

Big Data Analytics

In-Database Analytics Drivers



The use of Big Data is becoming a leading means of competition and growth for companies. As the pace of business quickens, and competition stiffens, the ability to generate better and faster decisions will be what differentiates the market leaders from the laggards.

Companies that compete with analytics have already made significant investments in predictive analytics, and have realized significant business benefits including: streamlined operations; accelerated revenues; improved profitability; and better risk management.

In-database analytics takes predictive modeling and analysis to the next level by providing an efficient and cost-effective way for organizations to enhance their data exploration capabilities through an optimized analytical processing environment—within the database—to accelerate decision making and identify opportunities they can act on.

Data integrity, analytical performance, scalability and enhanced data security are all enabled by embedding analytical capabilities and processing logic completely within the database itself—eliminating the excessive time and labor needed to extract data and load it into a separate analytical environment for building and deploying predictive models.

With in-database analytics, issues such as inefficient data movement processes, data duplication and data security are addressed, allowing organizations to quickly and effectively apply analytics against massive amounts of customer and business data—Big Data.

Benefits of In-Database Analytics

Using the Angoss In-Database Analytics driver, provided as an optional module, analysts can connect directly to a parallelized and optimized enterprise data warehouse to perform data preparation, data profiling, Decision Tree analysis, advanced modeling and strategy design.

The many improvements enabled by in-database analytics include:

Performance

- Leverage your enterprise data warehouse investment and massive parallel processing by running data mining algorithms on managed databases deployed on your most powerful servers
- Build models using extremely large data sets created within your enterprise data warehouse

Integrity

- Maintain a central data repository for analytics, promoting standardization for modelers and BI users
- Reduce delay between data acquisition, preparation and analysis by keeping everything within a single environment

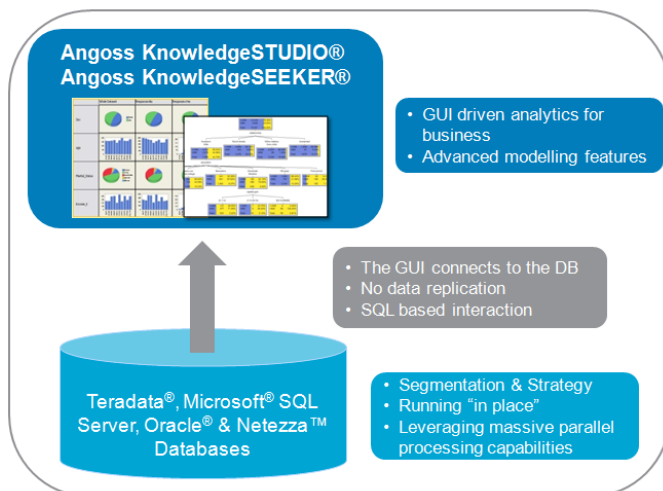
Security

- Control data access security to prevent data loss and misuse

Angoss In-Database Analytics Driver

Angoss' in-database analytics driver is available for Angoss' KnowledgeSEEKER® and KnowledgeSTUDIO® software products. It gives analysts the same robust analytics experience as working with data locally in a separate analytical processing environment.

The driver supports Teradata®, Microsoft® SQL Server, Oracle® and Netezza™ databases, with native ODBC drivers required for each respective database. In-database analytics with Angoss software products are available in desktop and client-server mode, with concurrent or named-user licensing.



Supported In-Database Analytics Features

Angoss' in-database analytics driver supports all KnowledgeSEEKER functionality and core KnowledgeSTUDIO features including data preparation, data profiling, Decision Trees and model analyzer—as well as Strategy Trees. It also supports in-database execution of Decision Tree models and strategies.

Data Preparation

KnowledgeSEEKER and KnowledgeSTUDIO include data preparation capabilities enabled in the in-database mode that allows users to easily manipulate and transform data to prepare it for modeling within the database. Data preparation wizards increase productivity and efficiency while eliminating the need to write code.

Dataset-level and row-level manipulations can be performed using the Data menu wizards: join; append and aggregate datasets; and remove duplicate records.

