MATILLION

The Four Stages of the **Data Journey**
(and how to get ahead)

Where is your organization on the data journey? Most of us are steadily moving toward the cloud, with most businesses planning to migrate to the cloud or expand their cloud footprint within a few years. But how do you get where you want to go? Take a look at the different stages of the end-to-end journey and learn what it takes to get to the next level.

Rollover/click to learn more

At this stage, your organization may still have your data in on-premises data warehouses. But you are thinking ahead to the cloud.

At this stage, you're committed to the cloud. You've chosen a cloud service provider and probably a cloud data warehouse. But where is your data?

At this stage, you're a relatively old hand at working with data in the cloud. Your goal: Move faster and do more

You aren't just a cloud data veteran; you are leading the pack. How far can you go?

Cloud Discovery

At this stage, your organization may still have your data in on-premises data warehouses. But you are thinking ahead to the cloud. You may even have some data in the cloud, but you're looking to both expand your cloud presence, and see what you can do in a more powerful, scalable environment. In terms of architecture and data, you know what you have, what you want, and definitely what you don't want going forward.

Questions to consider

- What is your desired business outcome?
- What is your budget and anticipated ROI?
- What are your concerns about moving to the cloud?
- What are the different use cases for data in the cloud across your business?
- What are your infrastructure security needs?

Main drivers

- Replace what you have and reduce costs
- Reduce time and resources spent on data management
- Gain faster data performance and more current and accurate data
- Speed up reporting times

Barriers and concerns

- Actually finding your data
- Security and governance
- The size and cost of a cloud data initiative
- Disruption to your business
- Knowledge of the cloud
- Vendor lock-in for your current data stack

Timeline

- Have a cloud presence in 3 months – 1 year

Get to the next stage

- Define your goals and desired business outcomes from a cloud implementation
- Think about a possible cloud architecture. Hybrid? One cloud? Multi-cloud?
- Assess cloud vendors
- Research cloud data warehouses
- Run proofs of concept. Try all the things.
- Consider finding a vendor or consulting partner to help you plan and implement

Cloud Data Migration

At this stage, you're committed to the cloud. You've chosen a cloud service provider and probably a cloud data warehouse. Your next big milestone: loading your data into the cloud. You'll need to tally up your data sources, connect to them, and move that data in an efficient, but considered, manner. You're also thinking about analytics, possibly beginning to set up and design the infrastructure you'll need to run business intelligence in the cloud.

Questions to consider

- How many data sources do you have?
- How are you going to get your data into the cloud?
- Do you want all data in the cloud?
- What do you want to do with your data once it's there?
- Do you need both a cloud data warehouse and a data lake?
- If you want to do analytics, what's your plan for data transformation?
- What are your data security requirements?

Drivers

- Do business faster
- Make better decisions
- Modernize your business
- Speed up reporting
- Improve business analytics
- Centralize data
- Set up a scalable infrastructure
- Build a future-proof platform for analytics

Barriers and concerns

- Employees have on-premises data skills but need cloud data skills
- The size and agility of your business
- The complexity of your migration
- Old technology that may be holding you back
- The cost of your current infrastructure
- How you will manage data during the transition from on-premises to cloud

Timeline

Want to move data into the cloud as soon as possible, but need to do it in a considered way. Need to show ROI and value as soon as possible to gain wider support.

Get to the next stage

- Identify the data sources that you want to migrate or replicate in the cloud
- Identify use cases that have the most business value
- Create a plan for migration
- Determine whether you need a pipeline or a transformation tool
- Choose ELT and Analytics vendors that are built for the cloud
- Do proofs of concept for tools you are considering
- Start small, achieve success, then expand

Cloud Data Maturity

At this stage, you're a relatively old hand at working with data in the cloud. You have an established cloud data warehouse and are probably already doing cloud-based business intelligence and analytics. You most likely have some sort of ETL process, but is it keeping up with the rest of your technology stack? You might be using a traditional ELT product that has been retrofitted for the cloud. Or you may be using a pipeline tool, but are spending a lot of time hand-coding transformations. Now's the time to explore where you can achieve efficiencies and cost savings, and really leverage the power of the cloud.

Questions to consider

- Am I getting the right kind of insight to make informed business decisions?
- Are there data sources I'm not able to leverage?
- What does data transformation look like in our business?
- How much time do we spend hand-coding to prepare data for analytics?
- What can we automate? What should we leave alone?

Drivers

- More and better data
- Speed to insight
- Technology that takes advantage of the speed and power of the cloud
- Compatibility with other products in the technology stack

Barriers and concerns

- After loading data, no way to transform it
- ETL technology designed for an on-premises data warehouse
- Lack of automation features
- Hand-coding is too time-consuming and costly
- Lack of time for experimental data analytics

Timeline

- Achieve faster time to insight and speed up ETL processes within 3-6 months

Get to the next stage

Look at hardening infrastructure with safeguards for:

- High availability
- Disaster recovery
- Git integration
- Versioning

Have a transformation product that

- Automates some transformation steps
- Supports CDC/audit logging
- Eliminates most hand-coding but still allows for flexibility
- Facilitates faster loading of more data
- Plays well with the cloud data warehouse and cloud analytics products

Cloud Data Leader
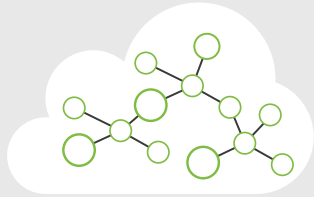
You aren't just a cloud data veteran; you are leading the pack. Your organization is data-driven from the ground up, with data science and analytics informing decisions in many parts of the business. You are regularly transforming data for business intelligence and analytics in the cloud. You may be experimenting with advanced analytics such as machine learning (ML) and artificial intelligence (AI). "Data as an asset" is a driver for business and product development. Data may even be your product. How do you stay at the forefront of data innovation?

Questions to consider

- How do I do feature engineering for AI/ML models?
- How do I load and transform more data for advanced analytics?
- How do I enable data self-service across my business?
- How do I ensure proper security and governance for widespread data access?

Drivers

- A 360-degree view of the customer, informed by as much data as possible
- Prescriptive or predictive analytics (next best action, lookalike modeling)
- Access to analytics-ready data for everyone in the business who needs it
- Product innovation/data as the product

Barriers

- Scalability
- Enough data for AI or ML
- Lack of governance/security policies and data stewardship
- IT resistance to democratizing access to data
- The learning curve and skills for advanced analytics

Timeline

Substantial innovation of business operations and outcomes within a year

Get to the next stage

- Get educated on data best practices for ML
- Ensure that you have the right data and data sets for advanced analytics
- Ensure widespread data literacy across your organization
- Get agreement and buy-in between IT and the business on what's needed for data self-service
- Put proper governance in place for data security and access
- Ensure that you have the right tools for the right roles. For example, some roles may only need data loading, while IT handles centralized data transformation.