Word2vec in Spark ML to identify top 25 closest synonyms to "ggplot2".
- Devised SQL database for Yelp restaurants; built ETL pipelines to predict ratings from reviews, including n-gram model with Bayesian smoothing.

▼ Publications (Optional)

If you have them, publications can be great thing to add to your resume. You can include:

- Papers
- Blog posts
- A link to your Google Scholar profile (if applicable)



If you're a co-author of a paper with other authors, bold your name in the author list.

▼ Pro Tips

▼ Customize your resume for each position.

- Don't turn in the exact same resume for each position.
- To customize you can reorder and include different things based on what is most relevant to the position to which you are applying.
- You can also change around your keywords to match the position better. 

▼ Use keywords (buzzwords).

- Keywords should appear multiple times in the resume. You can use them in both:
 - Professional Experience
 - Technical Skills
 - The exact technical skills will vary from job to job. Make adjustments to this section both in what you include and what you put first to best suit each job description.

▼ Here is an example of how you can extract keywords from a job description and use them in your resume:

▼ Job Description

Data Scientist - Consumer Product - Metrics Team

SAVE 



Twitter
San Francisco, CA

[Apply on Jobilize](#) [Apply on Ladders](#) [Apply on Jora](#) [Apply on Jobrapido.com](#)

6 days ago Full-time

Data Scientist - Consumer Product - Metrics Team Locations Seattle, San Francisco Company Twitter serves the public conversation by encouraging people all over the world to connect, learn, debate and solve problems together.

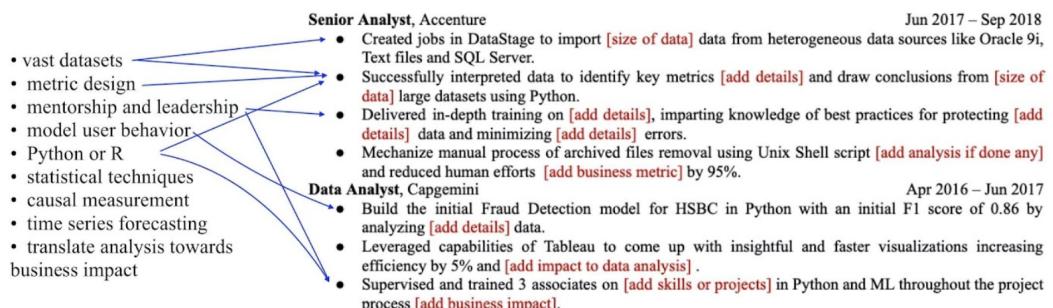
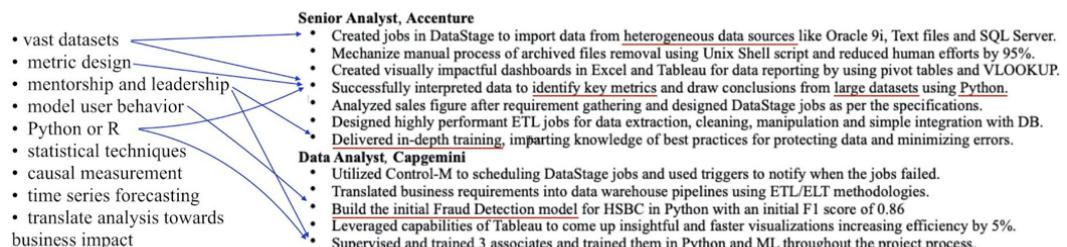
We believe conversation can change the world, and that's why Tweeps (that's what we call Twitter employees) come to work every day. Job description You will be a key member of the Metrics Data Science team working closely with our partners in engineering, product, design, user research and operations to provide cross team insights to help unify efforts across the company and craft the product business strategy for Twitter. As a Data Scientist on the Metrics team you will:

- Conduct analyses to derive strategic insights from vast datasets across different teams.
- Apply sophisticated statistical techniques to model user behavior, identify causal impact and attribution, build and benchmark metrics, and forecast future performance.
- Communicate findings to executives and cross-functional product teams by translating statistically rigorous analyses towards business impact.
- Improve the reproducibility and efficiency of analysis work through code review and other established processes. Qualifications
- Advanced degree in quantitative analysis field and 2+ years of experience (or 4+ years of experience).
- Focus on learning on the job, taking initiative, and can thrive within a large team where you seek to mentor others.
- Own impactful projects from start to finish and influence organizations through cross functional efforts.
- Track record of mentorship and technical leadership
- Experience with causal measurements, time series forecasting, and metric design
- Experience with analytical programming languages such as Python or R. Additional information Job opportunities should be equal.

We don't discriminate.

Period.

