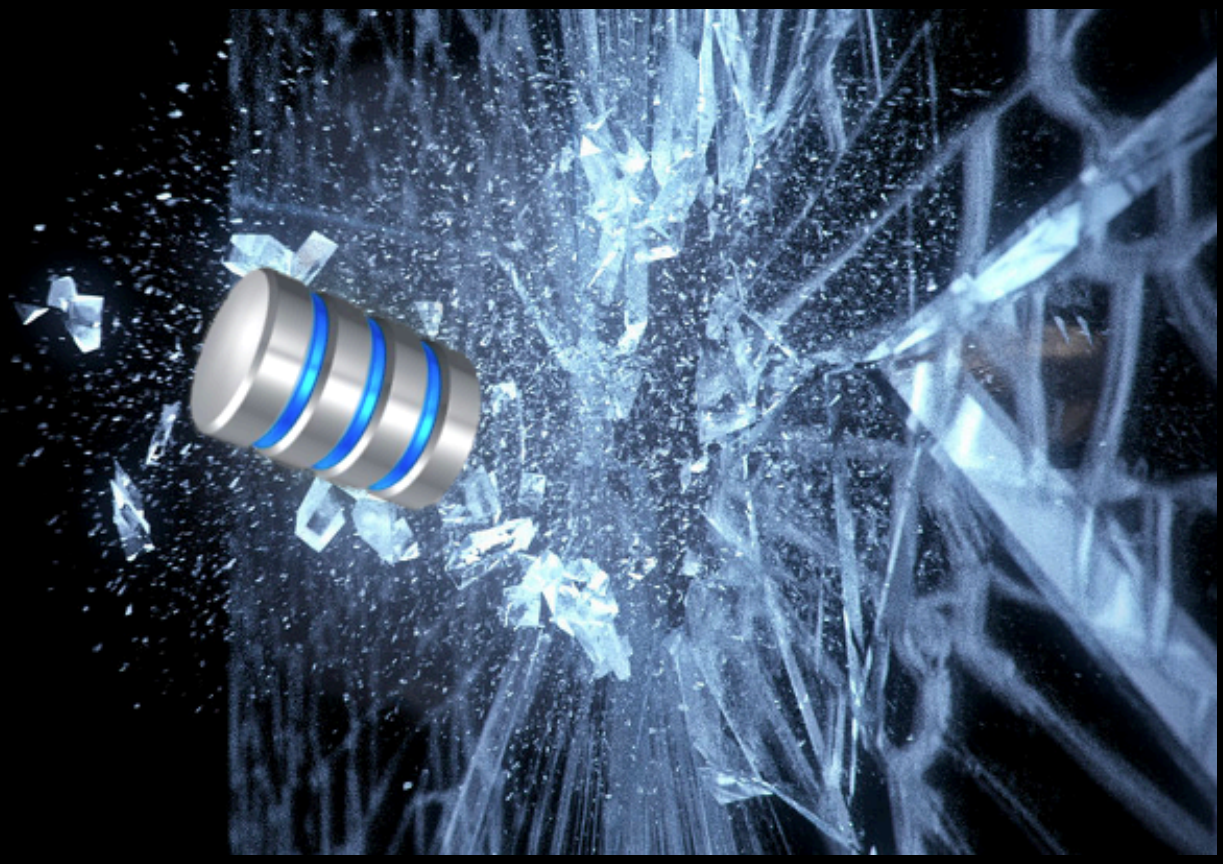
Breaking Into Data Science

Part 1



25 Tips on #BreakingIntoDataScience by Michael Green

"I'm really interested in #datascience.

How can I break into this field if my #career has been in a completely different domain thus far?"

This was my dilemma few years ago, when I was working as an engineer and discovered my true passion in #data.

I had to decide between continuing on the path of least resistance, or making a career change to become a #datascientist.

I chose Option 2.

Did you know that you can break into the field of Data Science, even if you've started down a completely different career path?

Did you know that you can do it at your own pace?

Did you know that you can do it for free?

In fact, if you play your cards right, you may even get paid to do it.

I invite you to join me for a series of posts in which I share a few tips and tricks for mid-career professionals who want to get serious about #datascience.

