HIRING PATH FOR (ASPIRING) DATA SCIENTISTS

BOOST YOUR CHANCES TO BECOME A DATA SCIENTIST

Analytics Vidhya
publications@analyticsvidhya.com

The unbalanced state of data science industry & need for this guide

The data science industry is growing and it is growing fast! McKinsey quotes a shortfall of 140k – 190k data scientists in the U.S. alone by 2018. A well respected expert I met recently estimates this number to be in millions. Only time will tell what is the accurate representation.

Most of the people would hear these estimates and believe / dream that people with data science skills will be hunted by employers like this:

However, the ground reality is actually very different. In reality, most of the companies look for one of the two profiles while hiring data scientists:

1. People completing Ph.D. / Graduation / Post graduation from Tier 1 institutes in a quant discipline
2. People with experience in data science industry.

So, all the people out there, aspiring to be a data scientist, who do not fall in the above 2 categories – I am sorry to bring you out of your illusion – but that is reality!

Does this mean that the estimates you have heard are wrong?

No they are not!

Confused? Let me explain!

While the estimates are broadly right, most of the companies hiring data scientists have stuck to the two profiles of rare data scientists mentioned above.

This has created an imbalance in the industry. On one hand, we have companies who find it hard to find these scientists and on the other we have an army of data science aspirants with technical knowledge, waiting for a chance to show their skills.

Now that you see the larger picture, how can this guide help?

This guide aims to provide you everything you might need to get a break as a data scientist. **We assume that you already have the technical knowhow, but are still struggling to get your first break**. If you are gaining the technical knowledge – even better. While we haven't provided the resources on technical subject, but you can start concentrating on things which matter equally (if not more) to get the first job.

But remember, the journey will not be easy!

If you see the road ahead, it might look something like the picture below.

We can't do the hard work for you. You have to do it for yourself.

But we promise that if you follow the guide, your first job and the journey there after are not far away.

All the best!

Table of Contents

The journey of a thousand miles begins with one step!

1 . The map – Understand the landscape and finalize plan of action

What is the most apt way to start this guide? You should have got it:

1.1 Why build a career in Business Analytics / Data Science?

Analytics, as an industry is set for exponential growth. With more and more data being available in digital form, need for smarter, faster, data based decisions is only going to increase. Consider following facts to substantiate what I am saying:

- According to Harvard Business Review (October 2012 edition), job of a data scientist is the sexiest job of 21st century.
- According to the McKinsey Global Institute (In a May 2011 report): "By 2018, the United States alone could face a shortage of 140,000 to 190,000 people with deep analytical skills as well as 1.5 million managers and analysts with the know-how to use the analysis of big data to make effective decisions."

Imagine what would be the number across the globe!

1.2 Should you become a Data Scientist?

Not everyone can become a data scientist! If they could, the shortfall would be all gone along with the premium salaries which these data scientists earn. You need a flair for number, problem solving attitude and a ton of logical reasoning.

So, if you are new to this domain and want to test your fit, you should check this app we have created in form of a very simple test. This test is based on the attributes every analyst should possess. You should score yourself against each of the questions (out of the score mentioned after the question) and then add your scores. A good analyst should score more than 70 and any one scoring below 50 should seriously re-consider a decision to be a data-scientist.

Test Your Fit as a Data Scientist

I understand Data Science will become big and I have tested my fit, but I don't know how to start…

1.3 How to start a career as a Data Scientist?

Like any other employment skill, there are 2 approaches to enter / start an analytics career:

Approach 1 – On the job training

Get hired by a company which trains you (on job / internal trainings) on the necessary skills. These would be companies which have Analytics in their DNA and use it for their day to day decisions. While this approach is better from long term perspective, it takes time and investment (especially if there is no structured training in the companies). Some of the companies known for using cutting edge Analytics are:

- **Technology leaders:** Google, Facebook, Linkedin
- **Banking, Financial Services and Insurance (BFSI):** Capital One, American Express, J.P.Morgan Chase, Citibank, HDFC, Barclays
- **Telecom companies:** Vodafone, Airtel, Three
- **Analytics Consultancies:** Fractal, Mu-Sigma, Absolutdata, ZS Associates

One of the alternate route to get into these companies can be internships. So if you have a 2 – 6 month break, give a try to becoming an intern in these companies.

P.S. *As mentioned before, this approach doesn't work for most of the people. If it does for you, consider yourself lucky and make the most out of it.*

Approach 2 – Get data science certifications and the required credentials

While these certifications provides you with the technical skills required, these would not be able to compensate for experience on the job. Following are some options of training available:

For people with work experience, various leading academic institutes / institutes run certification programs. You can start by enrolling into a course here.

For freshers, there are certification courses offered by SAS training institute, Udacity, Coursera, Udemy, Jigsaw Academy and several others. Getting these certification courses can increase your chances of getting hired in some of the best companies.

These certifications are just the starting point. In most cases, you don't get a job on the basis of these certifications. You need to have some hands on experience of working with data.

