

■ **В.Домрачев**



**Технології аналізу
даних.**

**Інструменти бізнес –
аналітики.**

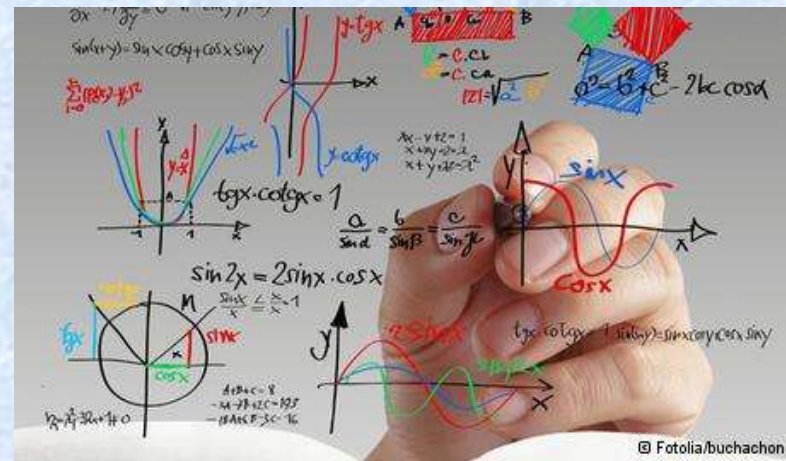


Київ – 2019 рік

Мета

Проаналізувати сучасні інструменти бізнес – аналітики, які застосовуються в процесі роботи фінансовими установами України.

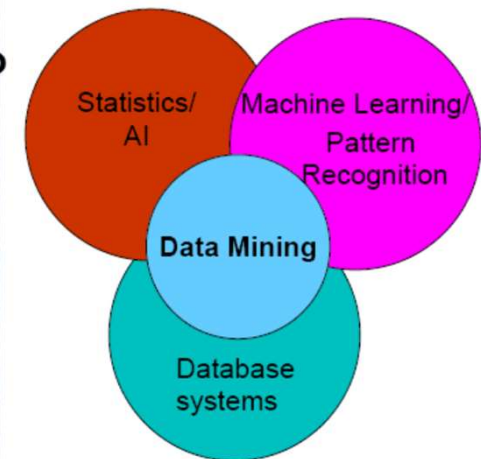
Обговорення можливих рішень з управління фінансами в Україні.





Що є бізнес - Аналітикою?

- Це діяльність компаній з моніторингу, аналізу, прогнозування оточуючого бізнес середовища. Для цього компанії використовують інструменти бізнес аналітики (Business intelligence tools), які дозволяють моніторити дані, аналізувати їх, розробляти стратегію поведінки на ринку та приймати найкращі рішення.

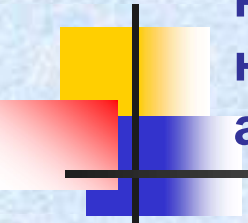


Що є інструментом бізнес -

Аналітики?

- Це ІС, які широко використовують дані, **статистичний і кількісний аналіз**, будують пояснювальні і **прогнознi моделі** для сприяння прийняттю рішень у бізнесі. Формують наочні **звіти**.

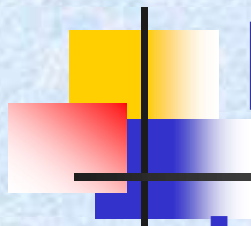




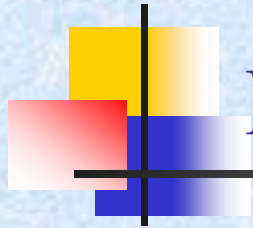
Банкам, компаніям, державі необхідно вчасно реагувати на зміни, особливо в умовах фінансової кризи. З'явилась необхідність вдосконалення інструментів бізнес – аналітики.

- До розробки систем бізнес – аналітики підключились такі відомі комп'ютерні компанії як IBM (Cognos Express, SPSS), Microsoft, Oracle, SAS та інші. Тому є актуальним проаналізувати ті риси таких систем, які дозволяють їм мати попит на сучасному ринку ІТ технологій.

Системи прийняття рішення

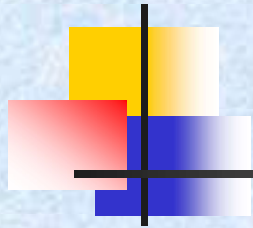


- **Визначення**
- **Види Рішень**
- **Планування та Ухвалення рішення**
- **Компоненти (методи) прийняття рішень, аналіз, моделювання, прогнозування**
- **Якість Рішення**
- **Стратегії прийняття рішення**
- **Процедура Ухвалення рішення**
- **Ризики**
- **Математичне Обґрунтування**
- **Стратегії Управління Ризиком**
- **Значення людського фактору у Плануванні та Ухваленні рішення**
- **Системи підтримки прийняття рішення. Експертні системи**
- **Приклади**



Що таке - прийняття рішення?

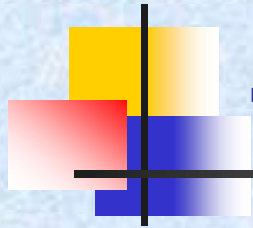
- Ми всі ухвалюємо рішення щодня, так що ідея що ухвалення рішення може бути швидше досвідченим мистецтвом, можливо, спочатку здаватися дивним. Проте, дослідження показали, що більшість людей набагато бідніша в інструментах ухвалення рішення, ніж вони думають. Розуміння того, що ухвалення рішення залучає БА, разом з декількома ефективними методами, допоможе прийняттю кращих рішень.



Що є бізнес - Аналітикою?

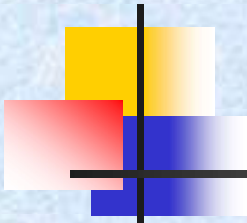
Аналітика - використовує:

- дані
 - інформаційні технології
 - статистичний аналіз
 - кількісні методи, і
 - математичні або комп'ютерні моделі
- щоб допомогти менеджерам поліпшити процес прийняття рішення відносно проведення операцій бізнесу



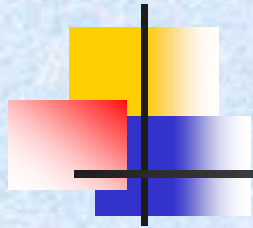
Застосування бізнес - Аналітики

- Управління відносинами до клієнта
- Фінанси та маркетинг
- Послідовність управління
- Планування кадрового ресурсу
- Оцінка рішень
- Стратегії гри команди
- ...



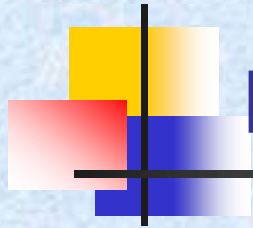
Інші можливості застосування інструментів бізнес - Аналітики

- Реклама в інтернеті
- Особисті відносини
- Охорона здоров'я
- Боротьба з шахрайством
- Підвищення безпеки
- Кібербезпека
- тощо



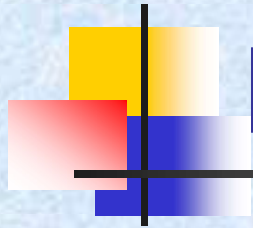
Важливість бізнес - Аналітики

- У зв'язку з імовірною природою ризику прогнозування прибутку стає дедалі гнучкішою задачею
- БА надає додаткове розуміння динаміки даних
- БА сприяє конкурентоздатності бізнесу
- БА вдосконалює систему звітності



Риси бізнес - Аналітики

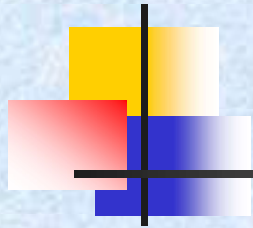
- Описова аналітика
 - використовує дані для розуміння минулих та поточних явищ
- Прогнозна аналітика
 - аналізує минулі явища
- Предписуюча аналітика
 - використовує методи оптимізації для оцінки параметрів



Риси бізнес - Аналітики

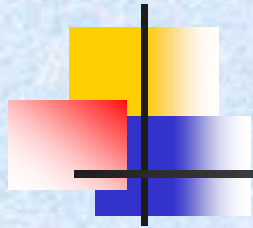
Приклад 1 Роздрібна торгівля

- Більшість супермаркетів розпродають сезонний товар, знижуючи ціни.
- Питання: Коли знизити ціну і наскільки?
- **Описова аналітика:** розглядає історичні дані для подібних товарів (ціни, кількість, витрати на рекламу)
- **Прогнозна аналітика:** передбачає продажі, засновані на певній ціні
- **Предписуюча аналітика:** знаходить кращі ціни і витрати на рекламу, які дозволяють максимізувати комерційний дохід



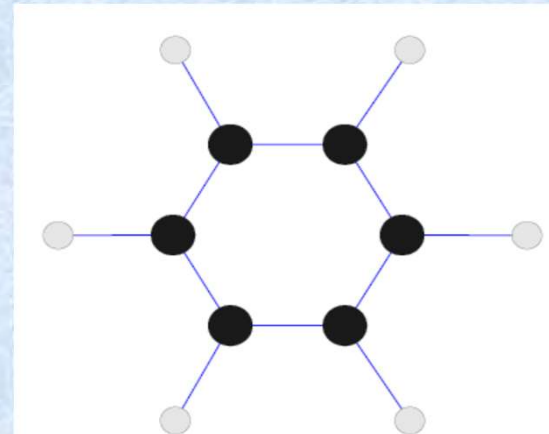
Дані бізнес - Аналітики

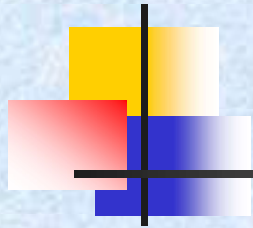
- DATA
 - зібрані факти і властивості
- DATABASE
 - комп'ютерні файли, що містять дані
- INFORMATION
 - результат аналізу даних



Дані бізнес - Аналітики

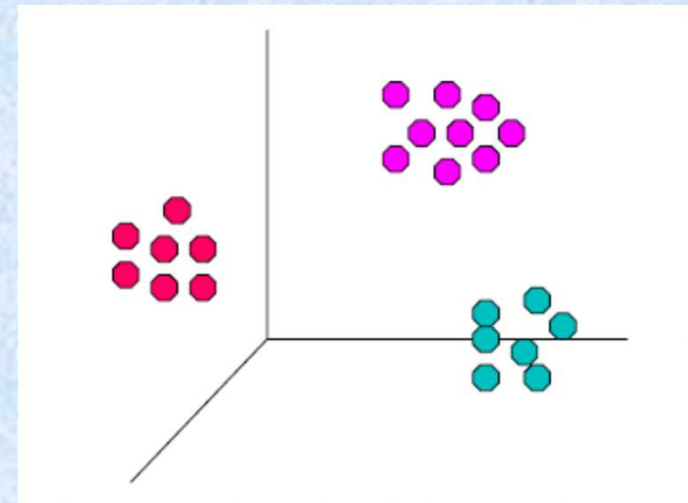
- Metrics використовуються для кількісного виміру.
- Measures є числові значення метрики.
- Discrete – дискретна метрика
- Continuous - метрика базується на континуумі

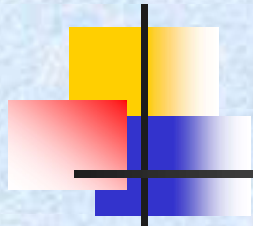




Дані бізнес - Аналітики

- Типи даних:
- Categorical (номінальні) дані
- Ordinal дані
- Interval дані
- Ratio дані





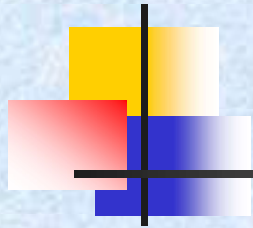
Дані бізнес - Аналітики

Категорні (номінальні) Дані

- Дані розміщено по категоріях згідно вказаній характеристики
- Категорії не носіть ніякого кількісного взаємовідношення один до одного

Приклади:

- країни походження товару (Америка, Європа, Азія)
- класифікація працівника (менеджер, службовець, викладач, інспектор, доктор)
- Колір очей



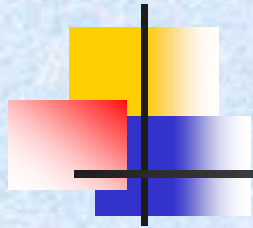
Дані бізнес - Аналітики

Впорядковані Дані (Ordinal)

- Дані які класифіковано або замовлено згідно деякому відношенню один до одного
- Відсутні одиниці вимірювання

Приклади:

- - ранжирування футбольних команд
- - відповіді студентів (погано, середня величина, добре, дуже добре, чудово)

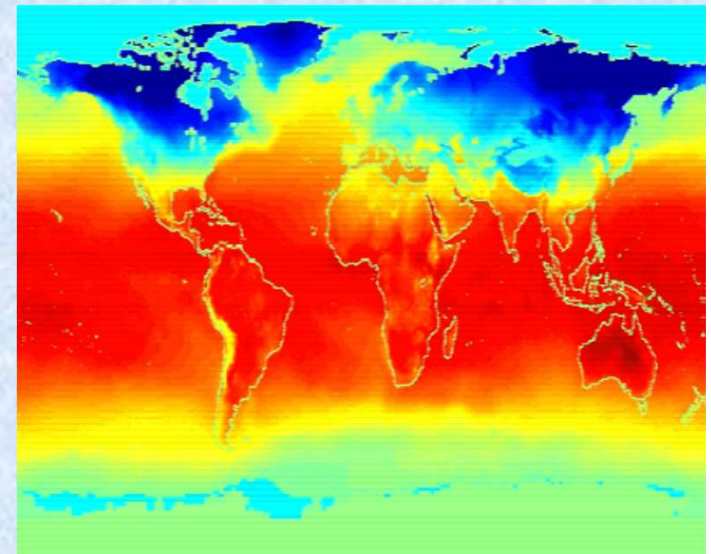


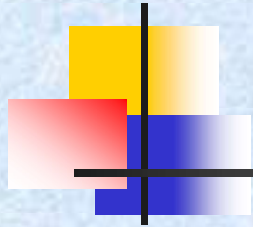
Дані бізнес - Аналітики

- Інтервальні Дані
- Впорядковані дані але без відмінності у інтервалі
- Відношення не має значення

Приклади:

- - температурні дані
- - скорингове значення
- - дати
- - розмір



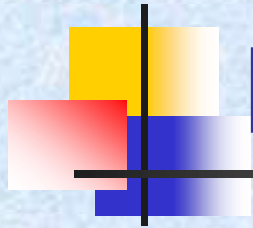


Дані бізнес - Аналітики

- Дані Співвідношення (Ratios)
- Непреривні дані і мають природне нульове значення
- Відношення є характерною рисою

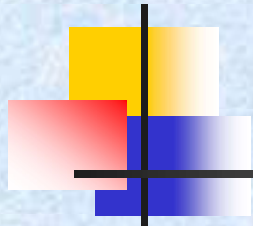
Приклади:

- - щомісячні продажі
- - часи постачання товару
- - довжина



Властивості (атрибути)

- Відмінність = ≠
 - Порядок < >
 - Складання + -
 - Множення * :
-
- Categorical дані Відмінність
 - Ordinal дані Відмінність та Порядок
 - Interval дані Відмінність, Порядок, Складання
 - Ratio дані Всі



Моделі прийняття рішень

Модель:

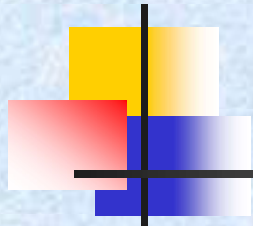
- Абстракція або представлення реальної системи, ідеї, або об'єкту
- Містить найголовніші особливості
- Може бути **записаним або усним** описом, **візуальним показом, математичною формулою**, або представленою у вигляді **електронної таблиці**

Моделі прийняття рішень

Приклад:

- Продаж нової продукції, як наприклад iPad або 3D телебачення, часто мають загальну тенденцію.
- Продажи, можливо, росли б у вигляді *S-форменої* кривої.
- **Математична модель** S-кривої може бути представлена у вигляді: $S = ae^{be^{ct}}$, де S продажі, t - час, e - основа натурального логарифму і a , b і c - константи.





