**Знайомство з Statistica. Робота з таблицями даних і побудова графіків**

Система «STATISTICA», розроблена компанією Statsoft, є однією з найбільш популярних статистичних програм для пошуку закономірностей, прогнозування, класифікації, візуалізації даних. Може застосовуватися в економіці, промисловості, медицині, наукових дослідженнях і інших сферах людської діяльності. Клієнтами Statsoft є найбільші компанії зі світовим ім'ям.

STATISTICA - це інтегрована система аналізу та управління даними. STATISTICA - це інструмент розробки користувальницьких додатків в бізнесі, економіці, фінансах, промисловості, медицині, страхуванні та інших областях. STATISTICA легка в освоєнні і використанні.

У системі існує можливість застосовувати класичні й новітні методи проведення аналізу даних: кластерний, факторний, кореляційний, дисперсійний аналіз, лінійну й нелінійну регресії, нейронні мережі й ін. Візуалізація вихідних, проміжних, вихідних даних може бути здійснена вибором з великої кількості різних графіків, піктографіків і діаграм.

Всі аналітичні інструменти, наявні в системі, доступні користувачеві і можуть бути обрані за допомогою альтернативного інтерфейсу. Користувач може всебічно автоматизувати свою роботу, починаючи з застосування простих макросів для автоматизації рутинних дій аж до поглиблених проектів, що включають у тому числі інтеграцію системи з іншими додатками або Інтернет. Технологія автоматизації дозволяє навіть недосвідченому користувачу налаштувати систему на свій проект.

Процедури системи STATISTICA мають високу швидкість і точність обчислень.

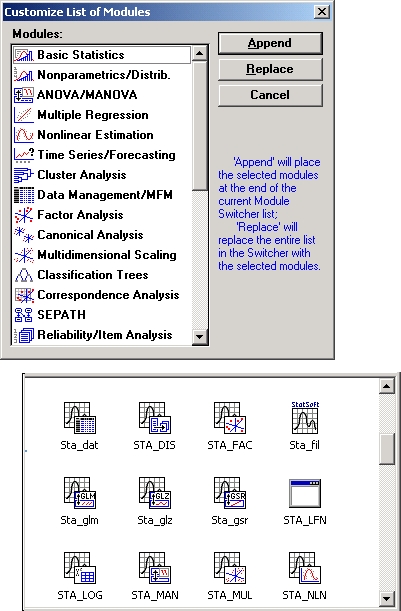
Гнучка і потужна технологія доступу до даних дозволяє ефективно працювати як з таблицями даних на локальному диску, так і з віддаленими сховищами даних.

Система володіє наступними загальновизнаними достоїнствами:

* містить повний набір класичних методів аналізу даних: від основних методів статистики до просунутих методів, що дозволяє гнучко організувати аналіз;
* є засобом побудови додатків в конкретних областях;
* в комплект поставки входять спеціально підібрані приклади, що дозволяють систематично освоювати методи аналізу;
* відповідає всім стандартам Windows, що дозволяє зробити аналіз високоінтерактівним;
* система може бути інтегрована в Інтернет;
* підтримує web-формати: HTML, JPEG, PNG;
* легка в освоєнні, і як показує досвід, користувачі з усіх областей застосування швидко освоюють систему;
* дані системи STATISTICA легко конвертувати в різні бази даних і електронні таблиці;
* підтримує високоякісну графіку, що дозволяє ефектно візуалізувати дані і проводити графічний аналіз;
* є відкритою системою: містить мови програмування, які дозволяють розширювати систему, запускати її з інших Windows-додатків, наприклад, з Excel.

STATISTICA складається з набору модулів, в кожному з яких зібрані тематично зв'язкові групи процедур. При перемиканні модулів можна або залишати відкритим тільки одне вікно програми STATISTICA, або всі викликані раніше модулі, оскільки кожен з них може виконуватися в окремому вікні (як самостійний додаток Windows).

При виконанні модулів STATISTICA як самостійних додатків в будь-який момент часу в будь-якому модулі є прямий доступ до «загальних» ресурсів (таблиць даних, мов BASIC і SCL, графічних процедур).



**Формати Дата і Час**. У файлах даних системи (які організовані як бази даних) формат відображення значень застосовується до всієї змінної, а не до окремих осередків (як в Excel). Тому значення, які в Excel були відформатовані як дати, у файлі системи STATISTICA будуть відображатися як цілі значення (наприклад, 34092 замість May 3, 1993), якщо для відповідних змінних не встановлений формат **Дата** або **Час** .

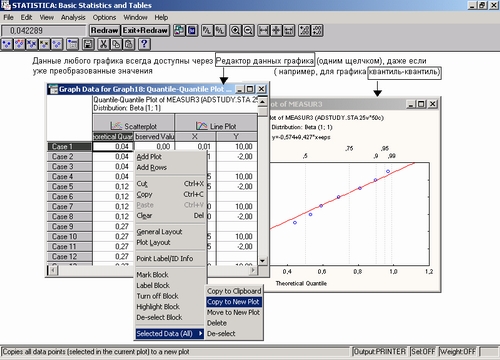
**Власний графічний формат STATISTICA.** Графічні файли системи STATISTICA мають розширення \*.stg. Їх основна відмінність від метафайлов і растрових зображень полягає в тому, що вони містять не тільки картинку, але і всю інформацію, необхідну для налаштування графіка та аналізу даних. Тут записані всі представлені на графіку дані, їх зв'язку, рівняння підгонки, параметри впроваджених об'єктів, зв'язку графіків і малюнків і т. п. Записані в такому форматі графіки можна згодом відкрити в будь-якому з модулів системи STATISTICA для продовження налаштування і аналізу даних. Крім того, їх можна роздрукувати в пакетному режимі за допомогою команди **Друк** файлів з випадаючого меню **Файл** . Графічні файли у власному форматі системи STATISTICA можна динамічно пов'язати з документами додатків Windows за допомогою методів OLE.

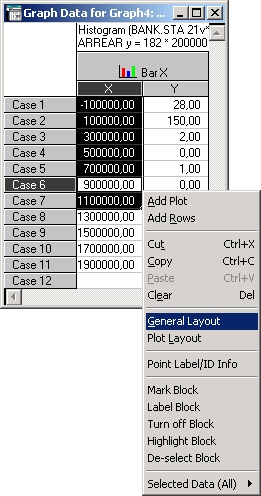
**Експорт через буфер обміну (вставка або спеціальна вставка методами OLE).** Використання буфера обміну - це найшвидший спосіб експорту графіка в інший додаток. При копіюванні в буфер обміну створюється три графічних представлення об'єкта: у власному форматі STATISTICA, у форматі метафайлу Windows і у форматі растрового зображення. Кожне з них може бути використано в інших додатках.

Графіки системи STATISTICA можуть бути присутніми в інших додатках (редакторах або електронних таблицях) як в якості пов'язаних, так і впроваджених об'єктів. При використанні методів OLE вони зберігають свій зв'язок з системою STATISTICA і, отже, можуть інтерактивно редагуватися в рамках інших програм.

