

Ein Leitfaden für Data Scientists für mehr Zeitersparnis

Als Data Scientist stehen Sie an der Schnittstelle von Wissenschaft, Technik, Business Intelligence und Mathematik. Sie haben viele Dinge in Einklang zu bringen – von der Codierung und Datenvisualisierung bis hin zum Bereinigen von Datensätzen und der Reaktion auf Ad-hoc-Verpflichtungen. Wenn Sie Ihre Zeit nicht genau pro Tätigkeit überblicken, werden Sie vielleicht überrascht sein, wie Sie diese tatsächlich einsetzen. In einer kürzlich von HP durchgeführten Umfrage unter 350 Data Scientists weltweit, gaben 48 % an, dass sie mehr Zeit mit der Organisation ihrer Daten verbringen als mit der eigentlichen Analyse der Daten.¹

Ihre Zeit ist wertvoll, daher sind Instrumente zur Effizienzsteigerungen ein wichtiges Handwerkszeug. Wer hätte gedacht, dass herausragende Leistungen in Ihrer Funktion genauso viel mit Zeitmanagement zu tun haben könnten wie die Beherrschung der Feinheiten von SQL? Im Folgenden finden Sie sechs Tipps zur Optimierung Ihrer Arbeitsabläufe und zur bestmöglichen Nutzung Ihrer Zeit.





1

Zeitersparnis durch proaktive Kommunikation

Um Zeit zu sparen, ist es wichtig, Kommunikationspunkte zu schaffen, die gewährleisten, dass Sie während des gesamten Projekts die richtigen Entscheidungen treffen. 40 % der befragten Data Scientists geben an, dass sie häufig mit der Bearbeitung von Daten beginnen, bevor sie die Geschäftsziele vollständig verstanden haben.¹ Dieser Mangel an Kommunikation führt häufig dazu, dass Manager unrealistische Erwartungen an das Projektergebnis haben.

Das Problem kann sich als erstaunlich facettenreich erweisen. Schließlich sind sowohl die Interessenvertreter der Unternehmen als auch Data Scientists relativ technisch veranlagt – sie verwenden lediglich unterschiedliche Terminologien und Kommunikationsstile. Das führt dazu, dass beide Gruppen manchmal aneinander vorbeireden. Ein wichtiger Unterschied: Business Stakeholder neigen dazu, in binären Ergebnissen zu denken, während sich die Datenwissenschaft in Nuancen der Ungewissheit ausdrückt. Es ist wichtig, dass Sie Ihren datengesteuerten Ansatz mit dem Ihrer Interessensgruppen in Einklang bringen.

“Business Stakeholder neigen dazu, in binären Ergebnissen zu denken, während sich die Datenwissenschaft in Nuancen der Ungewissheit ausdrückt.”

Unabhängig davon, wie fachbezogen das Thema ist, muss die Kommunikation so erfolgen, dass sie für die Menschen, die es umsetzen müssen, als sinnvoll verstanden wird. Ken Jee, Z by HP Ambassador² und Leiter der Data Science bei der Scouts Consulting Group, ist im Bereich Sportanalytik tätig. “Viele Sportler werden nicht bereit sein, eine Lösung zu implementieren, die sie nicht verstehen. Beginnt man mit einfacheren linearen Modellen, auch wenn die Ergebnisse vielleicht nicht so gut sind, können wir den Sportlern unsere Gründe für die Entscheidungsfindung sehr gut vermitteln.” Mit anderen Worten: Ihr Projekt kann schneller umgesetzt werden, wenn jeder Beteiligte den Wert des Projekts versteht.