So, what else can be done?

Here are a few smaller steps you can take to increase your chances of hiring (more on this in next section). In the meanwhile use following resources:

- **Continue reading on the subject:** Keep reading about the subject through various blogs and communities. Check out our resource finder for this.
- **Try internships:** Internships are awesome way to start your career. They provide you a flavor of the work before you take a plunge and the employer gets to assess you without significant investments.
- **Books:** Here is a good infographic on books related to Data Science.
- **Get your hands dirty in data science competitions:** Data Science competitions and Hackathons are one of the best way to test your knowledge, benchmark yourself and learn from your peers. In a matter of few days, you can learn more than entire year worth on the job training.
- **Build project portfolio and share it on GitHub**

1.4 How to become a Data Scientist?

After taking the test, if you are convinced that you are the hard working soul to become a data scientist, come with me a step further, now I'll teach you the best way to become a data scientist. My journey as a data scientist was awesome. The data science ecosystem has evolved a lot over these years. The facilities that you have now, wasn't there so openly 8 years back.

Here are the ideal steps to become a data scientist:

Step 1: Graduate from a top tier university in a quantitative discipline

Thankfully, this didn't change much for me. Education makes a huge difference in your prospects to start in this industry. Most of the companies who do fresher hiring, pick out people from best colleges directly. So, by entering into a top tier university, you give yourself a very strong chance to enter data science world.

Ideally I would take up Computer Science as the subject of study. If I didn't get a seat in Computer Science batch, I'll take up a subject which has close ties with computational field – e.g. computational nueroscience, Computational Fluid Dynamics etc.

Step 2: Take up a lot of MOOCs on the subject – but do them one at a time

This is probably the biggest change, which would happen in the journey, if I was passing out now. If you spend even a year studying the subject by participating in these open courses, you will be in far better shape vs. other people vying to enter the industry. It took me 5+ years of experience to relate to the power R or Python bring to the table. You can do this today by various courses running on various platforms.

One word of caution here is to be selective on the courses you choose. I would focus on learning one stack – R or Python. I would recommend Python over R today – but that is a personal choice.

Step 3: Take a couple of internships / free lancing jobs

This is to get some real world experience before you actually venture out. This should also provide you an understanding of the work which happens in the real world. You would get a lot of exposure to real world challenges on data collection and cleaning here.

Step 4: Participate in Data Science Competitions & Hackathons

You should aim to get at least a top 10% finish on Kaggle or Analytics Vidhya Hackathons before you are out of your university. This should bring you in eyes of the recruiters quickly and would give you a strong launchpad. Beware, this sounds lot easier than what it actually is. It can take multiple competitions for even the smartest people to make it to the top 10% on Kaggle or these hackathons.

Here is an additional tip to amplify the results from your efforts – share your work on Github. You don't know which employer might find you from your work!

Step 5: Take up the right job which provides awesome experience

I would take up a job in a start-up, which is doing awesome work in analytics / machine learning. The amount of learning you can gain for the slight risk can be amazing. There are start-ups working on deep learning, re-inforcement learning – choose the one which fits you right (taking culture into account)

If you are not the start-up kinds, join an analytics consultancy, which works on tools and problems across the spectrum. Ask for projects in different domains, work on different algorithms, try out new approaches. If you can't find a role in a consultancy – take up a role in captive units, but seek a role change every 12 – 18 months. Again this is a general guideline – adapt it depending on the learning you are having in the role.

Finally Few bonus tips!

Try learning new tools once you are comfortable with ones you are already using. Different tools are good for different types of problem solving. For e.g. Learning Vowpal Wabbit can add significant advantage to your Python coding.

You can try a shot at creating a few web apps – this adds significant knowledge about data flow on the web and I personally enjoy satisfying the hacker in me at times!

Few modifications to these tips, in case you are already out of college or hold work experience:

- In case you can still go back to college, consider getting a Masters or a Ph.D. Nothing beats the improvement in probability of getting the right job compared to undergoing a good programme from top notch University.
- In case full time education is not possible, take up a part time programme from a good institute / University. But be prepared to put in extra efforts outside these certifications / programmes.
- If you are already in a job and your company has an advanced analytics setup, try to get an internal shift by demonstrating your learning.
- I have kept the focus on R or Python, because they are open source in nature. If you have resources to get access to SAS – you can also get a SAS certification for predictive modeler. Remember, SAS still holds the majority of jobs in analytics!

People with goals succeed because they know where they are going

2. Getting ready for data science – building the technical know how

Since the entire journey starts by gaining the knowledge, it would be unwise to not provide a direction here. While training is not the mainstay of this guide, it is a critical element. Hence this chapter on how to choose right data science trainings.

2.1 How to choose the right data science training?

I aim to provide a framework to you to decide:

- Which tool to learn?
- Which techniques to focus on?
- How to learn?
- Where to learn?

You can apply it at various stages of your analytics career to find out what should you be learning next.