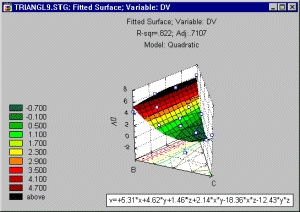
**Доступ до всіх даних графіка.** Дані, представлені на графіках системи, можна безпосередньо переглядати і змінювати незалежно від їх типу у вбудованому редакторі даних графіка. Це можуть бути вихідні дані, частини таблиці результатів або ряд розрахованих значень (наприклад, імовірнісний графік).

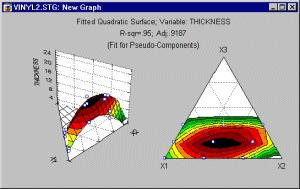
Для кожного графіка створюється пов'язане з ним «дочірнє» вікно Редактора, яке закривається разом зі своїм графічним вікном. Редактор організований у вигляді груп стовпців, що представляють окремі залежності даного графіка.



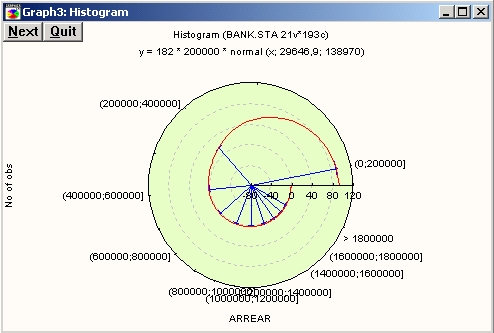


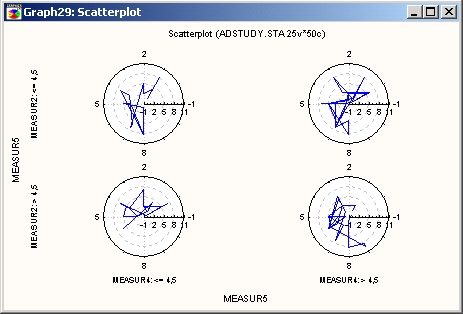
**Тернарні графіки поверхні і карти ліній рівня.** При виведенні результатів аналізу щодо складання сумішей в модулі Планування експерименту можна побудувати тернарние графіки у вигляді тривимірних поверхонь або карт ліній рівня.





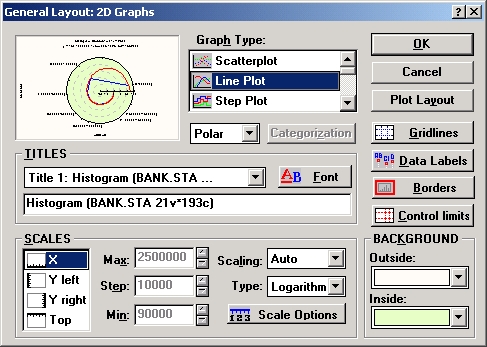
**Графіки в полярних координатах.** Деякі типи графіків можна побудувати в полярних координатах. До них відносяться графіки розсіювання, лінійні графіки і послідовні вкладені графіки з підміню Статистичні 2М графіки (викликається з випадаючого меню Graphics).





У полярних координатах можна побудувати і категоризувати графіки.

Багато графіків, намальовані у звичайній прямокутній системі координат, можна представити у полярних координатах. Для цього потрібно встановити відповідний перемикач в діалоговому вікні Загальна розмітка в положення Полярні.

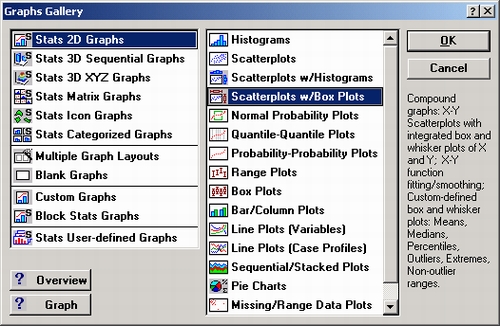


Для вставки будь-яких графічних об'єктів, сумісних з системою Windows, можна використовувати операції вставки за допомогою буфера обміну (включаючи зв'язування та впровадження методами OLE). Ці операції можна здійснювати над растровими об'єктами, метафайлами Windows, графіками у форматі STATISTICA, а також будь-якими OLE-сумісними об'єктами.

За допомогою описаних вище операцій з буфером обміну на графіки STATISTICA можна помістити дуже великий текстовий об'єкт (наприклад, звіт довжиною кілька сторінок). Цей текст редагується і змінюється у вікні Редактор тексту графіка системи STATISTICA або у відповідному додатку, який є сервером в методі OLE.

Всі описані в попередньому розділі операції вставки і використання буфера обміну застосовні до будь-якого сумісного з Windows графічного об'єкту, а операції зв'язування та впровадження виконуються для всіх об'єктів, що підтримують методи OLE.

**Галерея графіків STATISTICA**. За допомогою цієї кнопки відкривається діалогове вікно Галерея графіків STATISTICA. Ця кнопка присутня в діалоговому вікні кожного типу графіків.



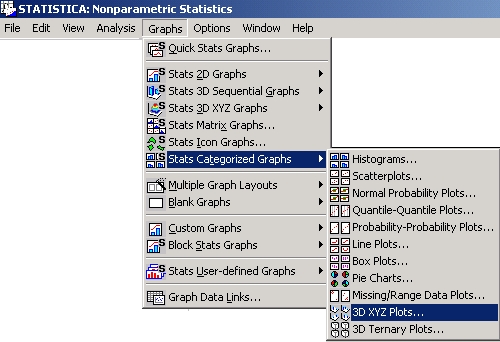
Звідси швидко і легко викликаються всі статистичні і призначені для користувача графіки, порожні графічні вікна та статистичні графіки користувача. Для цього потрібно виділити назву потрібного типу графіка і двічі клацнути на ньому (або натиснути кнопку **ОК** ).

**Користувацькі і статистичні графіки.** Окрім спеціалізованих графіків, які викликаються безпосередньо з підсумкового діалогового вікна будь-якої програми статистичної обробки, існують ще два основних типи графіків, доступних з меню або панелі інструментів будь-якої таблиці: користувацькі графіки та статистичні (і швидкі статистичні) графіки.

**Користувацькі графіки**. Користувацький графік дає можливість відобразити будь-яку задану користувачем комбінацію значень з таблиць вихідних даних або таблиць результатів (а також з будь-якої комбінації їх рядків і / або стовпців). У меню пропонується п'ять типів таких графіків: 2М користувальницькі графіки, ЗМ користувальницькі послідовні графіки, ЗМ користувальницькі діаграми розсіювання і поверхні, призначені для користувача матричні графіки і призначені для користувача піктографікі. При виборі одного з них відкривається відповідне діалогове вікно, де для відображення на графіку можна задати діапазон даних поточної таблиці. Вміст цього діалогового вікна залежить від обраного типу користувацького графіка. Початковий вибір даних для побудови графіка, пропонований в цьому діалоговому вікні, визначається положенням курсора в поточній таблиці. У кожному діалоговому вікні користувацького графіка при завданні параметрів передбачена можливість вибору певного виду графіка (в рамках основного типу). Вид графіка також можна підібрати і після побудови (за допомогою діалогових вікон Загальна розмітка або Розміщення графіка, які відкриваються при подвійному натисканні мишею на області фону графічного вікна або при виборі відповідного рядка меню **Розмітки**).

http://www.hr-portal.ru/statistica/4.gif

**Статистичні графіки.** На відміну від користувацьких графіків, які представляють собою засіб наочного відображення числових даних будь-яких таблиць (вихідних даних, результатів), статистичні графіки пропонують сотні заздалегідь визначених типів графічних представлень, що включають аналітичне узагальнення статистичних даних. Вони викликаються з діалогового вікна Галерея графіків, яке відкривається за допомогою однойменної кнопки панелі інструментів або з випадаючого меню **Графіка** .

