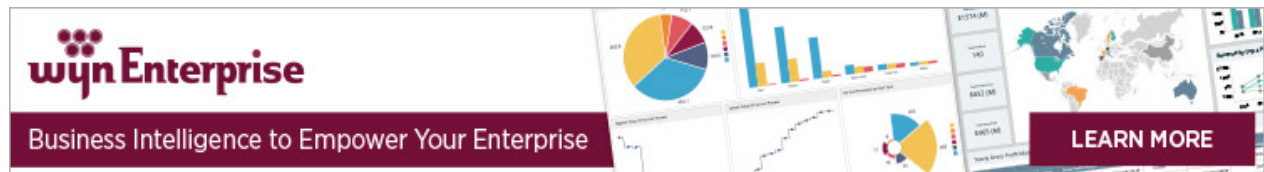


# KDnuggets

[Subscribe to KDnuggets News](#)




- [Blog/News](#)
- [Opinions](#)
- [Tutorials](#)
- [Top stories](#)
- [Companies](#)
- [Courses](#)
- [Datasets](#)
- [Education](#)
- [Events \(online\)](#)
- [Jobs](#)
- [Software](#)
- [Webinars](#)



[Wyn Enterprise provides scalable business intelligence with server-based licensing for a low TCO.](#)

Topics: [Coronavirus](#) | [AI](#) | [Data Science](#) | [Deep Learning](#) | [Machine Learning](#) | [Python](#) | [R](#) | [Statistics](#)

[KDnuggets Home](#) » [News](#) » [2020](#) » [Aug](#) » [Opinions](#) » These Data Science Skills will be your Superpower ( [20:n33](#) )

## These Data Science Skills will be your Superpower

[<= Previous post](#)

[Next post =>](#)

Like 50

Share 50

Tweet

Share

Share

76

Tags: [Communication](#), [Data Preparation](#), [Data Science Skills](#), [Data Visualization](#), [Mathematics](#), [Statistics](#)

Learning data science means learning the hard skills of statistics, programming, and machine learning. To complete your training, a broader set of soft skills will round out your capabilities as an effective and successful professional Data Scientist.



[The future of](#)

[enterprise data science is here.](#)

[Discover what's new in Domino 4.2](#)

---

By [Benjamin Obi Tayo, Ph.D.](#), DataScienceHub.

[comments](#)

Most academic training programs in data science are focused mostly on teaching hard skills. Time and time again, industry data, market trends, and insights from top business leaders highlight soft skills as a key component to success in the workplace. This article will discuss the essential hard and soft skills for success in data science practice.

## Hard Skills

### 1. Mathematics and Statistics Skills

Math skills are essential in data science and machine learning. For more about the basic math skills needed for data science and machine learning, please see this article: [How Much Math do I need in Data Science?](#)

### 2. Essential Programming Skills

Programming skills are essential in data science. Since Python and R are considered the 2 most popular programming languages in data science, essential knowledge in both languages is crucial. For more information on essential programming skills needed for data science, please see this article: [How Much Programming do I need in Data Science?](#)

### 3. Data Wrangling and Preprocessing Skills

Data is key for any analysis in data science, be it inferential analysis, predictive analysis, or prescriptive analysis. The predictive power of a model depends on the quality of the data that was used in building the model. Data comes in different forms such as text, table, image, voice, or video. Most often, data that is used for analysis has to be mined, processed, and transformed to render it to a form suitable for further analysis.