Моделі прийняття рішень

Модель рішення – це модель, яка використовується для розуміння, аналізу, та полегшення процесу прийняття рішення.

Входні параметри

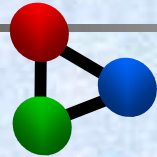
- - **дані**
- - **неконтрольовані змінні**
- - **змінні які використовуються у процесі управління**

Вихідні параметри

- - **критерії якості роботи, прибуток, ефективність**
- - **заходи поведінки**

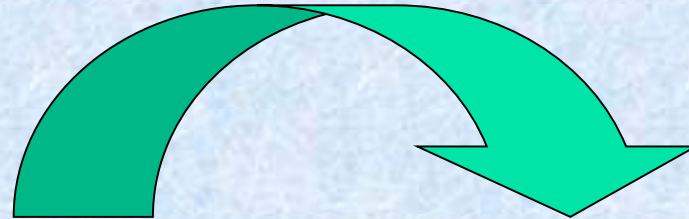
Моделі прийняття рішень

Історичні дані...



Модель

Аналітик



Управлінці



Microsoft BI Matrix

© OfficeBICentral.com



REPORTING



ANALYSIS



MEASUREMENT



PLANNING

Desktop Applications

- Excel 2007
- Outlook 2007
- PowerPoint 2007
- Visio 2007

Excel Print / PDF

Excel PivotTable/Chart

Excel What-If

Outlook Business Contact Manager

Excel List Table / Conditional Formatting / Data Bars / Color Scales / Icon Sets

Excel Data Mining

Excel / MSQuery external data – Access / Flat Files / SQL Server / OLAP

PowerPoint

Visio PivotDiagram

Server Applications

- SharePoint Server 2007 (MOSS)
- PerformancePoint Server 2007

SharePoint Excel Services

SharePoint Report Center / Key Performance Indicators / Dashboards

Analytic Chart / Grid

Scorecard Designer

Biz # Planning

- Office Live

Office OnLine – Office.Microsoft.com

- MapPoint 2006
- Win Mobile 6.0

MapPoint / Web Services / Virtual Earth

Office Mobile Excel / PowerPoint – SmartPhone

- SQL Server 2005

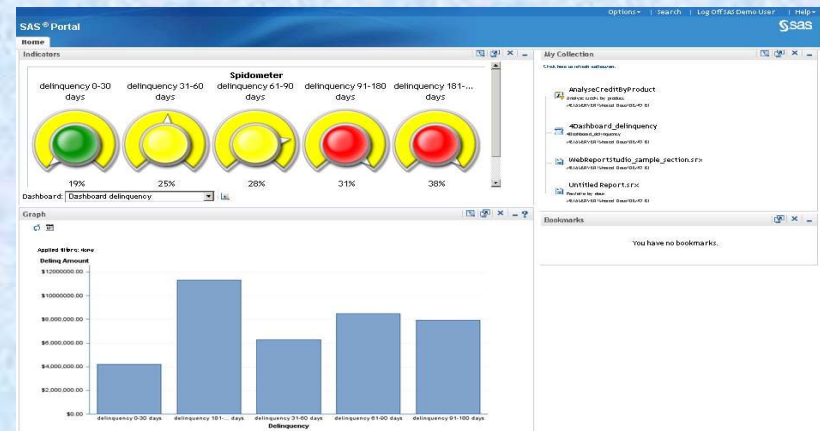
Reporting Services

Analysis Services OLAP / Key Performance Indicators

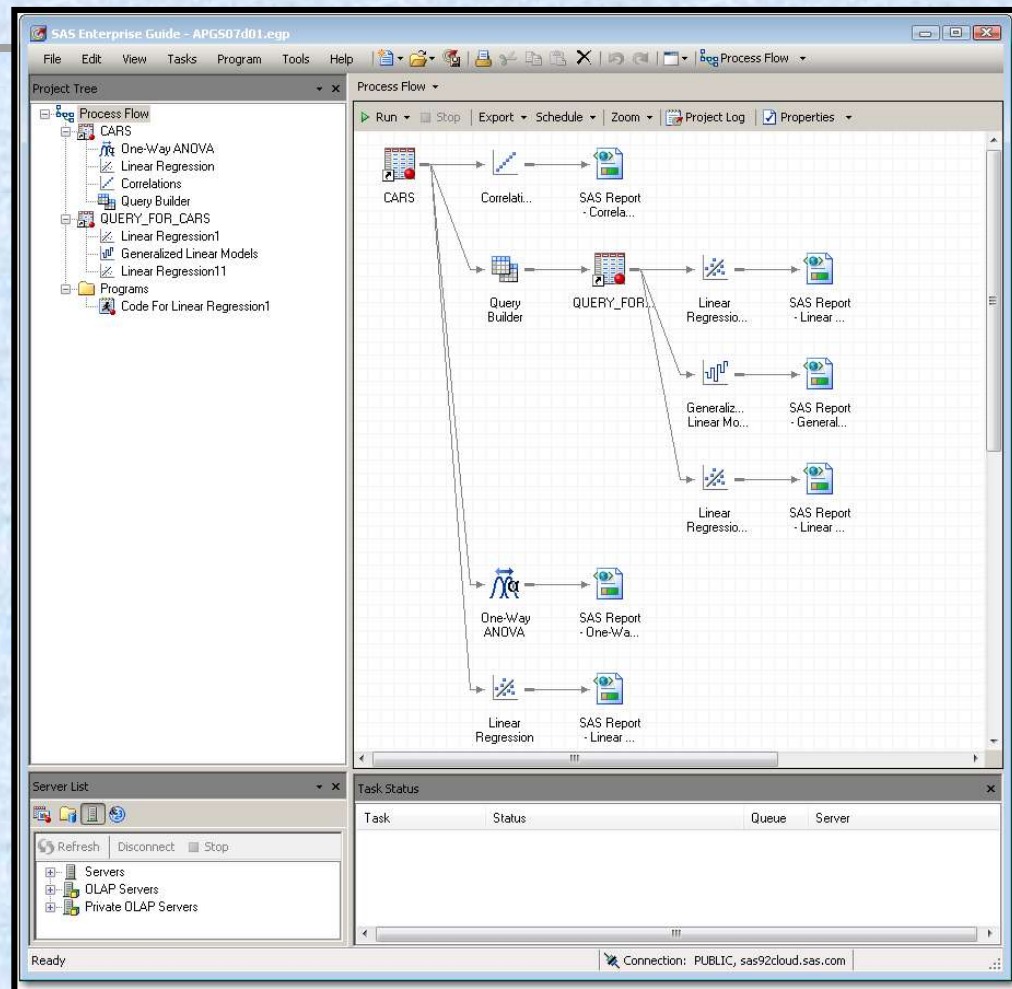
Integration Services ETL / Data Mining

Інструменти Business Analytics (SAS)

- Інструменти бізнес – аналітики SAS включають
 - SAS Enterprise Guide
 - SAS Enterprise Miner
 - SAS Forecast Studio.
- Розглянемо їх окремі риси.



SAS Enterprise Guide



SAS Enterprise Guide

- Обработка
та аналіз
даних

The screenshot displays the SAS Enterprise Guide interface. The main window shows a project tree on the left with a tree view containing 'Схема процесса', 'ua_data', and 'Линейная рег'. A menu is open over the 'Данные' (Data) section, listing various tasks such as 'Фильтр и сортировка...', 'Конструктор запросов...', 'Присоединить таблицу...', 'Сортировка данных...', 'Создать формат...', 'Создать формат из набора данных...', 'Транспонирование...', 'Разбиение столбцов...', 'Объединение столбцов...', 'Случайная выборка...', 'Ранг...', 'Стандартизация данных...', 'Атрибуты набора данных...', 'Сравнение данных...', 'Удаление наборов данных и форматов...', 'Выгрузить файлы данных на сервер...', 'Загрузить файлы данных на компьютер...', 'Импортировать файл JMP...', 'Импортировать файл SPSS...', and 'Импортировать файл Stata...'. The main plot area shows a residual plot titled 'Residuals for EXTRAVR' with a y-axis labeled 'Proportion Less' ranging from 0.0 to 1.0. The bottom status bar indicates 'Готово' (Ready) and 'Профиль не выбран' (Profile not selected). The task status window at the bottom shows a table with columns: 'Задача' (Task), 'Состояние' (Status), 'Очередь' (Queue), 'Сервер' (Server), and 'Server Type'.

Задача	Состояние	Очередь	Сервер	Server Type
--------	-----------	---------	--------	-------------

SAS Enterprise Guide

The screenshot displays the SAS Enterprise Guide interface for a project named "Проект2.egp". The main window shows a project tree on the left with "Схема процесса" and "ua_data" under "Линейная регрессия". A menu is open, showing options like "Данные", "Описательная статистика", "График", "Дисперсионный анализ (ANOVA)", "Регрессия", "Многомерный анализ", "Анализ живучести" (highlighted), "Возможность", "Контрольные диаграммы", "Диаграмма Парето...", "Временные ряды", "Интеллектуальный анализ данных", "OLAP", and "Шаблоны задач". A sub-menu for "Анализ живучести" is also visible, showing "Таблицы продолжительности жизни..." and "Пропорциональные интенсивности отказов...". The main plot area displays a residual plot with "Residual" on the y-axis and "Quantile" on the x-axis. The plot shows a linear trend with data points and a fitted line. The status bar at the bottom indicates "Готово" and "Профиль не выбран". The taskbar at the bottom shows various open applications including Total Commander, Lingvo, Adobe Reader, SAS Enterprise Guide, Microsoft Excel, and Microsoft PowerPoint.

СAS Enterprise Guide - Проект2.egp

RU Русский

Файл Изменить Вид Задачи Программа Сервис Справка

Дерево проекта

- Схема процесса
- ua_data
- Линейная регрессия

Журнал Выходные данные Результаты - HTML Результаты в формате PDF

Задача Экспорт Отправить Опубликовать Свойства

Анализ живучести

- Таблицы продолжительности жизни...
- Пропорциональные интенсивности отказов...

