

The 2021 SDS

Data Scientist

Learning Path

Your 99-Day Study Plan #SDS99





A cheatsheet with week-by-week instructions to get you started as a Data Scientist

Thank you for downloading the Study Plan for our brand new <u>Data Scientist</u> <u>Learning Path</u>.

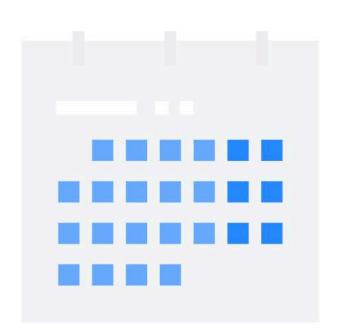
We put together our best courses, projects, and resources in order to provide you with a **complete blueprint for your first job in data science**. Working as a Data Scientist, you'll empower people in your organization to make smarter decisions through your analyses, visualizations, reports, and experiments.



Use this Plan together with your SuperDataScience monthly or annual subscription to get started as a Data Scientist in less than 100 days!



Plan Overview



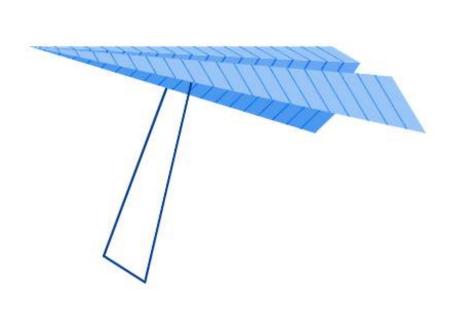
Completing the Learning Path should take **9 weeks of part-time work** (10-15 hours per week) **for the core curriculum,** and **3 more weeks for your specialization.** There are also two weeks designed for you to catch up and take a little breather!

	Week	Days		Topic	
	1	1-7	Intro to Data Science		
	2	8-14	Basics of		
	3	15-21	Data Visualization		
	4	22-28	Coding for Data Analysis		
a,	5	29-35			
Core	6	36-42	SQL & Databases		
	7	43-49			
	8	50-56	Applied Business Analytics		
	9	57-63	Intro to Machine Learning		
	10	64-70	Catch-up week		
rtion	11	71-77	Machine Learning	Data Visualization Wrangli	
alizo	12	78-84			
Specialization	13	85-91			vvianging
	14	92-98	Catch-up week		
	15	99	Celebration Day		

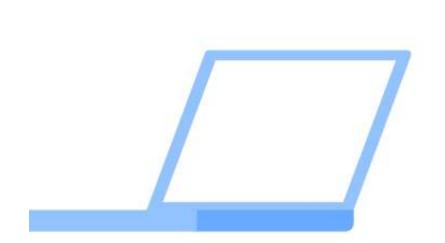
You can also take this journey at a slower pace, skip modules if you have the skills, or take more time and **complete more than one specialization** – how you implement our Study Plan is up to you! As long as you have an active SDS membership, you will be able to freely access **all of the content** in the path.



How to tackle your work



Try to set aside and schedule 10-15 hours of work during your week (early mornings, late evenings, or weekends – as you prefer!) for the following activities:

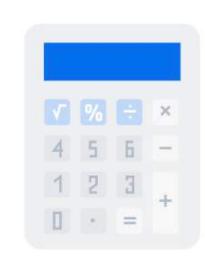


SuperDataScience Courses

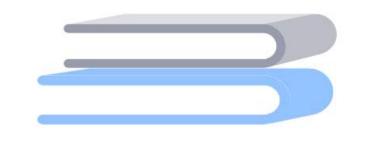
These are are the most important part of your learning. You should complete all practical assignments included in the courses, and take whole courses or only suggested modules.

Portfolio Projects

Every topic in the Learning Path features a **hands-on workshop** to complete after finishing your courses on SuperDataScience. Use these to fill up your **portfolio**!







These are optional **articles**, **books**, **and resources** to expand your learning. Most of them are **freely available**, but we suggest investing \$30-50 to buy some of the books if you're interested!

Career Prep Checklist

Starting from Week 1, we'll recommend a series of actions that will set up for success later in your job search. If you need guidance on these actions, you'll find it in our <u>Career Toolkit</u>.





Your Career Prep checklist



In order to enter the field of data science successfully, you'll need to treat your career preparation as rigorously and methodically as you approach your SQL or Python skills.

Some of the activities we suggest are **repeated** more or less frequently along your 99-day plan, since they are **a matter of habit** rather than one-off actions. We hope you'll consider these, and all of your checklist actions, as part of your **regular weekly practice** towards a career in data.

Many actions have links to **useful online resources** to support you – you can find more on our <u>Career Toolkit</u> page.



These actions usually involve finding **extra opportunities** for practicing and showcasing your skills in real-life scenarios.

Projects, competitions, datasets are a great opportunity to **shape your portfolio in a unique way** that suits your interests and makes you stand out.

Connect with the community

These actions help you **put your name out there**, reach out to potential **mentors** in the field, and get valuable and direct **feedback** on your work.

Most jobs in any industry are **found through referrals**, so nailing this part will maximize your chances.

Apply for jobs

These actions usually involve identifying the **right job offers**, preparing effective **applications**, and rehearsing your **interviews**.

While applying, don't be picky and **cast a wide net in terms of job titles** you are applying for: whether your first job in data is as a Junior Data Scientist, Data Analyst, BI Specialist, or something else, there will be plenty to learn and valuable experience to carry into your next career moves.

