ScaleCapacity

Data Lakes and Deep Analytics

While top executives have the ability to request data and reports from various departments, it typically takes hours, days, or even weeks to collect, aggregate, and synthesize the information into a holistic view of the business. Sometimes, by the time the data is gathered, it's already out-of-date. With the use of data lakes, however, information can be collected in real time from various sources, making informed business decisions the norm.

Data Democratization

Employees from the top down can make business decisions in real time based on access-controlled, collaborative data.

Scalability

Supply any amount of computing resources to any user or workload, resulting in dynamically changed compute cluster sizes without affecting running queries.

Data Silo Elimination

Eliminate data silos by consolidating multiple data types into a single, unified, and scalable platform.

Schema flexibility

Data lakes allow you to be schema free or define multiple schemas, enabling you to decouple schema from data, which is excellent for analytics.

Versatility

Capture batch and streaming data in a common repository with strong governance, security, and access control.

Advanced analytics

Unlike a data warehouse, a data lake excels at utilizing the availability of large quantities of coherent data along with deep learning algorithms.

sales@scalecapacity.com 1-833-777-2253 scalecapacity.com

ScaleCapacity

Flexible Tiers to Fit Your Needs

Tier 1

- 4 weeks
- Set up data lake with Lake Formation
- Ingest structured and semi-structured data (databases, logs)
- Demonstrate ETL capabilities
- Demonstrate ad hoc querying
- Demonstrate basic visualizations

Tier 2

- 8 weeks
- Set up data lake with Lake-Formation
- Ingest structured, semi-structured data and streaming data (databases, logs, clickstreams/log streams)
- Demonstrate ETL capabilities
- Demonstrate ad hoc querying
- Demonstrate basic visualizations
- Demonstrate basic QA/QC
- Demonstrate basic data governance

Tier 3

- 12+ weeks
- Set up data lake with Lake-Formation
- Ingest structured, semi-structured data, streaming data, and real time data (databases, logs, clickstreams/log streams, time series data)
- Demonstrate ETL capabilities
- Demonstrate ad hoc querying
- Demonstrate basic visualizations
- Demonstrate basic QA/QC
- Demonstrate basic data governance
- Demonstrate advance analytics (ML capabilities)

At ScaleCapacity, we offer years of experience combined with business and engineering expertise to help you become more efficient, agile, and scalable in the AWS Cloud.

To learn more about our process and chat with us about deeper business analytics through data lakes, visit us at scalecapacity.com.

sales@scalecapacity.com 1-833-777-2253 scalecapacity.com