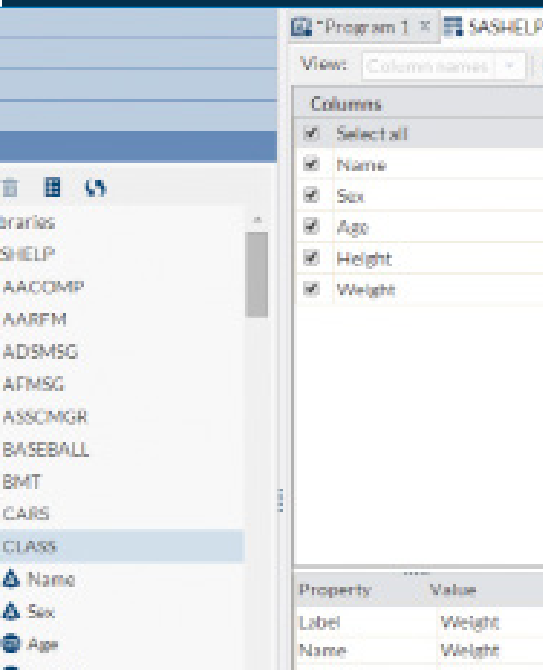


SAS® In-Memory Statistics for Hadoop

One environment for the entire analytical life cycle



What does SAS® In-Memory Statistics for Hadoop do?

SAS® In-Memory Statistics for Hadoop provides a single interactive programming environment for the entire analytical life cycle. It enables you to perform data management, variable transformations, exploratory analysis, statistical modeling using machine-learning techniques, integrated modeling comparison and scoring - all inside the Hadoop environment.

Why is SAS® In-Memory Statistics for Hadoop important?

It's a fast, powerful and customizable in-memory programming language that lets multiple users concurrently and interactively analyze large amounts of data stored in Hadoop. This results in greater analyst productivity and turn-on-a-dime creativity.

For whom is SAS® In-Memory Statistics for Hadoop designed?

It's designed for experienced statisticians, data miners, data scientists, engineers, researchers, biostatisticians and scientists who need to analyze large and complex data in Hadoop.

Increasing numbers of organizations are incorporating Hadoop into their data storage mix. Unfortunately, it's been difficult to perform the fast analytical processing needed to deliver instantaneous insights from this data. Organizations are also finding that multiple software products and rare data scientist talent are needed to manage each task in the analytical life cycle. Getting insights out of Hadoop in a timely manner requires a different approach.

SAS In-Memory Statistics for Hadoop gives multiple users the ability to simultaneously analyze huge volumes of data. This, combined with very powerful analytic techniques, provides an unprecedented way to tap into Hadoop data to quickly derive insights for fact-based analytical decisions. Dramatically reducing model development time means more models can be put in action sooner.

This solution also minimizes processing time and scales to meet enterprise needs. So you can handle more data, more users and even more complex issues.

Benefits

- **Delve deep into Hadoop for fast, accurate insights.** Apply proven state-of-the-art statistical algorithms and machine-learning techniques to find the best answers. You can explore and use multiple analytic approaches to reveal insights and make fact-based decisions.
- **Increase productivity for your data scientists.** Multiple users can concurrently and interactively analyze big data in Hadoop using the fast, in-memory analytical programming language. Prepare, manipulate, transform, explore, model, access and score data - all within Hadoop.
- **Take advantage of an end-to-end, scalable environment.** Until now, statisticians and data scientists had to piece together different programming languages or products to manage the variety of analytical lifecycle tasks in Hadoop. And when it came time to operationalize models, the software couldn't scale. No more. From data manipulation and exploration to model building and deployment, our solution is proven, tested and accurate - and can scale to your production environment.
- **Avoid unnecessary, multiple passes through the data.** Our in-memory infrastructure running on top of Hadoop eliminates costly data movement and persists data in-memory for the entire analytic session. This significantly reduces data latency and provides rapid analysis at lightning-fast speeds. Extensive analytical data manipulation capabilities make it easy for you to quickly prep data before analysis, which can typically take up to 80 percent of an analytic project's implementation time.

Overview

Personal sales collections

SAS In-Memory Statistics for Hadoop is a single, interactive programming environment for the entire analytical life cycle. Multiple users can simultaneously manage data, transform variables, perform exploratory analysis, build and compare models, and score models with virtually no limits on the size of data stored in Hadoop. And they are able to move quickly through each phase of the analytical life cycle. It uses

in-memory analytic processing so results are produced nearly instantaneously.

State-of-the-art statistical and machine-learning techniques, including regression, clustering, decision trees, random decision forests, etc., help identify new patterns, trends and relationships in complex big data stored in Hadoop. The interactive language and built-in actions enable you to explore multiple modeling approaches and reach optimal solutions faster than other

programming languages like MapReduce and R. This in turn drives better strategies for improving business processes.

Interactive programming

Move through the entire analytical life cycle in Hadoop with an extremely fast, multi-user environment. Users experience instant response times for complex problems on large amounts of data. The large level of interactivity means they have the flexibility to go on different paths and ask difficult what-if questions that can generate higher returns.