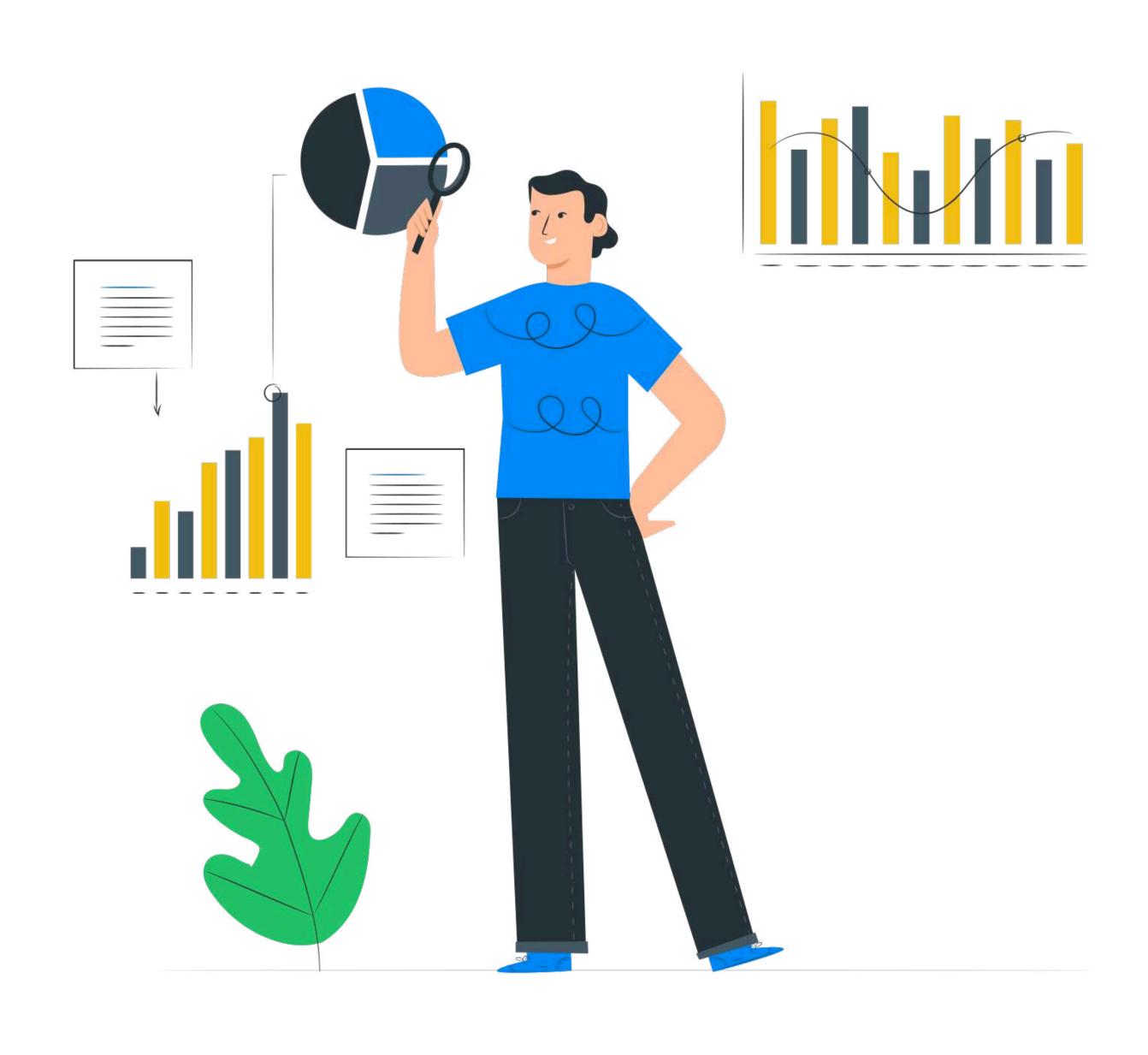
Reep up with data science

These actions help you keep in the loop of **constantly evolving data science knowledge** and practices.

Follow online sources, build a healthy diet of data science sources, and make sure you know where the industry is going.



Let's get started

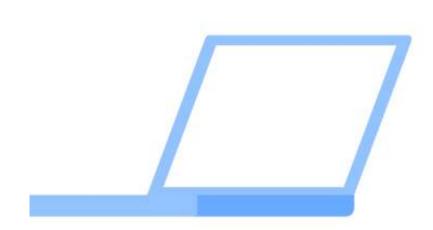




Intro to Data Science

You will get an overview of the **data science world and different careers** (including Data Scientist), the key **skills** to build, and best **resources** to learn.

Our instructors, Kirill Eremenko and Hadelin de Ponteves, walk you through the key questions and doubts of any beginner in the field (How much math should I know? Do I need a degree?), empowering your future choices in this path.



SuperDataScience Courses

Careers in Data Science A-Z (3.5 hours of lectures)



Portfolio Projects

Your first project will be... **Setting up your portfolio!** You can get creative about it, but at the very least you'll need a GitHub account and Tableau Public profile to showcase your work. Our <u>Career Toolkit</u> has guides and tips on how to do it.



Extra Resources

Pick one of these great books: <u>Confident Data Skills</u> by Kirill Eremenko or <u>Build a Career in Data Science</u> by Emily Robinson and Jacqueline Nolis.

Set a goal to finish them in your first 4 weeks of learning.

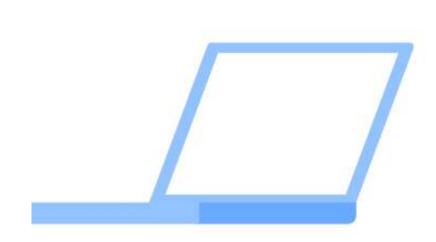


- □ Open a <u>GitHub portfolio</u> •
- ☐ Create your LinkedIn profile 🤝 💼



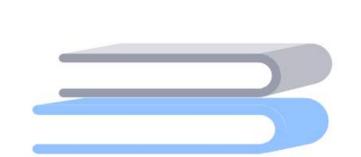
Basics of Data Visualization (I)

You'll learn how to **connect Tableau to a variety of datasets**; **analyze**, blend, join, and add calculations; and how to **visualize data** in the form of various charts, plots, and maps.