Overview of the framework

The answer to first 2 questions in this framework are in form of levels or steps. You start from level 0 and move one step at a time. So if you are a complete fresher start from Level 0 of tools and level 0 of techniques. But, if you are a fresher with statistics background, start with Level 1 of tools (assuming you know Excel) and Level 1 of techniques (move to level 2 if you know predictive modeling).

Once you have finalized the tools and techniques to learn, move on to step 3 and step 4 of the process.

Step 1: Which tool to learn?

Level 0: Excel

If you don't know excel, you should learn it first. You should be able to play with Pivot tables, do simple data manipulations and apply lookups in Excel.

Level 1: SAS / R / Python

This is going to be your work horse. You can choose any of these languages.

Level 2: QlikView / Tableau / D3.js

You should add up your repository with one of the visualization tools.

Level 3: Big Data tools

This in itself can be multiple levels – start with Hadoop stack – HDFS, HBase, Pig, Hive, Spark

Level 4: NoSQL Databases

Again, you can read an overview of NoSQL databases here and start by learning the most popular one – MongoDB.

Exception 1: If you come from MIS / reporting background, you can start from learning visualization tools like QlikView and Tableau (Level 2) and then go to Level 1

Exception 2: If you come from software engineering / web development and know one of the 2 languages – Java or Python, you can start from Big Data tools as well (level 3)

Step 2: Which techniques should you be learning?

Now that you know, which tool would you want to learn, let us look at the techniques to learn. Again the structure is similar:

Level 0: Basics of statistics – Descriptive and Inferential statistics

Level 1: Basic predictive modeling – ANOVA, Regression, Decision trees, Time Series

Level 2: All other remaining machine learning techniques except Neural nets

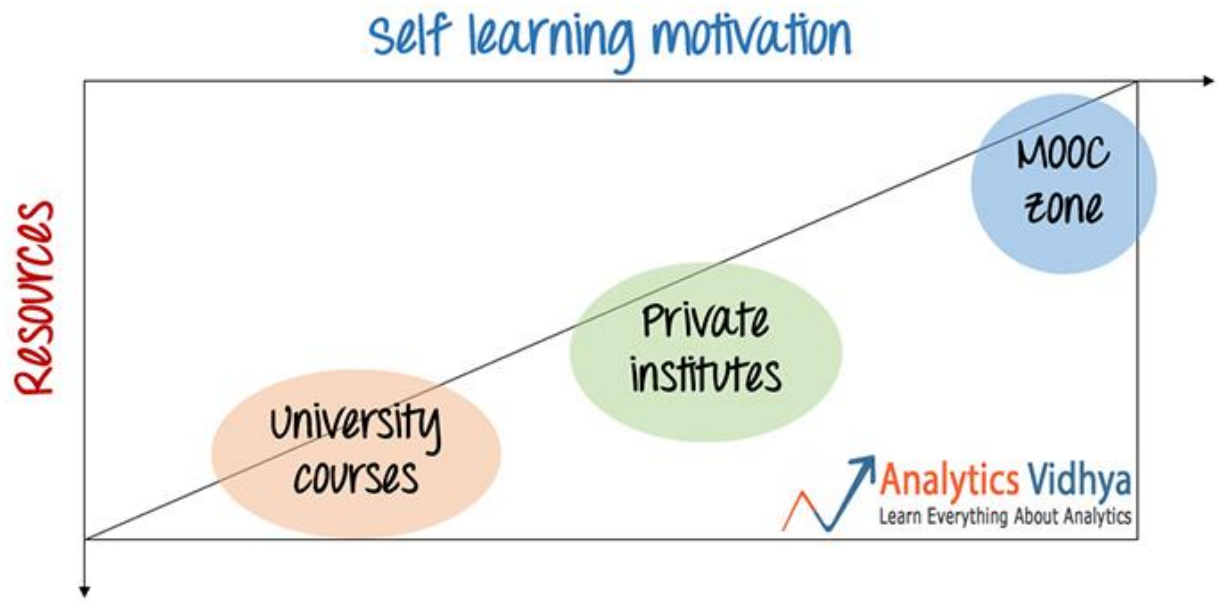
Level 3: Neural nets and deep learning

Step 3: How should you learn?

Your way of learning is dependent on 2 factors:

- Resources you can spend on learning; and

- Your self-learning motivation.



On one extreme, you have option to join open courses – where you spend low (almost zero) resources, but need high self-learning motivation. On the other hand, you have courses run by big universities like Stanford / MIT / North Western, where you will need to spend money and will get help and mentor-ship from experts over longer duration. You can choose the style of your learning depending on where you fit in.

Please note that irrespective of which method and blend you choose, you will need to aid these trainings by hands on projects and practice. No resources or trainings can cover that for you.

Step 4: Where to learn?

Now that you know, what to learn and how to learn, you can shortlist various options available. You should talk to people who have undergone that training / course and gather some reviews. You can also use our training listing page and apply filters to shortlist the trainings available for various tools and techniques. We have more than 300 trainings listed here and are in process of adding more trainings and courses.