Residual

Quantile

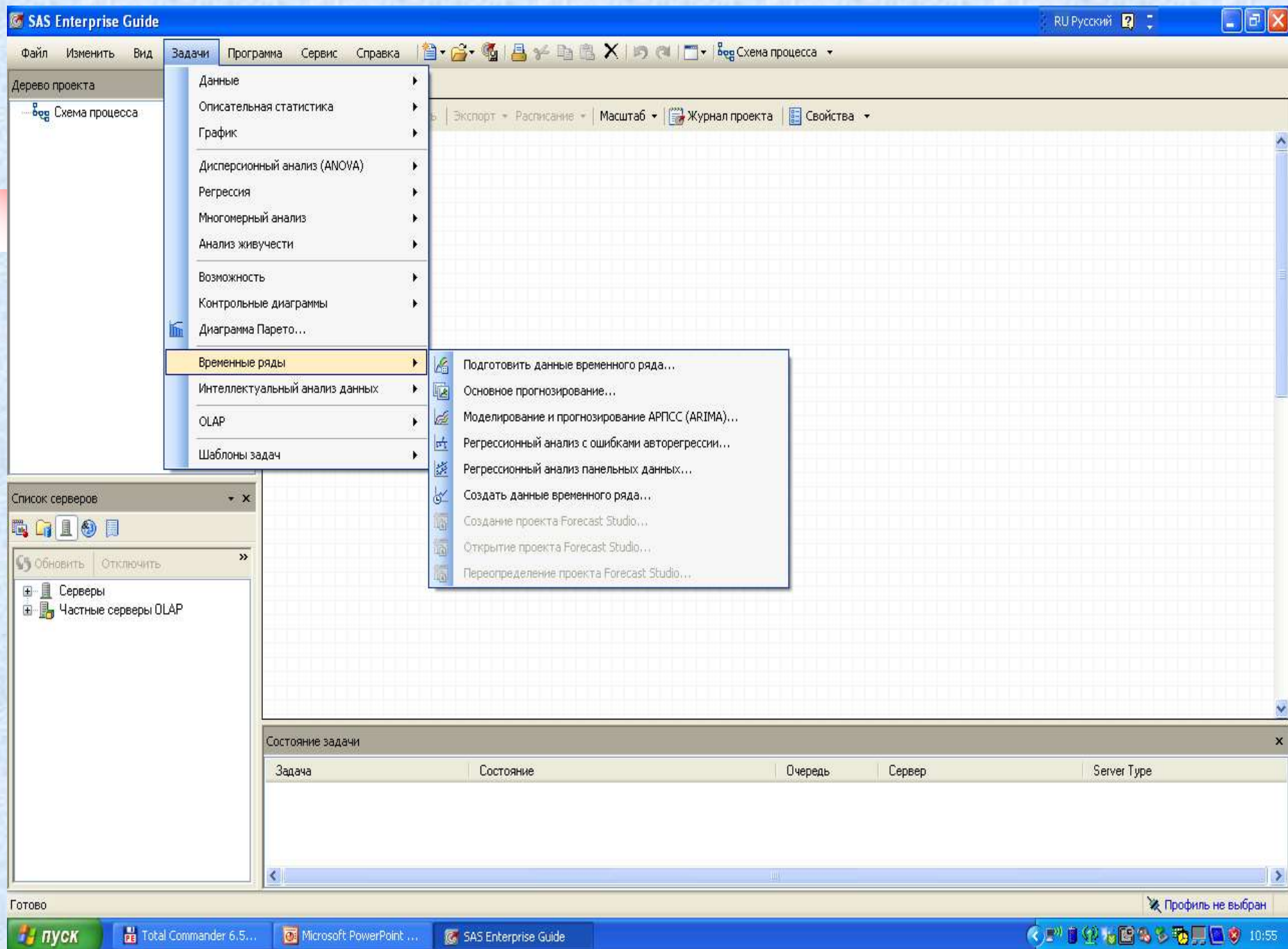
Состояние задачи

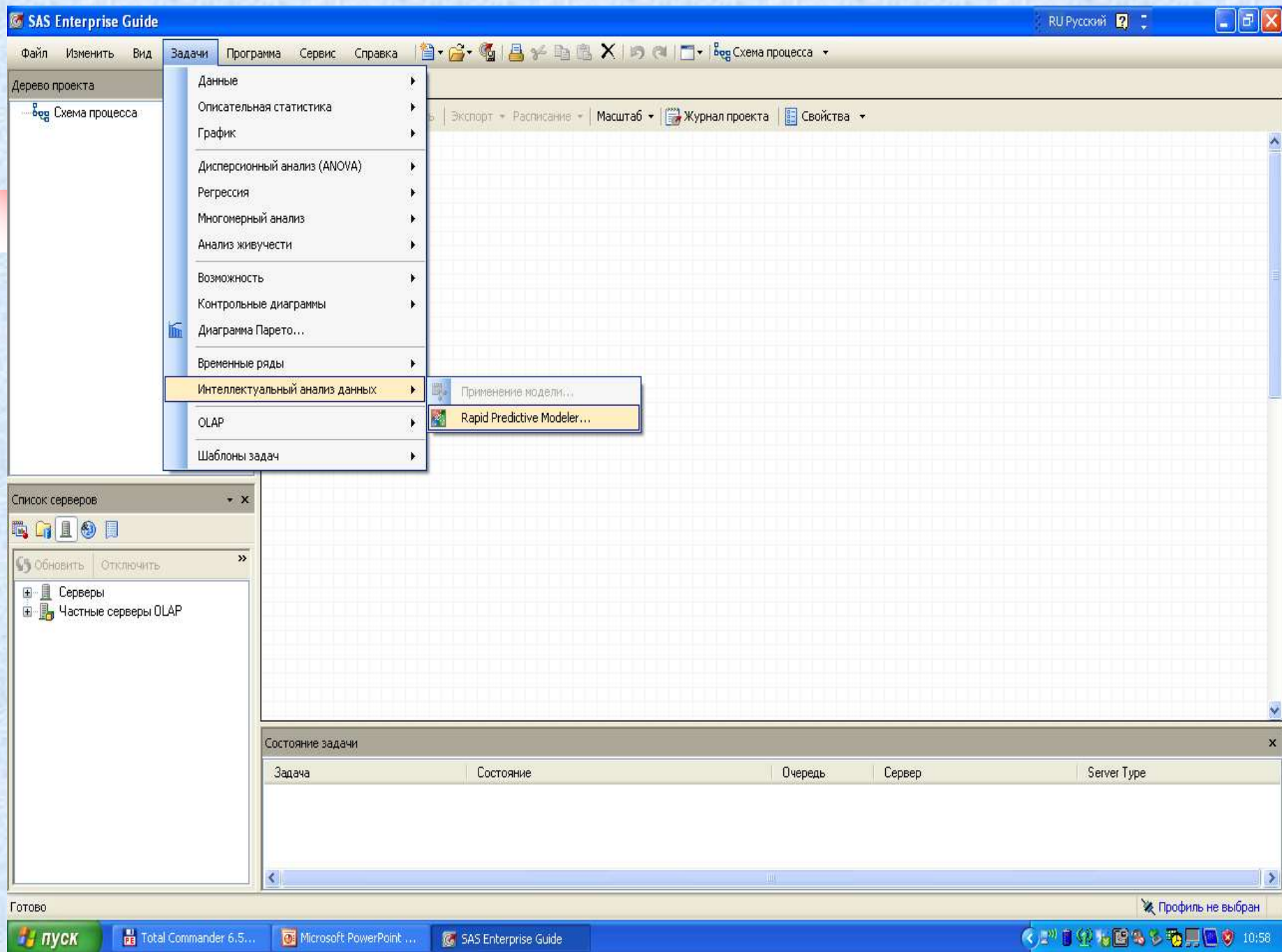
Задача	Состояние	Очередь	Сервер	Server Type
--------	-----------	---------	--------	-------------

Готово

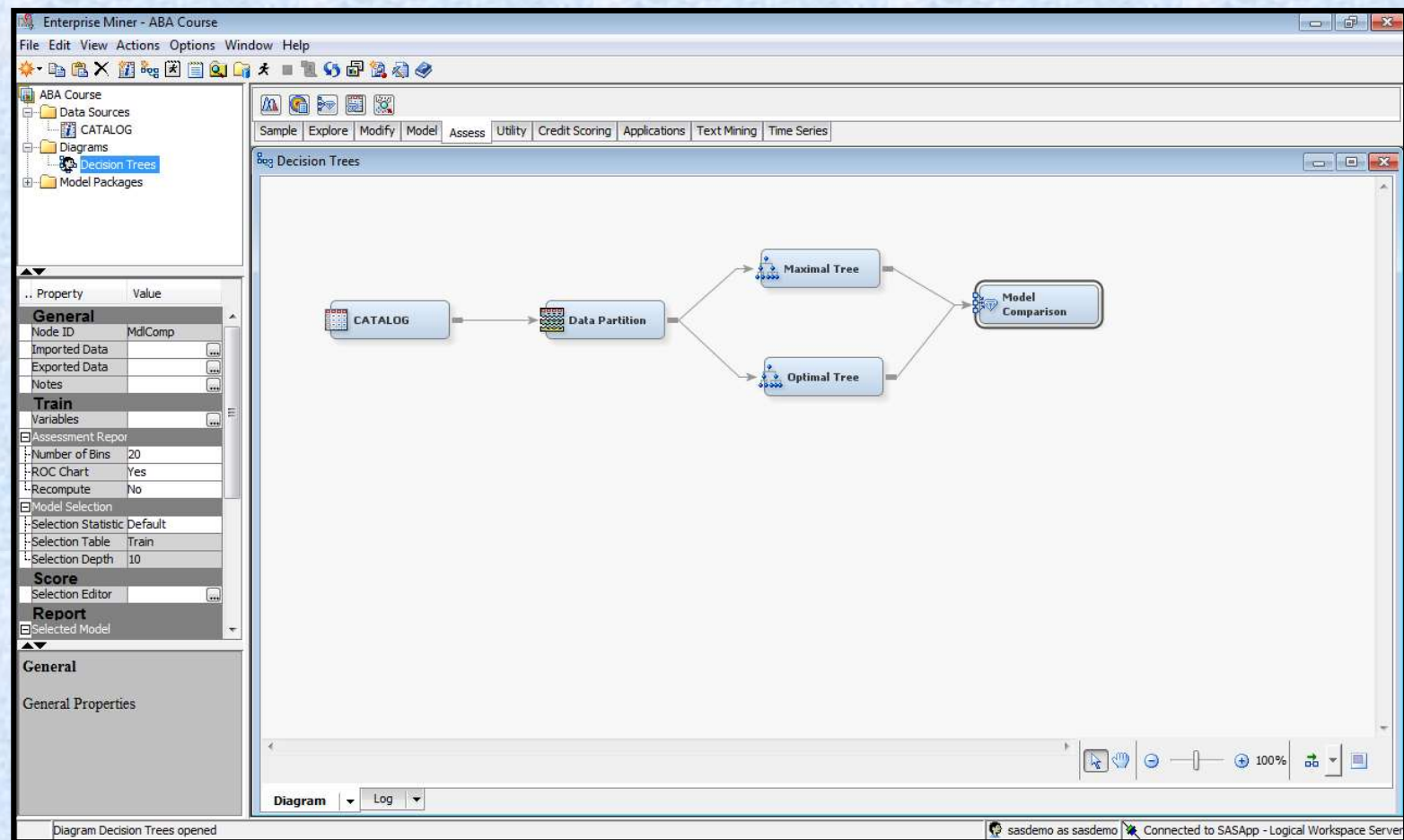
Профиль не выбран

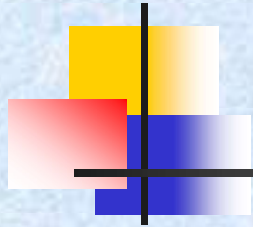
пуск Total Commander 6.5... Lingvo 5.0 Adobe Reader - [Dat... SAS Enterprise Guide ... Microsoft Excel - UA_D Microsoft PowerPoint ... 21:28





SAS Enterprise Miner



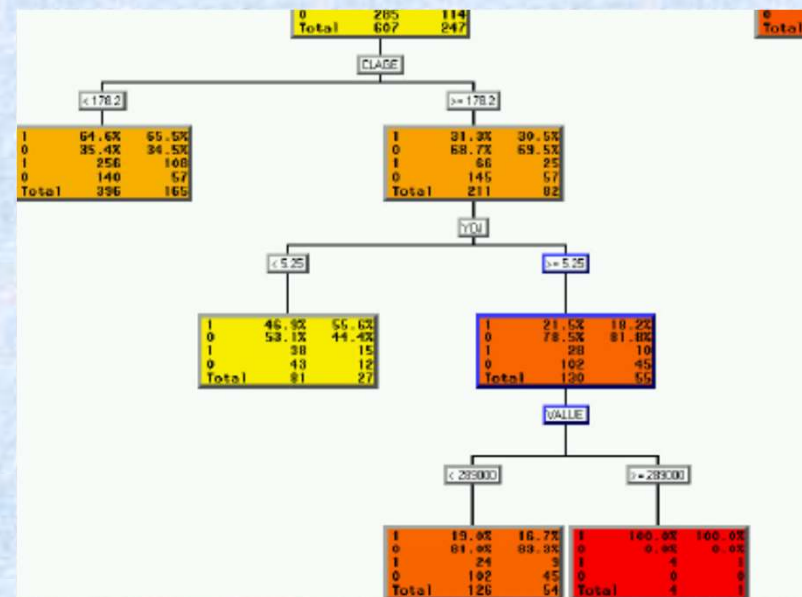


Data mining – це процес знаходження відношень (залежності) між даними з великих за обсягом та неструктурованих баз даних

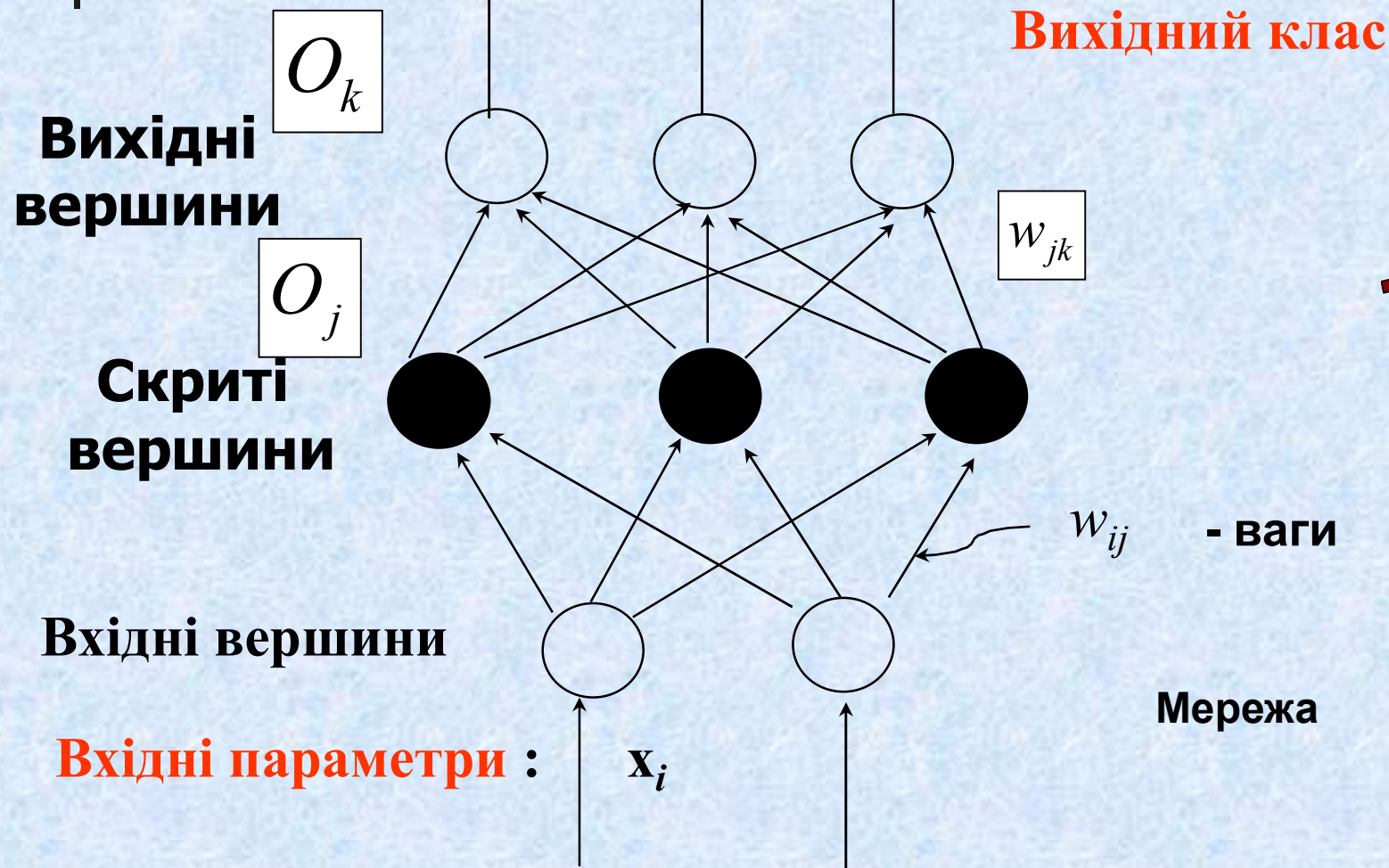
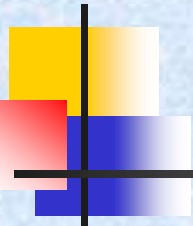
SAS Enterprise Miner:

- прогнозує часові ряди,
- використовує нейронні мережі у процесі моделювання,
- дозволяє будувати скорингові картки,
- тощо.

