

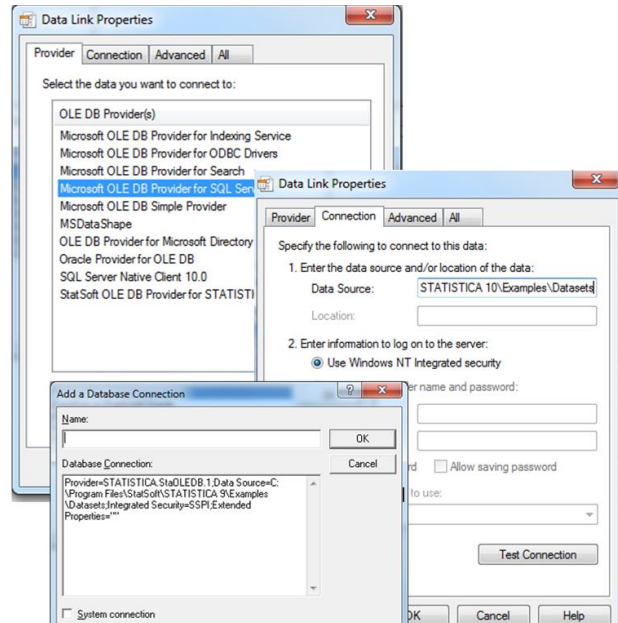
Připojení na datové zdroje

Software Statistica podporuje import dat prakticky ze všech zdrojů (např. text, SAS, Excel atd.), včetně databází (SQL, Oracle, Teradata, přes OLEDB [ODBC]).

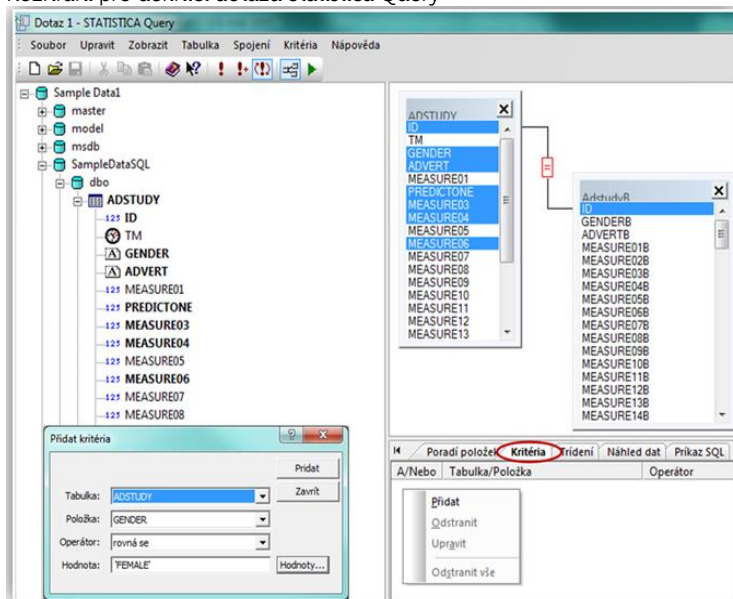
Ukázka rozhraní pro připojení na databázi

✓ Připojení na databázi

Statistica umožňuje uživatelům připojení do různých databází nebo datových skladů. Velkou předností software Statistica je to, že nepotřebuje instalaci oddělené databáze, nepotřebuje tedy žádnou vlastní databázi a je snadno integrovatelná prostřednictvím standardních připojovacích prvků, do stávajících datových struktur společností přes OLEDB, ODBC, ADO. Dotazy mohou být vytvořeny ve velmi přívětivém grafickém uživatelském rozhraní Statistica Query a Statistica Advance Query Builder (nové grafické rozhraní, které vylepšuje původní Query), viz následující obrázek níže.