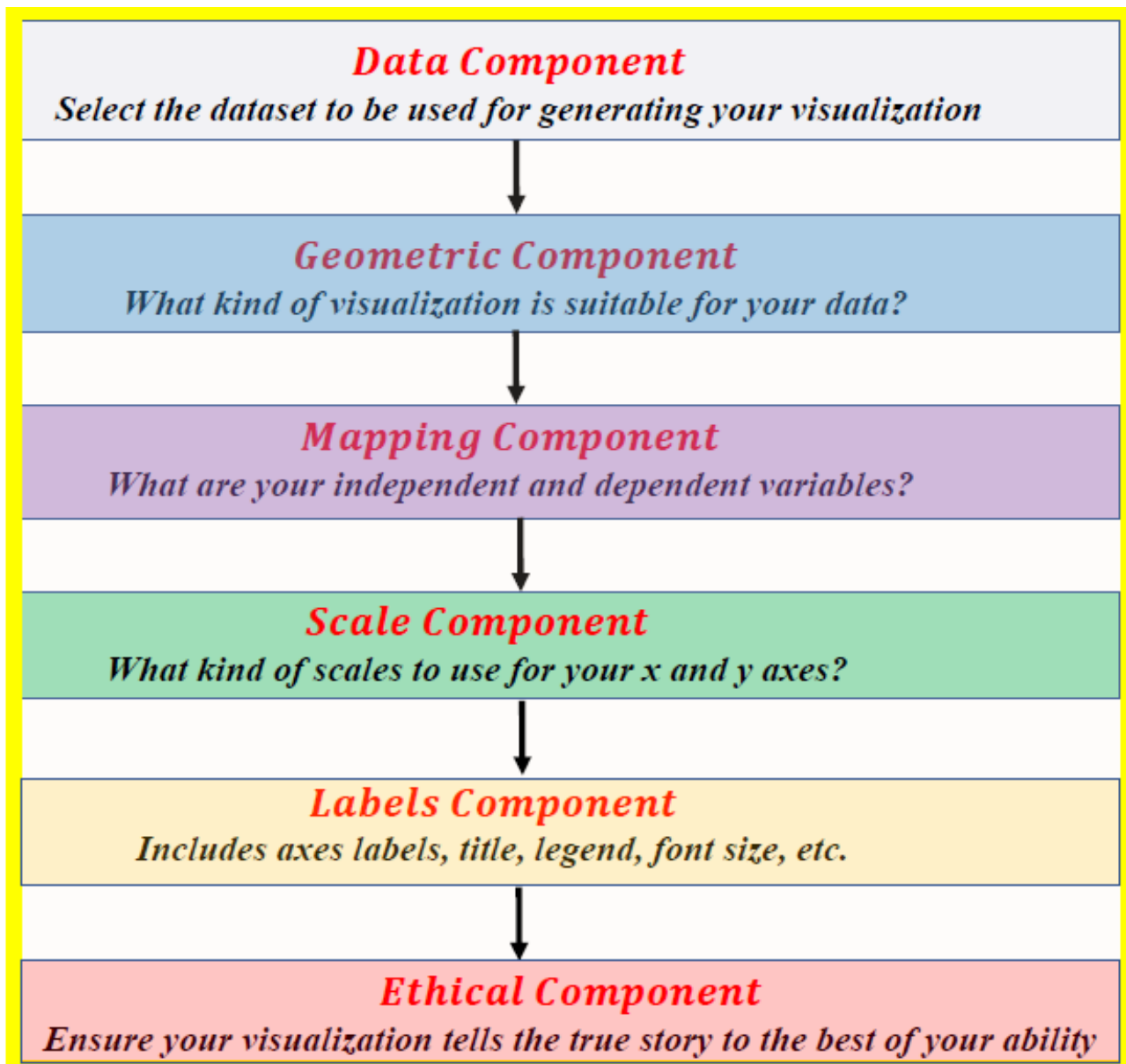
i) **Data Wrangling:** The process of data wrangling is a critical step for any data scientist. Very rarely is data easily accessible in a data science project for analysis. It's more likely for the data to be in a file, a database, or extracted from documents such as web pages, tweets, or PDFs. Knowing how to wrangle and clean data will enable you to derive critical insights from your data that would otherwise be hidden.

ii) **Data Preprocessing:** Knowledge about data preprocessing is very important and include topics such as:

- a) Dealing with missing data
- b) Data imputation
- c) Handling categorical data
- d) Encoding class labels for classification problems
- e) Techniques of feature transformation and dimensionality reduction such as Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA).