SAS Enterprise Miner будує дерева прийняття рішень



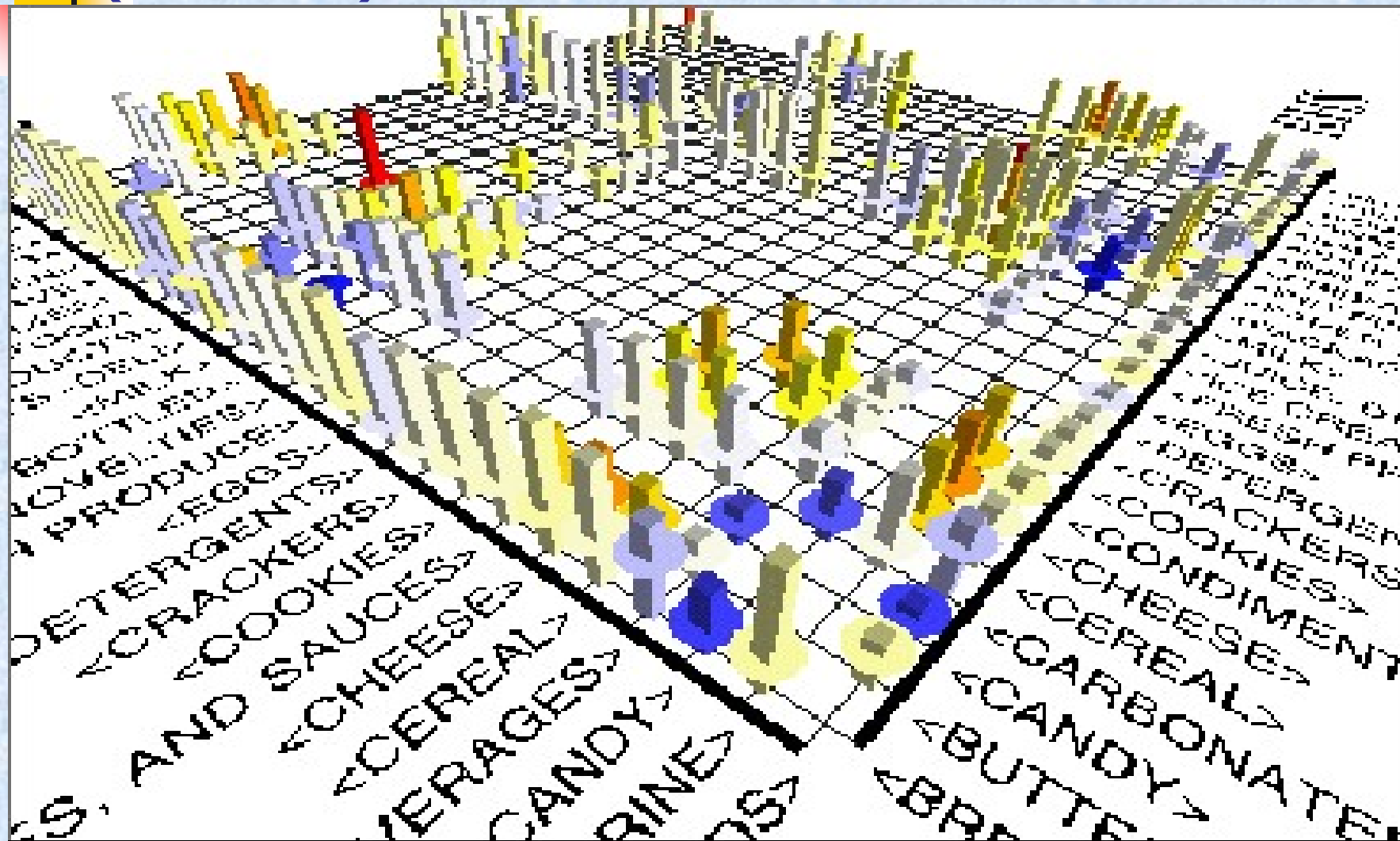
SAS Enterprise Miner будує нейроні мережі (моделі)



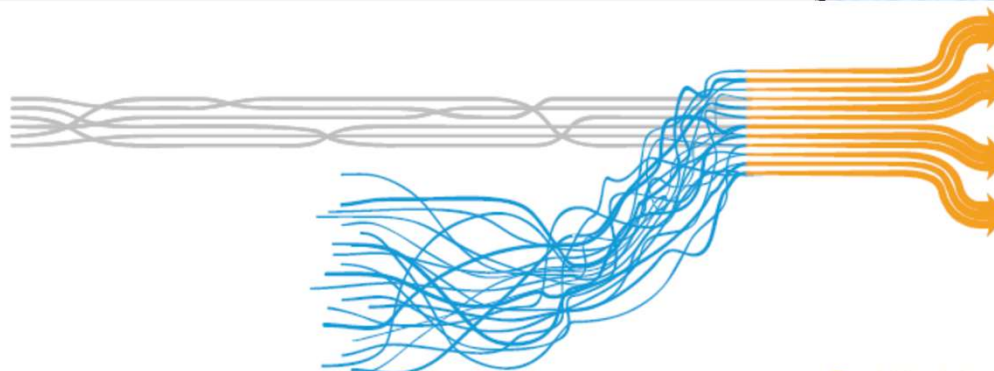
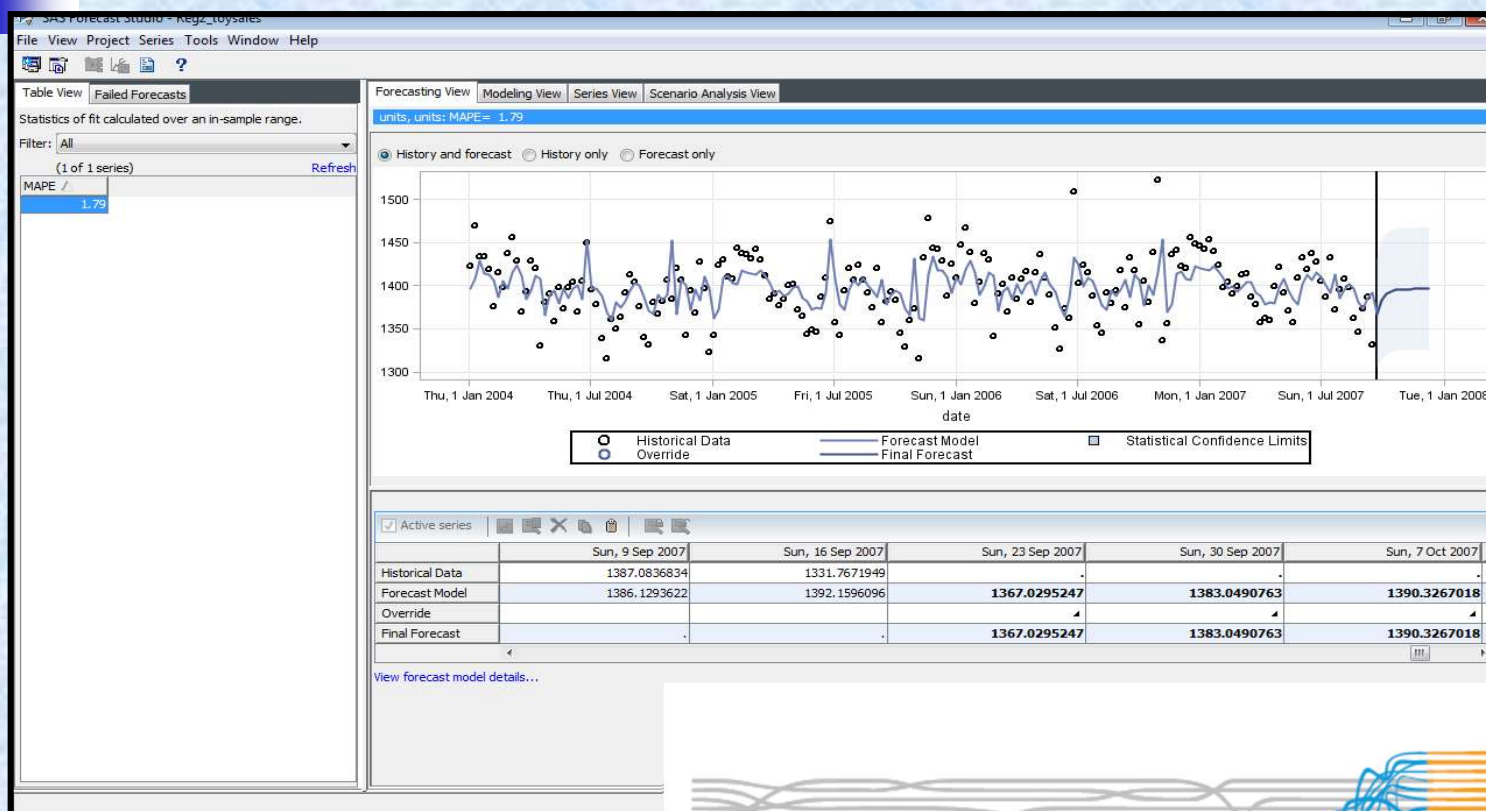
Data Mining: об'єднує багато дисциплін



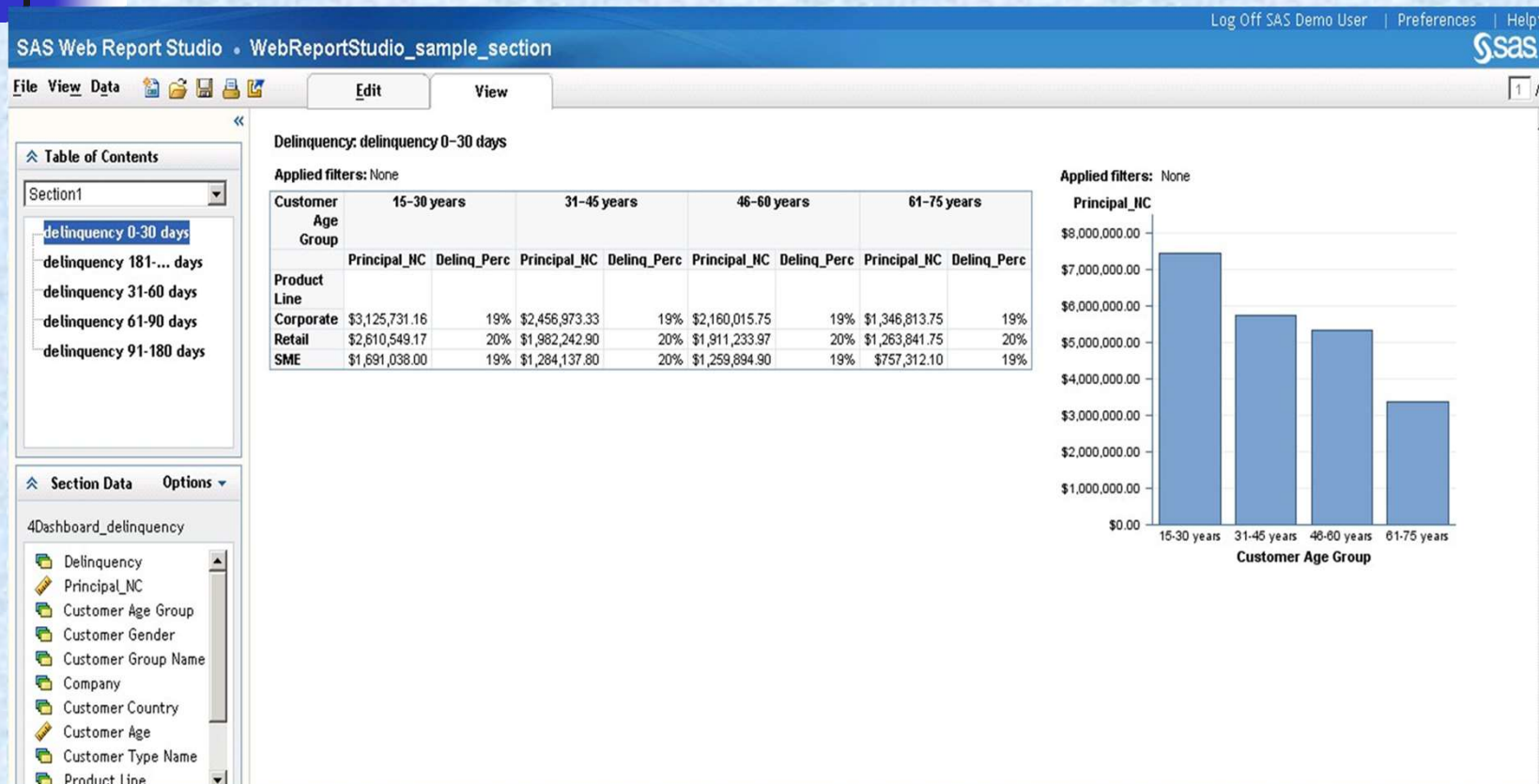
Візуалізація асоціативних правил (моделі)



SAS Forecast Studio: прогнозування



Звітність: Приклад побудови Web звіту у системі SAS Web Report Studio





ЦИТАТЫ ИЗ КНИГИ ЛЬЮИСА КЭРРОЛЛА «АЛИСА В СТРАНЕ ЧУДЕС»:

«ПЛАН, ЧТО И ГОВОРИТЬ, БЫЛ ПРЕВОСХОДНЫЙ; ПРОСТОЙ И ЯСНЫЙ, ЛУЧШЕ НЕ ПРИДУМАТЬ. НЕДОСТАТОК У НЕГО БЫЛ ТОЛЬКО ОДИН: БЫЛО СОВЕРШЕННО НЕИЗВЕСТНО, КАК ПРИВЕСТИ ЕГО В ИСПОЛНЕНИЕ»

ФОТО: ИЛЛЮСТРАЦИЯ ДЖОНА ТЕННИЕЛА К КНИГЕ «АЛИСА В СТРАНЕ ЧУДЕС»

АРГУМЕНТЫ
И ФАКТЫ aif.ru

Об'єднання зусиль – ключі до бізнес - аналітики



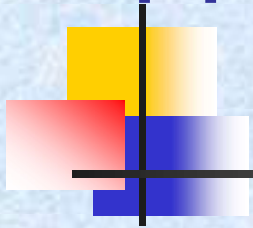
- Експерти
- ІТ фахівці
- Аналітики

Множина Даних

- ☐ **Визначення (Ідентифікація) джерел.**
- ☐ **Визначення (Ідентифікація) цілей.**
- ☐ **Визначення методів досягнення цілей.**



Дані фінансових інститутів

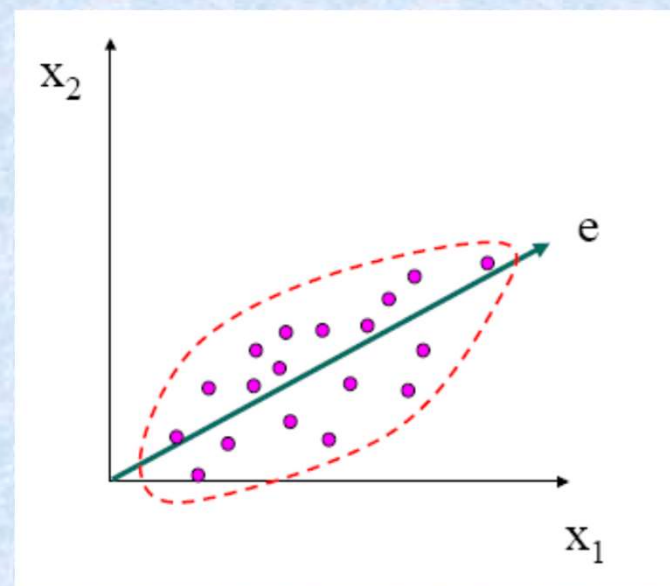


- Фінансові компанії працюють з досить точними даними з чіткою структурою даних. Зазвичай ці дані дозволяють побудувати розумну прогнозну модель. Однак дані характеризують велику кількість клієнтів з хаотичною поведінкою.
- До того ж різні дані мають різні, не завжди, сумісні формати представлення.

