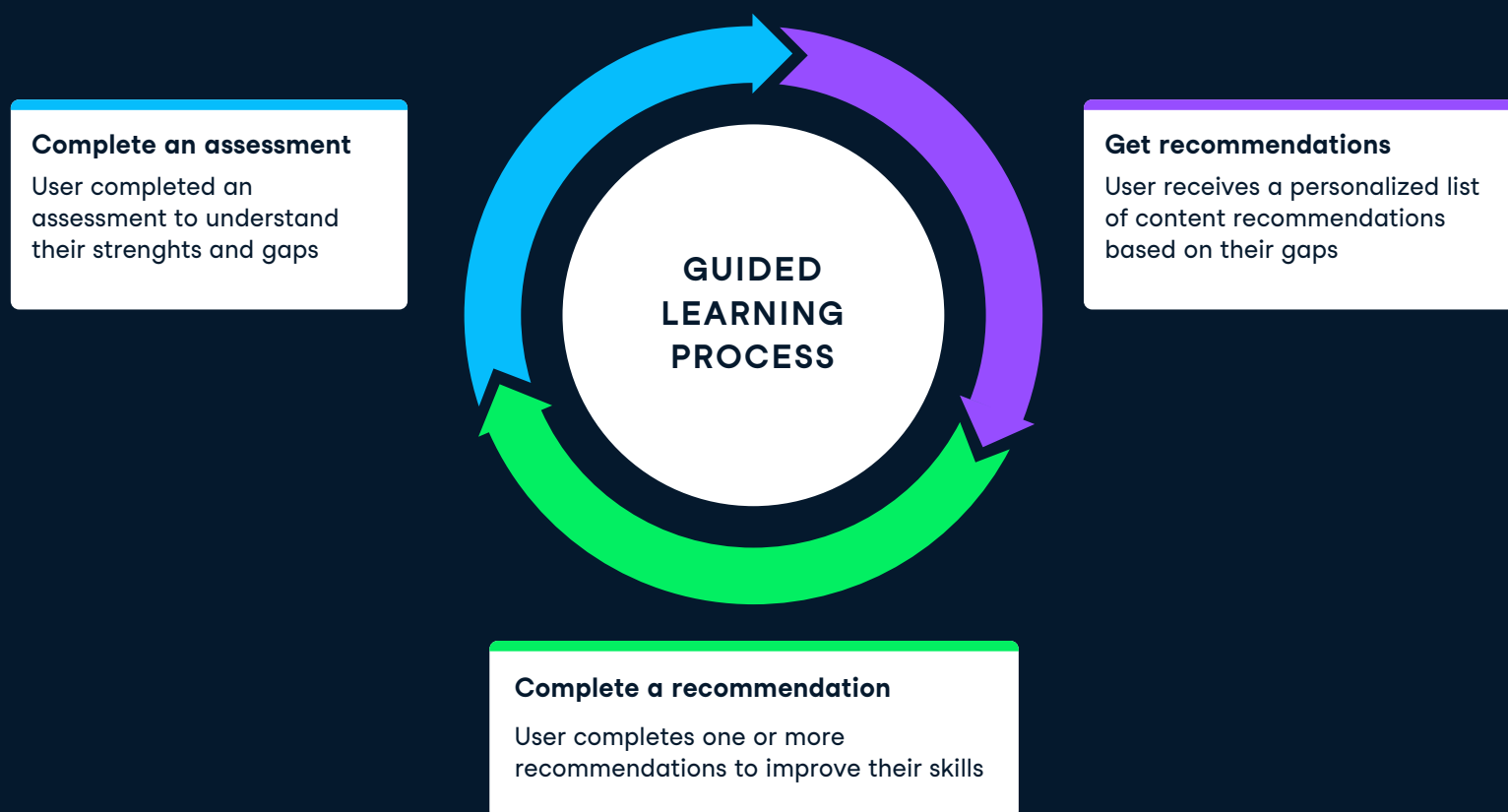**DataCamp Signal™**

An effective and reliable tool for data skill assessment

datacamp

What is DataCamp Signal™?