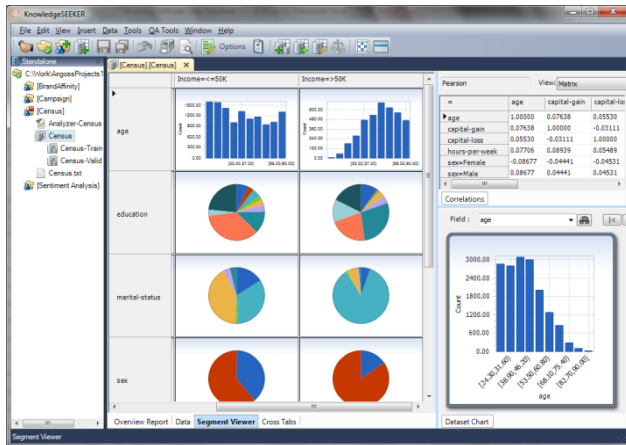
New variables can be defined using standard ANSI SQL expressions through a data editor wizard. Auto-generation of expressions using helpers is available for common tasks, such as:

- Binning of values and categories
- Logarithmic transformations
- Optimal binning using Weight of Evidence (WOE) aided by decision tree interface
- Substitution of missing values
- Calculation of weights for balanced samples
- Generation of dummy variables

Data Profiling: Understanding and Visualizing Your Data

Through an extremely powerful and easy-to-use graphical interface, in-database analytics from Angoss allows business users and analysts to examine data quickly and discern and interpret patterns, trends and business rules within the database.

All standard data profiling features of KnowledgeSEEKER and KnowledgeSTUDIO are available using the in-database analytics mode.



Segmentation and Modeling with Decision Tree Analysis

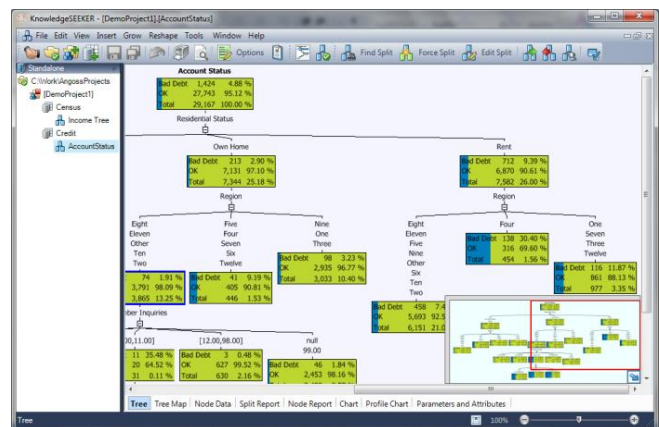
Decision Tree analysis allows users to segment and classify data, as well as quickly determine the best predictors of selected dependent variables. From this, users can implement business rules for implementation in a decision-support environment or plan further modeling and analysis processes.

The interactive Decision Trees in KnowledgeSEEKER and KnowledgeSTUDIO provide intuitive, color-coded visualizations of dependencies between variables. Trees can be grown manually by creating splits with custom settings, auto-grown using segmentation algorithms or grown using a combination of both methods.

With Angoss' in-database analytics driver, Decision Trees allow users to:

- Find, force, erase, copy and paste splits
- View or export node data
- View split statistics and the rank of independent variables, with respect to their predictive strength in a split report
- Copy and paste trees into Microsoft® Office applications
- Evaluate and compare the predictive performance of trees with model analyzer charts (e.g., lift and ROC charts)

- Automatically generate code for tree rules in multiple formats (e.g., SQL, SAS, SPSS) for deployment in other analytic environments
- Deploy a Decision Tree model directly in the database using in-database scoring



Strategy Authoring and Design with Strategy Trees

Angoss' Strategy Trees combine the usability of Decision Trees with a suite of enhancements for strategy developers, providing users with an innovative and unique toolset for strategy design, authoring and validation workflow.

For richer segmentation, Strategy Trees allow for the use of multiple target variables and provide more feedback and control with key performance metric calculations for each node or segment. Automatic code generation for deployment eliminates manual coding errors.

In-Database Model Execution

By providing in-database model execution in addition to model creation, Angoss enables complete analytics workflow within the database - from model development to deployment. Models can still be exported to other analytical environments using automatic code generation, direct in-database execution of Decision Tree models and Strategy Trees to save time and eliminate the need for other deployment environments.

About Angoss Software

As a global leader in predictive analytics, Angoss helps businesses increase sales and profitability, and reduce risk. Angoss helps businesses discover valuable insight and intelligence from their data while providing clear and detailed recommendations on the best and most profitable opportunities to pursue to improve sales, marketing and risk performance.

Our suite of desktop, client-server and in-database software products and Cloud solutions make predictive analytics accessible and easy to use for technical and business users. Many of the world's leading organizations use Angoss software products and solutions to grow revenue, increase sales productivity and improve marketing effectiveness while reducing risk and cost.

Corporate Headquarters

111 George Street, Suite 200
Toronto, Ontario M5A 2N4
Canada
Tel: 416-593-1122
Fax: 416-593-5077

European Headquarters

Surrey Technology Centre
40 Occam Road
The Surrey Research Park
Guildford, Surrey GU2 7YG
Tel: +44 (0) 1483-452-303
Fax: +44 (0) 1483-453-303