**Data Analyst Skills You Need for Your Next Promotion**

**Advice from the “older” generation**



Photo by [Simon Abrams](https://unsplash.com/@flysi3000?utm_source=unsplash&utm_medium=referral&utm_content=creditCopyText) on [Unsplash](https://unsplash.com/s/photos/office?utm_source=unsplash&utm_medium=referral&utm_content=creditCopyText)

**Introduction**

Let me start off by acknowledging that this subject has been written about and discussed repeatedly on many different platforms. So, what could I offer that hasn’t already been said?

Well, as it turns out, I’m a lot older than the average Medium author. Although some people politely refer to me as “mature” or “experienced,” the truth is that my perspective is shaped by the fact that I’ve been around long enough to remember when neon colors and gradients on bar-charts were cool. Therefore, what I can offer you is advice that stands the test of time.

**Data Visualization**

As an analyst, you will probably spend at least 10 times as much time analyzing the data as you will have to present the information to your audience. Therefore, graphics and other visual representations are important because it creates a better understanding, even for those who aren’t trained in analyzing datasets. By building visualizations, you can help your company’s decision-makers understand complex ideas at a glance.

So how do you improve this skill? Well, there are plenty of training options in the form of online courses and plenty of tools specializing in visualization. However, I’m hesitant to recommend specific technology or course because the world changes so quickly, and I want to offer advice that stands the test of time.

Therefore, make a habit of exploring other people’s work. Build a bookmarks folder called “Inspiration” and fill it with blog links. Dedicate 15 minutes weekly to browsing the blogs to fill your brain with possibilities.

I also recommend buying a couple of books. Edward Tufte is the grandfather of data visualization, and I’m also a fan of Nathan Yau over at FlowingData.com, where you can find books, blogs, courses, and tutorials. I keep both of his books, *Data Points* and *Visualize This,* on my bookshelf for inspiration.

The mechanics of how you will actually create the visualization will come with practice. Some of you will fall in love with R or python, and others with Excel or Tableau. The key is having your brain filled with possibilities, so you can imagine what you want before you build it.

**Data Cleaning**

Data Scientists are known for their claim that “80% of building a machine-learning model is preparing and cleaning the data.” However, this is also true for data analysts. Whatever the actual percentage, the truth is that a large proportion of time spent with data, in general, is spent cleaning it.

Cleaning data is important, because uncleaned data can produce misleading patterns and lead to mistaken conclusions. Early in my career, I was tasked with delivering the percentage of support calls that were resolved on the first try. As I walked my management through the results, one of them insisted that the numbers “didn’t make sense.” When I did more research, I discovered that the calls data had a column called “status” that sometimes was populated with an “X.” Apparently, the system would record an “X” for a test record, which should be ignored.

Data cleaning skills grow with hands-on practice and business expertise. This makes it difficult to accelerate. In general, sites like [kaggle.com](https://www.kaggle.com/) or [zindi](https://zindi.africa/) are not the best places to practice data cleaning because they are focused on data science, and the datasets are usually pretty clean already. On the other hand, government websites such as <https://data.ca.gov/> or <https://data.gov/> are a great place to find datasets that are messy. You can also follow the [TidyTuesday](https://github.com/rfordatascience/tidytuesday) project, even if you’re not an R user, to find interesting datasets and get familiar with the types of cleaning steps that occur in the wild.

**SQL**

As I mentioned, I am hesitant to recommend a specific tool or technology because of how quickly the landscape changes. It is probably safe to say that python and R are here to stay, but SQL is on another level. SQL is the language of databases; therefore, learning SQL will always be the most direct way you can manipulate and study datasets for your analysis.

If you work at a company that lets you download extracts in excel sheets — ask around to see if it’s possible to access the database directly with SQL. Once you are comfortable with database structures and writing SQL to get the data represented in the way you want, your efficiency will increase along with the quality of your work.

Many resources are available to help you improve your SQL game, [CodeAcademy](https://www.codecademy.com/learn/learn-sql), [Udemy](https://www.udemy.com/topic/sql/free/), and [Udacity](https://www.udacity.com/course/sql-for-data-analysis--ud198) are great free resources for finding hands-on courses. [SQL Generator](https://app.rasgoml.com/sql) and [SQL Beautifier](https://codebeautify.org/sqlformatter) are great links to keep handy that will help you learn. [Stack Overflow](https://stackoverflow.com/questions/tagged/sql?tab=Active) has a great community for answering technical SQL questions if you get stuck.

