Data Lakehouse: Combining the best of data lake and data warehouse

Over 50% of organizations use data warehouses and data lakes

Source: Databricks

Data Lakehouse: Combining the best of data lake and data warehouse

To get the most value from your data with data lakehouse architecture write to us at marketing@sigmoidanalytics.com

www.sigmoid.com

Comparison of architecture

<table>
<thead>
<tr>
<th>Data Warehouse</th>
<th>Data Lake</th>
<th>Lakehouse</th>
</tr>
</thead>
<tbody>
<tr>
<td>BI Reports</td>
<td>BI Dashboards &amp; Reports</td>
<td>BI Dashboards &amp; Reports</td>
</tr>
<tr>
<td>Data Marts</td>
<td>Data Marts</td>
<td>Data Science AI and ML</td>
</tr>
<tr>
<td>ETL</td>
<td>Structured, semi-structured, and unstructured data from internal and external data sources</td>
<td>Structured, semi-structured, and unstructured data from internal and external data sources</td>
</tr>
<tr>
<td>External Data</td>
<td>Data Lake</td>
<td>Metadata and Governance Layer</td>
</tr>
<tr>
<td>Internal/Operational Data</td>
<td>Data Lake</td>
<td></td>
</tr>
</tbody>
</table>

Data lakehouse can offer benefits such as:

- Cost-effective data storage
- Simplified schema
- Better compliance
- Data quality delivered through simplified schema
- Reduced data redundancy
- High scalability
- Seamless data governance and data security
- Efficient administration
- Faster interactive query coupled with data democratization
- Reduction of data drift

Source: Adapted from The Modern Cloud Data Platform Rise of the Lakehouse by Alice LaPlante

However, there are issues such as constantly changing datasets, vendor lock-in and advanced analytics limitations.