If I had six hours to chop down a tree, I'd spend the first four hours sharpening the axe – Abraham Lincoln

3. Pre-Interview Preparation

The quality of your pre-Interview preparations will define your success in them. Make sure you spend enough time making your CV crisp and clear. Make sure you sharpen your skills and knowledge. Basically, make sure you are in your best shape for the day which matters the most!

3.1 Incredible Tips to Prepare an Outstanding CV for Analytics / Data Science roles

Now comes the most critical task of pre-interview preparation. It is being said:

A good CV might not be sufficient to get you selected,

But, a bad CV can be sufficient to get you rejected!

Here, I reveal some of the best tips using which you can use to prepare an impeccable CV of yours. After all, CV creates the first impression on recruiter, even before your first personal interaction. Just 2 steps, nothing more.

Step 1: Get the basics right

Here are a few tips to get the basics in place. I am just listing the tips here as most of you would know these or can find better guidance on basics on the web.

1) Make sure the CV has no errors – either typos or grammatical. This is something which leads to a straight rejection.
2) The length of your CV should be ideally one page or a maximum of two pages. Anything more than that is likely to be skimmed through, rather than being read.
3) Make sure that every point you are mentioning in your CV is crisp and clear. Avoid using long, complex, technical statements.
4) Attach clear (business) benefit attached to each and every project you mention or call out why the project you are mentioning stands out.
5) The purpose of the CV is to showcase your work from past. Hence, you should emphasize on what you have done and not what you can do.
6) Highlight skills relevant to the role (more on this later)
7) Take review from a few experts in the industry. This is the acid test to check whether your CV is still up to the mark or not.

Step 2: Tailor your CV

Once a basic version of your CV is ready, you can start thinking "How to can tailor your CV depending on the role you are applying for?" Emphasize and highlight projects which showcase the skills required in the role you are applying for. There are broadly two classes of roles available in analytics industry currently:

Technical data science roles: These roles require you to spend time working on / leading data science projects. You need to slice and dice data to come out with insights for your customers / stakeholders. If you are applying to these roles, here are a few tips:

1) Focus on showcasing few key projects: Identify a select group of project (5 – 10) and focus on mentioning them only. Highlight the key benefits or uniqueness of each of these projects.

2) Illustrate depth of knowledge: There roles require you to possess deep knowledge on a topic.
3) Showcase technical skills: You should mention the types of tools you have worked on and your expertise with them.
4) Get (meaningful) certifications and add them to your profile – Lookout for courses running on platforms like Coursera, Udacity, edX and complete them along with their assignments.
5) Create a GitHub profile and share your work from past – Share whatever work you can share on GitHub (e.g. projects done while doing courses on Coursera, edX etc.), codes from Kaggle competitions, tutorials and blogs you have written in past

Non-technical data science roles: I refer to the roles, which are either client facing or related to sales of analytics solutions here. If you are applying for jobs like these, you don't necessarily need to showcase your expertise on a language.

What you need to showcase are the skills which are critical to that role. For example, a client facing analytics consultant would need to be very good at structured thinking. A pre-sales person would need skills to influence people basis data based stories. Here are a few tips to make your CV stand out for these roles:

1) Focus on breadth of knowledge rather than depth: These roles typically require you to know the entire spectrum of data science fields, so that you can understand the client requirement well and advice accordingly. So, you should focus on showing the breadth of knowledge rather than depth.
2) Emphasize the business impact from your previous engagements: This could be the biggest sales you made, the number of clients you consulted or any metric, which directly impact the top line (or bottom line in some rare cases).
3) If you have written blogs / views on the subject, you should include them in your CV.

Whatever you hold in your mind on a consistent basis is exactly what you will experience in your life

- **Anthony Robbins**

3.2 How should you prepare for an analytics interview?

I have met a lot of analysts, who are good analysts when you interact with them informally. But something happens to them, as soon as they enter into an interview!

As long as you know your subject, are a logical person and can stay calm – you can ace these interviews easily!

What is the employer trying to judge you on?

The actual skills, which the employer might be judging on, would vary from employer to employer, but it is likely a mix of the following skills:

- **Technical skills** – comfort and knowledge about various analytical tools
- **Knowledge of statistics** – whether you apply algorithms blindly or actually understand what they do?
- **Structured thinking** – Can you take ambiguous problems and put a framework around them?
- **Business understanding** – How well can you put on your business thinking hat?

- **Problem solving** – Can you provide (out of the box) solutions to problems?
- **Communication skills** – Can you communicate your thoughts clearly and crisply? Can you influence people?
- **Comfort with numbers** – How good are you at crunching them?
- **Attention to details** – Do you pay attention to small details and at them up to see the bigger picture

3.3 Types of Analytics Interviews

Analytics interviews can be divided in broadly three categories:

The preparation for technical analytics interviews happens over time. These interviews test how much time and efforts have you put, in learning your subject and tools.

If you are really good at what you do, these rounds should be a cake walk. If you are not, the best strategy is to be honest about what you know and what you don't and let your potential employer know.