DataCamp Signal™ is an online, adaptive skill assessment tool that uses state-of-the-art techniques from adaptive testing and psychometrics to accurately measure data skills.

Data skills are essential for any organization to remain competitive in the 21st century, but without a scalable means of assessing skills, it can be difficult to know where you stand. Signal solves this problem by helping you understand how skills are distributed within your organization—across departments, teams, and individuals.

It achieves this through a guided learning process (Figure 1). When a learner completes an assessment, they receive a score (0-200), a percentile (0%-100%), and an associated knowledge level (Beginner, Intermediate, Advanced) based on their performance (Figure 2). By diagnosing the learner's most prominent strengths and skill gaps, Signal is able to provide a list of personalized learning recommendations to address their gaps. As learners improve their skills over time by completing the recommendations, they can retake the assessment to measure their progress.

Complete an assessment
User completed an assessment to understand their strenghts and gaps

Get recommendations
User receives a personalized list of content recommendations based on their gaps

GUIDED LEARNING PROCESS

Complete a recommendation
User completes one or more recommendations to improve their skills

Guided learning consists of three steps: (1) assessing a learner's skills; (2) providing them with next steps; (3) completing at least one recomCended content item and repeating the cycle with another assessment as they make progress over time.
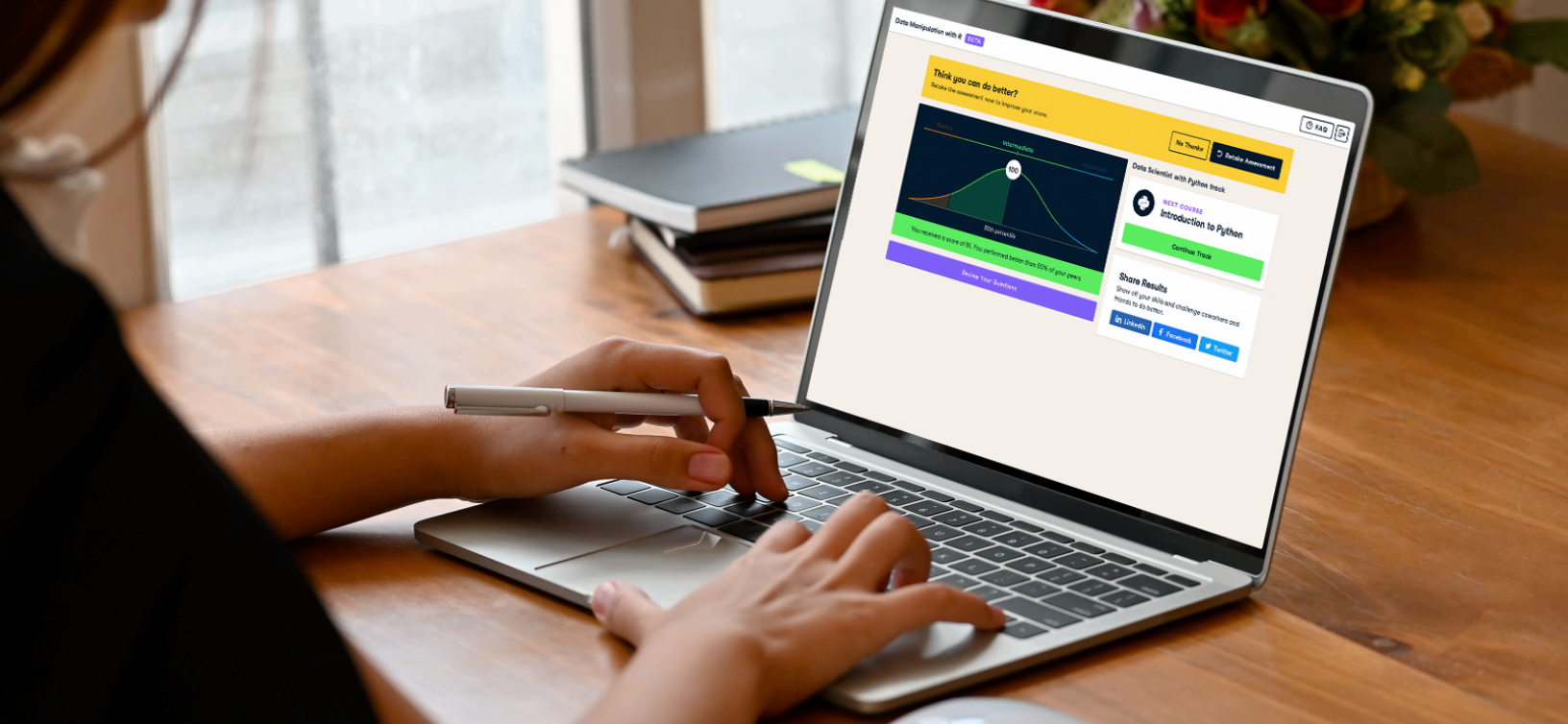
How does it work?