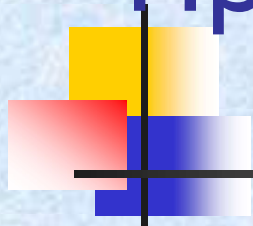
SAS Enterprise Guide Overview

■ SAS Enterprise Guide можливо використати для управління та аналізу даних:

- data exploration
- querying and reporting
- graphical analysis
- statistical analysis
- scoring



Проблеми Data Mining



■ Факти:

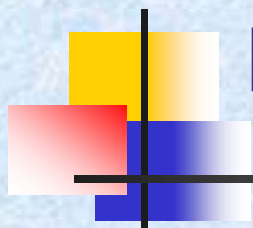
- a. Звичайно Великий обсяг «сирих» даних.
- b. Багато значущих змінних, що не дозволяє визначити чітку структуру даних.
- c. Деякі змінні конфліктують при побудові моделі.
- d. Певні сумнівні дані мають бути виключені при побудові моделі.

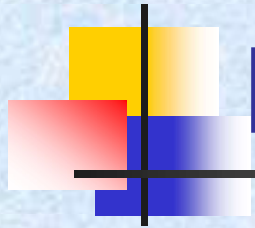


Побудова моделі

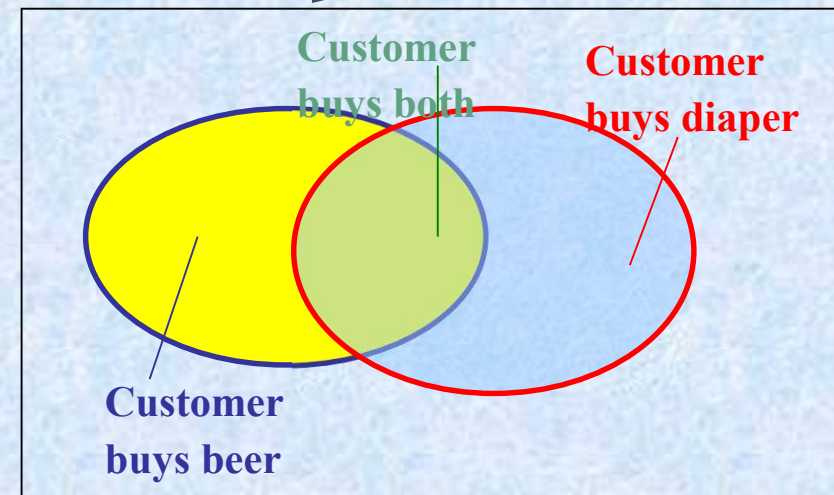
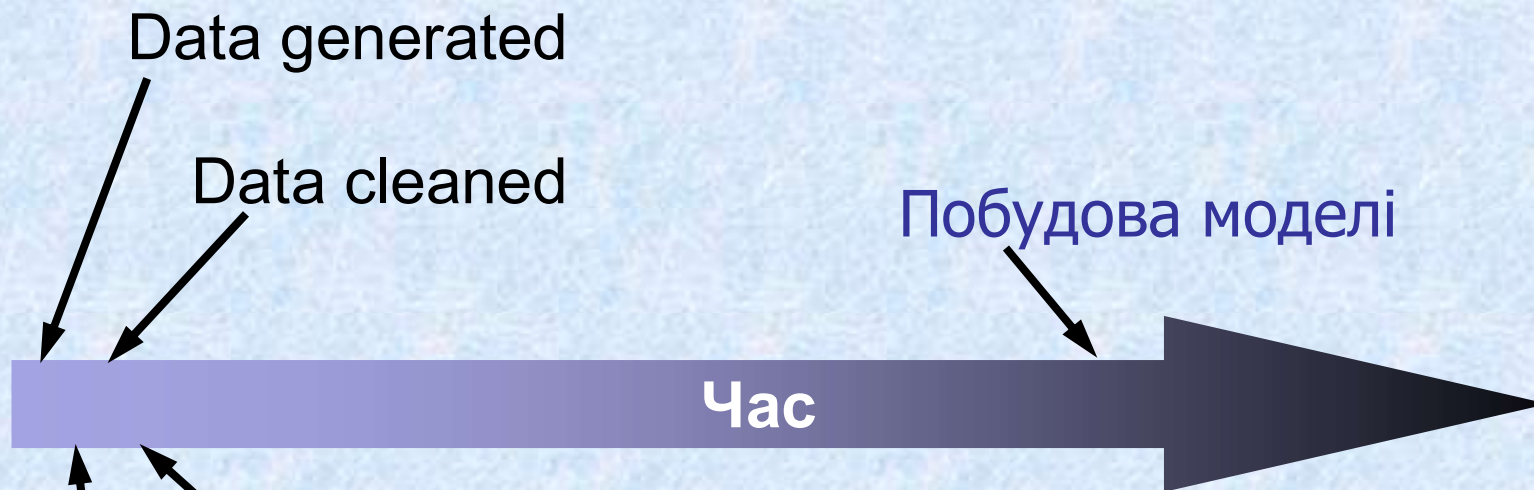
- Моделі обираються та порівнюються
 - Яка з них надає кращі прогнози?
- Прийняття рішення
 - Які дії будуть кращими у поточній ситуації?
- Подбор
 - Як застосувати результат прогнозу до поточної ситуації?

В процесі побудови моделі дані
розділяються на 3 частини





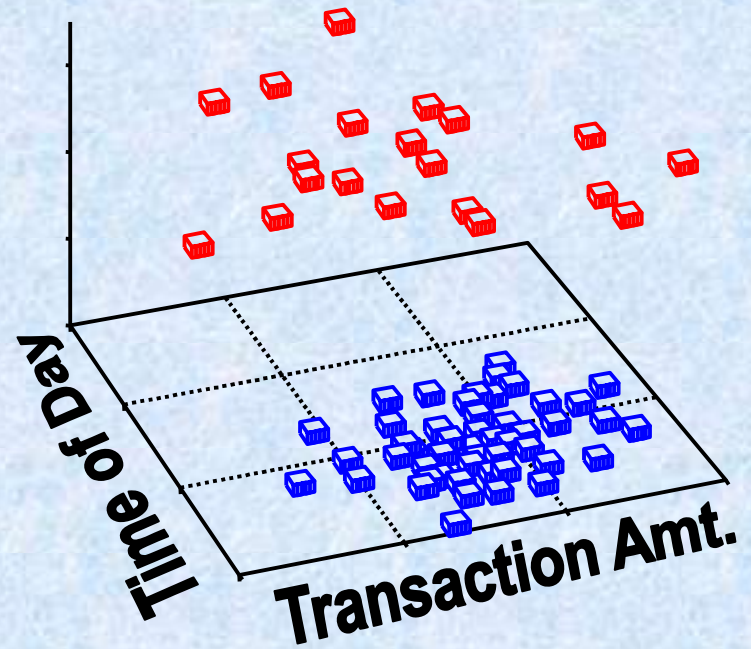
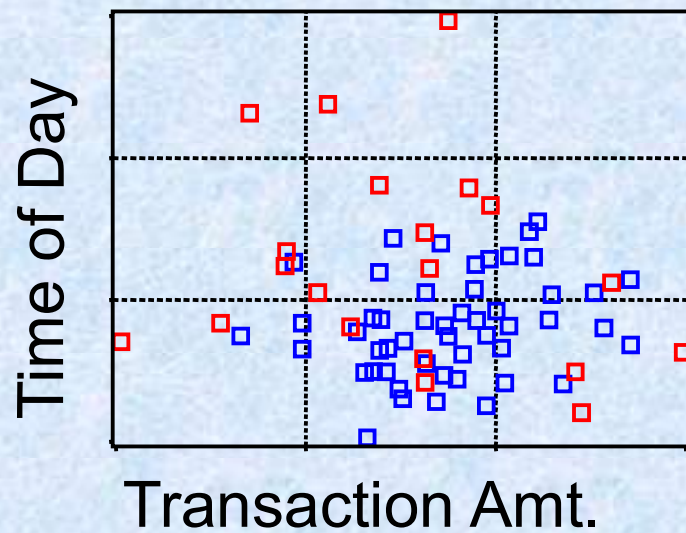
Побудова моделі



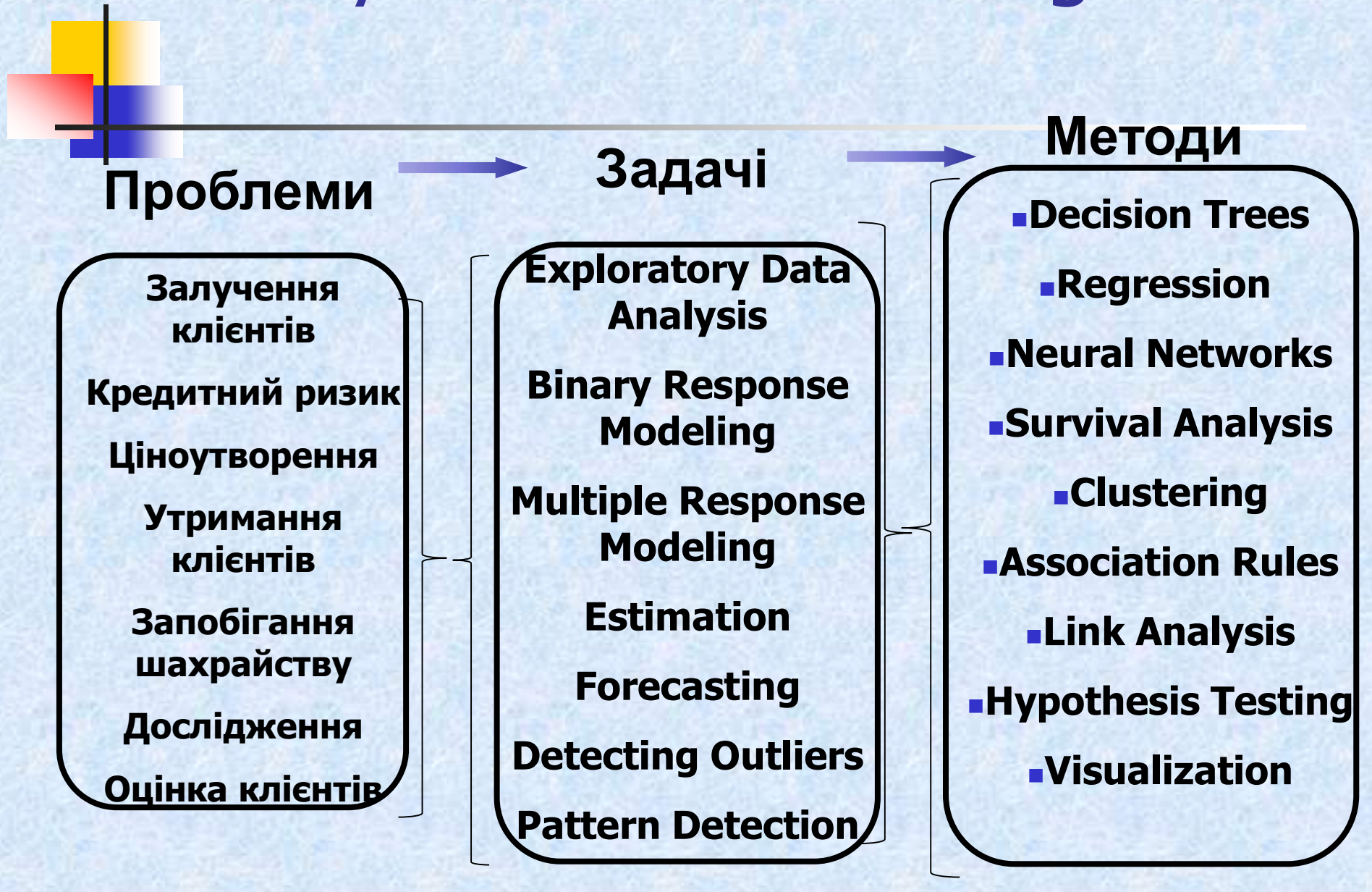
Приклад результату прогнозу

□ Fraud

□ OK

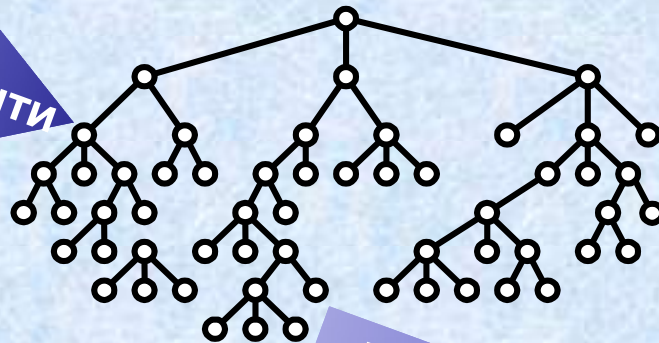


Застосування Data Mining



Виграшна стратегія: Очікуваний прибуток

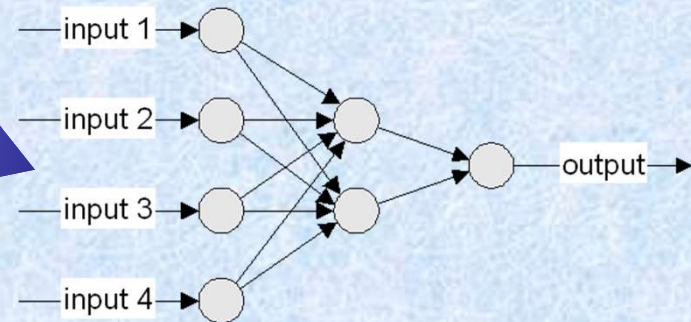
Потенціальні клієнти



Мета: Оцінити реакцію споживачів.

Дійсні клієнти

Мета: Оцінити значення реакції клієнтів:
 $dollars_{person}$ (only customers who respond).



Обчислити очікуваний ефект:
 $response_{person} * dollars_{person}$



Проблеми з даними

- Дані не чітко (точно) описують реальний стан.
 - Дані можуть бути нечіткими або відсутніми.
 - Приклади можуть бути не репрезентативними.
 - Категорійні змінні можуть мати дуже багато значень.
 - Чисельні дані можуть мати незвичайний розподіл.
 - Значення можуть змінюватись у часі.
 - Дані можуть суперечити одне одному.