SuperDataScience Courses

Tableau 2020: Hands-On Training, modules 1-5 (4.5 hours of lectures)



Extra Resources

Continue reading the book you've chosen on Week 1



FIND TIPS AND RESOURCES
IN OUR **CAREER TOOLKIT!**

Career Prep Checklist

- ☐ Set up your Tableau Public profile 💼
- ☐ Browse the <u>Tableau Authors</u> page for inspiration, then pick a project, duplicate it and explore it **
- ☐ Connect (meaningfully) with <u>5 "data voices"</u> (influencers, practitioners, and other people in the industry) ❖

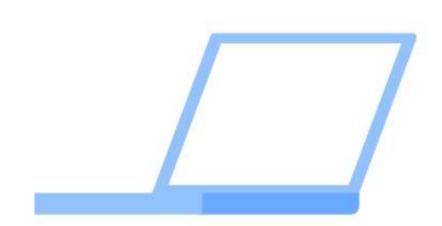
Why Tableau?

Tableau is the **most popular data analytics and visualization platform** in the world. You can easily download and install a free version, **Tableau Public**, and quickly learn how to master this powerful, industry-standard tool. Your Tableau skills are a **great currency for your job search**, but if later on you want to explore an alternative tool – Microsoft's own Power BI – on our platform <u>you'll find courses on that, as well!</u>



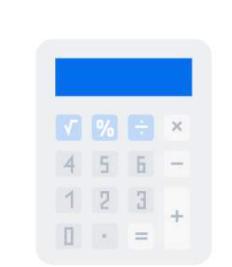
Basics of Data Visualization (II)

Time to put those data viz skills into work! You'll **complete your Tableau 2020 course** and tackle your **first portfolio project**.



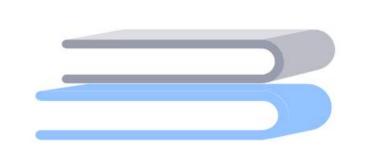
SuperDataScience Courses

Tableau 2020: Hands-On Training, modules 6-9 (4.5 hours of lectures)



Portfolio Projects

Console Games Sales Decline



Extra Resources

Continue reading the book you've chosen on Week 1



Career Prep Checklist

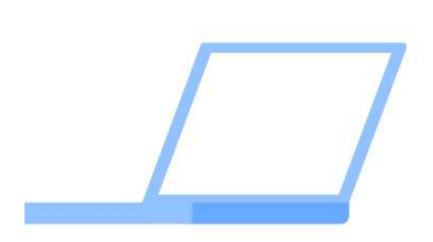
□ Research and <u>analyze 5 job positions</u> requiring Tableau
 □ Find and watch one online webinar/tutorial on Tableau
 □ Find and follow 3 data visualization influencers, blogs, or podcasts online



Coding for Data Analysis (I)



If you are entirely new to programming, we'll carefully teach you the **basics of the discipline** (variables, loops, conditional statements) before going into the **specifics for data science**: data frames, libraries, visualization techniques, and much more.



SuperDataScience Courses

If you choose the R track: R Programming A-Z (modules 1-4) If you choose the Python track: Python A-Z (modules 1-4) (each is roughly 5 hours of lectures)



Extra Resources

Continue reading the book you've chosen on Week 1

Python vs R - Which is Best? (web article)



FIND TIPS AND RESOURCES
IN OUR **CAREER TOOLKIT!**

Career Prep Checklist

- ☐ Start and document your own #30DaysOfCode challenge 🟋
- ☐ Research and understand 5 entry-level data scientist job description requiring R or Python ■
- ☐ Find someone with an entry-level data scientist job and ask them tips & questions about their job ❤

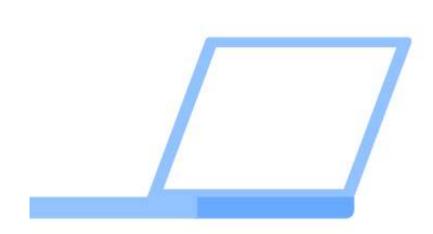
Python or R?

You can research this (we've provided some materials), but **don't overthink it!** If you're still undecided, **go for Python**. It's a little more flexible and perhaps has a slightly wider range of applications (outside of data science, at least), which may benefit you in the longer run. And don't worry: if you ever get a job that requires the one you didn't pick, **it should be relatively easy to pick it up** in a matter of weeks with the knowledge you've built in this journey.



Coding for Data Analysis (II)

Complete your introduction to your programming language of choice, highlight your projects, and continue to document your career development journey.



SuperDataScience Courses

If you choose the R track: R Programming A-Z (modules 5-8) If you choose the Python track: Python A-Z (modules 5-8) (both are roughly 5.5 hours of lectures)



Portfolio Projects

Your R and Python courses contain 5 different projects: pick at least one to add to your GitHub portfolio!



Extra Resources

Finish reading the book you've chosen on Week 1



FIND TIPS AND RESOURCES
IN OUR **CAREER TOOLKIT!**

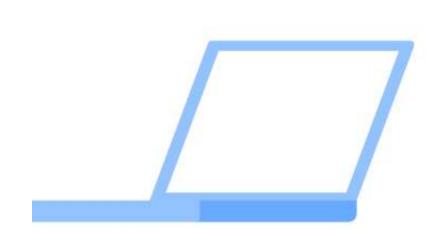
- ☐ Continue with your #30DaysOfCode challenge 溪
- Optimize your LinkedIn profile so that you are easily discoverable by recruiters
- ☐ Find one more person with an entry-level data scientist job and ask them tips & questions about what they do ❤
- □ Research and attend a Python / R meetup or webinar 🤝 🔎



Week 6 SQL & Databases (I)

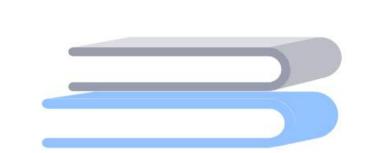
Understanding and mastering large databases is a key, if often overlooked, skill in data science. Through our SQL course you will learn the basics of this programming language, the fundamentals of database theory, and the most commonly used queries.

You'll be able to choose between two popular variants of SQL (PostgreSQL and MS SQL).