Signal skill assessments are built on a century of research into test theory, with a focus on building statistical models to answer three key questions:

1 Which questions yield the most information about a learner?

2 How can we adaptively select better questions as an assessment progresses?

3 How can we use data from many assessments to improve their overall effectiveness?

You received a score of 100. You performed better than 50% of your peers.
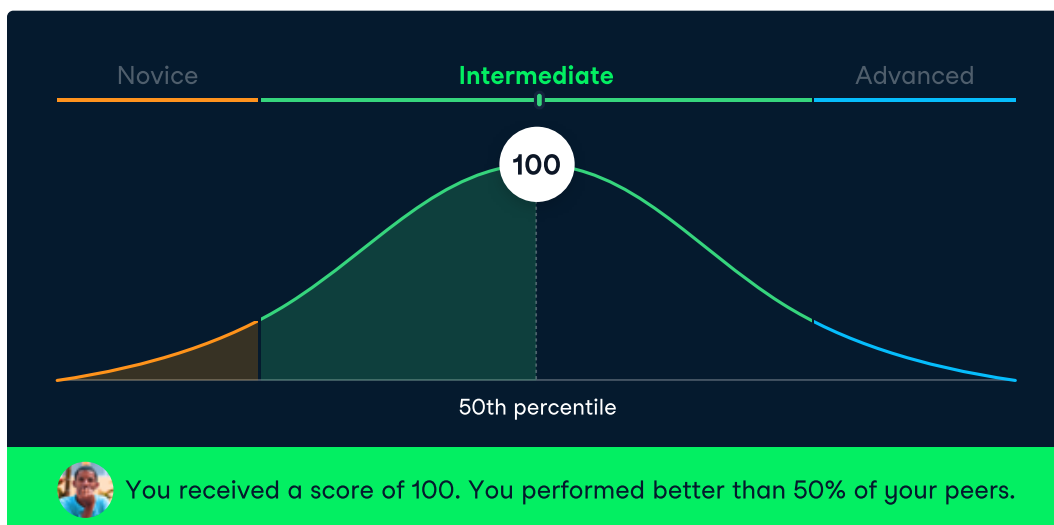
Assessment results include a score (0-200), a percentile (0%-100%), and an associated knowledge level (Beginner, Intermediate, Advanced).

datacamp

# ① Which questions yield the most information about a learner?

Signal uses Item Response Theory (IRT) to estimate how difficult questions are, how effectively they separate beginners from experts, and ultimately to estimate an individual learner's skill level. In IRT models, questions yield the most information when their difficulty matches a learner's skill level.

Unlike most assessment tools, Signal assessments are not entirely composed of multiple choice questions. Instead, our assessments include more hands-on challenges that require learners to write code and work with data to solve real-world data problems, which in turn gives us more information about their actual data skills.

② How can we adaptively select better questions as an assessment progresses?

Signal uses Computerized Adaptive Testing (CAT) to estimate an assessment taker's skill as they progress through an assessment. This allows Signal to select increasingly more informative questions as the assessment progresses. As a result, Signal's adaptive assessments require less than 10 minutes to obtain a high degree of information on a learner's skill level.