Comment #BreakingIntoDataScience below if you're interested in reading this series. 👇

I look forward discussing this topic with you!

#BreakingIntoDataScience Tip #1:

Your domain expertise is your superpower.

When I first made the decision to make a career switch into #datascience, I thought:

"If only I had studied statistics or computer science in college.

I spent years studying mechanical engineering and working as a product designer, only to start over again in my #career."

Looking back, I believe I had it all backwards, for three reasons:

#1) Domain expertise makes you more marketable as a #datascientist, not less.

If a manufacturing company needs help with their data science, they will hire someone with manufacturing experience and "just enough" data science knowledge, rather than a PhD in statistics without any manufacturing experience.

#2) Domain expertise gives you an advantage in problem solving.

Instead of viewing every dataset as "just a bunch of variables" for a regression problem, a domain expert will be able to understand the context behind the data.

#3) You got involved in your domain for a reason.

Your passion for solving problems within your domain will drive you to succeed where others might give up.

Therefore, I say that jumping into data science from another domain is one of the best ways to go about it.

I'm curious, what would you say is your unique "domain of expertise"?

#BreakingIntoDataScience Tip #2:

Being a #datascientist isn't really binary.

I'm not sure you can say, "Yesterday I wasn't a data scientist. But today I'm a data scientist."

Instead, the life of a data scientist is lived on a spectrum.

First, #datascience is a #hobby...

Which, at some point, turns into a #passion...

Which quickly grows into an #obsession...

And eventually becomes a #profession.

Spend enough time doing it, and it becomes a #career.

You probably know dozens of people on the data science spectrum.

Some are just beginning the journey, others are having some fun, still others are making a living of it.

Anywhere on the spectrum is 100% okay!

But if you're interested in getting serious about data science...

About getting paid to do the work you love...

You've come to the right place to talk about it.

Follow #BreakingIntoDataScience for helpful tips on... well, you know, breaking into data science.

I'm curious, where would you put yourself on the "data science spectrum"?

#BreakingIntoDataScience Tip #3:

Aspiring #datascientists can actually improve their job prospects by narrowing their focus.

When I first tried to make the #career switch from mechanical engineering to #datascience, I thought to myself:

"I will apply to ALL the data scientist jobs."

After a few hundred (thousand?) applications and barely a serious interview to speak of, I realized I needed a more effective approach.

I narrowed my scope to something like:

"I'm looking for a data scientist or analyst position in the #supplychain or #manufacturing space. I'd prefer the #ColumbiaSC area. If I can work within a new or growing #analytics department, I can continue to develop my skills alongside my team as it matures."

👆 That narrowed scope ultimately described my first job in data analytics.

The experience I gained from that position eventually made me much more qualified for my subsequent role as a data scientist at a fast-growing startup.

How about you?

Have you ever found that narrowing your #jobsearch focus actually increases your chances?

#BreakingIntoDataScience Tip #4:

Do something small each day.

At some point along your journey into #datascience, you'll want to take an inventory of what skills and experience you currently have.

And you'll compare that list to the skills and experience that are required for the job you're seeking.

This is called "gap analysis", and it allows you to identify what skills and experience you should be developing ("gaps") in order to get from A to B.

Once you are aware of your gaps (different for everyone), you might feel overwhelmed at first.

It can be tempting to try to work on all the gaps at the same time.

I tried this when I first decided to pursue a #career in data science.

I paid for an expensive 1-year course, tried to learn R programming overnight...

I burned out. I didn't finish the course. I didn't make much progress at all.

Months later, I tried a different approach.

Just a little bit each day. Something like...

✓ Reading an article about a #machinelearning algorithm
✓ Spending a few minutes learning a new function in R
✓ Starting conversations with data scientists on #LinkedIn

After a while, I started to see some progress.

I realized that becoming a #datascientist isn't a binary thing.

It’s taking 1000 small steps in the right direction. Agree?

#BreakingIntoDataScience Tip #5:

Avoid unnecessary competition.

Example:

Let's suppose you're seeking your first #datascientist role, and you're coming from a background in finance.