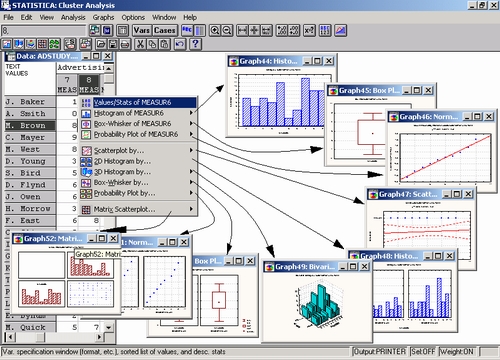


При побудові таких графіків використовуються значення безпосередньо з файлу даних, які не залежать від змісту поточної таблиці, виділення блоків і положення курсору. При цьому пропонуються або стандартні методи графічного аналізу вихідних даних (різні графіки розкиду значень, гістограми, графіки середніх значень, наприклад, медиан), або стандартні аналітичні методи досліджень (графіки нормальної щільності розподілу, імовірнісні графіки з виключеним трендом або графіки довірчих інтервалів ліній регресії) . При побудові статистичних графіків програма враховує умови вибору і ваги спостережень.

**Швидкі статистичні графіки** . Найбільш широко використовувані типи статистичних графіків (викликаються з меню **Графіка**) представлені в меню **Швидкі статистичні графіки**. Ці списки графіків не надають такий широкий спектр можливостей, як меню **Статистичні графіки**, але на відміну від останніх спрощують і прискорюють процедуру побудови графіка. Швидкі статистичні графіки:

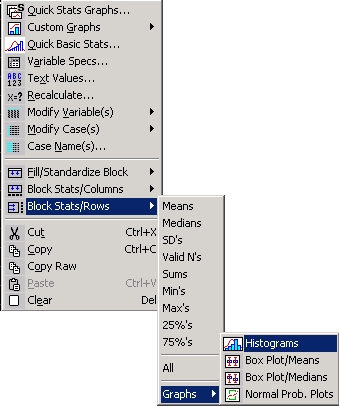
* викликаються з контекстних меню або за панелі інструментів будь-якої таблиці (зазвичай вони не вимагають звернення до випадних меню або діалогових вікон),
* не вимагають від користувача вибору змінних (цей вибір визначається поточним положенням курсору в таблиці) і проміжної налаштування параметрів (формат відповідних графіків визначається за замовчуванням).

При виборі пункту **Швидкі статистичні графіки** (за допомогою кнопки на панелі інструментів http://www.hr-portal.ru/statistica/5.gif контекстного меню або з меню **Графіка**) з'являється меню вибору статистичного графіка для поточної змінної таблиці, тобто тієї, на яку зараз вказує курсор.



Якщо курсор не вказує ні на одну із змінних, то перед побудовою будь-якого графіка з меню **Швидкі статистичні графіки** буде запропоновано вибрати змінну із списку.

**Блокові статистичні графіки**. Ці типи (для користувача) графіків викликаються з пунктів контекстних меню. **Статистики блоку за стовпцями** і **Статистики блоку по рядках** або з діалогового вікна **Галерея графіків.**



Будь-який з цих варіантів дає можливість побудувати підсумковий статистичний графік для виділеного блоку, щоб порівняти значення в рядках (Статистики блоку по рядках) або в стовпцях таблиці (Статистики блоку за стовпцями). Даний тип графіків схожий на ті користувацькі графіки, на яких відображаються дані поточного блоку таблиці.

**Інші спеціалізовані графіки**. Крім стандартного набору швидких статистичних графіків деякі таблиці дозволяють будувати і більш спеціалізовані статистичні графіки (наприклад, тимчасові послідовності в модулі **Тимчасові ряди**, піктографіки регресійних залишків, а також контурні графіки в модулі **Кластерний аналіз**). Як вже згадувалося раніше, спеціалізовані графіки, які пов'язані не з конкретною таблицею результатів, а з певним методом аналізу даних (наприклад, графіки апроксимуючих функцій в модулі **Нелінійне оцінювання** або середніх в модулі **Дисперсійний аналіз)** , викликаються безпосередньо з діалогового вікна з результатами аналізу (тобто з вікна, що містить вихідні параметри використовуваного методу обробки даних).

**Налаштування графіка до і після його побудови** . Будь-які зміни параметрів графіка в STATISTICA здійснюються з активного графічного вікна (після відображення графіка на екрані). Як правило, спочатку має сенс побудувати графік, прийнявши значення параметрів за замовчуванням, а потім уже вносити різні зміни. Однак у тих рідкісних випадках, коли побудова графіка займає надто багато часу (при створенні складних складових графічних зображень або обробці великих наборів даних), можна втрутитися в цей процес, щоб зробити необхідні настройки. Перервати малювання можна одним натисканням клавіші або клацанням миші в будь-якому місці екрану, а потім продовжити його після введення необхідних змін.

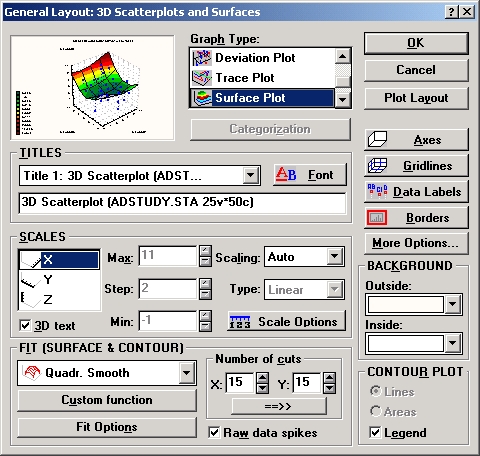
Передбачено два основні методи налаштування графіка - додавання та редагування власних графічних об'єктів, зміна структурних елементів графіка.

Незалежно від способу створення графіка для його налаштування і зміни можна використовувати будь-які можливості, передбачені в системі STATISTICA. До будь-якого графіку можна додати новий графік, об'єднати його з іншим графіком, помістити в нього зв'язаний або впроваджений об'єкт. Крім того, графік можна будь-яким чином змінювати, малювати на ньому і використовувати різні методи підгонки функцій. Ці ж методи настройки доступні при роботі з графіками, що були попередньо збережені і викликані з дискового файлу.

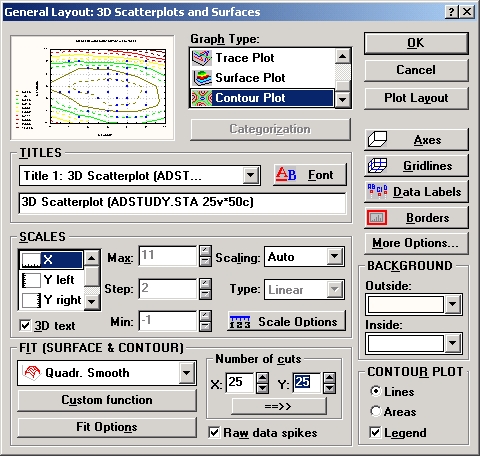
**Налаштування статистичного графіка до і після його побудови.** У розділі ***Як налаштувати графік STATISTICA*** показано, що більшість можливостей налаштування (сотні різних варіантів графічного представлення) доступні безпосередньо після побудови графіка. Для цього досить клацнути на конкретному елементі графіка або вибрати відповідний пункт у діалогових вікнах **Загальна розмітка** або **Розміщення графіка** , що викликаються з меню **Розмітки** .

