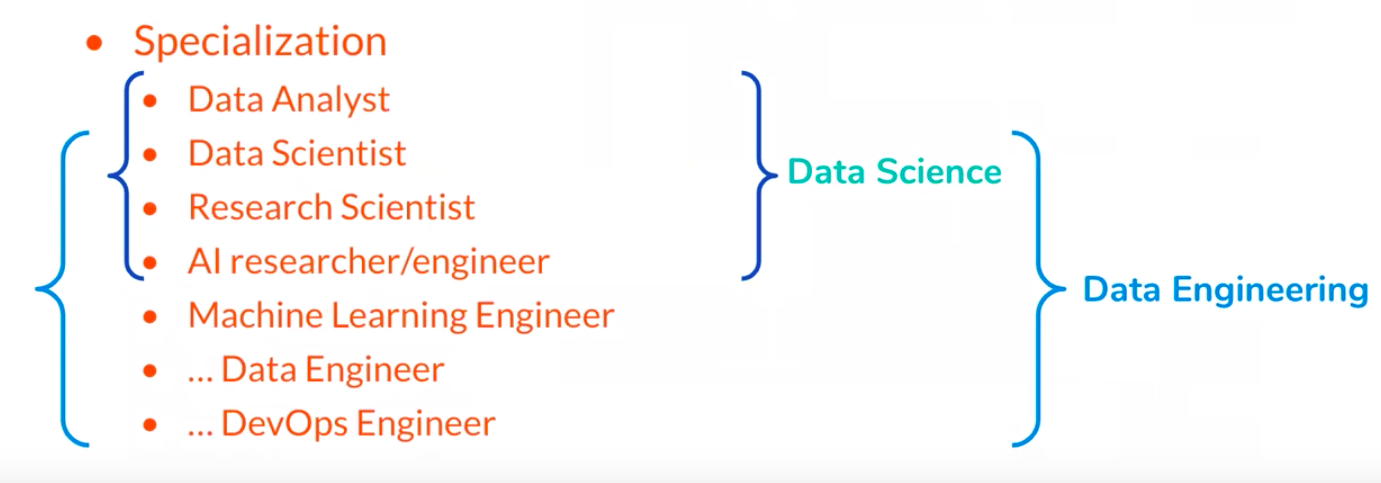
**Data Science interview tips**

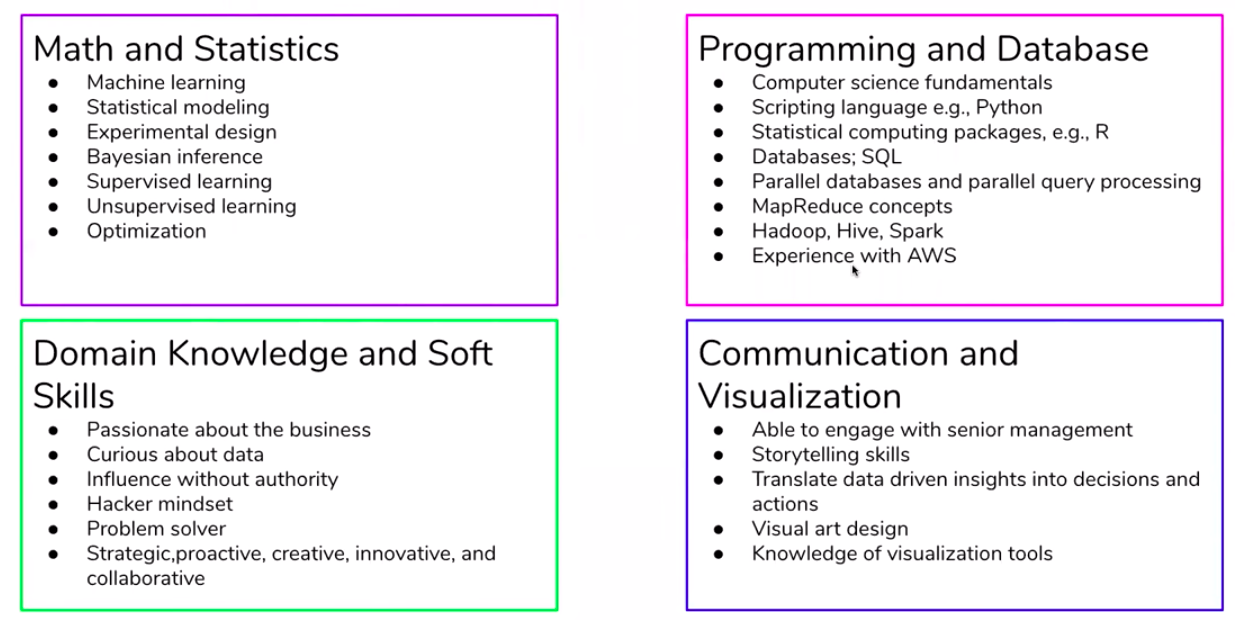
* Why PhDs?
* Tell me about your research
* How do you problem solve?
* Taking feedback
* Communicating the big picture
* Where do you fit in?