SuperDataScience Courses

SQL & Database Design A-Z, modules 1-7 (around 5 hours of lectures)



Extra Resources

<u>SelectStarSQL</u> – practical beginner guide to the most frequently needed part of all SQL knowledge



FIND TIPS AND RESOURCES
IN OUR **CAREER TOOLKIT!**

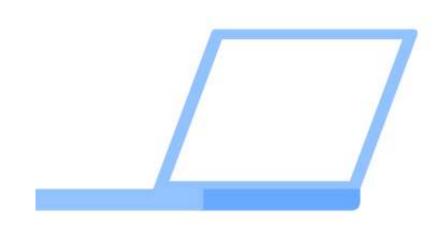
- □ Continue your #30DaysOfCode challenge with SQL (and add an extra hashtag #SDS99!) 🟋
- ☐ Research and understand 5 entry-level "data scientist" job descriptions requiring SQL skills •••
- □ Nurture your connections! Follow up with people you've connected with so far (share resources, give an update, follow up on their advice and how you applied it) >>



Week 7 SQL & Databases (II)



How is it going with your SQL? Complete your course load and dive into a very substantial portfolio project that is designed to challenge your skills.



SuperDataScience Courses

SQL & Database Design A-Z, modules 8 (around 1 hour of lectures) Module 9 is advanced and optional



Portfolio Projects

Product Data Analysis for Social Networking Platform Yammer



Extra Resources

SQL Questions – practical guide to the hardest, remaining 30% of SQL knowledge (only for the strong of heart!)



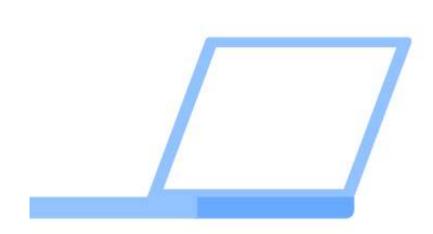
- ☐ Continue with your #30DaysOfCode challenge 溪
- ☐ Research and <u>reach out</u> to 3 <u>potential mentors</u> ❤
- ☐ Research & practice 10 typical <u>SQL interview questions</u>



Applied Business Analytics

As a Data Scientist, many stakeholders in a company will come to you expecting **strategic support in decision making** and a solid understanding of **key business metrics**.

Our Applied Business Analytics section contains a curation of key books and resources for you to design and track **business metrics, funnels, and A/B tests**. We'll also cover the **basic statistical concepts** you'll need, using a lot of intuition and practice, so you'll be ready to tackle all of your projects.



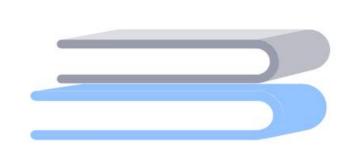
SuperDataScience Courses

Statistics in Business Analytics & Data Science A-Z, modules 1-3 (around 3.5 hours of lectures)



Portfolio Projects

A/B test results in Tableau: Analyzing Wikipedia's New Search Functionality



Extra Resources

A practical guide to A/B testing (web article)
Library of industry- and department-specific KPIs (platform)
Popular metrics frameworks (web article)
Lean Analytics (book)



FIND TIPS AND RESOURCES
IN OUR **CAREER TOOLKIT!**

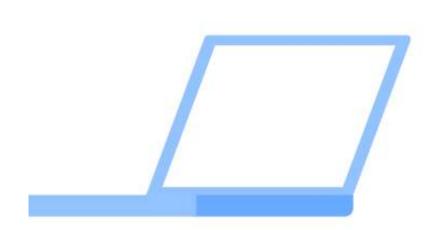
- ☐ Build your resume with recruiters' tips and <u>best practices</u> **●** ☐ When applying to big companies, make sure you CV passes
- ☐ When applying to big companies, make sure you CV passes bot filtering (using services <u>such as this one</u>) **.**
- □ Practice technical interview questions using online practice platforms (such as <u>this</u> or <u>this one</u>) 📜 💼



Intro to Machine Learning

Even if you land a first data job that doesn't require machine learning as part of your day-to-day tasks, you will be required to **have an understanding of the discipline** in order to spot opportunities and, when appropriate, **build your own simple models**.

This selection from our vast collection of machine learning content will help you do just that!



SuperDataScience Courses

Machine Learning A-Z: modules 1-7, 13-17, 24-26 (complete only R or Python lectures, depending on which language you chose – roughly 6 hours of content)



Portfolio Projects

The course features practical projects for each of the models covered. Make sure to **showcase them** in your portfolio!



Extra Resources

Machine Learning: A Visual Primer (web article)



FIND TIPS AND RESOURCES
IN OUR **CAREER TOOLKIT!**

- ☐ Are you keeping up with podcasts and articles? By now, you should be enjoying several each week! Find something that tickles your curiosity on real industry ML applications <a>>

- ☐ Research and reach out to 3 potential mentors ❤
- ☐ Research and critically compare 5 different data scientist job positions in order to understand them •
- ☐ Research <u>competitions</u>, datathons, and <u>open source projects</u> you can contribute



Catch-up Week (I)

Wow, congrats!

You've completed the main part of your Learning Path. Take a moment to breathe, look back, and celebrate your achievements.

Use this week to catch up with your study plan and focus on extra projects and career development.

Also, don't forget to reflect on your journey, see what you liked best, and make a choice on which specialization you'd like to tackle starting next week.