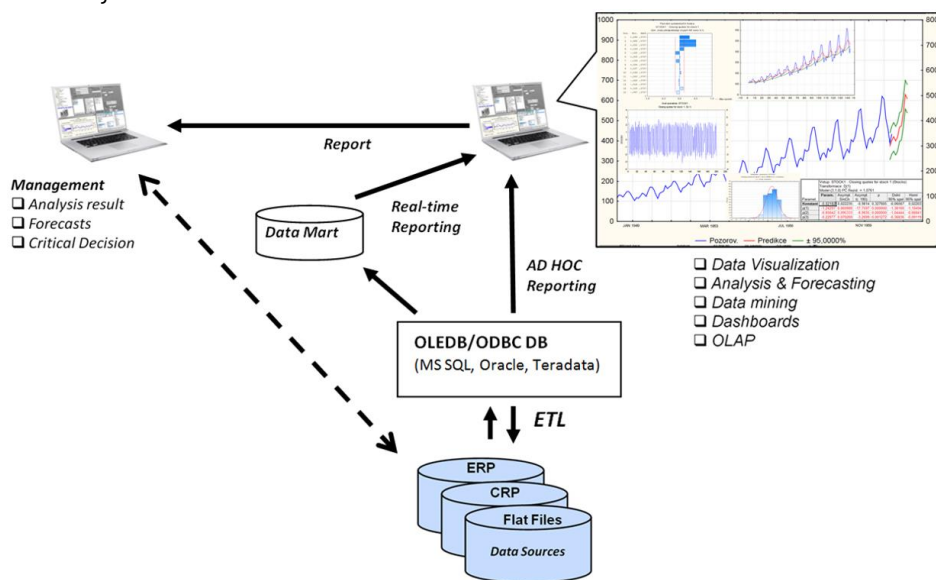


Rozhraní pro definici dotazů Statistica Query



- Na levé straně ve stromu jsou vidět konkrétní databáze na příslušném serveru a strukturu tabulek.
- Pravá strana potom ilustruje tvorbu databázového dotazu přes grafické rozhraní, kde výsledný dotaz bude složen ze dvou tabulek přes parametr ID a výstup bude omezen doplňkovým kritériem Gender.
- Dotaz lze v rozhraní definovat také pomocí standardní syntaxe Microsoft SQL (nebo MDX pro datové kostky).

Snadná integrace s různými datovými sklady a ostatními nástroji je dána také tím, že se software Statistica nespolehá na žádný proprietární skriptovací jazyk a využívá standardní (již zavedené) technologie pro skriptování. Princip práce softwaru Statistica nad datovým zdrojem ilustruje následující schéma:



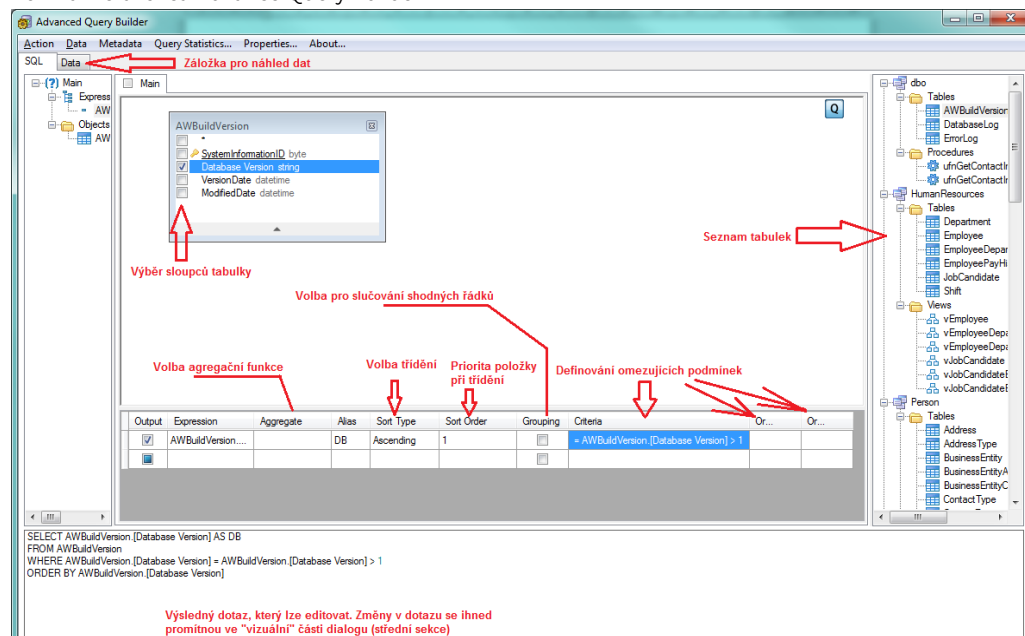
Analytický nástroj Statistica je přes příslušný standard připojen do databáze společnosti. Nad databází jsou pomocí integrovaného rozhraní Statistica Query definovány příslušné dotazy pro export dat do tabulky softwaru Statistica.