У той же час окремі параметри, які визначають джерело даних, потрібно задати до побудови графіка, наприклад, змінні, метод категоризації, значення міток, імена спостережень, мітки осей. У даному прикладі перед побудовою графіка потрібно вибрати змінні і метод, категоризації, а також при необхідності задати значення деяких параметрів за допомогою кнопки **Параметри** (яка тут не використана).

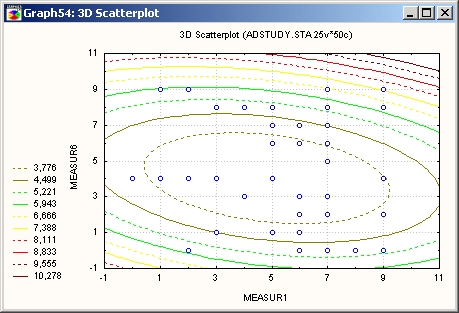
Після побудови графіка при клацанні на будь-якому місці фону графічного вікна з'явиться діалогове вікно **Загальна розмітка** , в якому регулюються параметри загального розташування графіка.



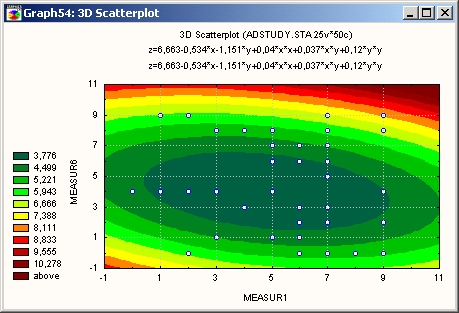
У цьому вікні можна змінити тип графіка і задати побудову карти ліній рівня (використовуйте для цього поле **Тип графіка** ). Крім того, можна змінити параметр Число перетинів з встановленого за замовчуванням зі значенням 15 x15 на 25 х 25 (цей параметр визначає точність побудови карти ліній рівня):



Після внесення змін натисніть **ОК** , і ви побачите новий графік:

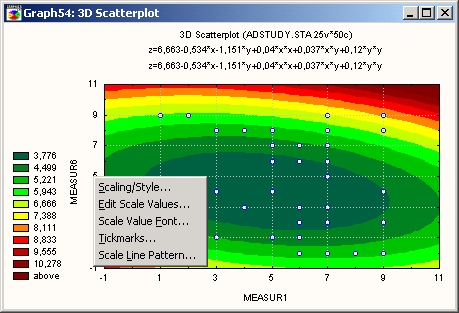


Для якнайшвидшого відображення та всебічного форматування рівнянь функцій краще використовувати діалогове вікно **Параметри** , яке викликається з діалогового вікна **Статистичні графіки.** Натисніть **ОК** , і ви побачите змінений графік:

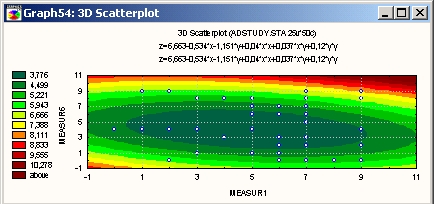


Найпростіший (і найшвидший) спосіб зміни параметрів будь-якого елемента - це подвійне клацання на ньому кнопкою миші. Крім того, за допомогою одного клацання правою кнопкою миші на даному об'єкті можна викликати відповідне йому контекстне меню.

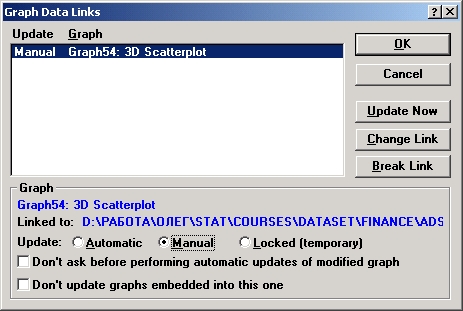
Наприклад, при натисканні правою кнопкою миші на одній з осей графіка з'явиться показане нижче контекстне меню, в якому пропонується вибір варіантів настройки для даної осі:



На показаному нижче графіку за допомогою кнопки панелі інструментів http://www.hr-portal.ru/statistica/6.gif підібрані інші пропорції графічного вікна, крім того, змінено статус умовних позначень з фіксованого на переміщуваний, а їх текст відредагований, упорядкований і переміщений на інше місце.



Графіки можуть автоматично оновлюватися при зміні файлу даних. Усі графіки зберігають зв'язки з таблицями вихідних даних, за якими вони побудовані. При цьому, якщо оновлення не відбувається вручну і зв'язки не скасовані, графік автоматично оновлюється при зміні вихідних даних. Для управління зв'язками є спеціальне діалогове вікно **Зв'язки даних і графіка** викликається з меню **Графіка**.



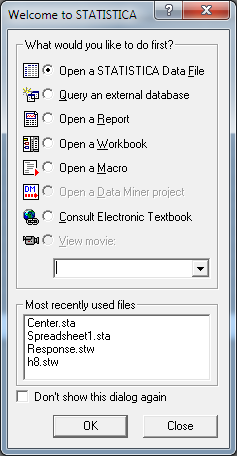
Тут можна встановити автоматичний режим зв'язку, коли графік автоматично оновлюється при зміні даних, за якими він побудований. Можна також задати режим **Вручну** або тимчасово заблокувати зв'язок. Крім того, можна встановити режим Зв'язок з поточним файлом даних і побудувати такий же графік або серію графіків для інших файлів даних. Спосіб зв'язку можна глобально змінити за допомогою команди меню **Сервіс**.

**Графічний формат STATISTICA.** Графіки і малюнки можуть бути збережені в графічному форматі STATISTICA у файлі з розширенням \*. stg. Для цього використовуються команди **Зберегти** та **Зберегти як ..** . з випадаючого меню **Файл** . Саме цей формат рекомендується для запису графічного файлу, якщо передбачається надалі знову відкривати його в системі STATISTICA чи приєднувати до інших додатків методами OLE. На відміну від інших графічних форматів формат STATISTICA зберігає не тільки саму картинку, але і Редактор даних графіка з усіма представленими на графіку даними, всі аналітичні параметри (рівняння підгонки, еліпси і пр.), а також інші параметри, що дозволяють згодом продовжити аналіз графічних даних . Цей формат найбільш зручний при зв'язуванні або впровадженні графіка в інший графік STATISTICA. Збережені в даному графічному форматі файли можна роздрукувати в пакетному режимі за допомогою команди **Друк файлів** з випадаючого меню **Файл**.