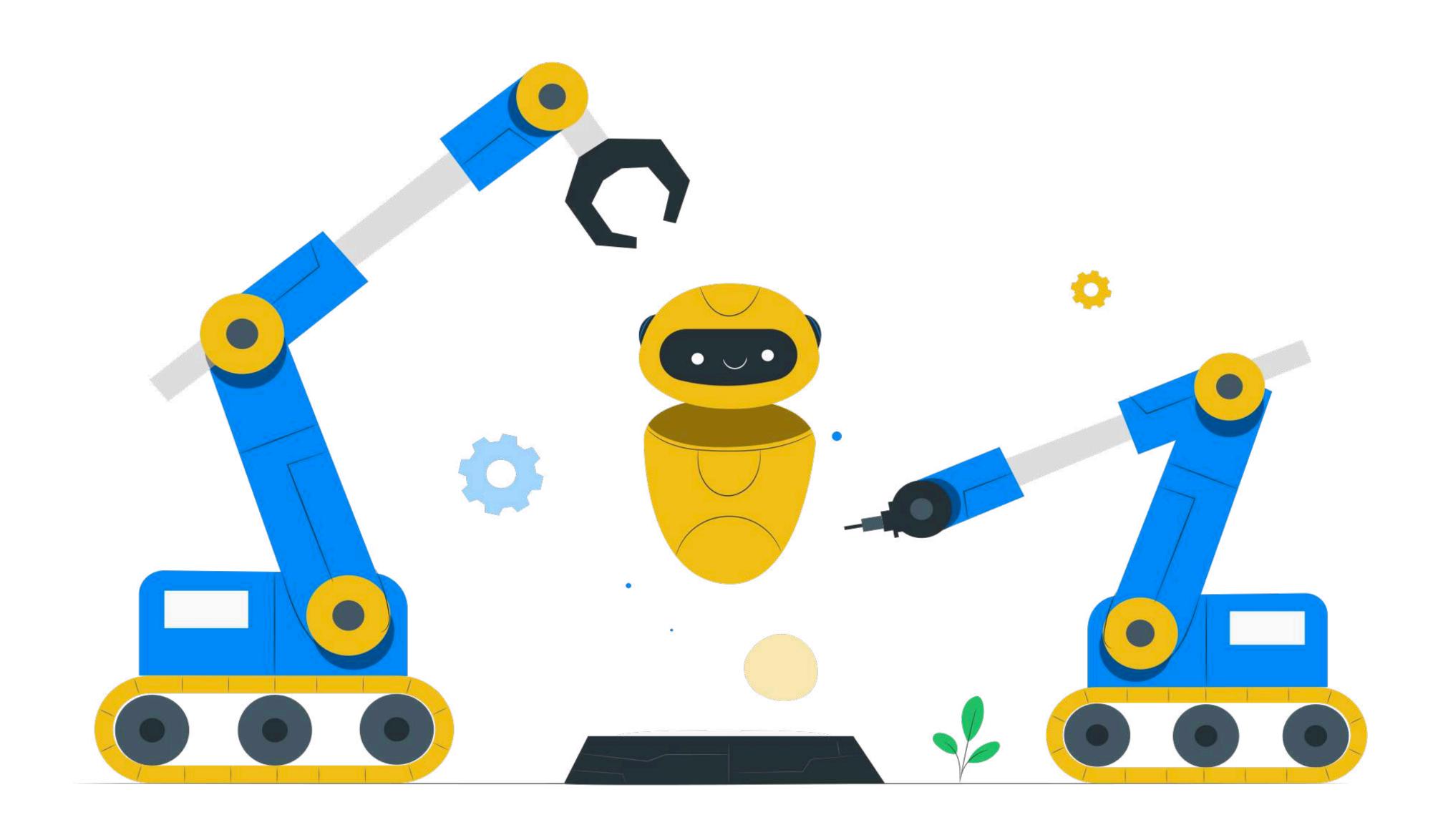


Write a blog post on your learning experience so far 💝
Make a list of 20 companies you could see realistically
yourself working for and research pros and cons about
them •
Reach out to 3 recruiters from your chosen companies 🤝
Create an Excel file (or use a Trello board, or tools such as
Scooby AI) to keep track of all
your applications and job-related network connections 💼
Research and apply to 10 job positions
Look for a dataset on a topic you are passionate about and
build a project, then add it to your portfolio 🐷



Specializations

- MACHINE LEARNING
- B DATA VISUALIZATION
- C DATA WRANGLING



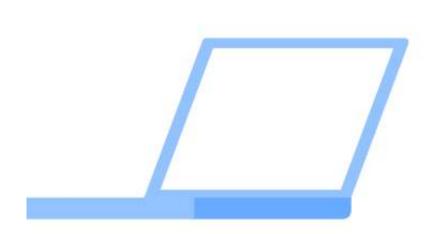


Week 11

Machine Learning (I)

If you want to delve deeper into the world of machine learning, through this specialization you will learn many different, more advanced models.

You will also further strengthen your **calculus and statistics** skills in order to understand ML algorithms more in depth.



SuperDataScience Courses

Machine Learning A-Z: modules 8-12, 18-23, 27 (complete only R or Python lectures, depending on which language you chose – around 7 hours of lectures)



Extra Resources

Linear Algebra, Multivariable Calculus, Gradient Descent refreshers (courses)

Introduction to Statistical Learning (free book)



FIND TIPS AND RESOURCES
IN OUR **CAREER TOOLKIT!**

☐ Apply to 10 more positions 💼	
☐ Reach out to 3 more recruiters from your top companies	
☐ Read <u>a guide</u> on how to prepare for a tech interview •	
☐ Practice writing a cover letter that stands out •	
☐ Schedule a session with a data science mentor who can	
help you critically review your applications (CV, portfolio,	,
LinkedIn. etc.) 🤝	

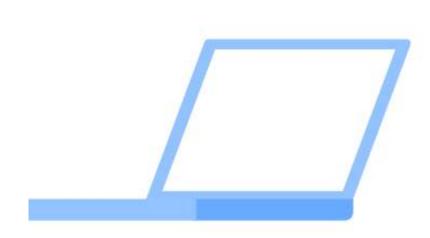


Week 12

Machine Learning (II)

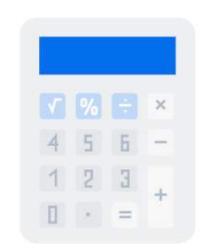
After covering all the most popular machine learning models, this week you'll explore **other key concepts** such as dimensionality reduction, model selection and boosting.

You'll also tackle your portfolio project for the specialization.