3.3.1 Skill assessment interviews

These are the deciding factor in most of the analytical hiring, and for a good reason – if a person has sound logical skills and can demonstrate good business thinking and logical skills – he can pick up technical skills easily! Since these interviews are aimed to assess various skills, what matters more, is that you demonstrate those skills. The actual answer and solution is irrelevant in most cases. Any hiring manager would prefer a wrong answer with a better approach rather than an accurate answer with bad approach.

Skill interviews, again can be categorized in 2 categories:

1. Guess estimates
2. Case studies and role plays

3.3.2 Guess-estimates

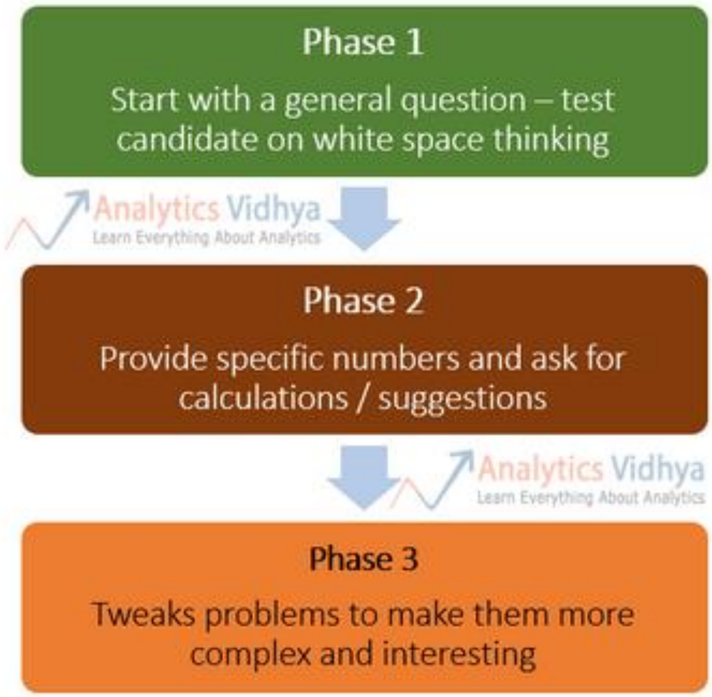
Guess estimates are puzzle like questions, where you are expected to estimate a figure by putting a framework to a question, creating segments, making assumptions and adding up the numbers to arrive at a number.

You can read details on how to ace a guess-estimate along with a few examples here. Here are a few tips I would recommend:

- It is the approach that matters – not the exact numbers. However, you should cross-validate numbers once you have them.
- Always go top down to solve a problem. Draw neat segmentation & diagrams to illustrate your approach.
- Keep a few common starting points / proxies on your finger tips. Population of your country, population across the globe, the GDP of your country are a few good starting points you should definitely remember.
- Analyze all possible uses of the subject. E.g. You should consider B2B & B2C markets, if you are asked to estimate market of tablets or smartphones.
- Call out assumptions and possible blind spots.

3.3.3 Case studies / Role plays

A case typically starts with a broad question providing a business scenario and then narrows down in a particular direction. Cases might also evolve and grow in complexity as the interview progresses. Here is how a typical interview evolves over time:



Following are some best practices to follow in a case study round:

- Case study is all about illustrating 3 things – Structure, structure and structure! Focus on putting framework to the problem provided, and you will be safe. Try deviating from it and you'll find yourself in trouble.

For example, when asked how can you increase Profits for a product company, you should not jump to conclusions like "I'll improve marketing or I'll cut costs". You should say Profits = Revenues – Costs. In order to increase profits, we can either increase Revenues or reduce costs. Revenues can be increased by increasing Sales or increasing the price. Costs can be reduced by doing ….

- Keeping a structure will not only help the interviewer understand you better, it will also help you make sure that you have not missed out anything.
- Call out assumptions, whenever you are making them. These could be assumptions about business or the sector in discussion.
- Lay out things neatly on paper, such that, they can be re-used later. Most of the times, case studies evolve over time. You will be asked to do similar questions, multiple times under multiple scenario. Keeping them handy can reduce calculation time!
- Think out loud – it is the thinking process, which matters. If you are not sure – ask the interviewer rather than staying quite!
- Communicate crisply and clearly – if you are not clear about your thoughts, take 2 minutes from the interviewer to arrange your thoughts and then communicate them nicely.