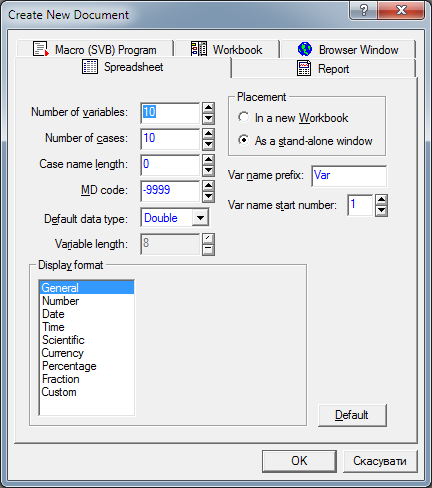
**Завдання 1. Робота з таблицями**

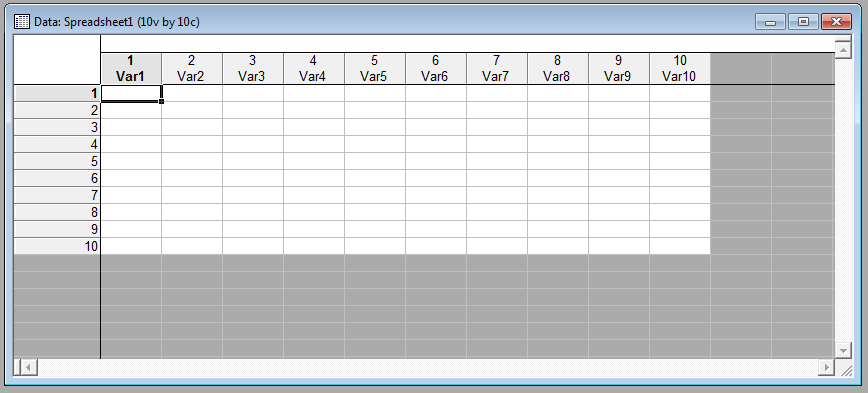
1. Створіть новий лист.

Запускаємо програму Statistica. Натискаємо Close у вікні, показаному на рисунку

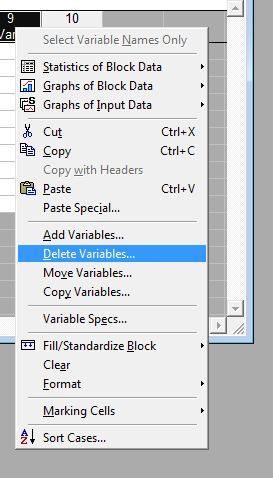


, після чого в меню обираємо File-New



Натискаємо ОК

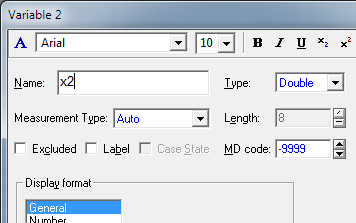
1. Змініть кількість рядків (випадків, cases) і стовпців (змінних, variables) таблиці за допомогою контекстного меню (Delete Variables, Case Management-Delete Cases)



Для додавання стовпців і рядків двічі клацаємо на сіру область таблиці

1. Змініть імена змінних var1, var2… на будь-які інші (наприклад x1, x2 і т.д).

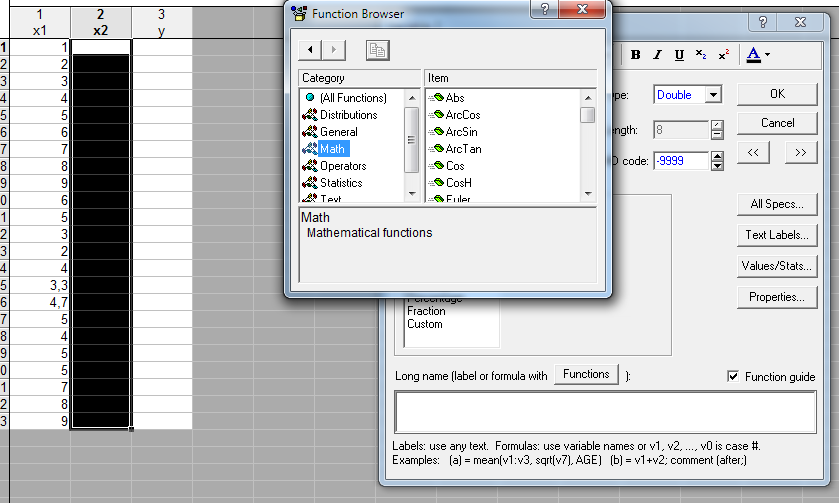
Подвійним клацанням на назву викликаємо вікно з властивостями змінної. В полі Name присвоюємо нову назву змінній.



1. Змініть ширину колонок таблиці.
2. Заповніть першу колонку таблиці довільними значеннями з клавіатури
3. Задайте значення другої змінної за допомогою формули

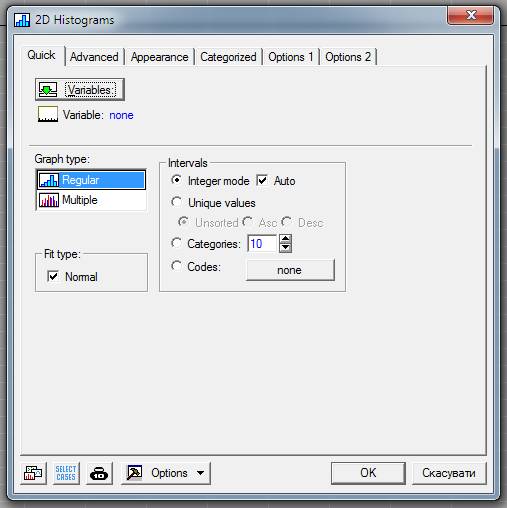
Відкриваємо вікно властивостей, прописуємо функцію в полі Long Name, натискаємо на Functions і обираємо будь-яку з стандартних функцій. v1, v2 і т.д. позначають номер рядка.

Наприклад =Cos(v1) розраховує косинус від значень в першому рядку.

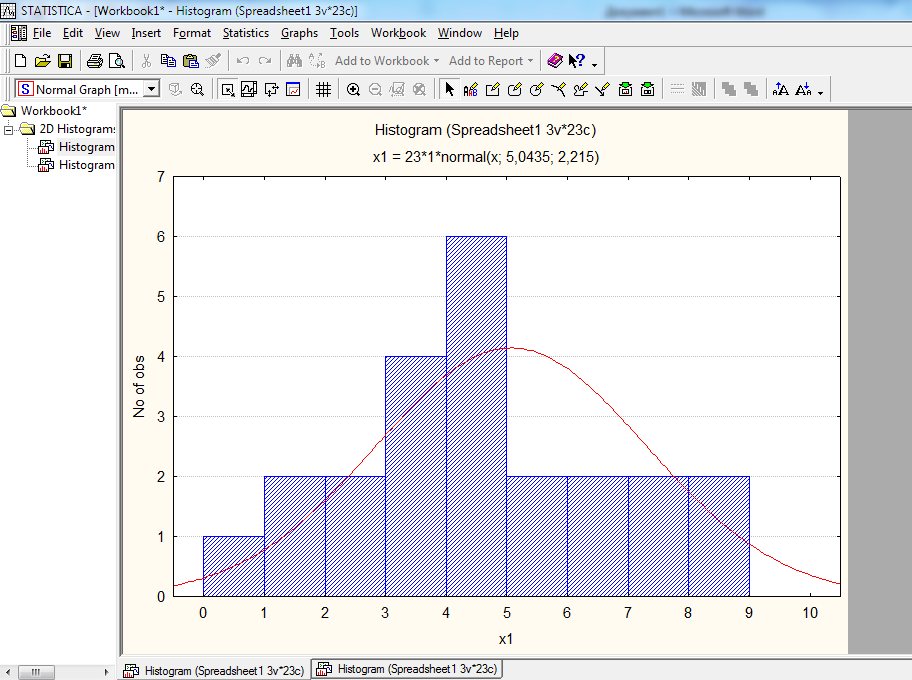


**Завдання 2. Побудова графіків**

1. Створіть нову таблицю з трьома змінними. Перейменуйте їх на x, y і z.
2. Заповніть таблицю довільними числовими значеннями.
3. Оберіть в меню Graphics-Histograms



