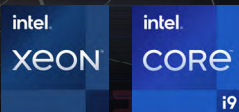
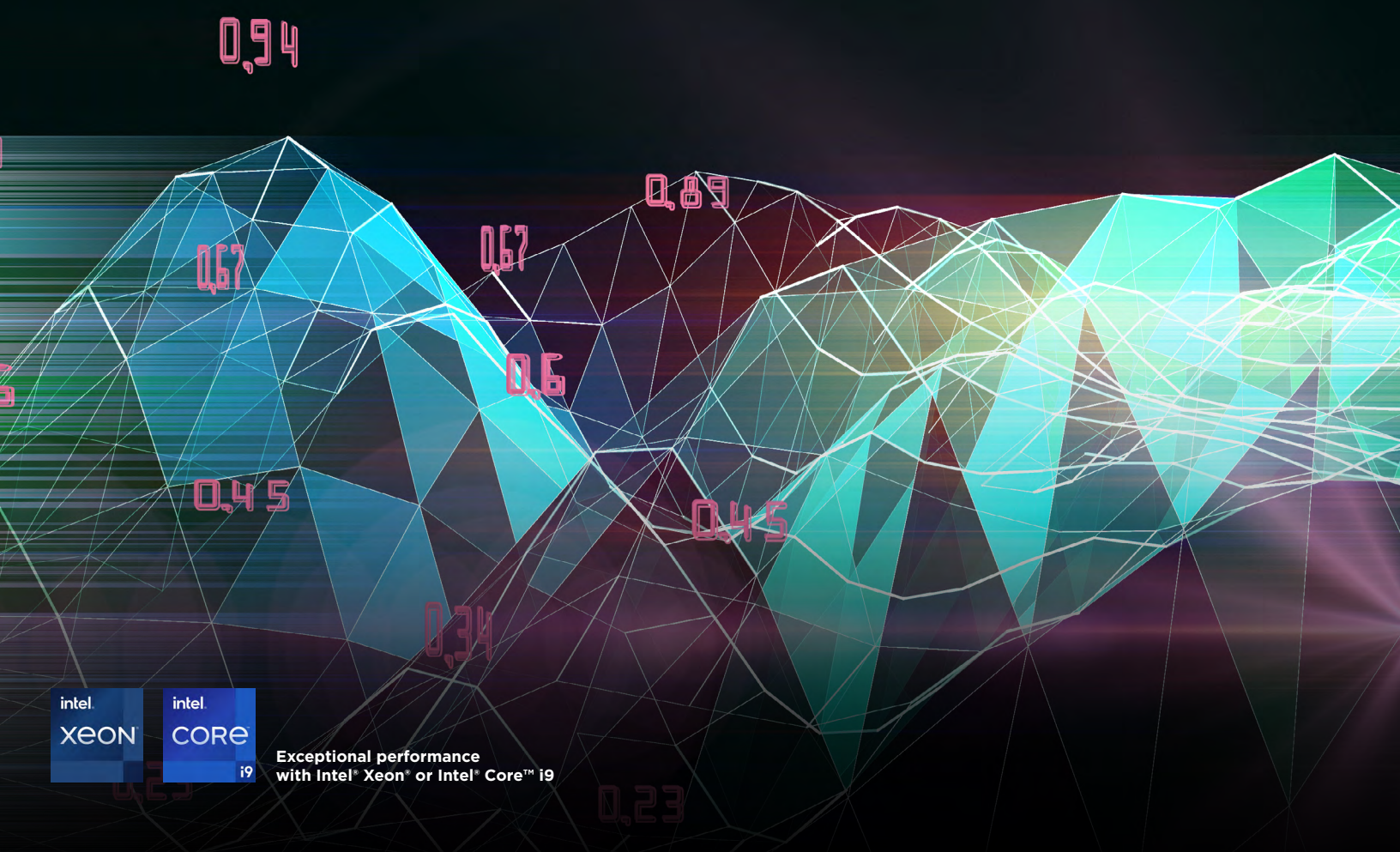


What your data science team wants you to know

Wouldn't the workflow of a data scientist improve if our peers really knew what it takes to produce valuable insights and overall, knew what they could do to be a better collaborator? According to a study conducted by Z by HP, nearly 40% of data scientists find it difficult to explain their work to non-technical stakeholders.¹ Here are a few things we're all thinking, but don't always have a chance to convey to the non-data scientists in the room.



Exceptional performance with Intel® Xeon® or Intel® Core™ i9



You need the data before the scientist

The old expression, “garbage in, garbage out”, conveys that the final results can be no better than the initial input. This expression is equally relevant in data science. It’s impossible to build good models and generate actionable, relevant recommendations without investing in the underlying data infrastructure first.

Companies should build out the data infrastructure before investing in data scientists. To allow data scientists to do their best work, there are some prerequisites to keep in mind: creating data capture systems, data engineering systems, and pipelines. Without these systems in place, data scientists are unable to complete the work we were hired to do.

Stakeholders don't need to be experts, but should understand the basics

An effective partnership goes both ways. As data scientists, we are consistently making sure to understand the business context of our work, so it only makes sense that leadership and other stakeholders do their best to understand the work we're doing. Luckily, Firat Gonen, a Z by HP Ambassador² and head of data science & analytics at Getir, contends that today's executives have easy access to learn the basics.

Gonen says, "The materials available today are top-notch and often essentially free. You don't need a very expensive education or a certification to learn how data science can help in industry X, whether it's hospitality, e-commerce, or something else."