### 4. Data Visualization Skills

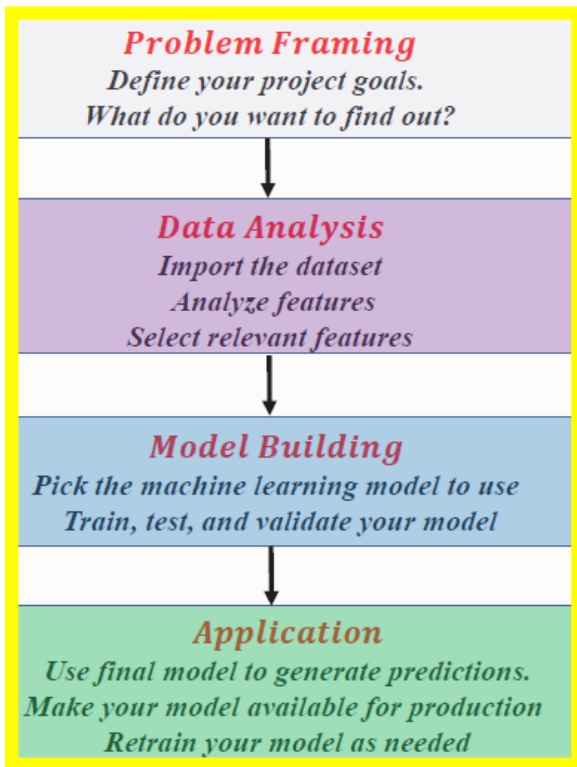
Understand the essential components of good data visualization (**see figure below**). Be able to use several data visualization packages, such as matplotlib, seaborn, and ggplot2.



*Typical workflow for a data visualization project. Image by Benjamin O. Tayo.*

## 5. Basic Machine Learning Skills

Machine Learning is a very important branch of data science. It is important to understand the machine learning framework: Problem Framing, Data Analysis, Model Building, Testing & Evaluation, and Model Application.



Typical workflow for a machine learning project. Image by Benjamin O. Tayo.

Find out more about the machine learning framework from here: [Machine Learning Process Tutorial](#).

## 6. Skills from Real World Capstone Data Science Projects

Skills acquired from course work alone will not make you are a data scientist. A qualified data scientist must be able to demonstrate evidence of successful completion of a real-world data science project that includes every stage in data science and machine learning process such as problem framing, data acquisition and analysis, model building, model testing, model evaluation, and deploying the model. Real-world data science projects could be found in the following:

- a) Kaggle Projects
- b) Internships
- c) From Interviews

## Soft Skills

### 1. Communication Skills

Data scientists need to be able to communicate their ideas with other members of the team or with business administrators in their organizations. Good communication skills would play a key role here to be able to convey and present very technical information to people with little or no understanding of technical concepts in data science. Good communication skills will help foster an atmosphere of unity and togetherness with other team members such as data analysts, data engineers, field engineers, etc.

### 2. Be a Lifelong Learner

Data science is a field that is ever-evolving, so be prepared to embrace and learn new technologies. One way to keep in touch with developments in the field is to network with other data scientists. Some platforms that promote networking are LinkedIn, GitHub, and Medium ([Towards Data Science](#) and [Towards AI](#) publications). These platforms are very useful for up-to-date information about recent developments in the field.

### 3. Team Player Skills

As a data scientist, you will be working in a team of data analysts, engineers, administrators, so you need good communication skills. You need to be a good listener too, especially during early project development phases where you need to rely on engineers or other personnel to be able to design and frame a good data science project. Being a good team player would help you to thrive in a business environment and maintain good relationships with other members of your team as well as administrators or directors of your organization.

### 4. Business Acumen Skills

**Most Popular**

1. [Know What Employers are Expecting for a Data Scientist Role in 2020](#)
2. [If I had to start learning Data Science again, how would I do it?](#)
3. [Netflix's Polynote is a New Open Source Framework to Build Better Data Science Notebooks](#)
4. [Must-read NLP and Deep Learning articles for Data Scientists](#)
5. [These Data Science Skills will be your Superpower](#)
6. [The List of Top 10 Lists in Data Science](#)
7. [I have a joke about...](#)

**Most Shared**

1. [If I had to start learning Data Science again, how would I do it?](#)
2. [Know What Employers are Expecting for a Data Scientist Role in 2020](#)
3. [Must-read NLP and Deep Learning articles for Data Scientists](#)
4. [The List of Top 10 Lists in Data Science](#)
5. [Netflix's Polynote is a New Open Source Framework to Build Better Data Science Notebooks](#)
6. [Top Google AI, Machine Learning Tools for Everyone](#)
7. [Metrics to Use to Evaluate Deep Learning Object Detectors](#)

**Latest News**

- [How To Decide What Data Skills To Learn](#)
- [Data Scientists think data is their #1 problem. Here's...](#)
- [The Most Important Data Science Project](#)
- [Book Chapter: The Art of Statistics: Learning from Data](#)
- [Design of Experiments in Data Science](#)
- [How to Evaluate the Performance of Your Machine Learnin...](#)

