×Hg

Data science in the real world

The Ubiquity of Data Science:

Data scientists are now part of virtually every major industry. Even traditional farming processes are being revolutionised through the deployment of data science and AI: measuring and optimising environmental conditions, seed genetics and even detecting pests. These same data principles can be applied to any core business processes. At Hg we've seen data science and data technology rapidly mature into a fundamental business tool that can drive a step-change in shareholder value.

At Hg we believe that harnessing data science and AI should now be core to how value is created in businesses.

Cloud infrastructure has given us computing power and storage that's quick to tap into and cheap to run. It takes weeks and months to create a rich, automated single customer view across a business, and deploy data science to then improve how we interact with and monetise customers. The impact we've seen on profit growth and business value is transformational.

This spurred us to establishing a Data Analytics & Insight unit in 2016, with the mission to help our Hg portfolio companies harness the potential of data. Evolving technology may have turned data science from alchemy to lego. But having impact 'in the real world', changing the behaviours of customers and tilting the way a business operates remains a complex craft. Now entering our 3rd year and after over 30 projects, we thought we'd share some honest perspectives on the daily real-world challenges we face — our 'top tips' for getting operational impact.

Each point here glosses over lots of nuance and some significant exceptions. We suspect many would also have lots to add to our list. But the common thread for us is that data science should 'just' be treated as another core business tool leaders have in their arsenal to drive value growth. So, don't be fooled by the hype (it's not magic!) nor be put off by it (it's not rocket science!). Deployed pragmatically to the core of the business it is now ready to drive transformative value impact.

> Modern tech allows the creation of a rich, automated single customer view in only a matter of months, and your first machine learning model may be only weeks. But to really drive value, a data science project also needs to be treated very pragmatically, like the change management effort it actually is.

Christoper Kindt,
Director and Head of Hg's Data Analytics & Insight team

Hg's 'top tips' for achieving operational impact from data science

Resist intellectual perfection

Simpler models are better in the real world. A linear regression is many times easier to explain and get adopted - and often only fractionally less effective. We've yet to see a sensible place to use deep learning.

Start with a simple crowd pleaser

There will be sceptics, and you have plenty to learn at the start about the data and how things really work. Bite off a small chunk that generates a very visible (if more modest) impact that then buys you the support you need for the real change management to come.

Even when you start small, don't forget to look ahead

You need to know where the data journey will go — and therefore what infrastructure is needed, what data sets might become useful, and the kind of team needed to operate it. So invest the time upfront to gather those broader data & AI opportunities and set a rough roadmap.

Avoid getting trapped in an IT silo

Unless the 'business' co-owns the data project, the implementation and change management effort becomes harder. And get their commitment — plan the operational changes upfront, define the KPIs to measure, and set targets.

Data is invariably imperfect

But that doesn't mean it doesn't contain valuable signals once it's scrubbed and distilled. And it's also invariably more challenging and skewed than you get told — be critical, spot what's bias and what's noise. The unglamorous secret is that data wrangling and cleaning is 80% of what we do.

Data Culture

Real 'data transformation' of a business is impossible without a 'data' culture. And that requires designing the data projects to be highly accessible — it's not just simpler models, but also intuitive tooling and offering training. If you can go beyond having just a small central team of data specialists to 'citizen data scientists' in each function you'll get real lift-off.

Embrace the cloud

Don't get trapped by tech that might already exist today, but is fundamentally not future proof. You _will_ regret it as the data work takes off, and the longer your leave it the more painful it gets!

Don't (yet!) believe the hype of automated data science

These platforms remain overly simplistic and narrow, they're no substitute for getting into the code. And getting your hands dirty is OK — it's not the painful, complex problem vendors might suggest.

AWS, Azure or GCP?

Though this might be heresy to some ... we think all 3 can be more than good enough (well, in most cases). Instead, what's key is how you set up these platforms: these have matured into broad, powerful and now complex ecosystems — ensure you have (or borrow!) the expertise needed to select the ingredients you need to concoct your future-proof data platform.

Invest in data governance

...to 'bottle' that initial data cleanse. Some core tools & validation processes will provide the basics, but ultimately it also requires creating a culture of data stewardship, driven by the business and its leadership. For example, best-in-class Hg CCOs bonus sales staff on how complete their CRM data is.

It takes months of live testing

...to refine the model and optimize how it's used. You're not done just because you've built it.

Be careful how you use external help

...(and that includes our Hg team!). These are core skills to have within any business. So, consultants are great to jump-start your journey, or help with specialist topics, but at least one data specialist should be on the ground when you start your first data project.