The amount of information an assessment captures is often evaluated using a measure called test-retest reliability, which is calculated by administering an assessment to learners twice, then calculating the correlation between learners' initial and second assessment scores (Figure 3). A high test-retest reliability means that an assessment yields consistent results.
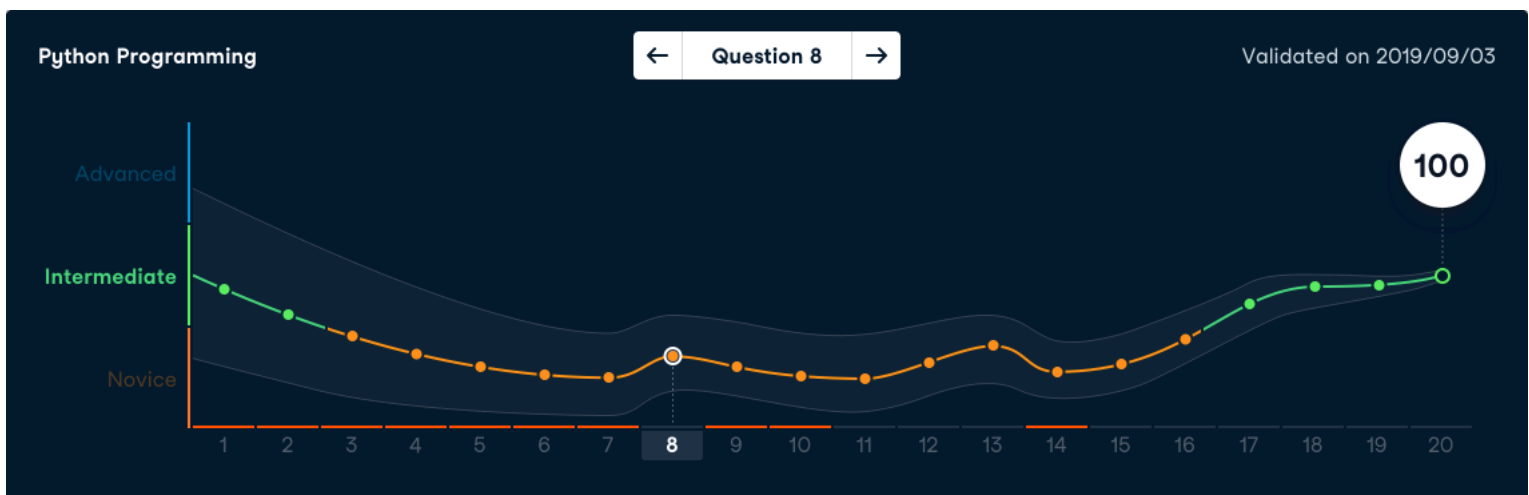
Evolution of a student's score across a test session. On each round, Signal's adaptive test re-estimates the student's skill level. This estimate improves over time, and results in progressively selecting more informative questions.

③ How can we use data from many assessments to improve their overall effectiveness?

Finally, by using assessment data from thousands of DataCamp learners, Signal is able to refine its estimates of how difficult and discerning of skill level questions are. Signal uses this information to flag questions that may not be indicative of skill and to report a learner's standing relative to other learners.