Firat Gonen

Give data scientists a seat at the table

A common theme among all data scientists is that communication with leadership is key. Inviting the leadership team to provide details about the business problem helps data scientists scope and prioritize work, as well as understand the data needs for the model. According to a survey, 40% of data scientists state that they often get started on a project prior to fully understanding the business objectives.¹ Instead, data scientists often benefit from having an open conversation with their non-technical partners: Is this a model that's going to go under production, that's going to be making decisions effectively, autonomously, with a little bit of human oversight or no oversight? Understanding the bigger picture and risk tolerance of the business is key to getting the most out of data science projects.

Data science is about probability, not prediction



Heng Cher Keng

As much as data science has the power to make incredibly accurate forecasts, it's important to keep in mind that no single model is a silver bullet. Stakeholders must understand that data science is about calculating probabilities – an expectation of 100% precision is unlikely, seeing as imperfections are part of every model. The idea is to train and adapt models in such a way where they improve incrementally over time – but even the best data science work is never perfect.

Perhaps Heng Cher Keng, a Z by HP Ambassador,² Kaggle Competitions Master, and a data scientist in Singapore, puts it best: “We’re not magicians.”

Every experiment is productive—even the failed ones

The nature of data science is experimental and iterative. When leadership is understanding of this and provides data science teams leeway (especially at the start of a project), it can lead to substantially better results later on.

Projects are often the story of backtracking and trying different approaches. This means there may be mistakes along the way and some misses along with the wins. In fact, there are times when a specific project might need to be abandoned because the available data and modeling methods simply aren't yielding the desired results.



Ken Jee

“Projects are often the story of backtracking and trying different approaches.”

“I don’t think data scientists or project managers really comprehend how long it takes to get results, a lot of the time,” says Ken Jee, head of data science at Scouts Consulting Group. “The nature of data science work can be variable. I may go down a rabbit hole and then realize, ‘Oh, no, this did not work.’” Stakeholders and peers should be aware that there are a lot of reasons why this can happen. The data might not be as clean as initially expected, or there’s too much variance to get reasonably good results with the first few tries.

Data science isn’t software engineering

Although software engineering and data science are often conflated, they only share a few similarities: coding, creating pipelines, and getting data from one place to another. However, the similarities end here. Where other disciplines often have a finite ending to a project, data scientists devise a model and continually retrain it so that it remains relevant based on new data coming in. Jee states it best that this phenomenon, “is something that you don’t really see very often in a pure software engineering domain,” and that it is, “something that’s very difficult to articulate or to conceptualize” to peers and leadership.

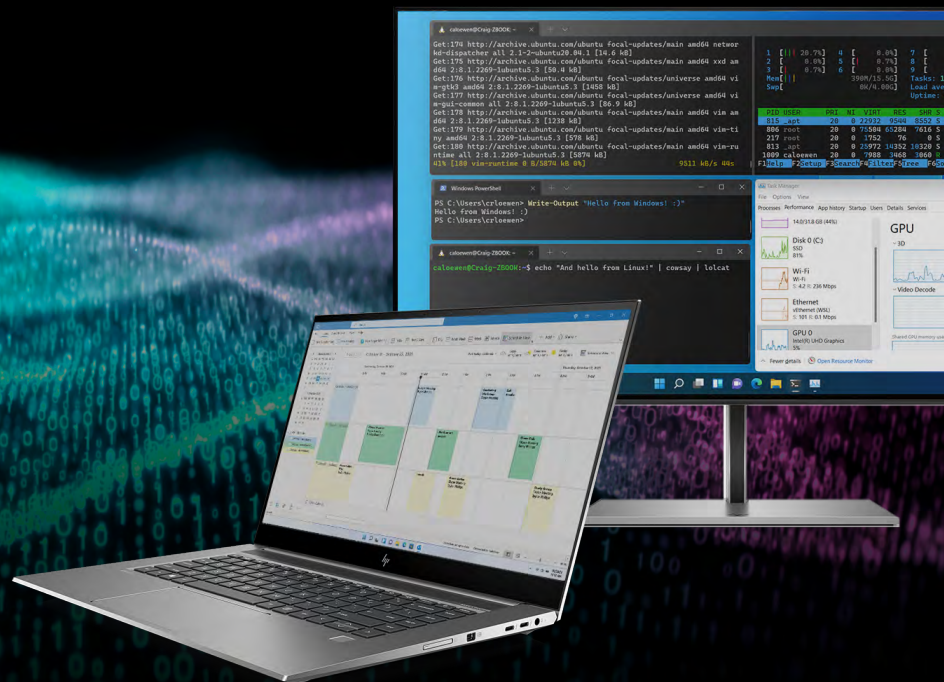
Data science requires a specialized setup

Over a third of data scientists feel they don't have the right technology to do their job efficiently.¹ Working without the proper tools can adversely affect the efficiency and flexibility of data science teams.

So, it's important to provide workstations that are built for data scientists. Take, for example, Z by HP – which delivers the ultimate data science solution. Z by HP workstations come preconfigured with the most popular data

science tools to save you time and are packed with enterprise-class components certified by partners like NVIDIA and Intel.

These features, “save me a lot of time compared to the other options I've used,” says Jee, who is a Z by HP Ambassador.² As much as the details may be hard for non-technical partners to understand, saving time is a universal goal anyone can get behind.



In conclusion

Ultimately, when leadership and peers are aware of what data scientists wish others know, it creates a more productive, successful atmosphere for all. From fostering open communication about business goals to truly understanding the processes of data science, anyone or any organization can implement these teachings. As a result, data scientists will bring more valuable insights to the business, helping the team achieve business objectives.

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

² Disclaimer: Firat Gonen, Heng Cher Keng and Ken Jee are part of the Z by HP Data Science Ambassador program and are provided with products.