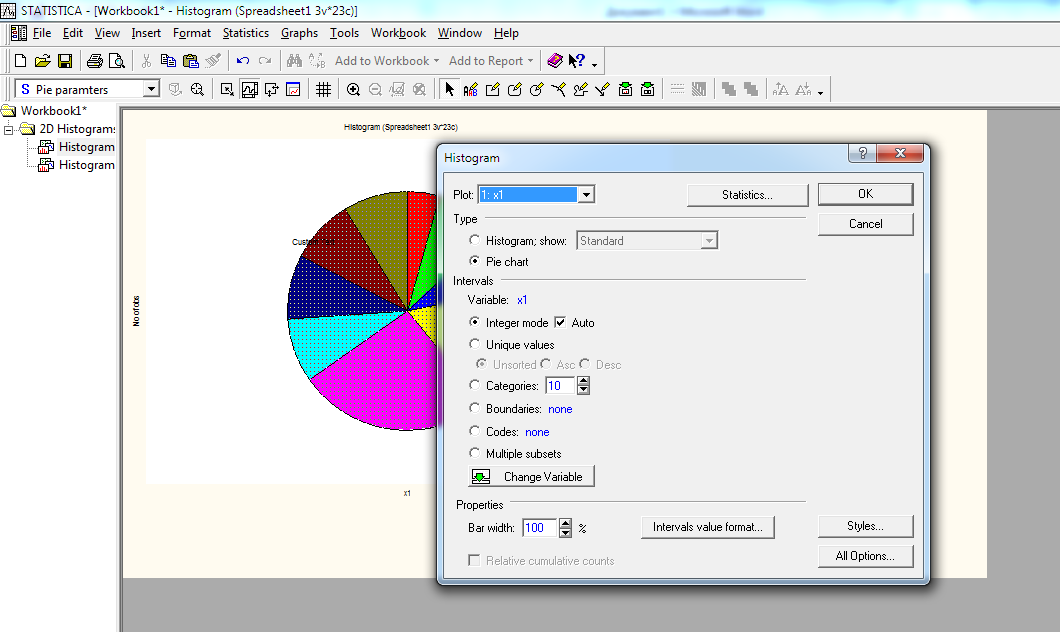
1. У новому вікні, що відкрилося, оберіть змінні за якими будете будувати графік (Кнопка Variables)
2. Натисніть OK



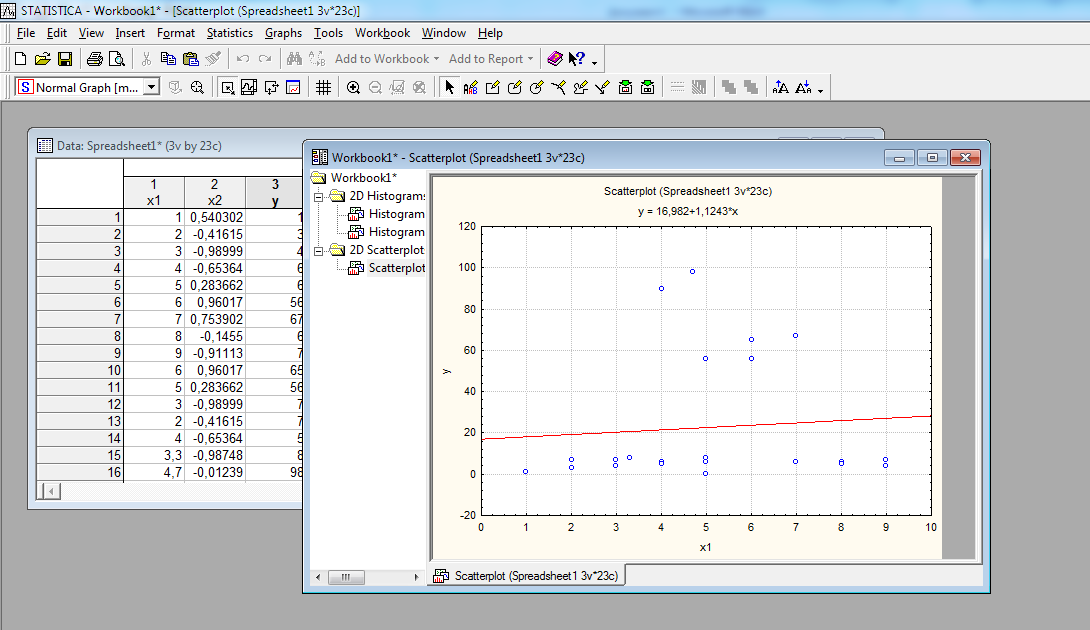
1. Змініть властивості графіка, такі як колір, товщина лінії і т.д.

Подвійним плацанням на область графіка викликаємо вікно з налаштуваннями властивостей графіка.