▼ Professional Experience



▼ Technical Skills

SKILLS

- Programming & Database: R, Python (NumPy, Pandas, Sci Kit Learn), SQL, AWS EC2, R, SAS, MySQL, MATLAB.
- Data Visualization: Tableau, Power BI, QlikView, Cognos.
- Machine learning: Classification, Regression, Clustering, Deep Learning, PCA, Arima, CNN, RNN, Random Forest.
- Tools & Techniques: DataStage, Tableau, R studio, JuPyteR, Excel, Informatica, KNIME, JIRA, TensorFlow, Toad



SKILLS

- Programming language: Python (Numpy, Pandas, ScikitLearn), R
- Statistics: time series forecasting, ARIMA
- Machine learning: Classification, Regression, Clustering, PCA, CNN, RNN
- Database: SQL
- Data Visualization: Tableau, Power BI, QlikView, Cognos

▼ Checklist

Once you are done drafting your resume, you can use this checklist to perfect it.

- ! No Typos.** A typo suggests a lack of attention to detail.



Use [Grammarly](#) to make sure there are no typos.

- [Grammarly Chrome Extension](#)

- Have all required sections.**

- Minimize abbreviations.**

- Do not risk the recruiter and hiring manager not knowing what you mean. Spell out terms like Machine Learning rather than putting ML to ensure that recruiters understand.

- All links are working.**

- Well-formatted pdf.**

- Send your resume as a PDF instead of an editable format like Google Docs. The editable format might not render perfectly on other people's computers.

- Data Visualization: Tableau, Power BI, QlikView, Cognos.
- Machine learning: Classification, Regression, Clustering, Deep Learning, PCA, Arima, CNN, RNN, Random Forest.
- Tools & Techniques: DataStage, Tableau, R studio, JuPyteR, Excel, Informatica, KNIME, JIRA, TensorFlow, Toad

PROFESSIONAL EXPERIENCE

Senior Analyst, Accenture, India

- Created jobs in DataStage to import data from heterogeneous data sources like Oracle 9i, Text files and SQL Server.

Jun'17 – Sep'18

▼ Action Items

1. **Make a copy of the Resume Template!**

DS Resume Template

FULL NAME Tel: | Email: name@gmail.com | LinkedIn: | Address EXPERIENCE

SUMMARY Senior Data Scientist with 5+ years of industry experience working on high-performance projects on fraud detection, payments, and risk management.

🔗 <https://docs.google.com/document/d/1QrlzU6xleL6xzSOi0g9ILFww2gOFGZ0jXNngMNy0Z18/edit?usp=sharing>

• Designed and developed a Python program that collects user input through JSON, constructs various types of items with embedded options and filtering rule, and calculates portfolio level risk as the output
 • Reduced 1000+ lines of code, implemented unit tests, and deployed programs using Python to serve daily usage for the model's vertical.
 Portfolio Company (Co. Name)

• In charge of the experiment design and result analysis leading to \$50K incremental annual revenue (the company leading to \$1M data purchase)

• Collected and analyzed loan quality data from XX, and built regression models to ascertain quality of debt and estimate incoming revenue (loan amount) before company's decision to purchase millions of discharged loans

TECHNICAL SKILLS

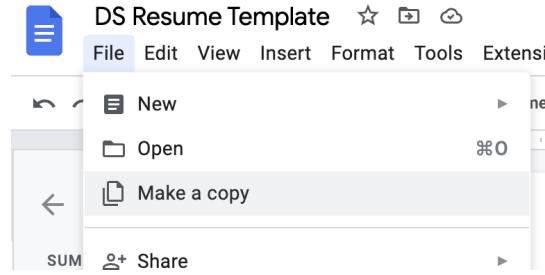
• Language/Tools: Python, R, SQL, C++, Go, LaTeX
 • Data Science: Data Mining, Machine Learning, Big Data, Pandas, NumPy, Beamforming, MapReduce
 • Machine Learning: Covariance matrix optimization, Classification (Random Forest, KNN, SVM), Regression Modeling (Linear, sparse, logistic, regularized), Principal Component Analysis (PCA, PCR, sparse PCA), clustering (K-means, hierarchical)
 • Stats & Experimentation: Time-Series Analysis (OLS, GARCH, ARIMA, MLE), hypothesis testing, Monte-carlo simulation, statistical testing, Covariance and correlation modeling

Metric Right ↗
(Note to Editors: You can select the icons and use Command + K to replace the links.)

Month Year - Month Year

SELECTED PROJECTS

• 50+ raw datasets are first collected with web-scraping, then cleaned by removing noises and normalizing the data, time aggregated to have consistent frequency and range



2. Update your resume using the template!