V softwaru Statistica je také možné dotazování do více typů různých datových zdrojů v rámci jednoho dotazu, což umožňuje výrazné zjednodušení práce při potřebě pracovat z informacemi uloženými různými způsoby.

✓ **Advanced Query Builder**

Nový nástroj Statistica Query builder, který je součástí základní licence Statistica, je novým rozhraním pro pokročilou práci s databází.

Rozhraní Statistica Advance Query Builder



Původní rozhraní, které bylo dostupné v starších verzích softwaru Statistica je dostupné i v současné verzi, stávající uživatelé tak mají 2 možnosti, jak k datům v databázi přistupovat.

Více informací se dočtete zde:

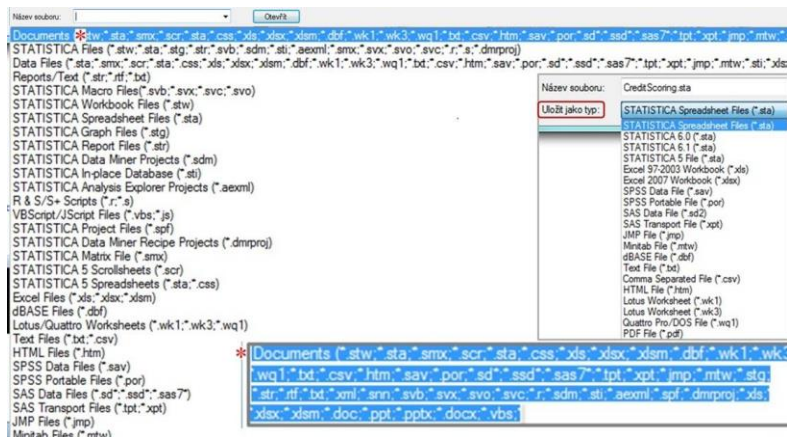
http://www.statsoft.cz/file1/PDF/newsletter/2014_05_28_StatSoft_Kladivo_na_data.pdf

✓ I

✓ mport/Export dat - načtení souborů všech typů

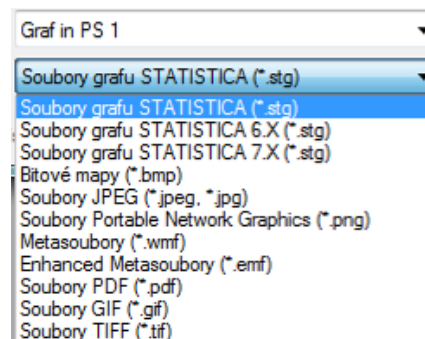
Software dále umožňuje načtení širokého množství formátů včetně Excelu, dále obsahuje nástroje pro přípravu dat před vlastní analýzou (úpravy a slučování tabulek, filtry atd.). Upravená a sloučená data ve výsledné tabulce lze ukládat do celé řady formátů, podporované formáty pro načtení a uložení zachycuje následující obrázek:

Podporované formáty Export/Import výstupních tabulek a datových souborů



Možnosti uložení grafických výstupů ilustruje následující obrázek.

Je potřeba zmínit, že proměnné pro analýzy a grafy jsou vždy sloupcové – proměnné bychom tedy měli mít ve sloupcích. Řádkům se ve Statistice říká *případy*. Pro představu, máme pacienty a u každého měříme několik atributů. Pokud chceme dělat běžné analýzy bez dalších úprav, pak potřebujeme mít údaje o jednom pacientovi v jednom řádku.



Načtení 3 souborů z různých zdrojů v modulu Statistica Data Miner:

