

Пакети опрацювання статистичних даних для студентів гуманітарних спеціальностей педагогічних університетів

Зміни, що відбулись останнім часом в Україні, стосуються насамперед освіти, основне завдання якої - забезпечення інтеграції освітнього, наукового та виробничого просторів та підготовка фахівців відповідно до потреб сучасного суспільства. Вони зумовили цілу низку протиріч між збільшенням обсягів даних з предметних галузей у професійній діяльності майбутніх педагогів-гуманітаріїв та обмеженими можливостями їх статистичного опрацювання; між необхідністю збільшення інтенсивності навчальної діяльності студентів гуманітарних спеціальностей з використанням комп'ютерних технологій статистичного опрацювання експериментальних даних та недостатньою розробленістю теорії і методики опрацювання даних у задачах педагогічного змісту; між зростаючими вимогами до рівня сформованості професійних компетентностей педагога та обмеженістю міжпредметних зв'язків між педагогікою, суспільними та іншими гуманітарними дисциплінами з одного боку та інформатичними дисциплінами – з іншого [1].

Одним з найважливіших завдань у педагогічних, соціологічних та інших дослідженнях є статистичний аналіз даних. Нині існує досить велика кількість статистичних пакетів опрацювання даних, використання яких дозволяє ефективно вивчати, аналізувати та досліджувати дані.

У статті зроблено порівняльний аналіз найбільш поширених статистичних пакетів опрацювання даних, які доцільно використовувати, як при навчанні, так і при проведенні наукових досліджень студентами гуманітарних спеціальностей педагогічних університетів.

Для порівняння були вибрані універсальні статистичні пакети, що працюють під управлінням ОС Windows і розроблені різними виробниками програмних продуктів, такі як SAS for Windows (SAS Institute Inc.), SPSS (SPSS Inc.), S-Plus (Mathworks), Systat (SPSS Inc.), NCSS (NCSS), STATA (Stata corp.), Statistica (Statsoft Inc.) і Statgraphics Plus (Manugistics, Inc).

Дані статистичні пакети є досить поширеними, доступні широкому загалу користувачів, прості та зручні у використанні, не вимогливі до потужності та характеристик комп'ютера. Завдяки зазначеним властивостям вони є досить популярними, часто використовуються у навчальній та науково-дослідницькій діяльності.

Таблиця 1

Пакети опрацювання статистичних даних	Доступ до файлів інших форматів	Кількість підтримуваних форматів	Наявність готових додаткових підпрограм
Statistica	Lotus, Quattro, Excel, Paradox, ASCII	5	Командна мова, макроси
STATA/ StatTransfer	Lotus, Access, dBASE, ASCII, Excel, Gauss, MatLab, Paradox, Quattro, SAS, S-Plus, SPSS, STATA, Systat	14	Розвинена командна мова
STAT GRAPHICS+	Statgraphics*, Execustat, dBASE, DIF, Lotus, Excel	6	-
S-Plus	ASCII, Excel, Lotus, dBASE, Paradox, Quattro, SigmaPlot, Systat, SAS, SPSS, STATA, MS Access, Gauss, MatLab, S-Plus*	15	Розвинена командна мова, скомпільовані додаткові модулі
MINITAB	Minitab*, Excel, Quattro, Lotus, Symphony, dBase, ASCII	7	Командна мова, макроси
SAS	SAS*, dBASE, Excel, Lotus, ASCII	5	Макроси, розвинена командна мова
SPSS	SPSS*, SYLK, Excel, dBASE, Lotus, Systat, ASCII	7	Макроси, розвинена командна мова

NCSS97	Access, ASCII, BMDP, Clone, dBASE, Excel, Gauss, Lotus, NCSS*, Paradox, Quattro, SAS, SigmaPlot, SOLO, S-Plus, SPSS, STATA, Symphony, Systat	19	-
SYSTAT	Excel, DIF, Dbase, Lotus, BMDP, Symphony, ASCII, SPSS	9	Розвинена командна мова

* – Застарілі або транспортні формати

У таблиці 1 наведено характеристики статистичних пакетів опрацювання даних такі як:

- доступ до файлів інших форматів, що дає змогу визначити коло інших програмних засобів, які можуть використовуватися з даним файлом відповідного статистичного пакету;
- кількість підтримуваних форматів, що свідчить про можливість доступу до даних з використанням різних статистичних пакетів опрацювання даних;
- наявність додаткових модулів підпрограм, призначених для встановлення чи створення нового програмного продукту.