4. Training Time: The Brain Training For Analytical Thinking

4.1 Perfect Way to Train Your Brain

The trick is quite easy for you to practice. If you make calculations on daily basis, your calculations become more reflexive and accurate. An average working person in weekday spends 25-30% of his time sleeping, 40-60% of his time working , 10% of time eating and 15-25% idle. In this busy world more than 50% of our idle time is spent on road. You can use this particular time to develop sharper reflexes on numbers. Check out the Perfect Way to Train Your Brain

4.2 Training

Time for rigorous training now. Put your brain to work. I've compiled some questions to challenge your intellect: Must for Data Scientist – Brain Training

5. Practice Time: Commonly Asked Interview Questions

5.1 Puzzles

Here's a compiled list of most commonly asked puzzle questions in interviews. Surely, they are tricky, but aren't impossible to solve. Make sure you don't miss out on any before going on any interview: Puzzles 1, Puzzles 2

5.2 Guess Estimate Questions

Time to practice some guess estimates now. It has been seen, as you clear basic rounds of interview and move up a level, you are likely to find guess estimate questions blocking your path. So, after you've learnt the logic behind solving these questions, here your challenge: Guess Estimate Questions

5.3 Case Studies

After you've survived guess estimates, finally, the candidate is tested on case studies to evaluate If he/she can think actively on real life business cases and possess the discretion to make decisions. Not every is able so solve case studies perfectly, here's your chance to sharpen these skills, by practicing dual level case studies:

a) Level 1

Interview Case Study Part 1, Interview Case Study Part 2, Interview Case Study Part 3

b) Level 2

Analytics Interview Case Study

5.4 Tools Related Questions

Now, you should practice on tools related specific questions. For now, I've listed the commonly asked questions on SAS and R.

a) SAS

[4 Tricky SAS Interview Questions – Part I]

[4 Tricky SAS Interview Questions – Part II]

b) R Programming

[4 Tricky R Interview Questions]

Important Factors Responsible For Your Success in Interviews

Many people new to the industry or who is trying to enter this industry are typically not aware and hence ignore these finer aspects. Hence, I thought I'll create a list of these small things, which are typically overlooked by people coming for an interview. You can almost treat this list as a checklist of things you need to avoid while you are being interviewed.

Hygiene factors

- You reach late for an interview.
- The folder containing your documents is all messed up. When someone asks you for the CV, it takes you more than 10 seconds to pull it out.
- During the case study round (which normally happens for analytics interviews), not writing your solution in a structured fashion. This is one point which so many people overlook. The idea is that whatever you write during the case-study should be re-usable at a later point. Typically, case studies evolve as you go through them. So you would need the same structure, multiple times during the same interview. You cannot be fiddling with papers when asked to re-calculate profits in the new scenario!
- If it is a coding assignment, not indenting and commenting your code. Same principle as the last point. It also shows, how much of your work would be re-usable by others.
- If you have done something wrong, defending it rather than accepting it and correcting it.
- Error / typo in a CV or a non-updated CV

Structured Thinking

- Not putting a structure / framework to your answers is a big red flag. An ideal analyst would always have his thoughts structured – starting from when he / she is telling about himself / herself to when they are solving a case study.
- Not stepping back to see the bigger picture. Quite often, being good with details can back fire. If there is a business problem / case you are solving, you should always take a step back and think about the bigger picture. Don't get into the details unless you are clear about the business problem and its impact.

Numerical Abilities

- Using calculators / excel too often – even when they are not needed. 33.33 times 6 is 200, you don't need a calculator for this.
- Not taking shortcuts when they are available. If two factors compensate each other, you can directly knock them off.
- Not triangulating numbers or making sense out of numbers once they have been calculated. Do they tie back to the original problem?

Curiosity / Motivation

- Not asking enough questions or good questions during the interview rounds or outside.
- Not asking for clarity in an ambiguous / amorphous situation
- Lack of energy at any point in time during the process.
- Discomfort in doing white board discussions during the interview.