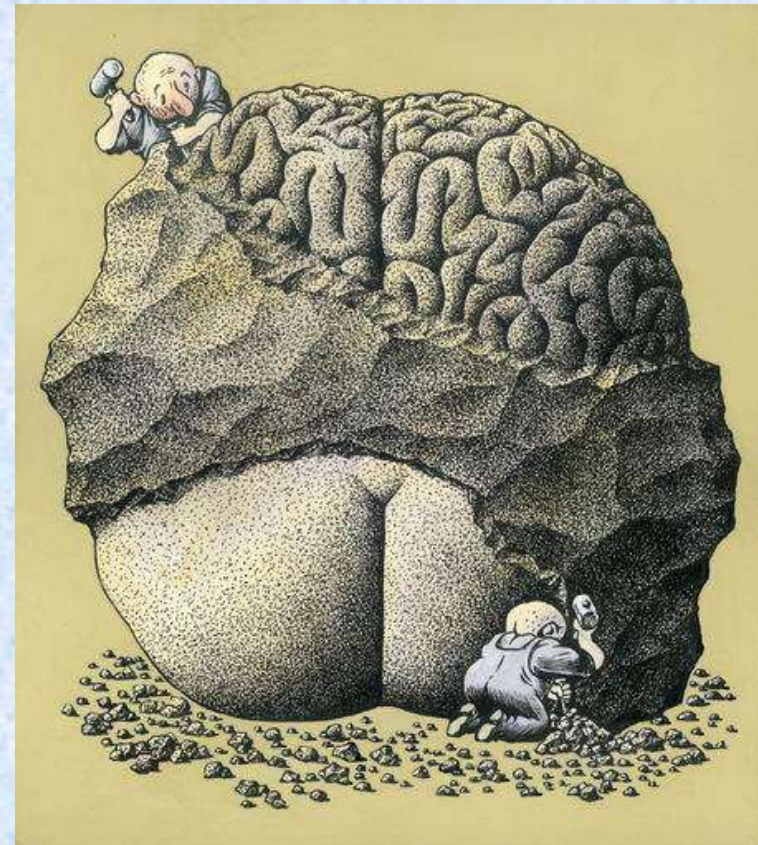
Training

Validation

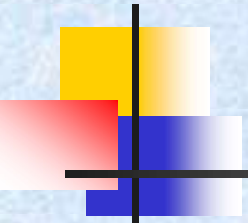
Test

Методів побудови моделі багато – треба обрати найкращий

- Регресії
- Дерева побудови рішень
- Пошук кластера
- Правила асоціації
- Нейронні мережі
- Генетичні алгоритми
- Алгоритми теорії ігор
- Аналіз на графах
- тощо



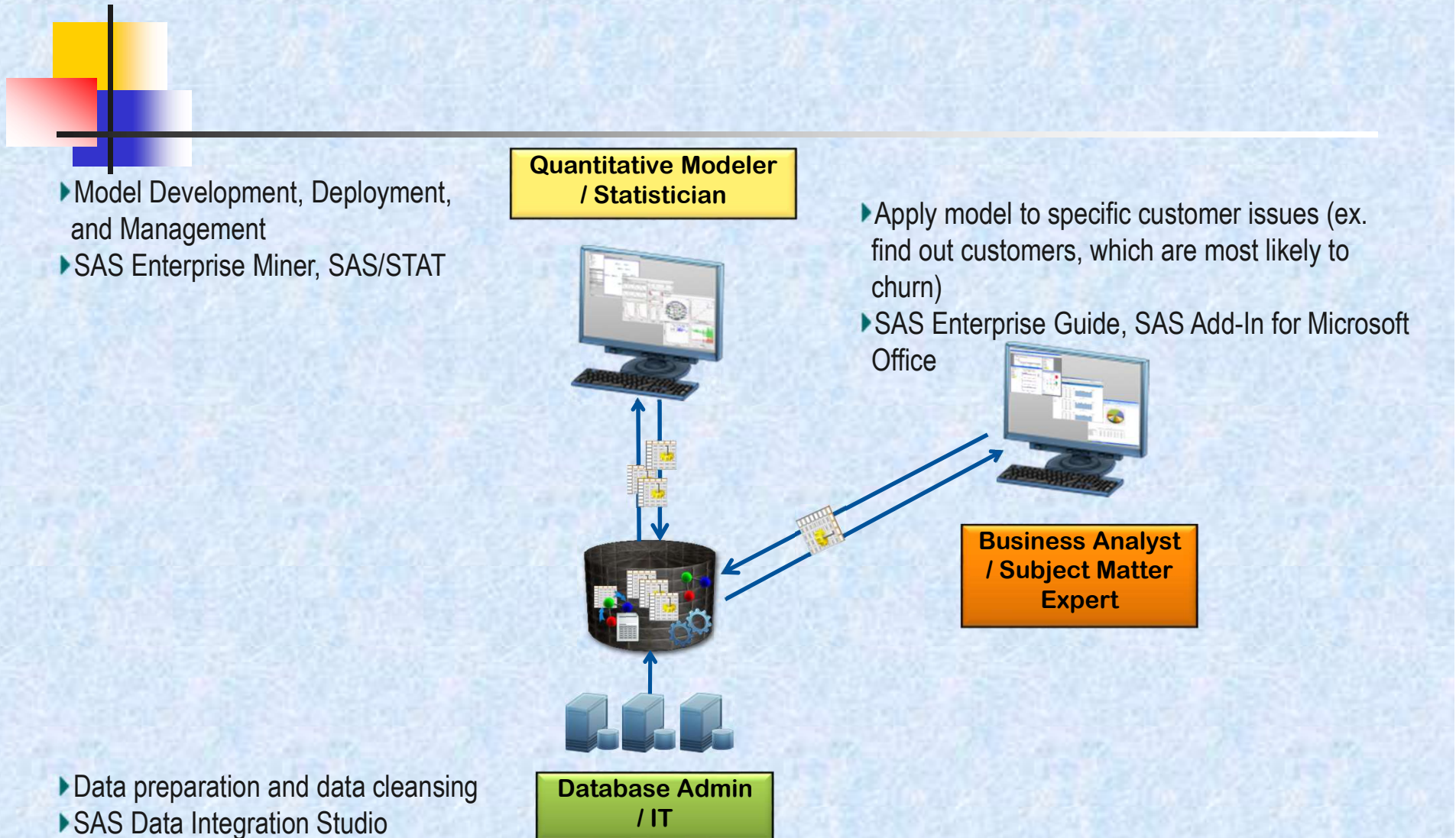
Порівняння моделей



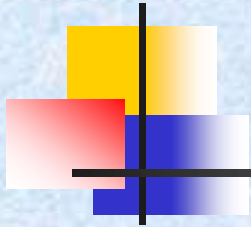
Моделі, які побудовані з використанням техніки data mining порівнюються на відповідність на наявних даних.



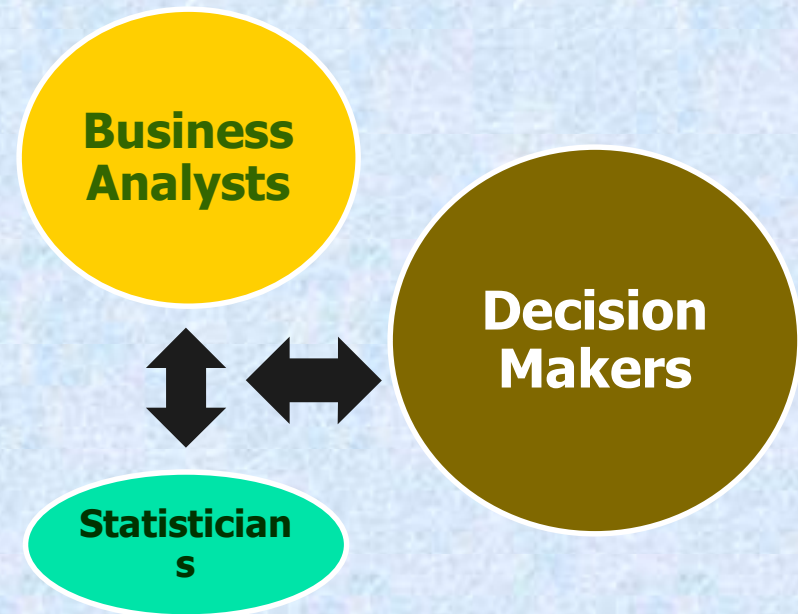
Data Mining and Predictive Analytics



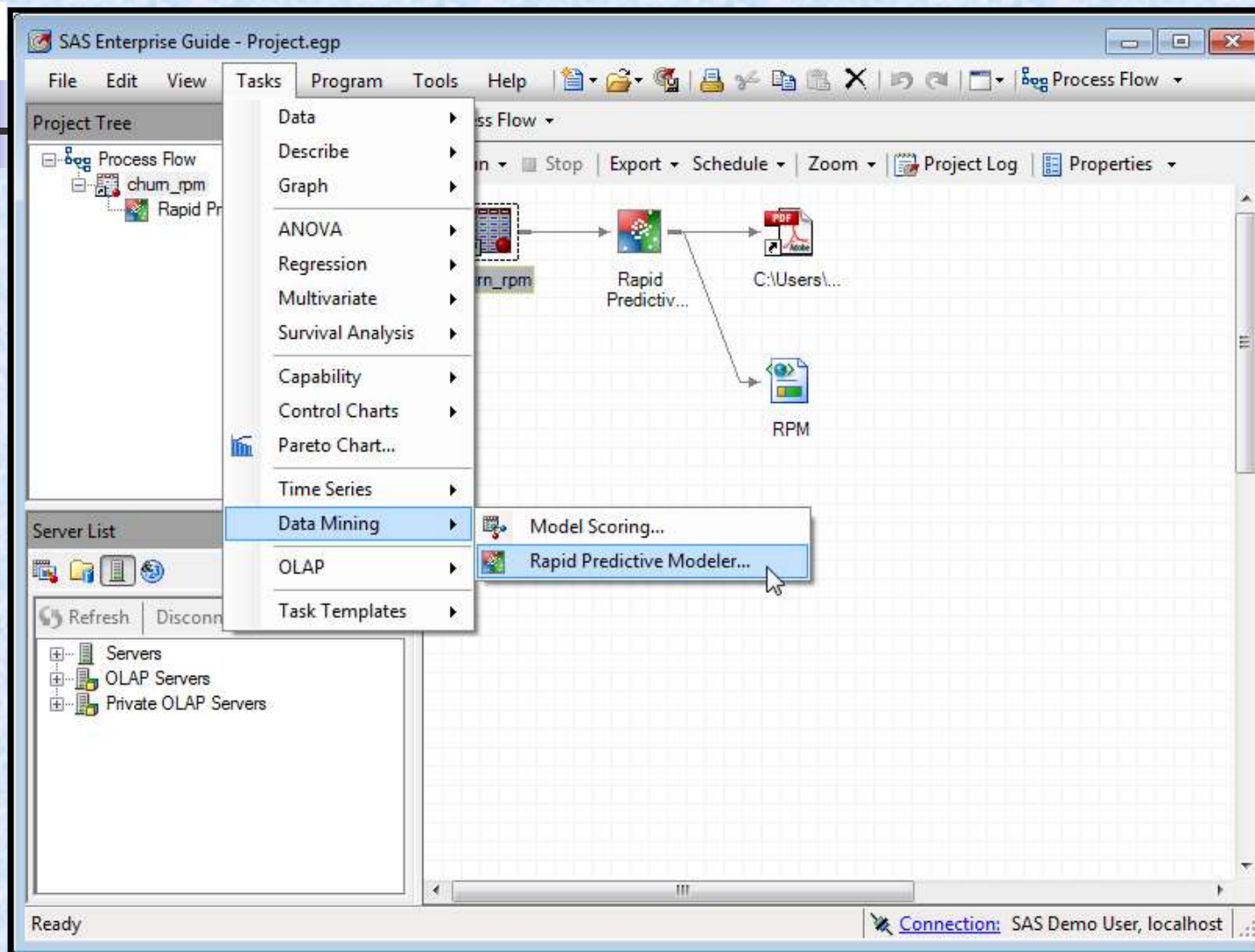
Система SAS Rapid Predictive Modeler



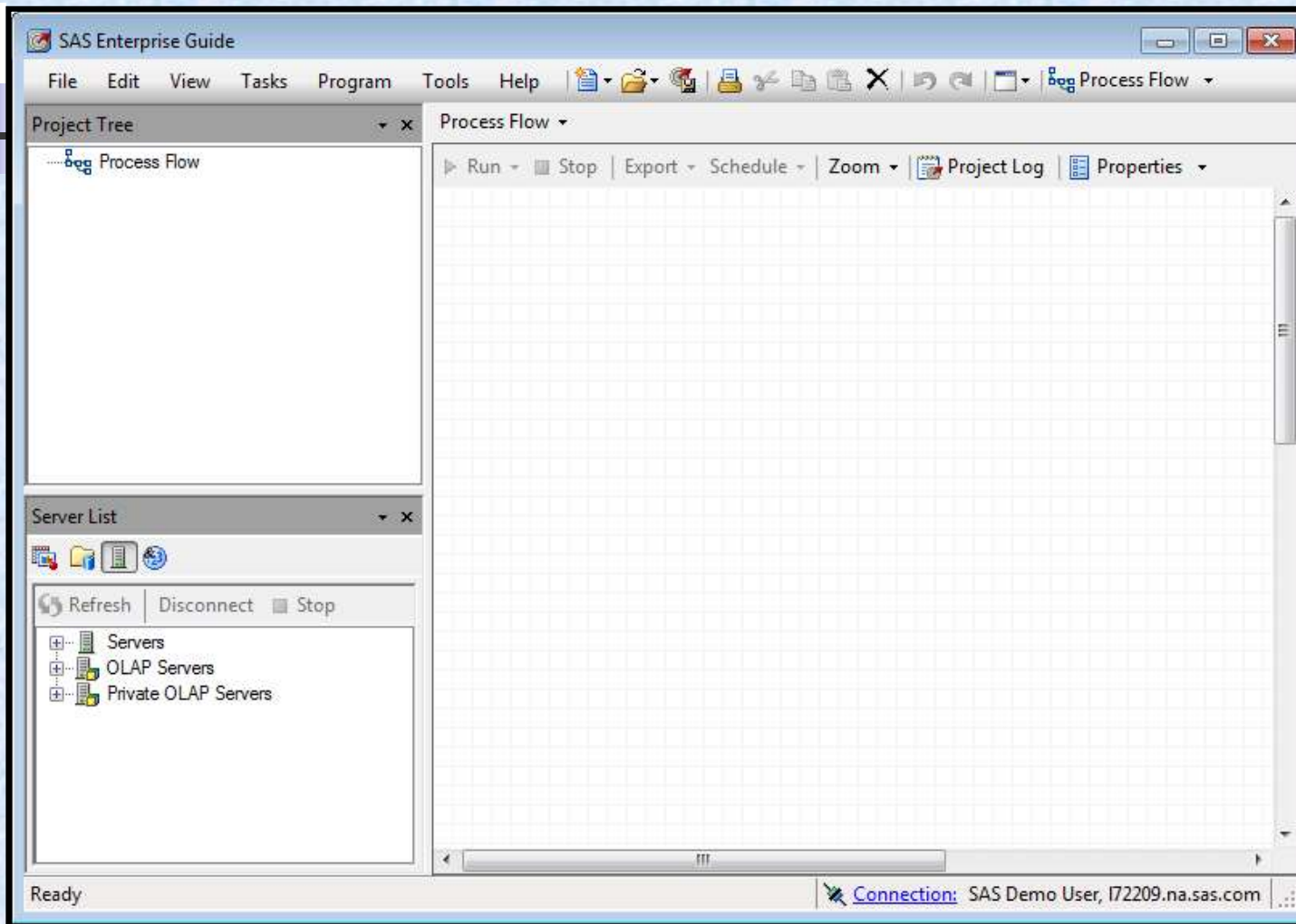
- Будує модель швидко та якісно.
- Модель має прогнозне значення.
- Генерує зрозумілий звіт та діаграми.
- Інтегрує аналітику у процес прийняття рішення.