**Top Stories  
Last Week****Most Popular**

1. [If I had to start learning Data Science again, how would I do it?](#)



2. [Must-read NLP and Deep Learning articles for Data Scientists](#)
3. [How to Optimize Your CV for a Data Scientist Career](#)
4. [These Data Science Skills will be your Superpower](#)
5. [Know What Employers are Expecting for a Data Scientist Role in 2020](#)
6. [Working with Spark, Python or SQL on Azure Databricks](#)
7. [4 ways to improve your TensorFlow model – key regularization techniques you need to know](#)

**Most Shared**

1. [4 ways to improve your TensorFlow model – key regularization techniques you need to know](#)
2. [The NLP Model Forge: Generate Model Code On Demand](#)
3. [DeepMind's Three Pillars for Building Robust Machine Learning Systems](#)
4. [Beyond the Turing Test](#)
5. [Working with Spark, Python or SQL on Azure Databricks](#)
6. [Data Science Tools Illustrated Study Guides](#)
7. [How to Optimize Your CV for a Data Scientist Career](#)

**More Recent Stories**

- [How to Evaluate the Performance of Your Machine Learning Model](#)
- [10 Things You Didn't Know About Scikit-Learn](#)
- [Top tweets, Aug 26 – Sep 01: A realistic look at the ...](#)
- [What Is Data Enrichment And How It Works](#)
- [Computer Vision Recipes: Best Practices and Examples](#)
- [Which methods should be used for solving linear regression?](#)
- [Here is What I've Learned in 2 Years as a Data Scientist](#)
- [PyCaret 2.1 is here: What's new?](#)

- [Top Online Masters in Analytics, Business Analytics, Data Scie...](#)
- [Showcasing the Benefits of Software Optimizations for AI Workl...](#)
- [eBook: Vocabularies, Text Mining and FAIR Data: The Strategic ...](#)
- [A Curious Theory About the Consciousness Debate in AI](#)
- [Top Stories, Aug 24-30: If I had to start learning Data Scienc...](#)
- [Linguistic Fundamentals for Natural Language Processing: 100 E...](#)
- [Accelerated Computer Vision: A Free Course From Amazon](#)
- [Data is everywhere and it powers everything we do!](#)
- [Beyond the Turing Test](#)
- [Microsoft's DoWhy is a Cool Framework for Causal Inference](#)
- [Explainable and Reproducible Machine Learning Model Developmen...](#)
- [4 ways to improve your TensorFlow model – key regularization techniques you need to know](#)

[KDnuggets Home](#) » [News](#) » [2020](#) » [Aug](#) » [Opinions](#) » These Data Science Skills will be your Superpower ( [20:n33](#) )

© 2020 KDnuggets. | [About KDnuggets](#) | [Contact](#) | [Privacy policy](#) | [Terms of Service](#)

[Subscribe to KDnuggets News](#)



X



A very important skill set that is essential for practical applications is business acumen. Business acumen is the ability to draw out meaningful conclusions from a model that can lead to important and cost-saving data-driven decision making. Acquiring business acumen skills is therefore essential for practical data scientists.

## 5. Ethical Skills in Data Science

Understand the implication of your project. Be truthful to yourself. Avoid manipulating data or using a method that will intentionally produce bias in results. Be ethical in all phases from data collection to analysis, to model building, analysis, testing, and application. Avoid fabricating results for the purpose of misleading or manipulating your audience. Be ethical in the way you interpret the findings from your data science project.

## Summary and Conclusion

In summary, we've discussed several essential skills needed for practicing data scientists. While academic training programs do a good job to teach hard skills, soft skills are essential for success in the real world.

[Original](#). Reposted with permission.

### Related:

- [Top 5 must-have Data Science skills for 2020](#)
- [The Most In Demand Tech Skills for Data Scientists](#)
- [Which Data Science Skills are core and which are hot/emerging ones?](#)

What do you think?

## Important Update

When you log in with Disqus, we process personal data to facilitate your authentication and posting of comments. We also store the comments you post and those comments are immediately viewable and searchable by anyone around the world.

Please access our Privacy Policy to learn what personal data Disqus collects and your choices about how it is used. All users of our service are also subject to our Terms of Service.

Proceed

---

[<= Previous post](#)  
[Next post =>](#)

---

## Top Stories Past 30 Days