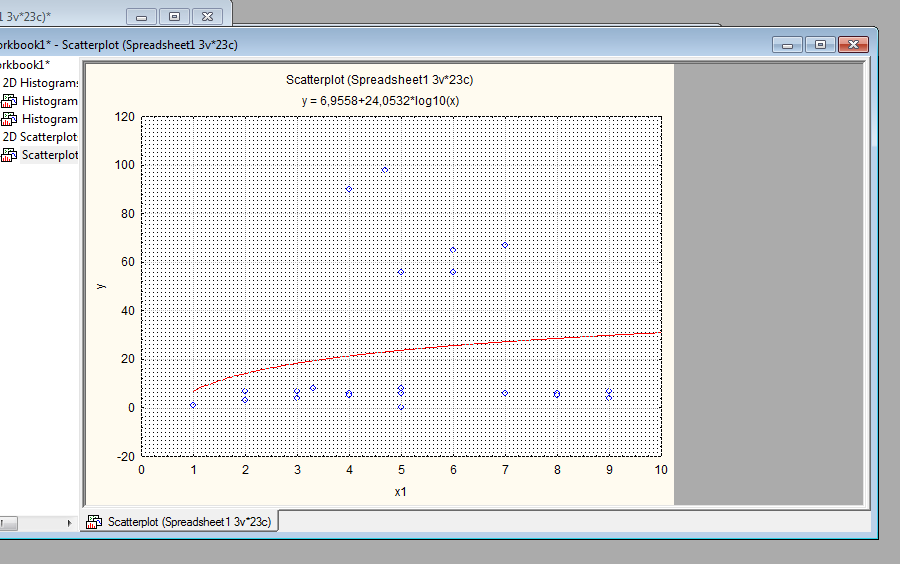
1. Змінюємо тип графіка на PieChart



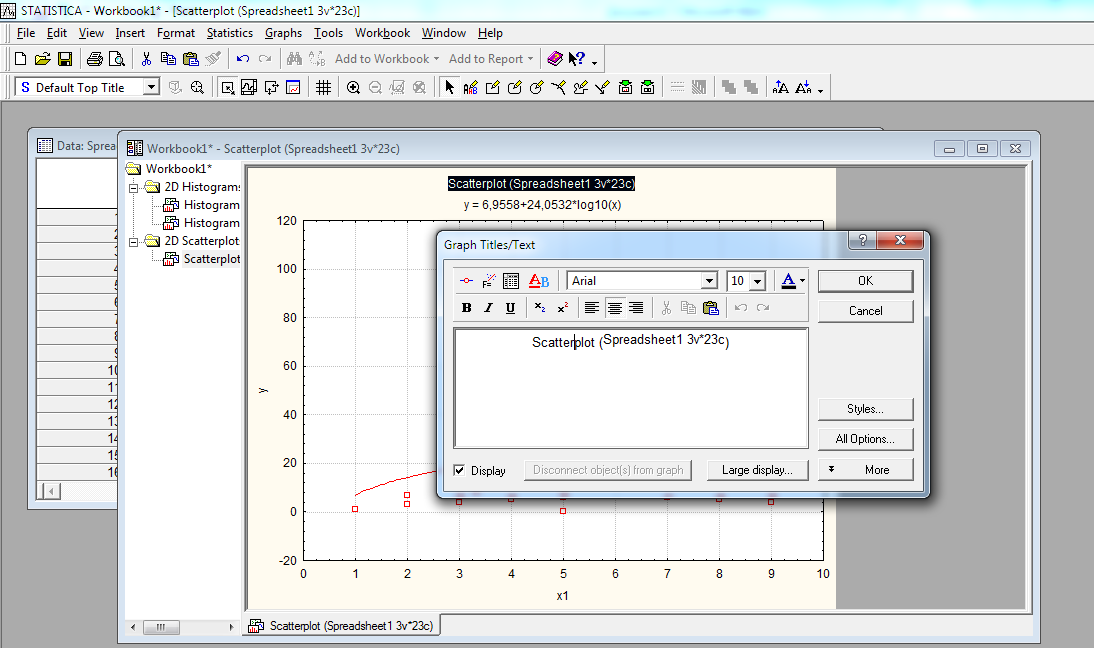
1. Аналогічно будуємо інші типи графіків (Graphs-2D Graphs-Scatterplots) та ін.

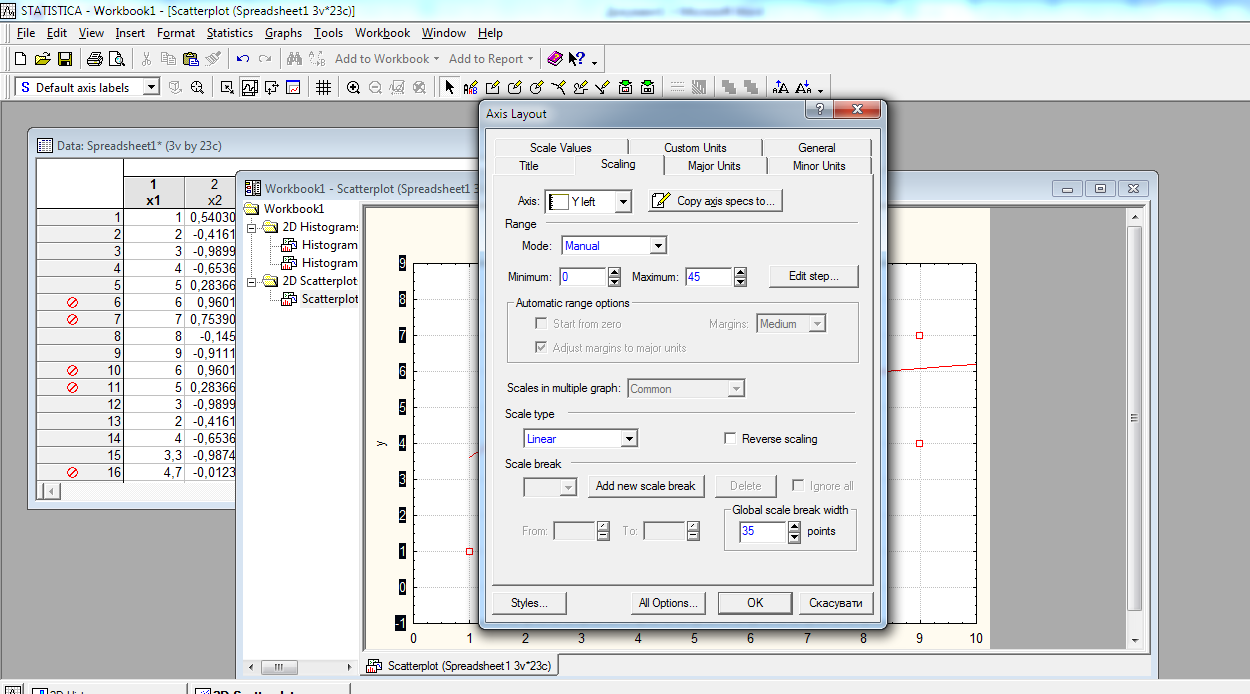


1. У графіку Scatterplot змінюємо тип функції з лінійної на будь-яку іншу нелінійну.



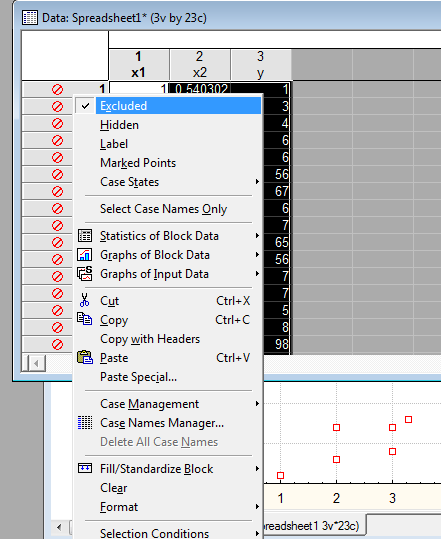
1. Змінюємо назву графіка і діапазон відображуваних значень осей x і y. (Axis:Scaling - Range – Mode - Manual)