SuperDataScience Courses

Machine Learning A-Z: modules 38-44 (complete only R or Python lectures, depending on which language you chose – around 4 hours of lectures)



Portfolio Projects

If you chose R as a programming language:

Predicting Crime Rates in R

If you chose Python as a programming language:
Predicting Car Prices in Python



☐ Apply to 10 more positions 💼
☐ Practice non-technical interview questions: compile a list of
common questions and ask a friend to do a mock interview
with you 🏋 💼
☐ Reach out to 5 people in your network to inquire about
job opportunities 🤝
☐ Implement feedback from your mentor(s) in your new
applications 💼

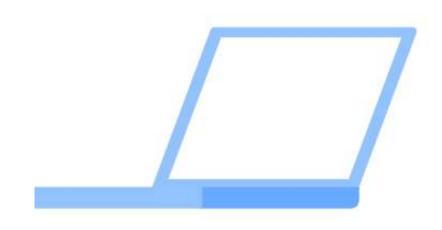


Week 13

Machine Learning (III)



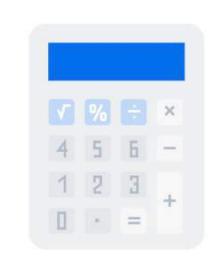
We are flying high this week! Diving deep into **Natural Language Process (NLP) and Deep Learning** to complete our tour of Machine Learning techniques.



SuperDataScience Courses

Machine Learning A-Z: module 34-37 for Python track, 34-36 for R track (around 6 hours of lectures)

Modules 28-33 are advanced and optional



Portfolio Projects

If you have the bandwidth, at the end of this Specialization, it would make sense for you to **participate in a collaborative or open source ML project online**, working together with people of different skills and levels.

Our Career Toolkit has some suggestions on **free platforms** where you could find such challenges.



Career Prep Checklist

□ Apply to 10 more positions
 □ Record your next interview and review it afterwards, looking for possible improvements
 □ Participate in a webinar, conference, or meetup and try to connect with at least 5 people
 □ Review content from previous weeks and find 1 or 2 topics to refresh. Use both SDS material and external sources to

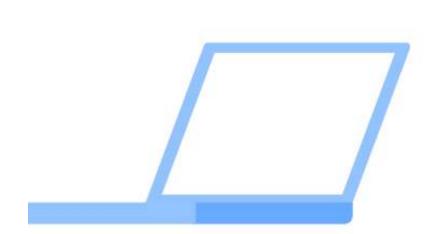
reinforce your knowledge through different sources 🟋 🔎



Week 11 Data Visualization (I)

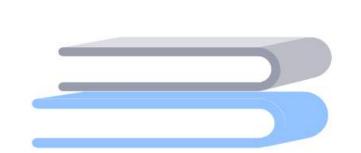
Were you fascinated by data visualization and you want to become a wizard of charts, bars, and advanced visual techniques to convey meaning through data? Look no further!

Our advanced Tableau courses will cover **animations**, complex graphs, and enhanced visualization techniques to make your reports stand out.



SuperDataScience Courses

Tableau 2020 Advanced: modules 1-4 (roughly 6 hours of lectures)



Extra Resources

Storytelling with Data (book)
Aim to finish this book in the 3 weeks of your Specialization!



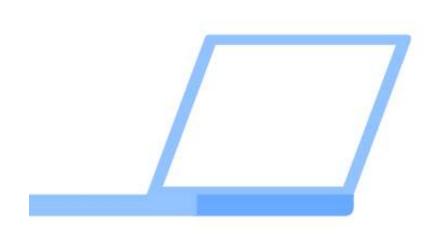
FIND TIPS AND RESOURCES
IN OUR **CAREER TOOLKIT!**

☐ Apply to 10 more positions 💼
☐ Reach out to 3 more recruiters from your top companies 🦠
☐ Read <u>a guide</u> on how to prepare for a tech interview •
☐ Practice writing a cover letter that stands out ●
☐ Schedule a session with a data science mentor who can
help you critically review your applications (CV, portfolio,
LinkedIn, etc.) 🤝



Week 12 Data Visualization (II)

This week you'll complete your overview of advanced Tableau skills (Level of Detail calculations, mapping techniques, animations) and explore complex visualization techniques.



SuperDataScience Courses

<u>Tableau 2020 Advanced</u>: modules 5-8
<u>Tableau Expert: Top Visualization Techniques</u>: modules 1-5
(roughly 6.5 hours of lectures)



Extra Resources

Continue reading the book you started on Week 11



Career Prep Checklist

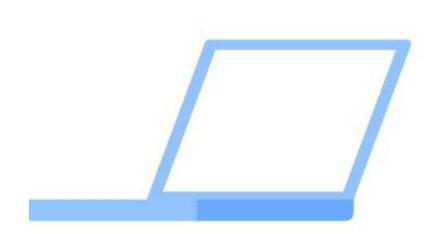
□ Apply to 10 more positions
 □ Practice non-technical interview questions: compile a list of common questions and ask a friend to do a mock interview with you
 □ Reach out to 5 people in your network to inquire about job opportunities
 □ Implement feedback from your mentor(s) in your new applications



Week 13

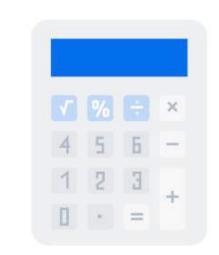
Data Visualization (III)

After mastering complex visualization techniques (such as the Sankey Diagram and Likert Scale), it's time to crown your portfolio with a **beautiful and complex Tableau project**.