Як будуються моделі у SAS?

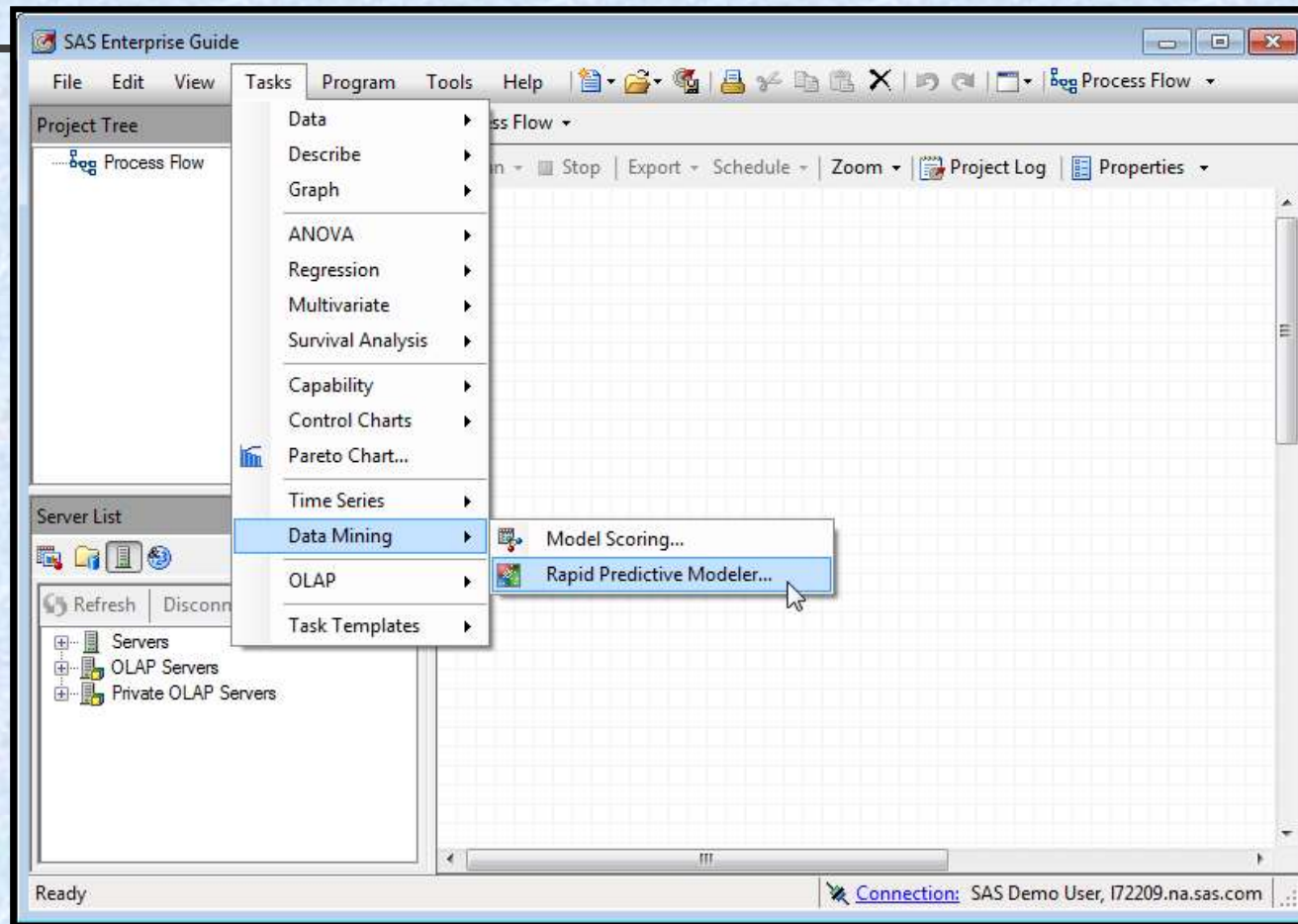


Відкриваємо SAS Enterprise Guide

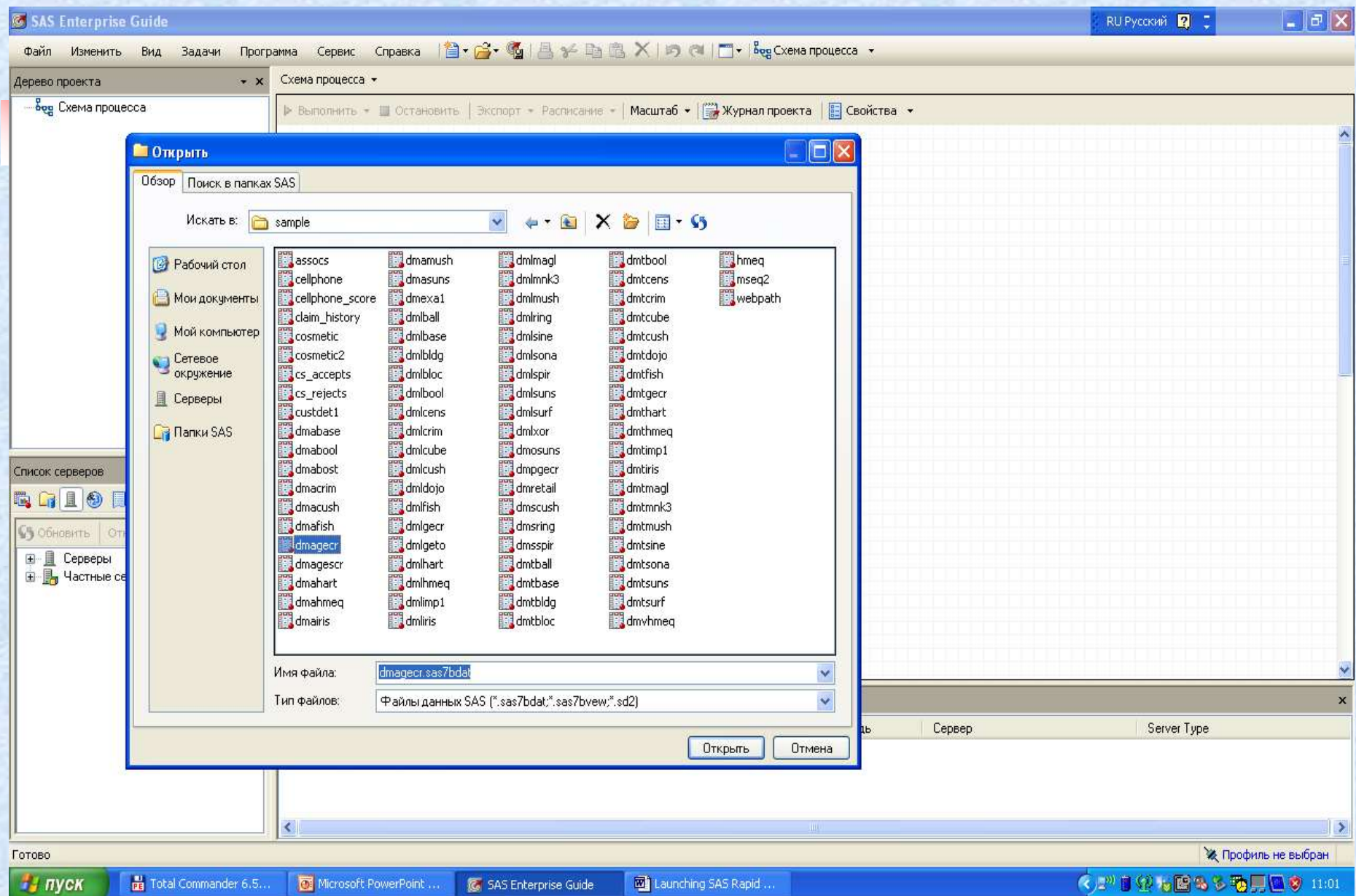


Обираємо в меню SAS Rapid Predictive Modeler

Tasks ⇒ Data Mining ⇒ Rapid Predictive Modeler



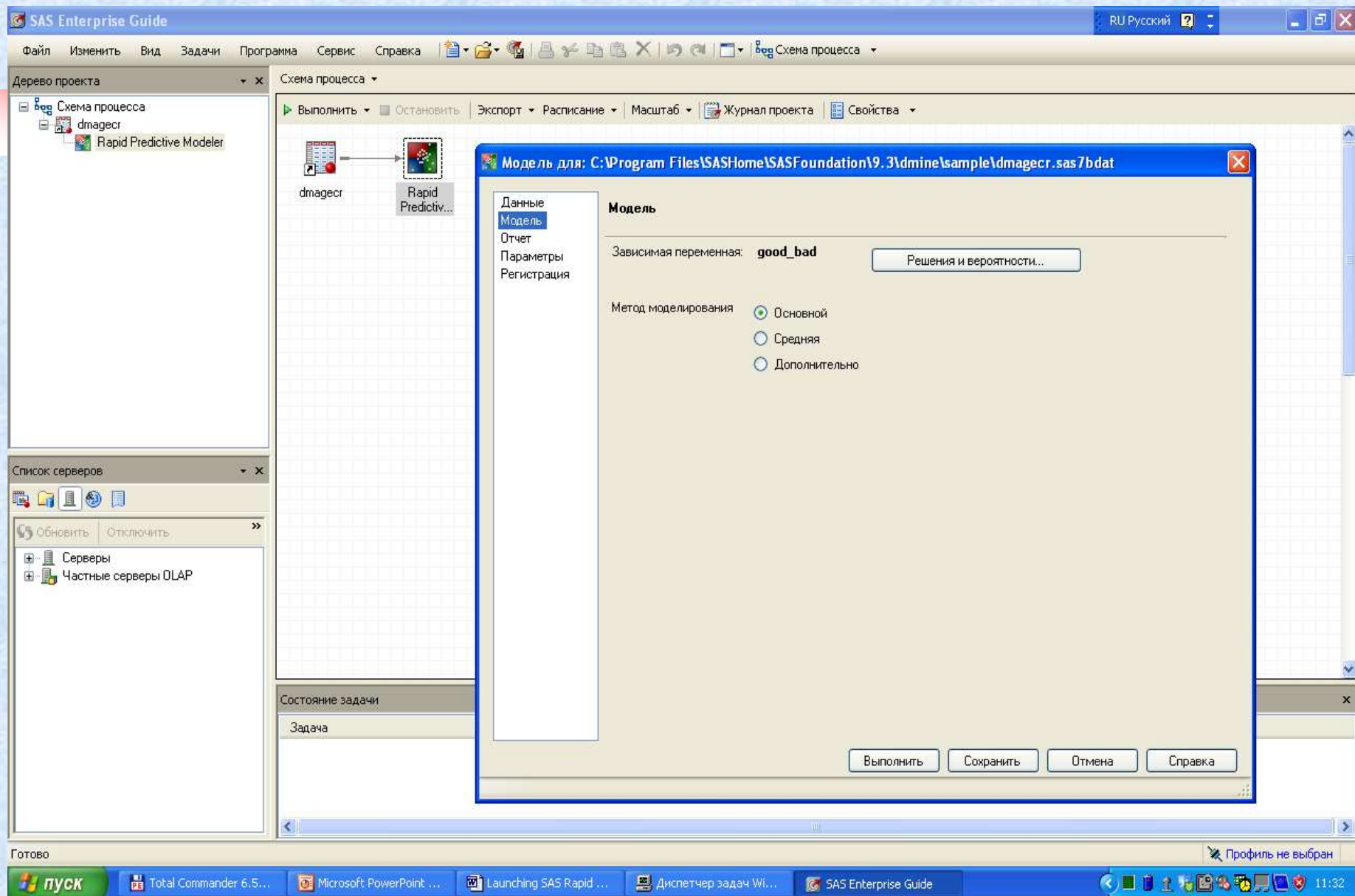
Як будуються моделі у SAS? Спочатку вводяться дані.



