**DW Report #1 – Websites Search**

|  |  |  |
| --- | --- | --- |
| Rank | Website | Description |
| 1 | [BigQuery.com](https://cloud.google.com/bigquery/) | The BigQuery was part in the data warehousing game in mid-2010. The BigQuery is the serverless data-warehouse that can handle over petabytes of the data. The BigQuery is deployed over cloud. It supports SQL for the querying. Now it also has BI features that handles analysis, reports, sheets and dashboard. Lately they have added ML over the cloud supporting the SQL queries. The other reliable feature is the auto backup and restore so the organization doesn’t have to worry about the loss of data. |
| 2 | [Cloudera.com](https://www.cloudera.com/) | Cloudera has raised its name in the field of Data Warehousing and the Machine learning services. They have established their own operating system where organization can load their database which is further saved to the cloud and carry out the Machine Learning and Artificial Intelligence. Cloudera has outgrown in each sector like Finance, Manufacturing, Tech, Retail, Public, education etc. Despite being on the cloud the services are smooth and fast execution over Hadoop. They have information assured with respect to database. |
| 3 | [Pentaho.com](https://www.hitachivantara.com/en-us/products/data-management-analytics/pentaho-platform.html) | Pentaho is one of the best well know tools for the DW and analytics. Pentaho is best described as the Business Intelligence software that supports the DW task like Data Integration, OLAP, Reporting, Graph Analysis, Extract, Transform, Load. There are three major service departments in Pentaho which are known as Data Management and Analytics, Infrastructure and Cloud services that can handle any size of the data. It has well established GUI that handles user requirements of the visualizations and dashboard for analysis. |
| 4 | [Vertica.com](https://www.vertica.com/) | Vertica is uprising BI software over the cloud. It is deployed over Hadoop for faster execution of the SQL queries. The primary focus of the vertica is Analytics, Machine Learning solution for prediction and analysis and the data security over the cloud. The eye catching feature that is unique: Time Series Analytics and Iot Analytics. Every solution is based on the SQL optimization techniques. |
| 5 | [Panoply.io](https://panoply.io/data-warehouse-guide/) | The websites start with the basic introduction of Data Warehouse concepts, architecture, and solutions. Further it defines the modelling and schemas. The next it involves the ETL process from the data mart. Where the data is processes step by step. The primary explanation holds the security and backup for the data warehouse. The data is integrated with any virtualization tools. The collection of things are easy to execute with dashboard and querying in the SQL workbench. |
| 6 | [Intellipaat.com](https://intellipaat.com/blog/tutorial/data-warehouse-tutorial/) | This website starts with the best tutorial video series stating with all the database technologies out there. After the brief explanation of the technologies, it defines the differences between the OLTP and OLAP and describes the DW architecture. Further it characterizes on the star and snowflake schema. As each technology has life cycle, the DW is dependent on the Business Intelligence Life Cycle. Lastly, it gives good explanation about the ETL process and tools that are helpful to carry out the process. |
| 7 | [SAP HANA](https://www.sap.com/products/hana.html) | SAP Hana is an in memory, column oriented and RDBMS deployed over Linux operating system. The primary objective defines the storage of the column-oriented data for easier execution. It is basically the combination of OLTP and OLAP. They have established analytics system, graph capabilities and ETL execution over the data to data warehouse. |
| 8 | [Databricks.com](https://databricks.com/product/unified-data-analytics-platform) | Databricks is known as the Unified Data Analytics Platform. The product is primarily designed for Data Scientist, Analyst, ML Engineers. The data services have reliable data lakes and efficient data pipelines. It has intuitive visualizations and dashboards with BI integration like Tableau, Qlik and Power BI. It supports all type of programming languages like Python, R, SQL, Scala and Java. ML libraries like Conda, Tensorflow and Pytorch. |
| 9 | [Abinitio.com](https://www.abinitio.com/en/) | It is multipurpose tool for data warehousing. It handles data analysis, manipulation and ETL processes. It has good GUI but slow query processing. It is deployed over UNIX system. |
| 10 | [Snowflake.com](https://resources.snowflake.com/data-warehousing) | This website holds really good information about the latest trends and changes in the data-warehousing world. The primary information provided in the website is webinars, conferences and reports on Data warehousing. |