SuperDataScience Courses

<u>Tableau Expert: Top Visualization Techniques</u>: modules 6-9 (roughly 2.5 hours of lectures)



Portfolio Projects

<u>Analyzing and visualizing flying delays and cancellation for Flying High Magazine</u>



Extra Resources

Finish reading the book you started on Week 11



FIND TIPS AND RESOURCES
IN OUR **CAREER TOOLKIT!**

- ☐ Apply to 10 more positions •
- ☐ Record your next interview and review it afterwards, looking for possible improvements •
- □ Participate in a <u>webinar</u>, <u>conference</u>, or meetup and try to connect with at least 5 people >> P
- □ Review content from previous weeks and find 1 or 2 topics to refresh. Use both SDS material and external sources to reinforce your knowledge through different sources 🖫 🔎

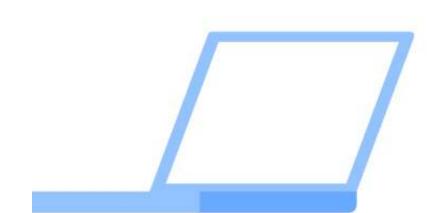


Week 11

Data Wrangling (I)

By now you'll have learned that **data cleaning and manipulation** takes up most of a data practitioner's time – learning to do it like a pro can help you speed up the process and free up time for your analyses.

You'll take a closer look at how data infrastructures work, learn how to scrape data from the web, and much more.



SuperDataScience Courses

Python track: <u>Data Manipulation in Python – A Pandas Crash</u> <u>Course</u>, modules 1-7 (about 6.5 hours of lectures)

R track: <u>Advanced R Programming</u>, modules 1-3 (about 4 hours of lectures)



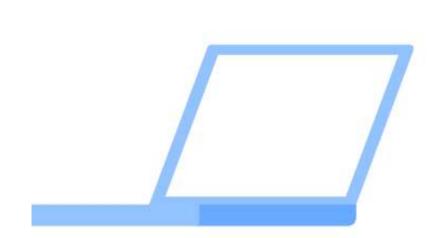
☐ Apply to 10 more positions 💼	
☐ Reach out to 3 more recruiters from your top companies	5
☐ Read <u>a guide</u> on how to prepare for a tech interview ●	
☐ Practice writing a cover letter that stands out •	
☐ Schedule a session with a data science mentor who can	
help you critically review your applications (CV, portfolio,	
LinkedIn, etc.) 🤝	



Week 12

Data Wrangling (II)

Time to complete your advanced Python and R courses and get into web scraping.



SuperDataScience Courses

Python track: <u>Data Manipulation in Python – A Pandas Crash</u> <u>Course</u>, modules 8-9 (about 2.5 hours of lectures)

R track: <u>Advanced R Programming</u>, modules 4-5 (about 2 hours of lectures)



Extra Resources

Guides to Web Scraping in R or in Python (web articles)



FIND TIPS AND RESOURCES
IN OUR **CAREER TOOLKIT!**

- ☐ Apply to 10 more positions •
- ☐ Practice non-technical interview questions: compile a list of common questions and ask a friend to do a mock interview with you 🖫 💼
- ☐ Reach out to 5 people in your network to inquire about job opportunities ❤️
- ☐ Implement feedback from your mentor(s) in your new applications •

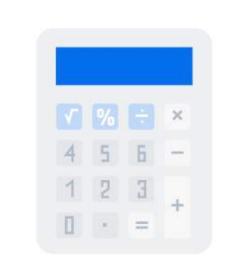


Week 13

Data Wrangling (III)

In smaller companies, a Data Scientist may have to design and maintain a small data infrastructure with help from backend developers; but even if you work in larger contexts, an **understanding of how the data ecosystem of an organization works** is crucial to build awareness in your practice.

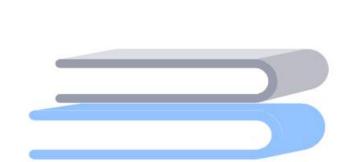
We'll look into that as we wrap up our Specialization with **portfolio projects** requiring some serious data preparation!



Portfolio Projects

Python track: <u>US working visa analysis</u>

R track: zoo and stringr



Extra Resources

The stages of data infrastructure (web article)



FIND TIPS AND RESOURCES
IN OUR **CAREER TOOLKIT!**

- ☐ Apply to 10 more positions •
- ☐ Record your next interview and review it afterwards, looking for possible improvements
- □ Participate in a <u>webinar</u>, <u>conference</u>, or meetup and try to connect with at least 5 people >> P
- □ Review content from previous weeks and find 1 or 2 topics to refresh. Use both SDS material and external sources to reinforce your knowledge through different sources 🖫 🔎



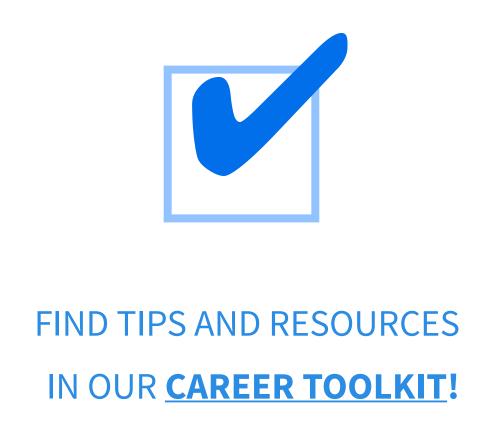
Catch-up Week (II)

You reached the end of your Learning Path! Can you believe it? Congratulations.

You may have applied for dozens of jobs by now – did you get any interviews? How far did you go? **Don't discourage**: you've learned the basic skills (both technical and not) to keep your momentum going. Take this week to catch up with the curriculum and **make more smart moves** in your career development strategy.

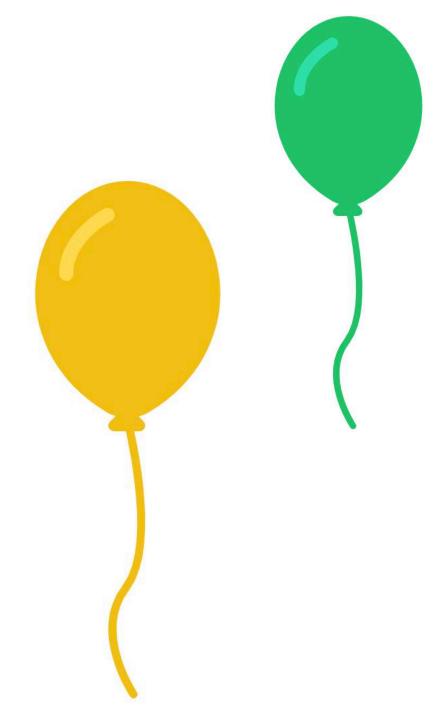
Continue to **nurture your network**, find inspiration in other people, **keep learning new things** and expanding your portfolio, reflect on who you are and what you want to achieve: you're already living the life of a successful data science practitioner!