R Programming Assessment: Test-retest Scores
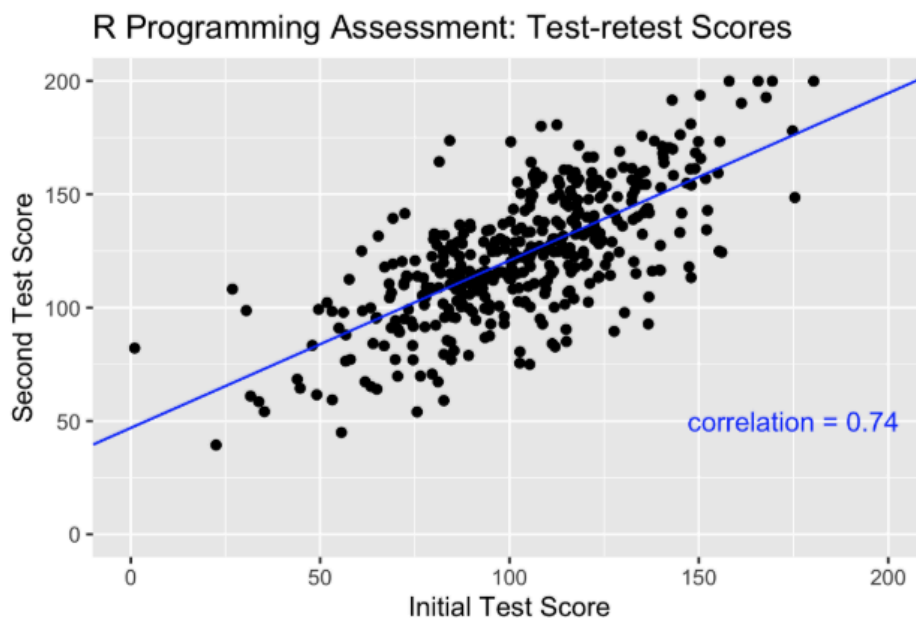
Test-retest scores for the R Programming assessment. The reliability of an assessment measures how stable scores are between learners (e.g., does a learner who receives a high initial score—relative to other learners—also receive a high second score?). As calculated by the correlation between initial and second scores, reliability for this assessment is 0.74. The closer this value is to 1, the more consistent scores are, with a value above .7 often cited as an acceptable standard[1].

[1] **See the following articles:**
Nunnally Jr, J. C. (1970). Introduction to psychological measurement. McGraw-Hill.
Matheson, G. J. (2019). We need to talk about reliability: making better use of test-retest studies for study design and interpretation. PeerJ, 7, e6918.

What skills does it measure?

Signal focuses on core data skills to take the guesswork out of what your teams need to know. For example, we offer assessments that cover all aspects of the data science workflow in the two most popular data science technologies, Python and R, as well as an assessment in SQL:

- Programming
- Importing & Cleaning Data
- Data Manipulation
- Data Visualization (Coming Soon)
- Probability & Statistics (Coming Soon)
- Machine Learning
- Data Analysis

How can it help my organization?

One of the techniques our customers find to be most useful is taking a broad inventory of their in-house data skills using what's called a skills matrix. This comes in many different forms, but generally involves visualizing strengths and skill gaps at either the organization, department, team, or individual level (Figure 4).

Customers looking for a scalable and reliable way to evaluate data skills across their organization use Signal results as an input to their skills matrix. We're currently laying the groundwork to automate this process so this information is available to Enterprise admins on demand.

	Programming	Data Visualization	Probability & Statistics	Machine Learning	...
Data science	INTERMEDIATE	INTERMEDIATE	ADVANCED	ADVANCED	...
Engineering	NOVICE	ADVANCED	INTERMEDIATE	INTERMEDIATE	...
Product	ADVANCED	INTERMEDIATE	NOVICE	NOVICE	...
Sales	ADVANCED	ADVANCED	ADVANCED	INTERMEDIATE	...
Support	INTERMEDIATE	INTERMEDIATE	NOVICE	NOVICE	...

An example skills matrix showing strengths and skill gaps for each member of a team.

Future direction

Since its release in 2019, more than 60,000 DataCamp Signal assessments have been completed. While Signal is already helping organizations around the world improve their data skills, we're also thinking about the future. Continuing to expand our assessment curriculum is a top priority. Some of the areas we're exploring include:

1. Role-based assessments (e.g., data engineer, data analyst)
2. Additional technologies (e.g., Tableau)
3. Advanced topics (e.g., deep learning, geospatial analysis)
4. Conceptual topics (e.g., statistical theory)

Additionally, we'll continue to build more useful tools for Enterprise admins, including:

1. A dedicated dashboard in DataCamp for Enterprise with information on how skills are distributed throughout your organization—at the department, team, and individual levels
2. Ability to benchmark your organization's skills against industry peers
3. Tools for measuring ROI on data skills training (i.e. skills improved vs. dollars spent)