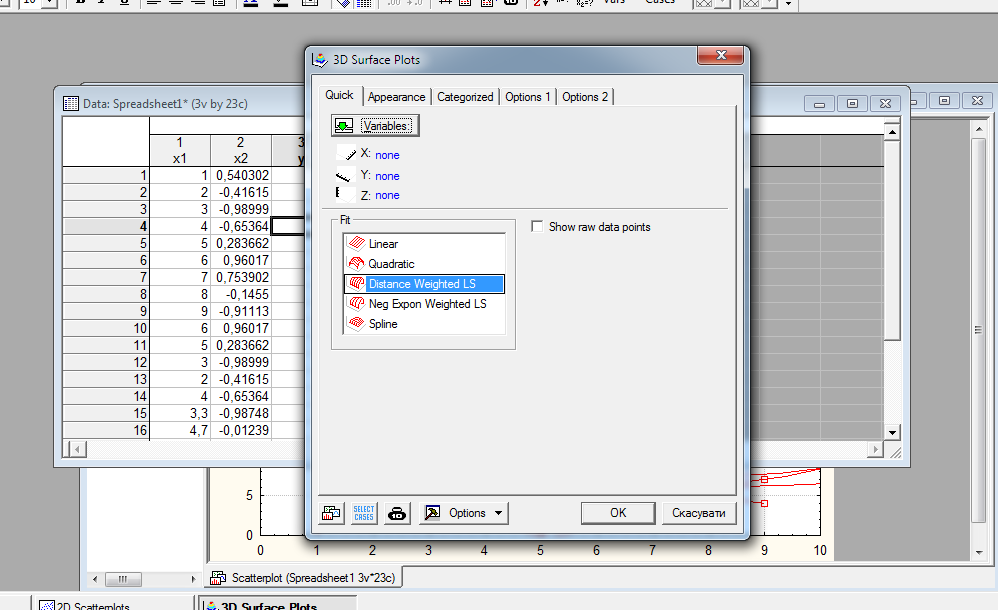


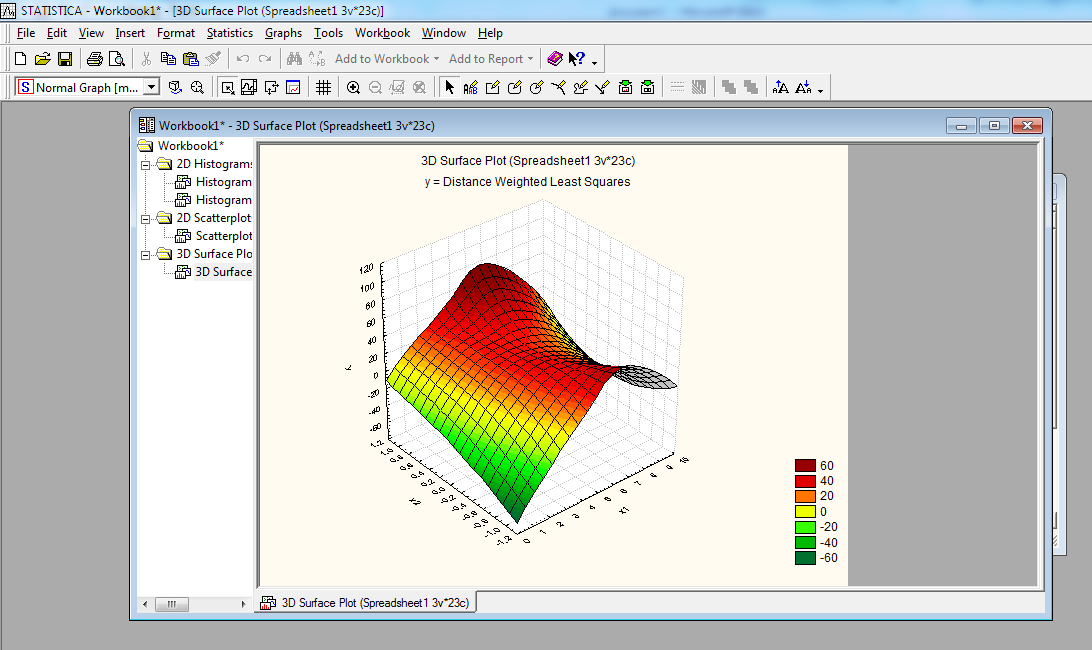
1. Використовуючи інструмент Brushing (Значок з прицілом) приховуємо з графіка, виключаємо з розрахунку деякі значення.

Приховувати, виключати з розрахунку і показувати значення можна безпосередньо в таблиці, як показано на рисунку

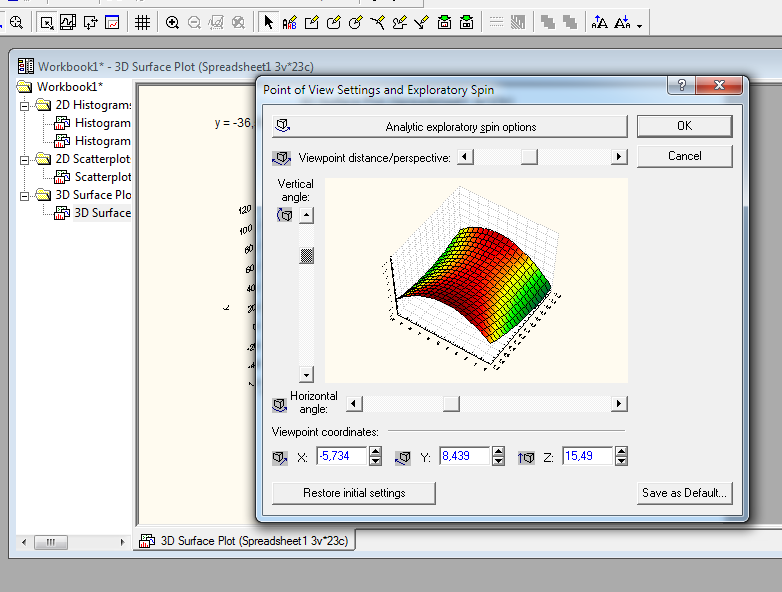
.

1. Будуємо 3D графік за трьома змінними (Surface Plots)



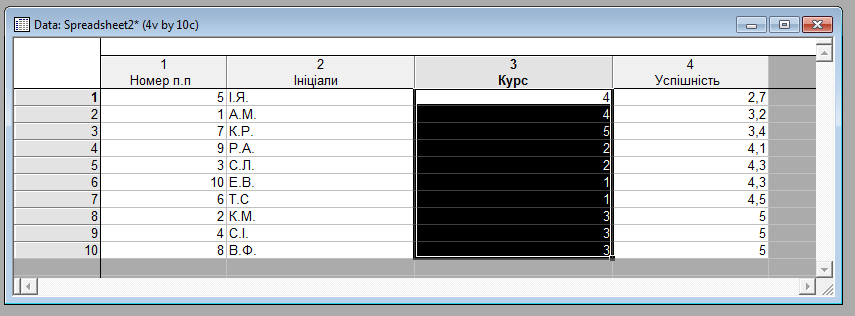


1. Змінюємо властивості графіка, зокрема кут нахилу і повороту



**Завдання 3.**

1. Створіть нову таблицю з чотирма змінними
2. Перейменуйте змінні var1, var2 та ін. на Порядковий номер студента (або Номер залікової книжки), Ініціали студента, Курс, Середній бал (Успішність).



1. Заповніть таблицю вигаданими значеннями.
2. Відсортуйте значення в алфавітному порядку, за середнім балом, за номером і т.д.
3. Побудуйте графік, що відображає залежність середнього балу від курсу.
4. Змініть властивості графіка, функцію кривої з лінійної на будь-яку нелінійну
5. Незважаючи на те, що значення є довільними і вигаданими, спробуйте проаналізувати графік і описати залежність.
6. Побудуйте гістограму для змінної «Успішність», змініть функцію кривої Normal на будь-яку іншу (через вікно властивостей графіка).

**Контрольні запитання**

1. Що таке система «STATISTICA» і для чого вона потрібна?
2. Особливість графічного формату \*.stg
3. Як експортувати графік в інший додаток?
4. Як помістити на графік системи STATISTICA графічний об'єкт з іншої програми?
5. Як помістити текст на графік STATISTICA (звіти, таблиці тощо)?