You see two job postings:

1) Data Scientist at Microsoft in California. Candidate will join a global team of 50 Data Scientists and Data Engineers.

2) Data Scientist at a small financial firm in your hometown. They have a growing team of 2 Data Scientists and 1 Data Analysts.

Which should you probably apply for?

In the case of the Microsoft position, it's pretty likely that you'll be competing with PhD's in statistics with 5 yrs experience in #datascience.

The second position is interesting for two reasons.

First, with your background in finance and your existing ties to your hometown, you may in fact be a very competitive candidate from the hiring manager's perspective.

In addition, it's unlikely that all the PhD's who applied to the Microsoft position will be competing with you for the spot on the smaller team in the more obscure city.

You don't need to be the most qualified person in the world.

You just need to be the best fit that your prospective hiring manager can find.

Boost your #career by avoiding unnecessary competition.

#BreakingIntoDataScience Tip #6:

If you want to stand out in a competitive job market, don't do what everybody else is doing.

(Apologies in advance. This post will be controversial and oversimplified.)

How many times have you seen someone post about their #100daysofcode challenge?

How many times have you seen someone post about their Titanic project on #kaggle?

How times have you seen a notification that one of your connections has competed a #MOOC course on #AI?

If you've been following #datascience for more than a few weeks, you've probably noticed that LOTS of people are doing these things.

To be clear, these are all good things that can help you learn.

But none of these things sets you apart as a competitive candidate for the #datascientist role you want.

To really stand out in the job market, you will need to do something different.

Something better.

We will discuss this further in upcoming posts, and the truth is, it will look different for everyone.

But in order to kickstart your #career in data science,

I'd recommend paying attention to what everyone else is doing...

And, (where it makes sense for you), not doing that. Agree?

#BreakingIntoDataScience Tip #7:

Not everyone in #datascience is a #datascientist.

Here is a very non-exhaustive list of other hot jobs in the data science domain right now:

Data Engineer
Data Analyst
Data Architect
Data Evangelist
Business Intelligence Developer
Software Engineer
Algorithm Engineer
Research Engineer

...and there are many others.

This is good news for 2 reasons:

#1) There are lots of opportunities to practice data science, even without the exact title "data scientist". If you know how to spot these jobs, your chances of getting hired will increase.

#2) As you nail down exactly how you want to contribute to the data science space, you may find that one of the above job titles better matches your desired "flavor" than the traditional data scientist job description.

I'm curious, what other job titles come to your mind, that I may have missed on the above list?

Bonus points if you can think of one that nobody else has mentioned. 👇

#BreakingIntoDataScience Tip #8:

If you want to make a #career change, your #LinkedIn profile and resume should be marked by simplicity and clarity.

I get it. You're unique. And your path into data science has probably been a bit complicated.

You have a specific set of experiences which set you apart from everyone else entering #datascience.

It can be tempting to fill your resume with ALL of your past experiences to show your uniqueness, but this comes at the risk of muddying the water and confusing the hiring manager.

Hiring managers want you to hand them an easy-to-digest narrative, showing why you're best qualified to solve their particular set of #datascience problems within the context of a particular job description.

✔️ Data science experience/training
✔️ Relevant domain experience
❌ Additional jobs or details unrelated to the above

By tailoring your profile/resume to the job you want, and eliminating the less relevant details from your personal journey, you'll dramatically increase your chances of getting hired.

Agree? 👇

#BreakingIntoDataScience Tip #9:

"What if I have zero job experience in #datascience?"

No problem! You can create your own experience.

Here are two things that have worked for me:

#1) Take initiative in your current job to solve a problem, no matter how small, with data science.

Most likely, nobody will ask you to do this.

Your coworkers may not understand why you'd even want to attempt it.

But if you end up getting a good result, everybody on your team will be glad you did it!

Even if you don't end up with life-changing results, the process you followed has given you some experience, and it will give you good talking points on your resume and in your interview.

#2) On your own time, work on a personal project that you care about.

For example, a few years ago, I was really interested in learning about patents.

I taught myself how to write a few lines of R code, in order to scrape publicly available patent data from the web.