Upskill your company with DataCamp for Business
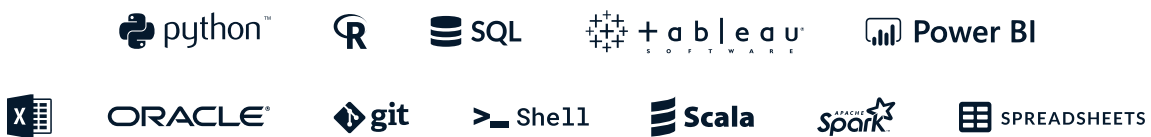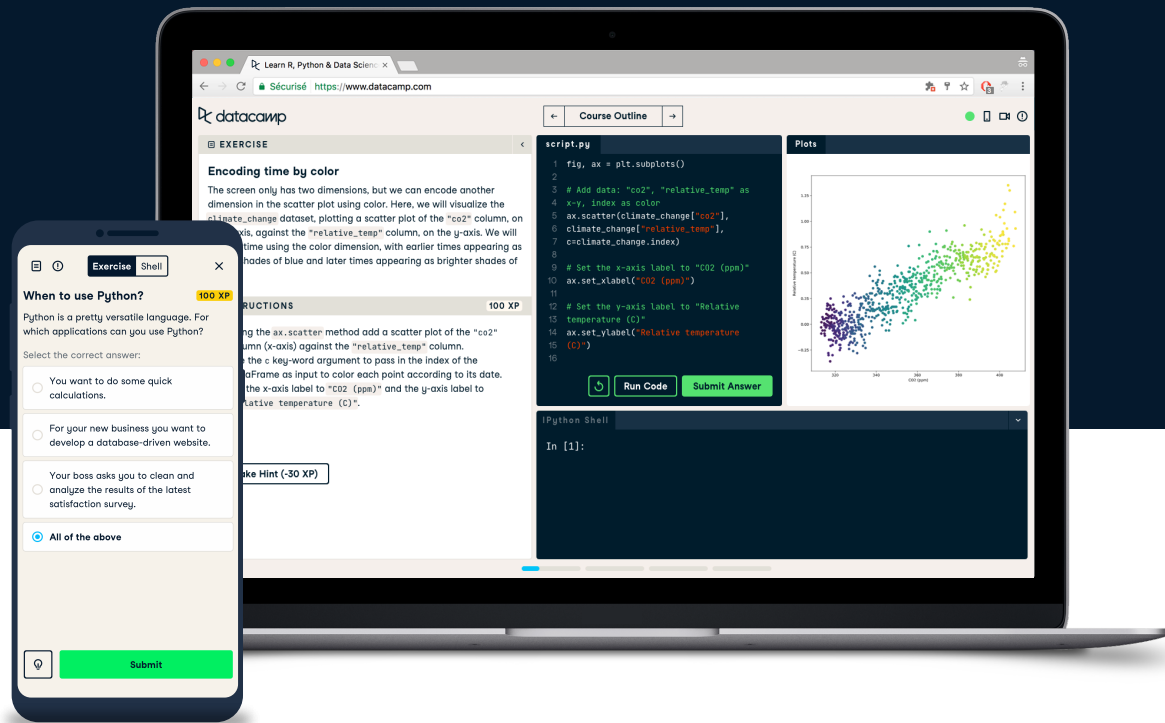
DataCamp for Business provides individual teams or your entire company with an online learning program scaled to your data training needs. It's easy to implement and manage for teams of any size, with advanced analytics and insights, custom learning paths, and seamless SSO and LMS integrations. We are constantly expanding our curriculum to keep up with the latest technology trends and to provide the best learning experience for all skill levels in Python, R, SQL, Scala, Tableau, Oracle, data engineering, and more. And learners stay engaged through hands-on learning, which means our course completion rates significantly exceed those of traditional online courses. We have more than 5 million learners around the world—and we're just getting started.

CLOSE THE SKILLS GAP

Learn More