Here at Rasgo, we’ve been investing heavily in giving back to the data community. Our most recent project was the launch of our free [SQL Generator](https://app.rasgoml.com/sql) that generates the SQL syntax needed for specific data transformations. We found people were searching on Google and Stack Overflow for required SQL syntax — wasting a lot of time that could be used for data analysis. The SQL Generator is a template for a SQL query, which lets you customize column names and table structure, choose the operation that you want to do, and then it constructs the syntax for you in a variety of different “flavors” of SQL. Never stress about the subtle differences between DATEDIFF() vs. DATE\_DIFF() again! Read [this post](https://towardsdatascience.com/top-5-bookmarks-every-data-analyst-should-have-547a2c9ad1fe) from earlier this month, for some other helpful online tools.

**Critical Thinking**

Critical thinking is one of the hardest things to grasp as there are not many courses or a one size fits all approach to mastering this skill. Critical thinking is a conscious effort to challenge the automatic mental processes that rule over us.

Critical thinking isn’t something you naturally have or something that once you get it, it’s always “on.” Instead, critical thinking is something that you consciously activate whenever you have an important decision or analysis. It is a purposeful effort to challenge, question, and confirm hypotheses.

While there are many different resources available to help you improve critical thinking skills, the one I want to focus on is asking questions. If you get into the habit of writing down lots of questions, you will engage in critical thinking.

Let’s pretend that you were asked to look at customer churn and whether it has been improving recently. Stop to ask yourself questions, such as:

* How is churn calculated? Why that way? Are there other ways?
* What is churn, from the perspective of the customer
* Is it always a customer’s choice?
* What factors might matter? Can I measure those?
* What is the goal of this task? Prove, disprove, demonstrate, support?
* What could I be overlooking? Are there known assumptions I can’t validate/prove?

Notice how this line-of-questioning forms layers upon layers. Answering a question might lead to more questions. That is perfectly fine. That’s the point!

For example, when I was an analyst at a Telco, I was asked to look at potential reasons why churn suddenly increased in the prior month. (it was currently April). I used critical thinking by writing down a list of questions similar to the above list.

When I got to the question about how churn was calculated, I challenged myself to consider if there were other ways to measure churn. Long story short, I discovered that:

* The company used “end-of-month” numbers to calculate churn
* The month of March was actually a “fiscal” month that covered February
* February was the shortest month of the year
* Using an alternative formula for churn showed that the rate did *not* increase in March

Basically, the formula for churn was such that it would “spike” during the shortest month of the year, which was called March, just simply because of how the math worked.

The moral of the story is that critical thinking sometimes takes you to conclusions you would not expect. It is also something that you can decide to improve upon today; it is not some mythical soft-skill that you need to be born with!

**Communication**

Communication on the other hand, *is* a soft skill. You might be the most talented and insightful data analyst the world has ever seen, but it won’t do any good if you can’t communicate with others. And by others, unfortunately, I mean non-technical people.

As an analyst, you straddle two different worlds. In one world, you must address technical points with your peers and other data experts. In the other world, you are a translator for business-centered decision-makers. You need to provide clear, high-level explanations in a supportive rather than confusing way.

One way to start exercising this soft skill is to get involved in a community. There are many online communities, websites, and forums for data analysts to get involved in. I’m a big fan of slack and discord channels as a way to interact with others, but those are known to die off quickly. I have found [Locally Optimistic](https://locallyoptimistic.com/) and [DataTalks.Club](https://datatalks.club/) to be a good consistent place for analysts to hang out. You should also practice your communication skills by starting a blog. [Medium](https://medium.com/) is a great place to get started, where you can exercise your creativity and practice, practice, practice.

**Conclusion**

It is still important to follow the latest trends and skill-up on things that make you marketable. For example, 20 years ago VBA, macros, DAX, and ASP were great skills to give you an edge towards promotion. These days, those skills are less likely to matter. Hopefully I’ve given you some helpful advice about skills that *haven’t* changed in the past 20 years, so you can avoid getting lost in the weeds. If you ever want to reach out to me and discuss the old days, you can find me hanging out in [Locally Optimistic](https://locallyoptimistic.com/community/) and [DataTalks.Club](https://datatalks.club/), and you can always ping me directly at [Rasgo](http://rasgousergroup.slack.com/).