I then taught myself how to write a few more lines of code, to clean and analyze this data.

There was no objective except satisfying my own curiosity.

But in the process, I had fun learning some coding with R.

And you can bet that in my next job interview, I used this experience as a talking point to get the job offer! Thoughts? 👇

#BreakingIntoDataScience Tip #10:

"#R vs #Python?"

It doesn't really matter.

To get started in #datascience, just take a free intro course on either one.

No big deal, it's not a huge commitment. You can't go wrong.

Then pick a problem that you want to solve or a dataset that you want to explore.

Then start googling "how can I do ___ with X" where X is the language you're starting with.

Very quickly, you'll begin to learn the patterns involved with coding, which are largely shared between both languages.

You'll also begin to understand the strengths and weaknesses of the language you're using. If you need to switch languages at any point, you'll know.

And with the experience you've built up along the way, the switch will be pretty easy.

I started using R a few years ago. Then I started learning Python a bit later. I use both nowadays.

The truth is, each employer will prefer you to work with different #tools.

You can't predict ahead of time, which tool will really be best for you long-term.

Just start gaining experience in whatever tool you want to try.

Learn to be flexible and find the information you need, when you need it.

And don't sweat "R vs Python", even for one second. Agree?

#BreakingIntoDataScience Tip #11:

#Storytelling is so important.

When I first set out to make a #career pivot into #datascience, I would send my resume to dozens of hiring managers, and never hear back from any of them.

After a while, I realized that the problem was the way in which I was presenting the facts of my career journey.

Like many job seekers, I was filling my resume with the technical details of my previous roles, 95% of which were irrelevant to the job I wanted.

I decided to try a different method called "storytelling".

Instead of stuffing as much information as I could into my resume, I would focus on the 5% of my experience that was relevant.

I would craft a compelling narrative around the problems that I had set out to solve, the methods that I had used to get the job done, and the exciting results that my team and I were able to achieve in the end.

And I came to a surprising realization.

It didn't matter if only 5% of my previous experience was relevant to the job I wanted.

If I could make that 5% shine with some thoughtful storytelling...

That 5% was more than enough to land the job offer.

I'm curious, have you used storytelling to your advantage in the job search?

#BreakingIntoDataScience Tip #12:

Your #data wrangling skills are more marketable than your #model building skills.

There exist software tools today, which can automatically generate #machinelearning models better than what ~90% of #datascientists can build, and 100x faster.

These tools are becoming increasingly popular in the business environment, but they only work if they are provided a clean dataset for a well-formulated business problem.

A rising trend in #datascience is to focus the #human efforts on these key areas, and let machines do most of the heavy lifting of model development.

Humans are much better at:

✓ Cleaning data: Taking messy, disparate sources of information, understanding the context, and transforming the data into a single, organized table at the appropriate level of detail or aggregation, which can then be used for machine learning.

✓ Formulating the problem: Understanding the true business need, properly posing the right data science question, and gathering the appropriate data for the machine learning problem.

Over the next couple of years, demand for people who are proficient at the two skills mentioned above, will far outpace demand for model builders.

I'm curious, are you seeing any similar trends in the data science job market?

#BreakingIntoDataScience Tip #13:

If you want to get hired in the competitive #datascience job market, you've got to short-circuit the #interview process.

Let's face it. If your best shot is uploading a faceless resume to a company's online job portal (read: black hole for resumes), you've already lost.

What if the hiring manager were to invite you to apply for the #datascientist position?

Then, the rest of the process would become a formality. They'd already be interested in hiring you from the start. Everything would just fall into place, kind of like magic. Or #AI. But I repeat myself.

Anyways, let's talk about how to get a personalized invitation from the hiring manager.

People only want to hire people that they know and trust.

How can you build trust with a #hiring manager you've never met?

Simple.

1. Go find the hiring manager on #LinkedIn.
2. Start an honest conversation. Ask a lot of questions.
3. After you've gotten to know them, ask if they think you'd be a good fit for the position they've posted.

If they say "yes", 75% of the interview process is in the bag. Just show up and shine.

If they say "no", that job portal wasn't going to help you anyway.

I'm curious, have you ever used an approach like this to get a job offer?

#BreakingIntoDataScience Tip #14:

If you want to get #hired in #datascience, you need to deeply understand the persona of your "ideal hiring manager".

Your ideal hiring manager is the imaginary person who would consider you to be the best possible fit for their team.

Think through these (and many other) questions to develop your ideal hiring manager's persona:

• What industry do they serve?
• What domain-specific problems are they trying to solve right now? (Prevent customer churn, reduce scrap rates in manufacturing, detect fraud, etc)
• Which area of data science are they most heavily involved in? (NLP? Image classification? Customer segmentation? Reinforcement learning? Etc...)
• How mature is their #datascientist team?
• What unique value can you offer them?

Once you have an idea of the type of hiring manager you want, start #LinkedIn stalking.

Go and find people that look, talk, and act like your ideal hiring manager.

Study their profiles. Send connection requests. Start conversations with them.

If you spend enough time getting to know your "ideal" hiring managers, they'll start to rub off on you.

The goal isn't to ask them for a job.

It's to learn how they think.

It's to build relationships.

Play the long game. In time, the job offers will come to you.

#BreakingIntoDataScience Tip #15:

While others are talking, build.

There are three types of people in the #datascience community:

#1) Those who talk, but don't build
#2) Those who build, but don't talk
#3) Those who build and talk

Most of the hype comes from people in the first camp.

They write blogs and attend #AI conferences. But they don't know how to build a solution to a data science problem. Or maybe even formulate a data science problem.

Don't be like them.

The second camp is full of either (a) highly specialized people who become isolated from the larger community, or (b) newcomers to data science, who are learning and building their skills.

It's totally OK to be part of the second camp. For a while at least. Much better to be in group 2 than group 1.

Finally, there is the third camp. These are the true practitioners of data science. They have meaningful substance behind what they say. Sharing their work with the world is just icing on the cake, a way to help others while realizing their true potential.

I think everyone should aspire to be in Group 3.

Even if it means you spend a while building while others are talking.

Short term, this means that others get the attention while you work hard.

Long term, you will win. Agree?

#BreakingIntoDataScience Tip #16:

The best way to start your #career as a #datascientist is to infuse your current job with #datascience.

A few years back, I first encountered #machinelearning and thought,

"I want to make a career out of this. But I don't have any formal training in #statistics. And I don't have a strong #coding background. Not to mention, my current job has very little to do with data science. How can I get THERE from HERE?"

I decided to try applying data science to my day job.

Never mind that I had no clue what I was doing at first.

I started close paying attention to the business problems around me.

Listening to the pain points of my coworkers.

Asking myself, "can this problem be solved with machine learning?"

Whenever I caught a whiff of a #data problem, no matter how small, I offered to help.

And I realized that almost every industry has data science problems.

I started small, building proof-of-concept models in #Excel.

As I gained confidence, I started to scale things up with a few lines of #R code.

Before long, I was competing for full time data scientist jobs in earnest.

It didn't happen overnight. But it happened. And I don't think I'm the only one.

If you have done this, or are doing it now, would you share your perspective as well?

#BreakingIntoDataScience Tip #17:

The best kept secret of #datascience? Paid apprenticeships.

One of the most common sources of frustration for an aspiring #datascientist is that most companies seem to be looking for experienced data scientists.

It's a classic chicken and egg scenario.

How to get a #job without experience?

How to get #experience without a job?

This is especially difficult for those who have already begun their #career journey on a completely different path.

Fortunately, there are a number of innovative companies (such as DataRobot) offering paid apprenticeships to aspiring data scientists.

These companies realize three truths about mid-career professionals:

1. With a little on-the-job training and mentorship, they are capable of ramping up quickly and making significant contributions to the organization.

2. While they may not have formal training or experience in data science, they bring invaluable domain knowledge to the table.

3. Importantly, these mid-career professionals rely on a steady income.

Paid apprenticeships make the transition into data science easier by allowing mid-career professionals to learn, while solving real problems, and while maintaining an income.

Are you aware of any other companies offering data science apprenticeships?

