

> Analyse large datasets and improve performance

PASW® Statistics* is a suite of products for statistical analysis and data management. It enables analysts and researchers to access, prepare, manage, analyse and report on data in order to solve business and research problems in the context of the analytical process. PASW Statistics Server* features client/server architecture, which is the combination of two powerful products:

- PASW Statistics client software for in-depth data exploration, analytical reporting and modeling
- PASW Statistics Server software, which scales up from handling the analytical jobs of a single department to handling jobs for hundreds and even thousands of users across an organisation.

This combination of the strength of world-class analytical tools and techniques with the flexibility and speed of server functionality delivers a powerful solution that can support better decision making throughout the enterprise.

Delivering better performance for organisations with distributed offices

When PASW Statistics clients are configured to connect to PASW Statistics Server, all the analysis is done on the server machine, which is typically co-located with the data in a central data centre. This eliminates the need to transfer data over the network, which saves time and improves analysts' productivity. The overall performance of applications over the network also improves due to more efficient use of the available bandwidth. Because it can take a significant amount of time to transfer data among offices, PASW Statistics Server is commonly used by organisations that conduct analysis in multiple locations.

Increasing productivity

Because analysis is performed on the server, PASW Statistics desktop software is available for other tasks. So analysts can work in their application of choice without any disruption while waiting for a long-running job to complete and can initiate multiple jobs at the same time. All of this can significantly improve analysts' efficiency and productivity. In addition, administrative controls in PASW Statistics server support assigning priorities so that resources can be reserved for high-priority users.

Achieving faster performance

PASW Statistics Server is designed to ensure optimal performance when working with large datasets having multiple predictors. For a detailed list of the features in PASW Statistics Server that improve performance, see 'Faster performance' on page 3.

Automating repeated tasks

PASW Statistics Server provides a batch facility (StatisticsB) that makes it easy to automate and schedule repeated tasks – for example, reports that need to be generated at regular intervals or data preparation involving large datasets – that can be performed during off-peak hours. By automating such time-consuming tasks, your organisation obtains information efficiently while making better use of your analysts' and administrators' time. Additional automation capabilities are available through PASW® Collaboration and Deployment Services* (see page 2).

* PASW Statistics, PASW Statistics Server, PASW Modeler, and PASW Collaboration and Deployment Services, formerly known as SPSS Statistics, SPSS Statistics Server, Clementine®, and SPSS Predictive Enterprise Services™, are part of SPSS Inc.'s Predictive Analytics Software portfolio.

Improving security and standardisation

Analysis is often performed on data that are sensitive, or are part of the organisation's intellectual property. In several industries, government regulations protect the confidentiality of privileged data. With PASW Statistics Server, data are typically stored in a central location rather than on a local computer where they could be compromised more easily. And, because data are stored in a central location, standards can be enforced to ensure that all analysts are using the latest versions of a syntax or data file.

Making better use of existing hardware

PASW Statistics Server is supported on a wide variety of operating systems including Microsoft® Windows® 2003 and 2008, Sun™ Solaris™, HP/UX, IBM® AIX® and Linux®.

Securing access for analysts on the go

Analysts working remotely or traveling may need to analyse data within your firewall and, naturally it is important to be sure that this does not compromise the confidentiality of the underlying data or performance. PASW Statistics Server supports Secure Sockets Layer (SSL) to encrypt the communication between a client and a server, and tunneling protocols and NAT are also supported. As well, remote users experience faster performance because the data do not have to be downloaded to end users' machines.

Leveraging your investment on PASW Statistics Server

To gain the greatest value from analytical initiatives, it is important to be able to efficiently leverage the intellectual capital within your organisation. When PASW Statistics Server is integrated with PASW® Collaboration and Deployment Services, all of your analytical assets are versioned and stored in a single repository, enabling analysts to collaborate on projects and establish and communicate best practices.

PASW Collaboration and Deployment Services also provides sophisticated automation, scheduling and deployment capabilities. With just a few clicks, an analyst can publish the results to a portal or custom dialogue that can be directly accessed by business users. Additionally, the entire process of running jobs and publishing results can be automated.

