

# Creating a Resume for Data Analysis and Data Science Jobs as an Undergraduate

**TLDR: Your resume is not a comprehensive account of your background and experiences, but rather a way to highlight the parts of your background and experience which are relevant to the job you're applying to. You should make it as easy as possible for a potential employer to connect the dots between their needs and your skills. For data analysis and data science jobs, that means going into a lot of detail about your experiences doing statistical analysis and coding and less detail about other areas.**

## **Things to include on your resume, if you have them:**

- Math/statistics/probability/programming classes, either in college or elsewhere (Coursera etc.) You can list this in the education section. If you have taken more than a few of these, prioritize statistical analysis, programming, and upper-level courses. For instance, list Data Structures and Program Design over Intro to Programming. List Statistical Computing in R over Introduction to Statistics. You're just trying to show you have some academic background with this, particularly for fields you didn't major in.
- Things that communicate you did well in college. Like, your GPA if it's fairly high, honors, that kind of thing.
- Non-academic college activities/leadership stuff/athletics/volunteer experience. I don't think this should be a big part of your resume, but you're potentially conveying a few things here: "I managed my coursework on top of major time commitments", "I have organized meetings/organized people/made budgets/handled responsibilities." If you have a lot, list the parts that best convey these types of things.
- Work experiences, even if they're not related to the kinds of skills we're going to focus on below. Again, not a big part of your resume, but having had a job where you had to get there on time and get along with people – that's a good thing to communicate.
- Anything (work experience, personal projects, coursework) that can convey the following skills: data analysis, programming, writing/other kinds of communication. This is what you're going to go into the most detail about, particularly data analysis and programming. This should be a substantial part of your resume. If you don't have much to put on this, seriously consider finding some opportunities to do the kinds of projects that you can then list on your resume.
- A link to your GitHub account, if you have a GitHub account. (And if you don't, seriously consider it!)

**You can think of this from an employer perspective: someone is looking at your resume, they have a type of project in mind that they want a junior person to do, and they're asking "is this someone who I can give this to and they'll be able to get up to speed on it fairly quickly and without a ton of oversight?" Part of that's related to general**

**competence/responsibility, part of that's related to specific skills. The purpose of your resume is to help them connect the dots between who you are and what they need.**

Something you did in a job or as a personal project is generally better than something you did in a class, but if what you have is classwork, talk about that.

Get very detailed! Here are some examples of lines on my resume:

- Presented on Type 1 vs. Type 2 error in the context of a machine learning binary classification model in order to enable the client to make decisions about where to set the model threshold
- Used the ArcGIS Python API and Python libraries like geopy, shapely, and geopandas for data analysis relating to address checking, geocoding, and assignments of zip codes to their closest offices
- Created user-friendly, well-designed maps for data visualization in Python
- Implemented hill climbing algorithms in Python for finding solutions to data analysis/optimization problem
- Wrote clean, documented, maintainable code, including test cases when appropriate, and which automate the creation of deliverables including graphs and charts for data visualization

**Here are some features that are present above which I recommend including when you talk about your data analysis work. Not all of these will be relevant for everything you do.**

- What language did you use? What packages?
- Conceptually, what were you doing? Descriptive statistics? A linear regression? Markov modeling? A random forest?
- What were the outputs, in plain language? A paper? A graph? A dashboard?
- What problem were you solving?
- Who was this for? Did you get something published?
- Is there anything about your processes that indicates 'good coding practices'? Like, did you use git? Did you write documentation? Test cases? Did your code go in a module? Are other people using it?

**Here's what I would leave off your resume, due both to space constraints and because it's not necessarily that useful:**

- Information that's only going to make sense to someone from your field for jobs that aren't in that field. Something like "published a paper" or "presented at a conference" is good to have, and you might even include the name of the paper, but don't use space on your resume for something that the people looking at aren't going to have the context to understand. Put it in language they will know, or leave it off.

- Lists of programming languages or software proficiencies. If you want to show you know a thing, find a way to talk in your resume about how you've used it. That lets you get more specific. There's such a huge range in what people mean by proficiency in a particular language that it's not very useful.
- Any kind of statement of purpose/who you are/what you're looking for. That might go in a cover letter, if there is a cover letter.

You may have more than one resume if you apply to different kinds of jobs. Like, if you apply to research jobs in your specific field, you'll want a version of your resume that goes into more detail about research and coursework in that field. You may also make light tweaks to your resume for different jobs – like, if you're applying to a job that specifically asks for experience with a tool, or with a particular field/data set, you can modify your resume to go into more detail about that. (For instance, adding that you've taken a health economics class if you're applying to a job involving health data, and going into more detail about the health aspects of any data-related work you've done.) Similarly, in general you should prioritize listing work with R, Python, and SQL over work in SAS, Stata, SPSS, and Excel - but if those tools are listed on a job application, even if it also lists R and Python, you may also want to write about your experiences with those.

**Finally, keep it to one page! Later on, you might have multi-page resumes, but not yet!**