#BreakingIntoDataScience Tip #18:

The best way to start building your credibility in #machinelearning?

Complete your first proof of concept (POC), no matter how small.

A POC is a mini-project which involves:

• Formulating your business problem as a machine learning problem
• Building a #model from your #data
• Using the findings of the model to drive meaningful action

A POC demonstrates that you can successfully deliver valuable outcomes in your domain using machine learning.

I think now would be a good time to let you know about a cool perk that my employer, DataRobot, is offering me this month.

As an employee, I'm allowed to hand-pick a few of my colleagues, who will receive free support in the development of their own machine learning POC.

This will include:

• Partnership with DataRobot's world-class #datascientist team to identify and formulate a valuable POC of your choosing in your domain
• Training on how to use DataRobot's cutting-edge automated machine learning platform to build highly accurate models in minutes (no coding required)
• Calculation of the business value of the POC, which you can use to demonstrate credibility to your management

If you would like to be considered for this program, please send me a private message and let's have a conversation. 💬

#BreakingIntoDataScience Tip #19:

Be patient with yourself.

Everyone's path into a #datascience career looks different. And that's ok.

There are a lot of things that a #datascientist must learn.
✓ Statistics
✓ Programming
✓ Domain knowledge

There are a lot of ways to get there.
❓ Degrees
❓ Online courses
❓ Apprenticeships
❓ Self-guided learning

No single approach works for everyone. No single timeline works for everyone.

It took me ~ 3.5 years to
• Decide that I wanted to become a data scientist
• Teach myself the basics of coding, statistics, and algorithms without a degree
• Apply my learnings to my engineering job
• Get an analyst job and learn even more
• Get the data scientist job

During those years, I saw many people move into the profession ahead of me.

I got stuck on a few things that other people picked up easily. I learned a few things more quickly than others. I've learned how important it is to stop comparing myself to others.

I'm happy with the progress I've made, the job I have now, and the things I've been able to learn and do during the transition.

I hope you also find the right #balance that works for you. It could take weeks, months, or years to get where you're going. Pace yourself, and enjoy the journey along the way. Thoughts?

#BreakingIntoDataScience Tip #20:

Don't memorize #code.

Just #google it as you need it.

Every hour spent memorizing code,

is an hour not spent on building projects.

Projects are where your growth happens.

Focus your efforts on solving practical problems.

Don't worry if you don't know how yet.

Think of what you need to accomplish with your code,

And then figure out how to get there.

A little just-in-time research goes a long way.

By the time you memorize something in #datascience...

It might already be obsolete.

Agree? 👇

#BreakingIntoDataScience Tip #21:

Focus the conversation on THEIR problems, not on YOUR qualifications.

I submit the following two #interview scenarios for your consideration:

SCENARIO #1

Candidate: "I'm a #datascientist, recently graduated from Bootcamp XYZ. I'm skilled in NLP, Time Series, and AI."

Hiring Mgr: "Cool, you sound like the other 25 people I've interviewed today."

SCENARIO #2

Candidate: "I noticed from your website that you're investing heavily in your supply chain right now. I'd venture to guess that maintaining healthy inventory levels is a challenge in your industry?"

Hiring Mgr: "Why yes, it's really hard to stock the right products when we have no clue what our customers are going to purchase next month."

Candidate: "Yes, demand forecasting can be quite challenging, especially with the large portfolio of products that your company manages. This reminds me of a time when I built a time series app to improve a client's forecast accuracy by 10%. Would you like to discuss how we could modify my previous approach to account for the particulars of customer behavior in your industry?"

Hiring Mgr: "Just take my money."

People hire data scientists because they have domain-specific problems that need solving.

Focus the conversation accordingly.

Agree?

#BreakingIntoDataScience Tip #22:

Build your professional network before you need it.

Have you ever heard the proverb below?

"The best time to plant a tree was 20 years ago. The second best time is now."

It's the same with #networking.

Meeting people and building relationships takes time.

If you're desperate to find a job within a month, it can be difficult to focus on forming genuine relationships with your colleagues.

Conversations can start to feel rushed.

You may be tempted to ask people for a job or referral,

before they've had a chance to get to know you.