In-memory analytical processing

Get fast analytic computations that are optimized for multiple passes across distributed clusters. This dramatically shortens model development time so more models can be developed in less time and put in action sooner. SAS In-Memory Statistics for Hadoop uses a "read once, analyze repetitively" in-memory computational paradigm for faster interactive ad hoc analysis to prepare the data for modeling.

Data persisted in-memory

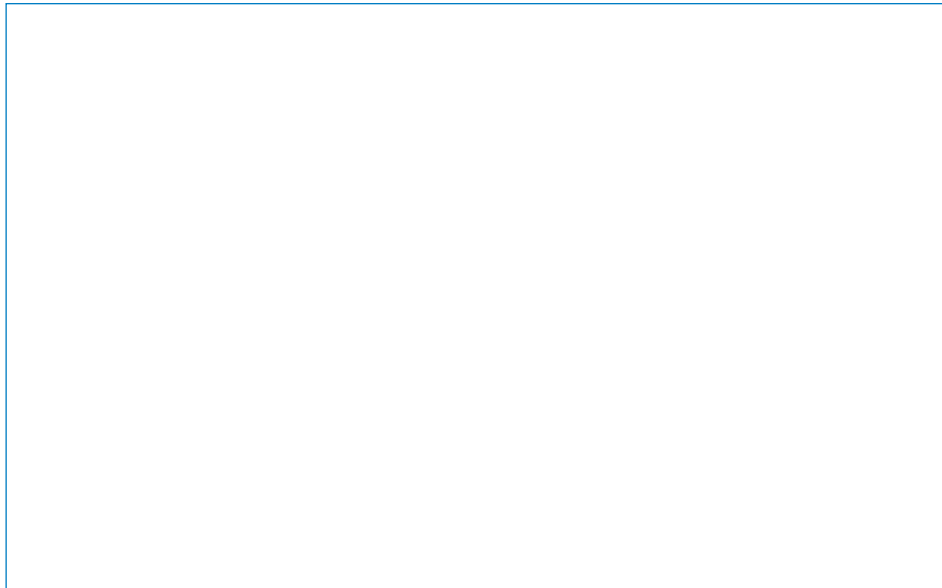
This solution requires loading the data in parallel only once to support truly interactive discovery and problem solving using a full spectrum of analytical methods. Gain speed and reduce latency because data is held in-memory.

Analytical data manipulation

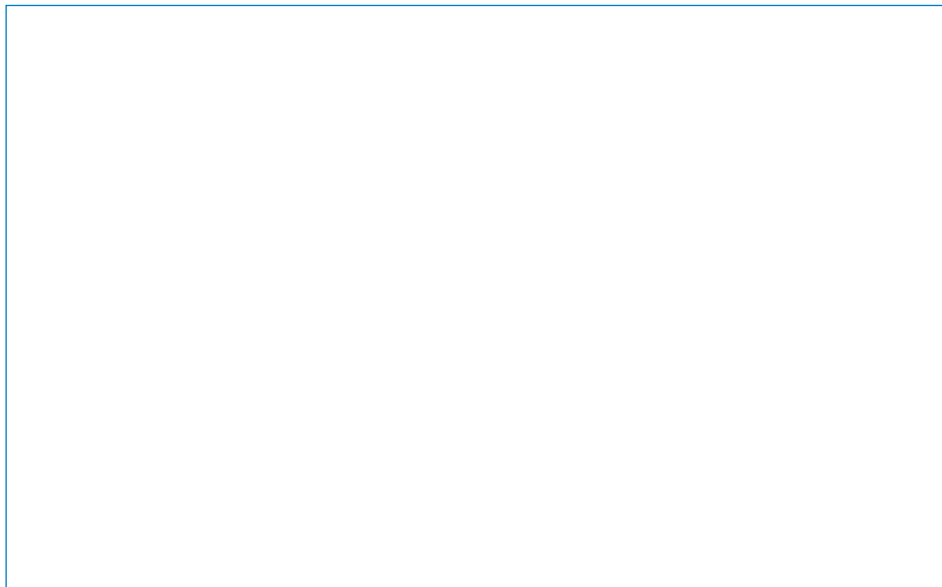
Prepare data for modeling with data integration, variable transformations and creation, and exploratory analysis. Users can explore and prepare their data more efficiently and are able to shorten the time for model development and deployment.

Statistical algorithms and machine-learning techniques

Uncover patterns and trends faster than ever with a huge breadth and depth of analytical techniques. This wide range of analytical algorithms and processes can reveal patterns, anomalies, key variables and



Screenshot caption



Screenshot caption

Key Features (continued)

To learn more about SAS text text text text
text text text text text text text text text
text text text text text text text text text,
please visit
sas.com/-----.

Subhead

- bullet
- bullet
- bullet
- bullet
- bullet
- bullet
- bullet
- bullet - last

Subhead

- bullet
- bullet
- bullet
- bullet
- bullet
- bullet
- bullet
- bullet - last

Subhead

- bullet
- bullet
- bullet
- bullet
- bullet
- bullet
- bullet
- bullet - last.

To contact your local SAS office, please visit: sas.com/offices