**Highlight examples of things that they’re looking for**

* **Make a table of the things they’re looking for and make columns for Fluidigm and columns for my project of when I’ve done that**
* **Sprinkle jokes in ☺**
* **Do the same with Alector**
* Job market
  + DS jobs are growing and the specialization of jobs are growing
  + Data engineering **uses some APIs to get big data**



* Who are academic scientists?
  + Plan/execute study takes months/years
  + Collect and clean data
  + Use programming and statistics/ML to discriminate signal and noise
  + Convey results to the scientific community
  + There’s a gap between what we get training for academia and what we want to show as a data scientist
* Who are data scientists?
  + **Plan and execute a study in week-long sprints (talk about how I did it over a period of months – show github)**
  + **Collect and clean data (API)**
  + **Use programming and statistics/ML to discriminate signal and noise**
  + Convey results to **team/company/investors (talking to internal team)**
  + **Make data-informed decisions that directly impact a product and business (e.g. data on whether they should launch a new product, C1 HT)**
* Why PhDs?
  + Transferrable skills – don’t need all of this but emphasize more the communication/visualization and soft skills
    - E.g. adjusting storytelling skills depending on whether I’m talking to scientists or people
    - ID the skills I’ve done here and talk about:
      * what you’ve done but don’t like
      * what you’ve done and like
      * **what you haven’t done but want to learn**



* Highlight things where the job fits you
* Highlight things that are important across all job skills. **E.g. are you a problem solver? What’s an example? (maybe have a couple of stories)**
  + Outdated RNA-seq pipeline (old reference, old tools)
  + How did you assess it? Did you create a feasible timeline? Was it completed and what were the results that provided impact?
  + Re-wrote it, provided new metrics, provided visualizations that are automatically generated so that people can assess their experiment and be able to make the next decision whether it’s an experimental optimization or whether it’s closer to show to marketing or beta-test customers, etc.
* **Identify what you want to learn. Know the breadth of some stuff, but then highlight what you want to specialize in. Be able to convey what you don’t know and then have something that you want to be fluent in.**
* For data science: focus on math and statistics box, especially ML
* For data engineering: focus on programming and database, have to know SQL
  + Large companies use tons of SQL
* Tell me about your research or tell me about your background
  + Behavioral part of interview
  + Will apply to insight or any job interview or talking to recruiter
* Describing your background effectively
  + Want it to be a very good 30 seconds; concise and clear
  + Want it to be specific to the job (e.g. provide data science relevant things to a data science job discussion)
    - If you talk about something, ground it into something that the average person can relate to.
    - Convey impact, why it’s interesting, and skills it highlights
  + Listen to directions and queues
  + Keep it conversational and interesting!
    - Remember a person is interviewing lots and lots of people – stand out!
    - **Use Toastmaster techniques in terms of vocal variety and body language**
  + Anything in my background is fair game – know it!
    - Turn this into an advantage
    - Focus on the projects that worked out the best or you know most about (e.g. talk about how TRseq has lead to sales); pick a “flashier” project
  + STAR(T)
    - Situation
    - Task (what task did you do; what did I specifically do?)
    - Action (what action did you take – similar to above)
    - Results
    - Takeaway (what’s the impact; saying it was in a good journal isn’t really that interesting – they what to know the DS impact)
* You’ll be asked: Tell me about a time when you had a conflict with a team member? Or conflict with adviser?
* How do you problem solve?
  + On paper a lot of people actually look similar, they use the same keywords, etc.
  + They want to know you solve problems
* Showing code effectively
  + Start at the top (high-level) – **use table of contents to your advantage**
    - Dig into most interesting aspects
  + Keep it conversational
  + Mind the time
  + Never go through line-by-line (very important!)
  + Basic script:
    - “Here’s a high-level problem I’m trying to solve”
    - “Here’s how I downloaded, cleaned data – considerations I had to make”
    - “Here’s the algorithm I looked into and why I chose it”
    - “Here are the most relevant features”
    - “Here are the final visualization that shows the main result”
  + Keep it conversational and keep it moving
    - The interviewer knows what you’re doing, e.g. splitting data, etc.
    - Touch on points that should be taken for granted (e.g. cleaning data, splitting, turning into categorical variables… don’t get stuck in the weeds)
* If you’re not showing code:
* How you would address a new scenario: Amazon sells millions of products per second. They might want to launch a new coupon. You’re the DS, who is supposed to advise them on what to do. You have 1 week… 1 month.
  + Data? (What kind do they have?)
  + Approach? (Can you get features of how the coupon was presented to customers, labeled with those that have bought, to build a model that can predict which coupon presentation is effective? Or specific industry? Or what metric would be used)
  + Evaluation? (How do you know it’s right? What graphs and visualizations would you create? Maybe do some clustering analysis.)
  + Validation?
  + Note: Facebook runs experiments internally (on their own employees) but Amazon may not want to launch too many experiments. What do you want to do if they **don’t** want to experiment? (They may want to put you in an intentionally challenging situation and see how you’d approach it.) You can dig into data more. (Maybe use historical data that was a similar situation. If you had a week, or month, or 3 months – what would you do?)
  + They want to see that you can outline a plan and provide a framework for it. They want to see this:
    - You can structure a problem
    - Creatively approach it in a way that’s feasible and uses existing data if possible; if not, what new experiments would you do?
* Taking feedback
  + How do I take feedback and learn from it
  + Demonstrate that you can give things a thought but also show that you can learn, work on a team, and let other people contribute/ask questions
  + Demonstrate that you can **iterate constantly (note that the theme of iterating is an Insight culture thing)**
  + Show that you can work with the person
  + Ask them constructively for feedback
  + Check in with the person who is asking for feedback
  + Shows that you’re a communicator and team player
  + Get out of the solo/academic ownership mindset
* Communicating the big picture
  + Product, not science
  + Product focuses on the end user, data-driven answers (an end-user could be an executive or decision-maker)
  + Science focuses on the novel method, the interesting exploratory problem
  + Get an answer, figure it out, move through, then try to get a better answer.
* Where do you fit in?
  + What type of data scientist do you want to be?
  + Have you evaluated the field?
  + **Look at blogs, job postings, companies, industries…**
  + Consider work-life balance, mobility in the company, etc.
  + Whatever experience you’ll get, you’ll be good at it… so think carefully about what you want to get experience in – analyst, visualization, engineering, etc. – it puts you on a certain trajectory
  + Where do you fit – what do you consider important?

Test git