Building #trust takes time.

Building #relationships takes time.

Give yourself, and your colleagues, the gift of time.

Start making friends and learning from them in the context of low-pressure, meaningful conversations.

Start building your network today.

Especially if you don't need anything in particular right now.

Your future self, and your newfound colleagues, will thank you.

Care to add your thoughts? 👇

#BreakingIntoDataScience Tip #23:

Help #recruiters help you.

Have you ever stopped and thought to yourself, how hard it must be for a recruiter?

Every recruiter I know:
✔️ Genuinely wants to help people get a #job they love.
✔️ Has a finite number of available positions.
✔️ Cannot possibly give the job to everyone.
✔️ Feels terrible about it.

They want to help everyone, but they can only help some.

Nevertheless, I find that many job applicants have unrealistic (unspoken) expectations for people who are #hiring.

"Surely if I want the job badly enough..."
"I am sure I'm best qualified..."
"Please view my resume/profile..."
"Please get back to me at your earliest convenience..."

Little do they know that there are 50 other people who have contacted the same recruiter, on the same day, about the same position, and whose resumes look EXACTLY the same. I hope you'll indulge me as I illustrate with a limerick. 🤣

There once was a recruiter named Bob,
who wanted to help people get a job.
He published a job post,
expecting 10 apps at most,
and was trampled by a 1,000-strong mob.

In order to make it easy for a recruiter to help you, you've got to get creative. You've got to make it painless. Give, not just take. That's what this series of posts is all about. Thoughts? 👇

#BreakingIntoDataScience Tip #24:

How to get buy-in from #stakeholders for your #datascience project?

Here are some best practices 👇

1. Let stakeholders talk for most of the meeting.
2. Listen intently.
3. Ask plenty of clarifying questions.
4. Say, "Here is what I think I'm hearing from you..." and restate their perspective in your own words.
5. Make it your priority to understand their expressed needs.
6. Perhaps equally important, help them feel understood.
7. Do your homework. Study the context. Experiment. Become an expert on your stakeholders' problems.
8. Form an educated (data-driven) opinion. Sometimes, this will be vastly different from the opinions of your stakeholders.
9. Say to stakeholders, "Here are some thoughts I have so far...",
10. Then add, "here are some factors that influence my thinking on this" and explain your position (clearly, logically, concisely, calmly).
11. Follow with, "what do you think?"
12. Truly listen to their response.
13. Take special note of any misalignment between your educated opinion and theirs.
14. Repeat steps 1-13 until you and your stakeholders are fully aligned.

A few iterations of this will help your stakeholders to trust you.

And you'll learn from them.

What would you add to this list?

#BreakingIntoDataScience Tip #25:

If a tree falls in the forest, but no one is around to hear it fall, did it really make a sound?

If you build an awesome data science project, but nobody knows about it, did it really happen?

Of course it did. You learned something, even if nobody else noticed.

But that's no way to become a #datascientist.

An important responsibility of a data scientist is communicating with stakeholders (or hiring managers).

Here's how to make sure you get credit for all your hard work on that project. 👇

DRAGON-SLAYING STORIES (DSS)

No project is truly complete without a DSS. Have you heard of these?

The DSS, coined by Liz Ryan, is a very effective tool for communicating the most important parts of your project to anyone. It has 3 parts:

1. The problem you set out to solve
2. What you did to solve it
3. The happy result of your work

With this approach, you can employ #storytelling to demonstrate how qualified you are. Without bragging about yourself. Example:

"A previous employer was struggling to forecast product demand. I built a time series model which improved forecast accuracy by 10%."

That's it. It's so simple, but many data scientists and job seekers are missing out on a great opportunity to get credit for their work. Thoughts? 💬

How are you enjoying the #BreakingIntoDataScience series so far?

I want to hear from you!

In the comments section below, please let me know if there is a particular topic that you'd like me to write about.

And if you have found this series of posts useful, can you help me in two specific ways?

1. Keep following the #BreakingIntoDataScience hashtag to read new posts

2. Tag a friend who would be interested in reading along with you

Thank you for being a part of the #BreakingIntoDataScience community!

- Michael