



MODERN ANALYTICS ECO-SYSTEM

Data is everywhere and becoming ever more prolific. No wonder the term “Data Lake” resonates on our blue planet where water is essential to all life. Using the water analogy, Excella has developed this simplified view of the modern analytics environment.

In the diagram:

- **The waterfalls represent multiple, different data sources.** Some common data types include financial, customer, social media and purchase data. While a waterfall implies that data is “streaming” into the lake below, data can also enter the lake in batches, much like a sudden rainstorm.
- **The lake represents a data repository storing massive amounts of all types of data — unstructured, semi-structured, and structured—in its native format.** Navigating this data can be difficult because it has not been cleaned, blended together or aggregated. Advanced analysts and data scientists with technical and subject matter expertise have the skills to voyage through this data successfully. Some training and knowledge of the source data, formats and query techniques are needed to be effective – just like a scuba diver needs training and experience to use scuba equipment to explore underwater.
- **The water filtration plant represents the data processing workflow where data is cleaned and transformed into a format that is easier to consume.** In the same way that water is made potable, data is also made to be safe and easier to use. This processing often involves joining data from different sources and applying business rules to standardize the data and make it consistent. The result is clean and clear data – just like the clean, clear water you expect every time you turn on the tap. Note that not all the data in the lake is cleaned – only the amount needed to replenish drinking water supplies. In the data world, the subsets of data that are frequently used are processed to make them easier to use. Data that is rarely requested stays in the lake and can be cleaned on demand if needed.
- **The final step shows how clean water/data is delivered for consumption.** Here, there are no training prerequisites to access the data, you simply turn on the tap.



Excella

2300 Wilson Blvd., Suite 630 Arlington, VA 22201
info@excella.com | excella.com | @excellaco