Architected for enhanced scalability

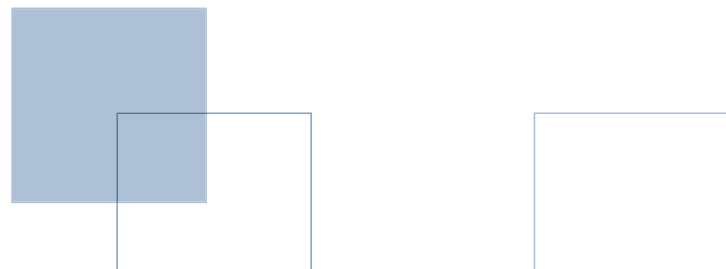
When integrated with PASW Collaboration and Deployment Services, PASW Statistics Server can be clustered to provide network load balancing and failover protection. This ensures that PASW Statistics Server can seamlessly scale from meeting the analytic needs of a single department to meeting those of hundreds and even thousands of users across the enterprise.

Using advanced analytics and scoring

PASW Statistics Server provides the Naïve Bayes algorithm and the Select Predictor procedure, both of which are designed to make it easier to build models efficiently and accurately with wide datasets (those with numerous predictors).

Also, PASW Statistics Server features a scoring engine that can be used to score new data. Users can open multiple XML models created in PASW Statistics, PASW® Modeler* or AnswerTree® and score new data with these models.

For more information about PASW Statistics Server, please visit www.spss.com/uk/statistics.



Features

Client/server architecture

- Reduce network traffic because data reside on the server and are not brought down to users' machines for analysis
- Analyse massive datasets faster using server-grade hardware
- Increase the speed of your analyses by letting your server do the heavy computation work, freeing your desktop for other activity

Faster performance

- No limit on the number of CPUs/Cores that an analytical procedure can use
- No limit on the number of threads for multi-threaded procedures
- Operations such as sort and aggregate can be performed faster by being pushed back to the database
- Support for data compression minimises time-consuming disk I/O and fits larger datasets into RAM
- Client/server architecture mitigates the need to transfer large files between offices
- Support for 64-bit hardware improves performance when working with large datasets
- Because analyses and data access are performed on the server, users' desktops are freed for other tasks
- Users can run multiple jobs from a single desktop without compromising performance

Advanced Analytics

- Naïve Bayes algorithm: predict classification of cases by treating each variable as independent or equal (available only with the server version of PASW Statistics Base)
- Predictor Selection algorithm: filter large amounts of irrelevant data to obtain only features relevant for modeling (available only with the server version of PASW Statistics Base)

Copy-free data access in SQL DBMS

- Sort and aggregate data inside the database before retrieval for analysis
- Easily read in data tables with the PASW Statistics Server Data Access Pack
- Read data stored in PASW Statistics (SAV) file format

Ability to launch multiple sessions

- Run multiple sessions of PASW Statistics simultaneously on the same desktop
- Access many datasets simultaneously by running multiple sessions from a single desktop of PASW Statistics client

Security

- Work efficiently within your vendor's security framework
 - Require password protection when clients access PASW Statistics Server
 - Set security levels and require passwords to access data sources
- Supports Open SSL
- Enable remote users to analyse data from off-site locations while keeping the data and PASW Statistics Server safely behind a firewall. Modern internationalised communication protocols are included with PASW Statistics Server to enable users to connect to PASW Statistics Server using:
 - Point-to-Point Tunneling Protocol (PPTP)
 - Level 2 Tunneling Protocol (L2TP)
 - Network Address Translation (NAT)
- Secure sensitive data by preventing analysts from seeing a data view

Client/server compatibility

- Move client freely between server and local mode
- Work in a multi-platform environment (for example, use a Windows client with a UNIX® server)

- Work in multiple locations (for example, Japanese and French PASW Statistics clients can be attached to a single English version of PASW Statistics Server)