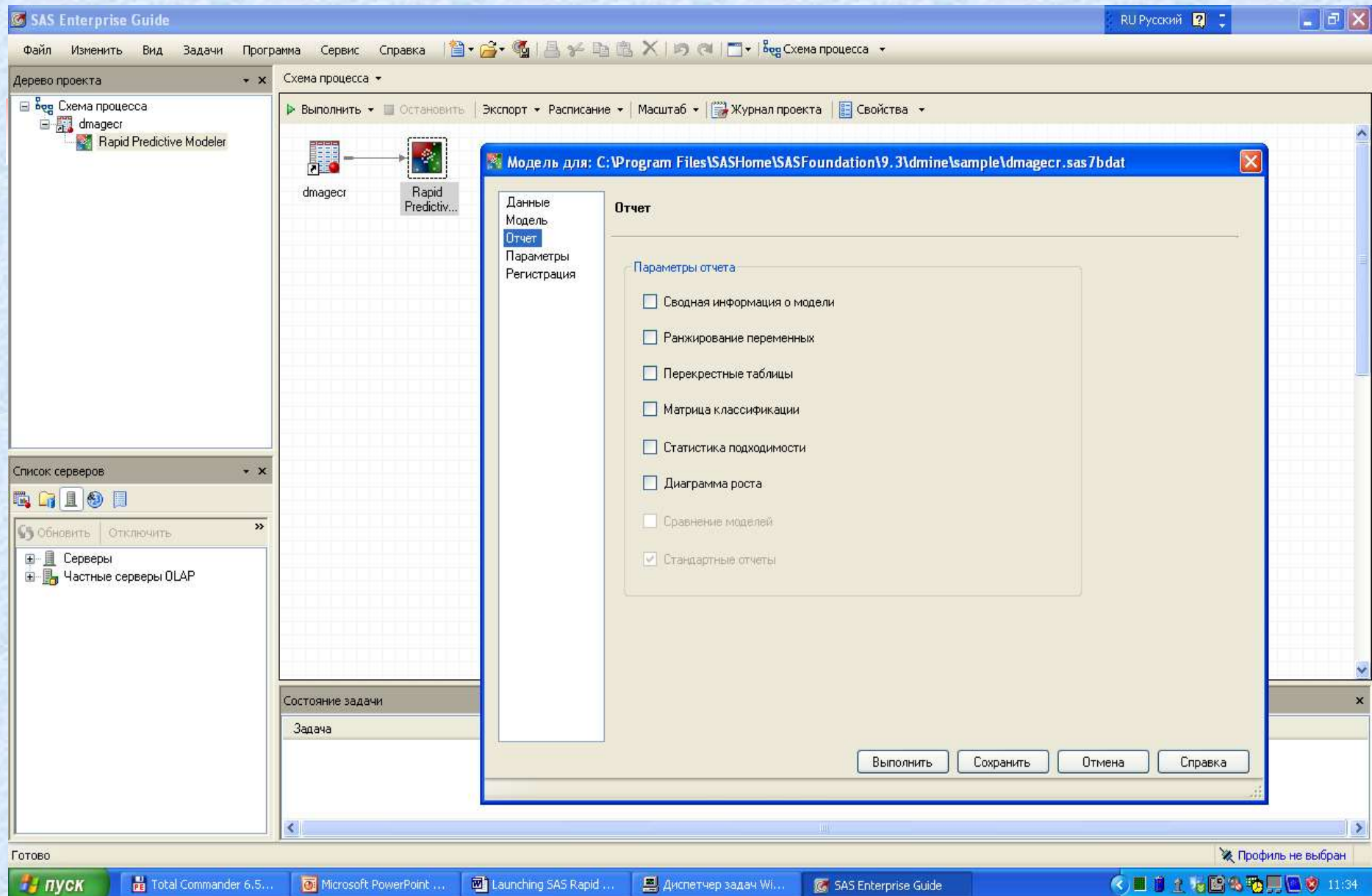
Як будуються моделі у SAS? Змінним призначаються ролі.

The screenshot displays the SAS Enterprise Guide interface. On the left, the 'Project Tree' shows a process named 'dimagecr' with a 'Rapid Predictive Modeler' icon. The main workspace shows a flow from 'dimagecr' to 'Rapid Predictiv...'. A 'Model Properties' dialog box is open, titled 'Модель для: C:\Program Files\SASHome\SASFoundation\9.3\dm\sample\dimagecr.sas7bdat'. The 'Data' tab is selected, showing the data source as 'C:\Program Files\SASHome\SASFoundation\9.3\dm\sample\dimagecr.sas7bdat'. The 'Input Variables' list on the left includes variables like 'purpose', 'amount', 'savings', 'employed', 'installp', 'marital', 'coapp', 'resident', 'property', 'age', 'other', 'housing', 'existor', 'job', 'depends', 'telephon', and 'foreign'. The 'Roles' section on the right, highlighted by a red circle, shows 'Зависимая переменная (Ограничен)' (Dependent Variable (Limited)) with the value 'good_bad' assigned. Other roles listed are 'Подсчет частот (Ограничение: 1)' (Frequency Count (Limit: 1)), 'ID', and 'Исключенные' (Excluded). The 'Table of Variables' button is visible at the bottom right of the dialog. The bottom status bar indicates 'Готово' (Ready) and 'Профиль не выбран' (Profile not selected).

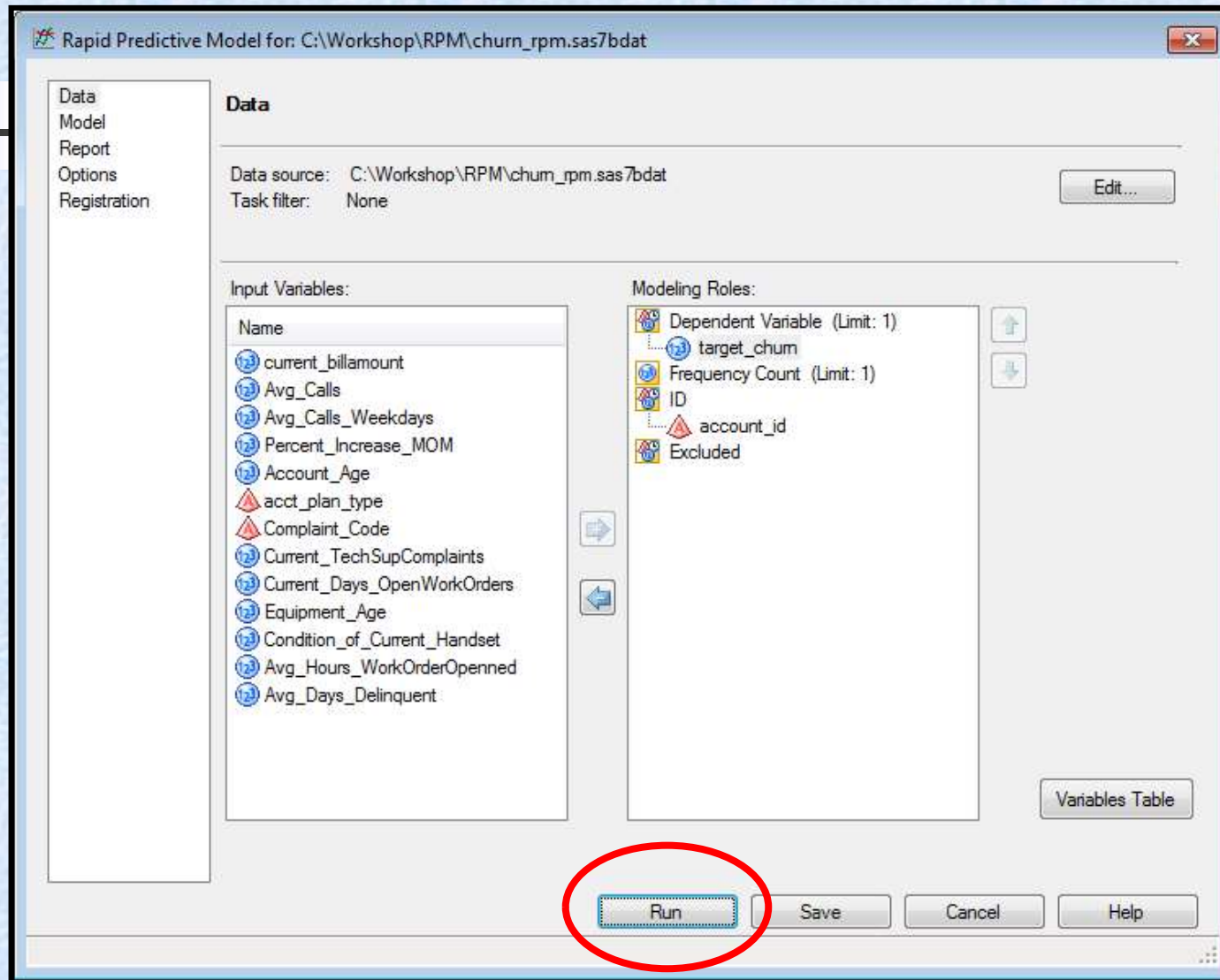
Як будуються моделі у SAS? Встановлюємо параметри моделі: метод побудови.



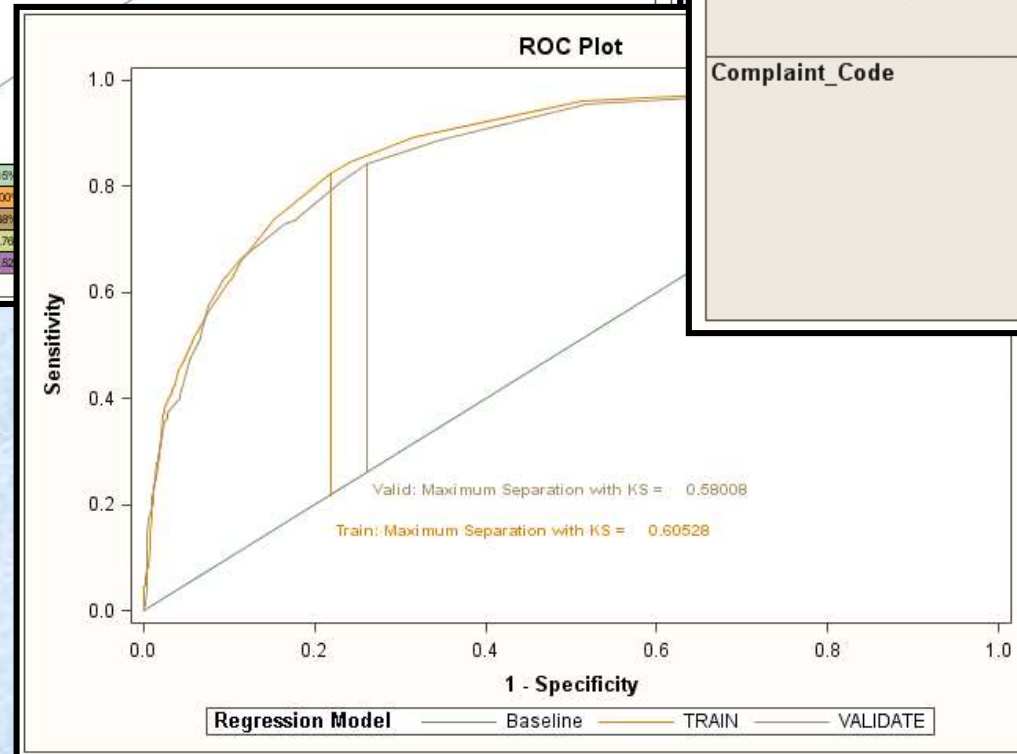
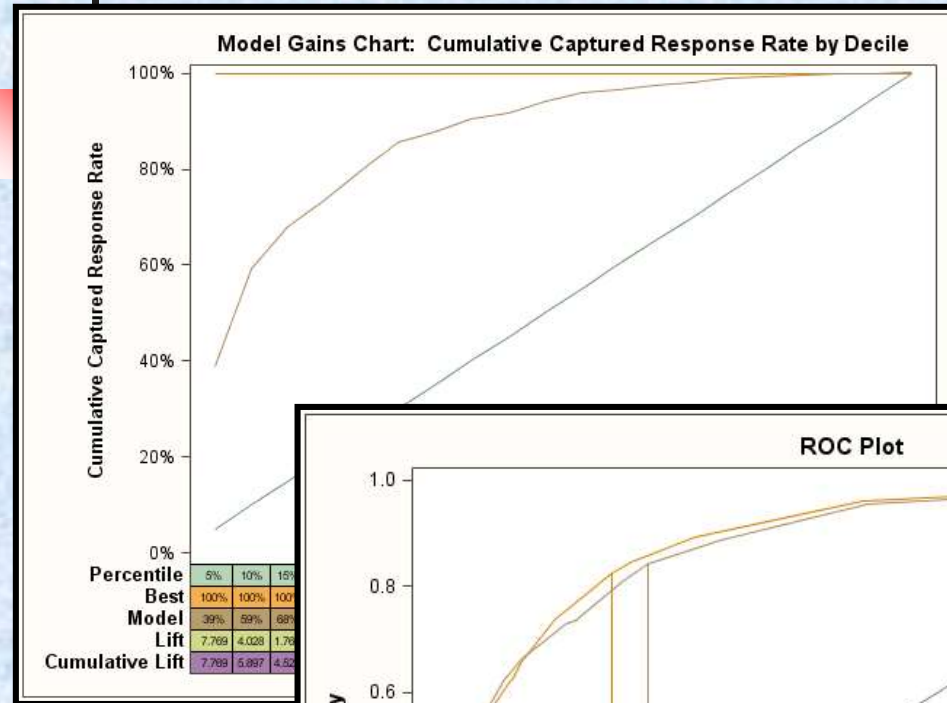
Як будуються моделі у SAS? Встановлюємо параметри моделі:
форму звіту, вигляд збереження моделі (SAS, C++, тощо).



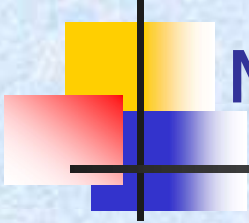
Виконуємо побудову моделі - Run



Аналізуємо результати

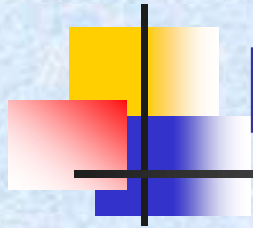


Scorecard		
		Scorecard Points
Account_Age	1: LOW - 21.5	72.00
	2: 21.5 - HIGH	0.00
Avg_Calls_Weekdays	1: LOW - 3532	112.00
	2: 3532 - 59769.833	0.00
	3: 59769.833 - HIGH	1.00
Avg_Days_Delinquent	1: LOW - 7.929	0.00
	2: 7.929 - 11.732	122.00
	3: 11.732 - 14.938	148.00
	4: 14.938 - HIGH	266.00
Percent_Increase_MOM	1: LOW - -0.511	225.00
	2: -0.511 - 2.422	0.00
	3: 2.422 - HIGH	8.00
Complaint_Code	BILLING PROBLEM	214.00
	CALL QUALITY	325.00
	CHECK ACCOUNT	0.00
	INACCURATE SALES INF	113.00
	MOVING	112.00
	PRICING	153.00



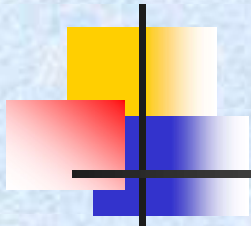
Як у SAS працює блок побудови моделі прогнозування?

- SAS Rapid Predictive Modeler автоматично обробляє аномалії (“виколотки”), нечіткі значення, нечасті явища, автоматично обирає змінні та метод побудови моделі.
- Аналітичні результати представляються у зрозумілому вигляді: scorecard, lift charts, списку змінних моделі.
- Аналітик має можливість ручного вдосконалення моделі (у SAS Rapid Predictive Modeler чи у SAS Enterprise Miner).
- Модель можна зберегти у вигляді SAS коду, або C++ коду для подальшого застосування у роботі.



ВИСНОВОК

- Використання систем бізнес-аналітики у банках та на підприємствах є необхідним. Це дозволяє розробити найбільш ефективну стратегію розвитку, оптимізувати процес документообігу, чітко визначити функціональні обов'язки різних підрозділів, тощо.



Дякую за увагу!



очікуємо на зауваження: vlad.mipt@gmail.com

www.domrachev.in.ua