Gewinnen Sie Zeit, indem Sie sich mit Ihren Daten im Vorfeld vertraut machen



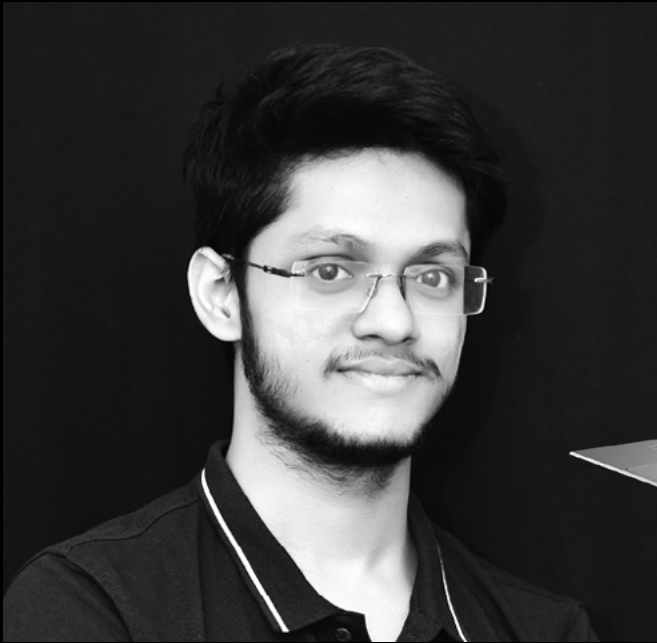
Louise Ferbach

Ihre Zeit ist nicht nur knapp bemessen, sondern häufig zwischen langfristigen und kurzfristigen Projekten aufgeteilt, so dass die optimale Nutzung Ihre Zeit von größter Bedeutung ist. Hier kann sich die Optimierung der Zeit am Anfang eines Projekts später enorm auszahlen. Ein häufiger Fehler, der Sie im weiteren Verlauf Zeit kosten kann, ist zu früh mit der Modellierungsphase zu beginnen, noch bevor Sie Ihre Daten tatsächlich verstehen. Wenn Sie mit der Arbeit an einem neuen Projekt beginnen, möchten Sie zweifellos mit der Modellierung beginnen – schließlich ist das der aufregende Teil der Arbeit, doch erfahrene Datenwissenschaftler wissen es besser.

Louise Ferbach, Z by HP Ambassador² und Data Scientist in Frankreich, erklärt, dass die Zeit, die hier aufgewendet wird, die fruchtbarste im gesamten Projekt sein kann: „Um Korrelationen zu erkennen,“ und, wie Frau Ferbach sagt, „Ihre Daten kennenzulernen“. Wenn man sich Zeit nimmt, und seien es nur ein oder zwei Tage, ist es möglich, Muster zu erkennen, die man in sein Modell einfließen lassen kann. Das kann sich letztlich als ein großer Zeitgewinn erweisen.

Die Dokumentation ist ebenfalls ein wichtiger Teil eines jeden datenwissenschaftlichen Projekts. Die frühzeitige Sicherstellung der entsprechenden Dokumentation ist ein entscheidender Schritt auf dem Weg zur Effizienzsteigerung und Zeitersparnis. Versäumen Sie es nicht, Ihren eigenen Code sorgfältig zu dokumentieren. Unzureichend dokumentierter Code ist eine schlechte Angewohnheit, die gleichermaßen Software Engineers als auch Data Scientists betrifft. Gehen Sie also niemals davon aus, dass Sie sich selbst nach einer Woche, geschweige denn nach einem Monat, an Ihre Absichten erinnern werden. Wenn Sie Zeit in die Dokumentation investieren, müssen Sie Ihren Code nicht jedes Mal entziffern, wenn Sie ihn erneut einsetzen.

Sparen Sie Zeit durch die Ausstattung Ihres Arbeitsplatzes mit geeigneten Tools



Paras Varshney



Bei der Auswahl der Hardware spielen CPU und GPU eine zentrale Rolle. Eine hohe Rechenleistung ist von Anfang entscheidend, um die Anforderungen Ihrer täglichen Arbeit zu erfüllen und um Zeit zu sparen.

Häufig sind es auch einfache Dinge, die täglich den entscheidenden Unterschied ausmachen. Die Auswahl vieler Tools basiert auf persönlichen Entscheidungen, die von Komfort, Nutzen und Vorlieben abhängen. Hier gibt es kein Patentrezept für Zeitersparnis und Effizienzsteigerung.

So hat beispielsweise jeder seine eigenen Vorlieben, wenn es um Mauszubehör geht. Paras Varshney, Z by HP Ambassador² und Datenwissenschaftler bei LogicAI in Polen, benutzt selten das Clickpad seines Laptops; stattdessen ist eines seiner Lieblingsgadgets seine Maus. Er findet, dass er mit der Maus viel schneller arbeiten kann. Ken Jee macht als Data Scientist ebenfalls intensiven Gebrauch von

seinem Trackball. Er arbeitet viel in Flugzeugen und ist begeistert, dass der Trackball nicht bewegt werden muss. Ein echter Zeitgewinn bei der Arbeit auf begrenztem Raum.

Paras konzentriert sich auf das Display: "Ich liebe die Nutzung einer geteilten Anzeige. Mit dem Curved-Monitor HP Z38c habe ich mehrere Fenster gleichzeitig geöffnet. So muss ich nicht zwischen Fenstern oder Tabs wechseln, da ich alle gleichzeitig offen habe." Die Arbeitsabläufe erfordern Multitasking in vielen geöffneten Fenstern, darunter Apps, Browser-Tabs und Dashboards.

Laptop-Bildschirme verlangsamen die Arbeitsabläufe durch ständiges Öffnen und Schließen oder bei der Verwaltung des Desktops. Mit dem richtigen Zubehör – von der richtigen Maus bis hin zu Curved Monitoren – können Sie Ihren Arbeitsablauf optimieren und Zeit sparen.

Sparen Sie Zeit durch das Erkennen Ihrer Produktivitätsmuster

Als Data Scientist liegt die Mustererkennung in Ihrer DNA. Ebenso, wie Sie Ihre Tage mit der Optimierung von Modellen verbringen, gibt es auch Möglichkeiten, Ihre täglichen Abläufe für mehr Produktivität zu optimieren.



Qishen Ha

Da kein Mensch dem anderen gleicht, gibt es nicht den einen optimalen Arbeitsablauf für alle. In einer Zeit, in der sich viele Menschen dafür entscheiden, von zu Hause aus zu arbeiten, hat Qishen Ha, Z by HP Ambassador² und Ingenieur für maschinelles Lernen bei LINE, für sich erkannt, dass die einsame Arbeitsumgebung nichts für ihn ist, da er sich zu leicht ablenken lässt. Stattdessen zieht er es vor, im Büro zu sein und umgeben von anderen Menschen zu arbeiten. "Man treibt sich selbst an, intensiver und konzentrierter zu arbeiten", sagt er.



Ken Jee

Ken Jee hat bei seiner Vollzeitarbeit von zu Hause aus viel über seinen Arbeitsablauf nachgedacht und Zeitblöcke für die Bearbeitung von E-Mails eingerichtet. Er hat die Ineffizienz erkannt, die sich beim Multitasking zwischen seiner Haupttätigkeit und dem Versand von E-Mails ergibt. Mit jeder gelesenen E-Mail büßt er bei der Erledigung seiner Hauptaufgabe wertvolle Zeit ein, weil es ihn jedesmal wieder Zeit kostet, sich in seine Aufgabe wieder einzudenken. In einer Zeit, in der von jedem erwartet wird, auf E-Mail oder Slack-Nachrichten praktisch in Echtzeit zu reagieren, kann man diesem Trend entgegenwirken, indem man Zeitblöcke für E-Mails einplant und so die Zeit optimal nutzt.

5

Sparen Sie Zeit mit den richtigen Tools und Konfigurationen

Einen neuen Computer zu konfigurieren, stellt immer eine Herausforderung dar. Bei einer Umfrage beklagten 42 % der Data Scientists, dass sie zu viel Zeit mit der Konfiguration ihrer Datenumgebung aufwenden, wobei durchschnittlich fünf Stunden pro Woche verschwendet werden.¹



Eine wichtige Möglichkeit zur Effizienzsteigerung ist der Einsatz eines Windows Subsystems für Linux® WSL 2³. Es ermöglicht Ihnen die direkte Ausführung von Linux-Tools, Dienstprogrammen und Anwendungen direkt innerhalb von Windows, ohne auf eine Dualboot-Konfiguration oder eine virtuelle Maschine zurückzugreifen. Jee berichtet beispielsweise, dass er nicht mehr auf seine Linux-Workstation aus der Ferne zugreifen muss, während er an seinem Windows-Desktop arbeitet. Eine Verbesserung, die Reibungsverluste reduziert und den Arbeitsablauf beschleunigt.

Ebenso haben sich vorkonfigurierte Software-Stacks als eine wahre Offenbarung für Data Scientists erwiesen, die das Vergnügen hatten, bereits damit zu arbeiten.

“Mir war nicht bewusst, dass es so etwas gibt, bevor ich damit konfrontiert wurde. Es hat meine Arbeit völlig revolutioniert”, berichtet Louise Ferbach. “Wenn Sie Ihren Data Science-Computer von Z by HP erhalten, verfügt er über einen sogenannten Data Science Software-Stack.” Bei diesem Software-Stack handelt es sich im Wesentlichen um eine umfassende Sammlung von Anwendungen und Umgebungen – allesamt vorinstalliert und mit automatischen Updates versehen. Dadurch werden die unvermeidlichen Software-Inkompatibilitäten und die Zeit für die Fehlersuche vermieden, die bei Routineeinrichtungen anfallen.

Sparen Sie Zeit mit Workflow-Automatisierungen

Je nachdem, wo Sie sich in Ihrer datenwissenschaftlichen Laufbahn befinden, haben Sie Ihren Arbeitsablauf wahrscheinlich bereits in gewisser Weise optimiert. Während Sie an Fähigkeiten und Erfahrungen wachsen, ist es gleichzeitig wichtig, auf zusätzliche Möglichkeiten zu achten, mit denen Sie Ihr Zeitmanagement weiter perfektionieren können. "Einfache Fertigkeiten wie das Aufteilen des Bildschirms können auch von Anfängern in der Datenwissenschaft erlernt werden", sagt Varshney und meint damit, wie man einen großen Bildschirm am besten nutzen kann.

Sie halten Ausschau nach einem anspruchsvolleren Instrument für Ihr Toolkit? Automatisierung! Jeder Data Scientist mit etwas Erfahrung ist sich des Werts der Automatisierung von Arbeitsabläufen bewusst. Man erkennt auch als Anfänger schnell, dass einige Aufgaben viel manuelle Arbeit erfordern und immer wieder durchgeführt werden müssen. Die Automatisierung dieser Aufgaben kann viel Zeit sparen. Mit nur wenigen Befehlen können Sie einen ganzen Prozess selbstständig ablaufen lassen.

Der Vorteil von Z by HP

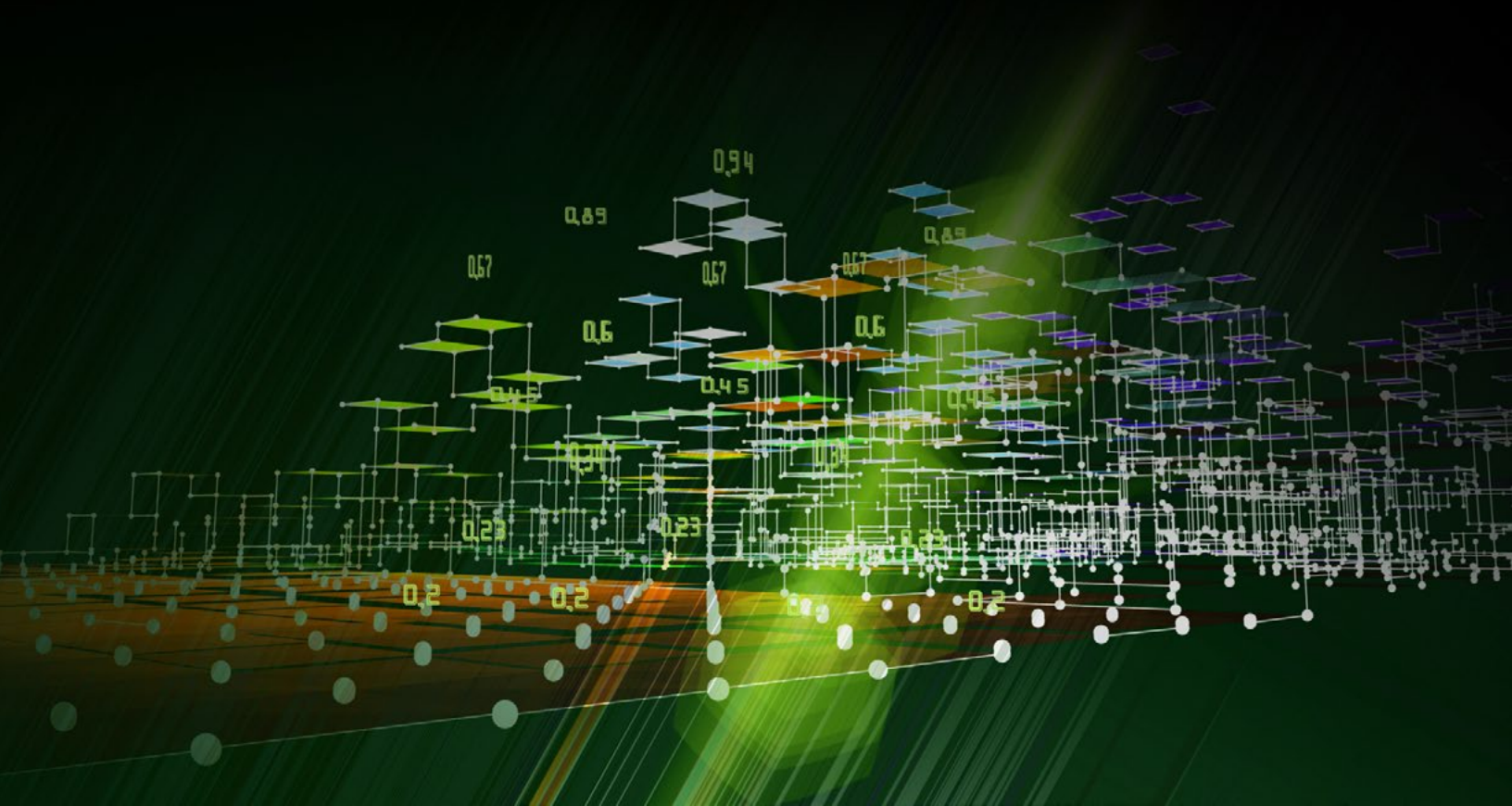
Wie die oben genannten Tipps zeigen, lassen sich eine Vielzahl von nützlichen Gewohnheiten für eine effizientere Arbeitsgestaltung in Ihrer Tätigkeit als Data Scientist aneignen. Selbst einige kleine Schritte können in der Summe zu einer erheblichen Optimierung führen. Die grundlegendste Zeitersparnis für Data Scientists wird jedoch immer von der Leistung ihrer Workstation bestimmt. Aus diesem Grund entwickelt Z by HP ständig Innovationen, um Data Scientists die High-Compute-Workstations, Bildschirme und Tools zur Verfügung zu stellen, die sie benötigen, um ihre Aufgaben so nahtlos wie möglich zu erledigen. Mit Komponenten der Enterprise-Klasse wie NVIDIA RTX-Grafikprozessoren können Sie die Wartezeit bis zum Erhalt wertvoller Erkenntnisse verkürzen und schneller zu Produktionsclustern übergehen. Informieren Sie sich über die Tools und Bibliotheken, die im vorkonfigurierten Software-Stack von Z by HP enthalten sind und für Zeitersparnis sorgen, oder erfahren Sie mehr über WSL 2, das auf ausgewählten Z by HP Data Science Workstations angeboten wird.



¹ HP proprietary research: Understanding Data Scientists, November 2021.

² Hinweis: Ken Jee, Louise Ferbach, Qishen Ha und Paras Varshney sind Teil des Z by HP Data Science Ambassador-Programms und werden mit Produkten versorgt.

³ WSL 2 erfordert Windows 10 oder höher, einen Intel Core i5 Prozessor oder höher und ist auf ausgewählten Z-Workstations verfügbar. Sie müssen Windows 10 Version 21H2 und höher (Build 19044 und höher) oder Windows 11 nutzen.



Jetzt beraten lassen!



Ihren persönlichen Ansprechpartner

Markus Hömig-Heidrich

erreichen Sie telefonisch unter:

+49 221.588 320 31

E-Mail:

hoemig@workstation-pro.com

WORKSTATIONPRO
Eine gute Entscheidung.

WORKSTATIONPRO GmbH
Vorgebirgstraße 59
50677 Köln
www.workstation-pro.com