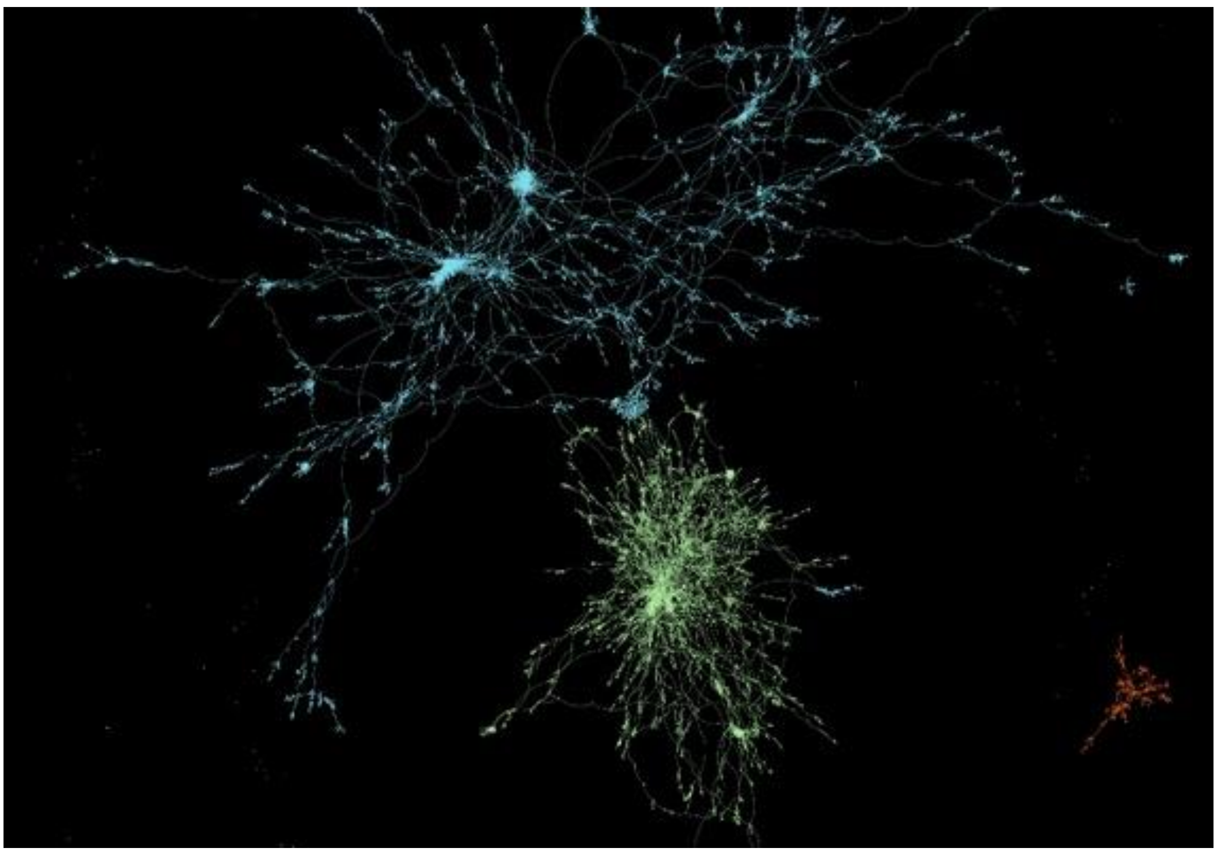
Problem Solving

Not suggesting a few crazy (read out of the box) solutions during the case studies

Suggesting too complex solutions to simple business problems

Thinking about the problem on a single track (usually data driven solutions) rather than holistically (e.g. Marketing, Operations, etc.)

With self-discipline, self-determination, self-endurance and constant curiosity for new knowledge, you can achieve any goal



6. Post Interview actions

Here's I'll tell you the best answer to recruiter's question, 'Do you have any questions to ask?' It is being said, 'All is well that ends well'. After, giving a fantastic performance throughout the interview, you should make sure, you survive this question intelligently.

Here are the 5 things that you should ask to the recruiter before taking up the job:

6.1 What is expected out of the role – Analysis or reporting?

You need to find out the expectation from the role. There can be various levels of expectation from a role (in order of increasing value addition):

- Creating business reports (pre- determined format) or refreshing dashboards with right data.
- Creating business report, looking at trends and providing a commentary (Note: I mention looking at trends and not Analyzing)
- Taking up a project (business problem), creating framework to analyze data, collate various hypothesis, mining data for insights and giving them to business
- Along with what is mentioned in previous point, owning the project implementation.

All these positions will be called Analysts! However, the learning in the last role would be far higher than the first one.

If you are still not clear, asking these questions from the interviewer can help:

- Please describe typical day for this role? What kind of problems will I work on day to day basis?
- What is the impact this role can create in the company?
- Who are the main customers (internal / external) for the role / team?

These should leave you with enough details to map the role correctly. If you are still not clear, next point might help.

6.2 Ask for a recent project done be person in the role

You should ask the interviewer to describe work done by the person in this role recently. This should give you a good flavor of the work which is expected out of the person. Following are the aspects you should specifically understand:

- How critical was the problem to the business? Was it creation of a report? Creation of a monitoring platform? or Segmenting the entire customer base to create strategy for the company?

- What kind of tools and techniques were used during the project? What was required – a complex data modeling? Clustering? Predictive modeling?
- What were the quality of insights that came out of the project? How much impact did the project create? Was the benefit monitored?
- If it was a strategic project, did the company test it before implementing it completely? Good companies will test out the insights before implementing them.

6.3 How big is the analyst community? Can you talk to them? Do you know someone?

Answer to this question would help you understand following:

- How much peer to peer learning can happen? Usually analytics is best learned through brainstorming with other analysts on the job.
- The perspective your customer and leadership would have towards analytics. The bigger the community, higher would be the mind-space in Leadership team

If you can connect to analysts in the company, that would help immensely.

6.4 Which function does the position / team report to?

While there might be exceptions to this, but the reporting function usually influences the kind of projects you work on. If the team reports to IT, the nature of projects would focus more on tools and dashboards (rather than customer insights). If it reports to Operations, you might get higher share of Operations Analytics. Typically, direct reporting of Analytics team to the CEO (or through Center of Excellence) brings the best mix of the problems.