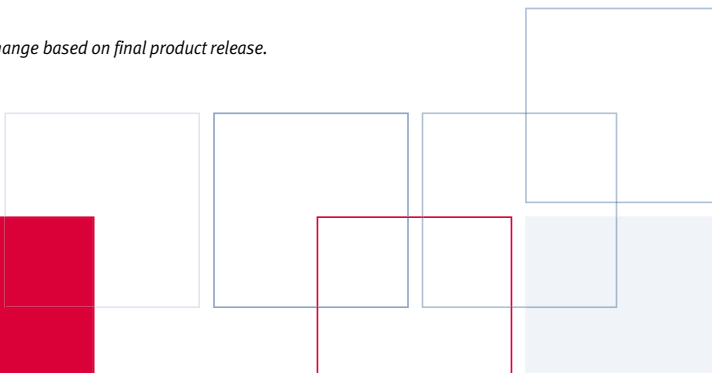
StatisticsB

- Automate production of PASW Statistics data preparation and statistical reports through command syntax files in a UNIX script or Windows batch files without requiring an active and connected PASW Statistics client
- Use the following output formats: Text, HTML, XML, SAV, SPV
- Save prepared data to the PASW Statistics (SAV) file format
- Manipulate the output using OMS commands

Administrator controls

- Work with a utility that assists the PASW Statistics administrator in monitoring and managing PASW Statistics usage on the server
 - Start/stop user sessions
 - Start/stop server processes
 - Log events
 - Assign priorities to individual users
 - Assign a unique disk to individual users for temporary files
 - Disconnect users
- Use a single Administrative Utility for working with PASW Statistics, PASW Modeler and PASW Collaboration and Deployment Services. With it, you can administer any combination of these three products, eliminating the need for separate administration tools.

Features subject to change based on final product release.



Maintain and increase functionality beyond that of the desktop version

- Leverage the same functionality as the client-only product: Graphical user interface (GUI) or syntax-driven capabilities are available in all areas, including statistics, graphics, OLAP report cubes, and data transformations
- Use scoring and analytical capabilities to work with wide datasets.
 - Open multiple XML models from PASW Statistics, PASW Modeler or AnswerTree and score new data using a scoring engine
 - Filter irrelevant data to obtain only features relevant for modeling by using the Predictor Selection algorithm. This algorithm supports categorical and continuous independent and dependent variables, and accepts very large sets of predictors (up to 100,000).

- Predict classification of cases by treating each variable as independent and equal with the Naïve Bayes algorithm. Predictors can be continuous or categorical, and the algorithm is best used when you have fewer than 200 predictors.
- Improve performance
 - Supports a client/server architecture where all the analysis and data access is done on the server
 - No limit on the number of CPUs/cores a procedure can access
 - No limit on the number of threads associated with multithreaded procedures
 - Support for data compression minimises time-consuming disk I/O by fitting larger datasets into RAM
 - Provides the ability to control in-database sort and aggregate functions
 - Large temporary files, which are often associated with sorting and aggregation, can be read and written more quickly by striping them over multiple disks (based on the administrator's settings).

System requirements

PASW Statistics Server 18

- Operating system: Windows Server 2008 or Windows Server 2003 (32-bit or 64-bit); Sun Solaris 9 or 10 (SPARC 64-bit machine); IBM AIX 5.3 or 6.1, 64-bit only (PowerPC); HP/UX 11i v3, 64-bit (Itanium); Red Hat Enterprise Linux 4x or 5 (32-bit and 64-bit), Advanced Platform (32-bit and 64-bit); or Advanced Server 4x (64-bit)
- Hardware:
 - Minimum CPU: two CPUs recommended, running at 1GHz or higher
 - Memory: 4GB RAM recommended
 - Minimum free drive space: 500MB for installation. Additional space is required to run the program (for temporary files)
 - Other: a network adapter running the TCP/IP protocol

Features subject to change based on final product release.

□ Symbol indicates a new feature.



To learn more, please visit www.spss.com.
For SPSS Inc. office locations and telephone numbers, go to www.spss.com/worldwide.

SPSS is a registered trademark and the other SPSS Inc. products named are trademarks of SPSS Inc. All other names are trademarks of their respective owners. © 2009 SPSS Inc. All rights reserved.
SSR18SPC-0709-UK