With enough **practice and consistency**, these small habits will compound and you will land your first job in data. **Take pride in what you have achieved, and never stop learning!**



☐ Write a new blog post on your overall learning experien	
☐ Take a look at your job application Excel file and review	
efforts: how many applications have you sent? Which o	ones
did you hear from, or even gave you an interview? Refir	
your strategy based on your outcomes (skills to streng	ıthen
new skills to add, type of companies to focus on, etc.)	
☐ Look for another dataset on a topic you are passionate	l ·
about and build a project, then add it to your portfolio	8
🗆 "Interview" 3 more people working as data scientists 🤝	•
\square Find a new topic to explore, and continue to learn \wp	





Day 99

Celebration Day

Rest is important, as is reflecting on your experience and celebrating your small (or big) wins.

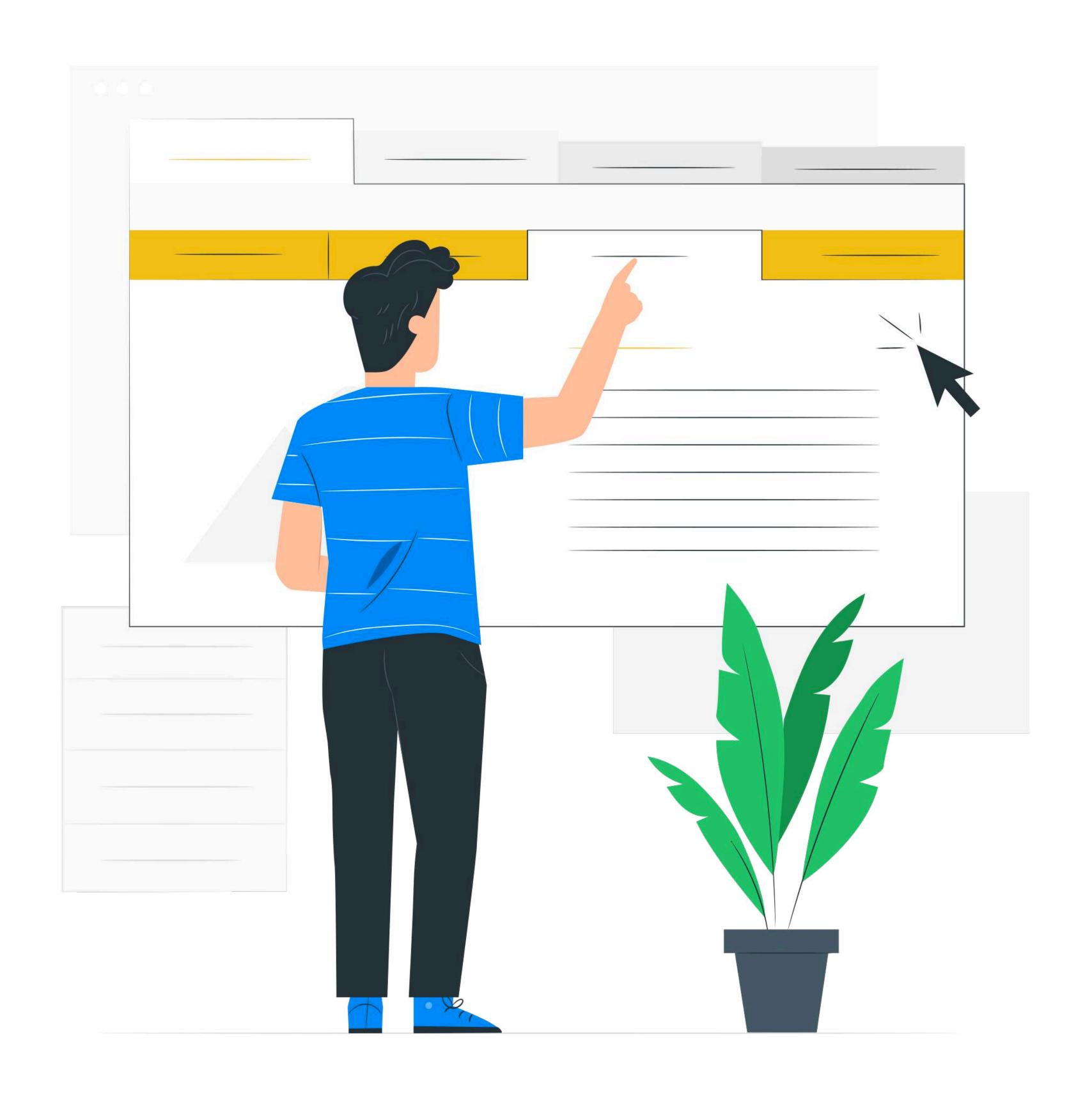
Take this day to do just that, in whatever way feels best to you.

Then, on Day 100, you'll be ready to **rinse and repeat**. No matter the outcome of your 99-day journey, **career development is a continuous effort**.

So, continue to **connect, practice, keep up, and apply**; explore the vast library of SuperDataScience courses on <u>Deep Learning</u>, <u>Python</u>, <u>Statistics</u>, and more; expand your portfolio with <u>practical machine learning case studies</u>, learn <u>new visualization tools</u>; or take an extra **specialization** and keep expanding your skillset.

If you have enjoyed the journey, get in touch at support@superdatascience.com: we want to hear your story, and keep improving this Learning Path and Study Plan for future students.

Happy analyzing!



Share your journey on social media, find other students like you

Use the hashtag #SDS99

Questions? Reach out at support@superdatascience.com