6.5 If it is replacement hiring, why did the last person leave?

This is a tricky question and you have to judge the right way and moment to ask this. An informal lunch / coffee post interview might be more suited than during the interview. But, asking this question might give you some good insights:

- Is the manager / interviewer comfortable while talking about it? If he is, he is likely to be comfortable having open and honest discussions later on.
- Does the reason sound reasonable? How did the Organization react to it? If the person left for higher education or starting his own venture, did he / she get the required support? If the answer does

not convince you, do not ask further, but try and understand it later on.

6.6 If it is a new team / setup, what is the vision / goals for the team?

For new setups, it is always worth understanding the vision / goals for the team. If the person is not clear about them or is unable to articulate it, consider it as Red flag. If the leader can't articulate / excite a new team member about the team's vision, chances that the team would be clear / excited about their role and responsibilities are bleak.

Asking these questions should give you a good read on fit between your expectation and interviewer's expectation. You should only join the potential employer when you get satisfactory answers to at least 4 of these questions (with full match on the first question.)


" The Finish Line is just the beginning of a whole new race "
- Unknown

7. Closing Notes – Life as a Data Scientist

By now, you have everything you need to start a career as a data scientist. But, have you thought, how would the life be once you become a data scientist? The world on the other side, appears luring enough to bring you towards it, but in reality, is a nasty world which expect you to adapt to dynamic changes.

Well! Here are the 5 habits of highly successful analyst that you should follow:

7.1 Habits of highly successful analyst:

Habit 1: Keep high bar on project delivery. Walk that extra mile and deliver your best

Quality thinking differentiates a high performing analyst from a low performer. Successful analysts provide enough quality "brain time" to any project they deliver. By brain time, I mean distraction-free time devoted to analytical problem solving. This is the time when you strive for going beyond what is expected. This is the time when an analyst asks some of these questions to himself:

- Is there a better way to structure this problem? Will that make the solution better or more intuitive for business?

- Is there a better way to present / summarize the findings of the project? How can I visualize the outcomes in best possible manner?
- Instead of simply highlighting the insights from a project, can I use these insights and chalk out business actionable? Can I size the impact from these actionable and tell them to business upfront?
- Is there any aspect of analysis I might have ignored?
- Is there any implicit assumption I have made which is impacting the result?

Remember, the value of analytics is recognized through the value it generates for the business and the amount of time spent asking these questions will have direct impact on the business value created.

There have been times when I have changed the presentation flow, performed additional analysis, verified and re-verified all the numbers / insights till the night before project presentation. All of that is done to make sure the project creates the impact it deserves. And, nothing beats the feeling you go through after creating that impact.

So, don't leave any stone un-turned and make sure there are no gaps in your thinking on every project you work on.

Habit 2: Segment, till you can!

Successful analysts never work on averages. Every time they see an average, they think if there is an underlying segmentation at work, which could explain things better? By not segmenting the average, there is value left on table. Successful analysts never do that.

Habit 3: Triangulate numbers, perform back of the envelope calculations and think what they mean for business

As an analyst, you deal with numbers day in and day out. You need to pick out that one cell in which the formula is wrong from a file containing thousands (if not millions) of formulas. The only way you can do it is by triangulating numbers and by making sense of what they mean for business.

While this might sound obvious, you will be surprised to see the number of times this is overlooked. While triangulating numbers deserves a post in itself, I'll briefly mention some of the questions I ask to triangulate numbers:

- Ask yourself, can I reach this number through a different framework / calculation? Do the numbers tie up or they are different by a magnitude?

- Are there process dependencies which can give you a sense of numbers? Can you issue 2000 credit cards every month, if you only get ~1800 applications every month?
- What do these number mean for business? Do they tie in with the infrastructure and resources business has?

Triangulation is like any other skill, it will look difficult to start with. But the more you practice, the better you become at it.

Habit 4: Test out your hypothesis (even if you think they make complete business sense)

There are times when you tend to overlook the need for testing. Just adding a live chat functionality to your website? Sounds like a good thing to do with no down side. Test it out and you will know. The customers might not like it! Here is another example:

One of the leading travel portal in India saw this in their data: 90%+ flights booked have departing location same as city from which tickets are being booked (determined by the I.P. address). They thought of making this location pre-populated (obviously with an option to change). This sounds like a nice idea which would help provide a better customer experience. Thankfully, they tested it out. Booking conversion dropped by double digit percentage within weeks of making this change. Possible reason: The customers are used to filling to and from location. Removing one of them adds to the confusion.

So, next time if you are implementing results from any analysis or hypothesis, test it out!

Habit 5: Learn something about analytics everyday

While this habit will not bear immediate results like other habits mentioned above, you will benefit the most from this in long run. Analytics is a dynamic and evolving field. A new tool / technology / update arrives almost every 2 – 3 months. Being up to date with latest updates in industry helps not only stay on top of it, but creates a huge gap from analysts who don't stay updated.

Some of the topics I enjoy reading whenever I get time are:

- What are the latest developments in Big data?
- How to analyze unstructured data from Social media? How can we make visualizing this better?
- What are the statistical concepts behind the algorithms used by various tools?
- How can you design and analyze a design of experiments?

The list is endless. Start reading and you are bound to run out of time!