Statistica (Рис.1). Дана система розроблялася як повна статистична система для користувачів персональних комп'ютерів, які не звикли до роботи в пакетному режимі перших версій SAS або SPSS. З самого початку цей пакет мав розвинений графічний інтерфейс і спирався на підтримку високоякісної графіки для аналізу даних. Система складається з ряду модулів, які працюють незалежно. Кожен модуль включає певний клас процедур (наприклад для кластерного аналізу). Майже всі процедури є керованими, тобто для запуску опрацювання даних необхідно вибрати з відповідного пункту меню команду і відповісти на ряд питань системи. Це дуже зручно для студентів-початківців. Інтерпретатор командної мови (*Statistica Command Language* доступний з будь-якого модуля. Майстер команд (*Command Wizard*) суттєво полегшує створення програми, зводячи програмування до вибору відповідних пунктів меню. За допомогою інтерпретатора опрацьовується мова макропослідовностей, тобто автоматизуються дії користувача. Вбудована мова програмування (*Statistica Basic*), не дивлячись на назву, за своєю структурою схожа на Паскаль, не набула значного поширення, тому додаткові підпрограми, створені третіми сторонами в ній практично відсутні. Жорстка структура також не дозволяє використання додаткових модулів.



Рис. 1 Вікно програми Statistica

Statistica є відносно невеликим статистичним пакетом, і має одну з найкращих систем підказок. Можливості експорту і імпорту даних розвинені достатньо, але без особливих доповнень. Працювати з графікою в цій програмі зручно і легко.

Statistica має дуже широку палітру статистичних методів і є однією з найпростіших для непідготовленого користувача систем і вдалим набором графічних можливостей. До недоліків системи можна віднести її малу розширюваність, відсутність модулів інших фірм і користувачів, а також недостатньо ефективну командну мову [2].

Stata. Stata є досить розвинутою системою статистичного опрацювання даних, яка працює під управлінням всіх основних операційних систем-Windows, UNIX. За своєю суттю вона є інтерпретатором мови програмування статистичних завдань. Звідси виникають всі позитивні і негативні сторони даної системи. До позитивних відносяться поширюваність, наявність великої кількості програм, написаних користувачами системи (служба технічної підтримки фірми випускає журнал, що містить програми, написані як співробітниками корпорації, так і користувачами), повна сумісність процедур, створених на різних платформах і легкість програмування власних статистичних програм.

Треба зазначити, що оригінальна версія системи не має користувацького інтерфейсу, а повністю управляється за допомогою командної мови. Для полегшення користування Stata студентами була розроблена оболонка StataQuest, в якій до системи було додано меню і вікна, що дозволило здійснити простий доступ до ряду статистичних процедур. Однак, оскільки StataQuest розроблялася для студентів, вона включила доступ лише до основних процедур: основні види множинної регресії, дисперсійного аналізу, непараметричної статистики та кореляційного аналізу. Крім того, в Stata вбудовані досить повні графічні можливості. Існує велика різноманітність графіків, які можна використовувати для візуалізації та візуальної оцінки даних, включаючи різні варіанти гістограм, Box-Plot, точкових і стовпчикові діаграм, діаграм розсіювання і т.п. Графіки зазвичай бувають двох типів – високих роздільних характеристик (при побудові яких використовуються графічні модулі системи) і низької роздільної здатності – коли для малювання використовуються звичайні текстові символи. Слід зазначити, що графіки можна зберігати в інших програмах, наприклад MS Word або MS PowerPoint (Рис. 2).

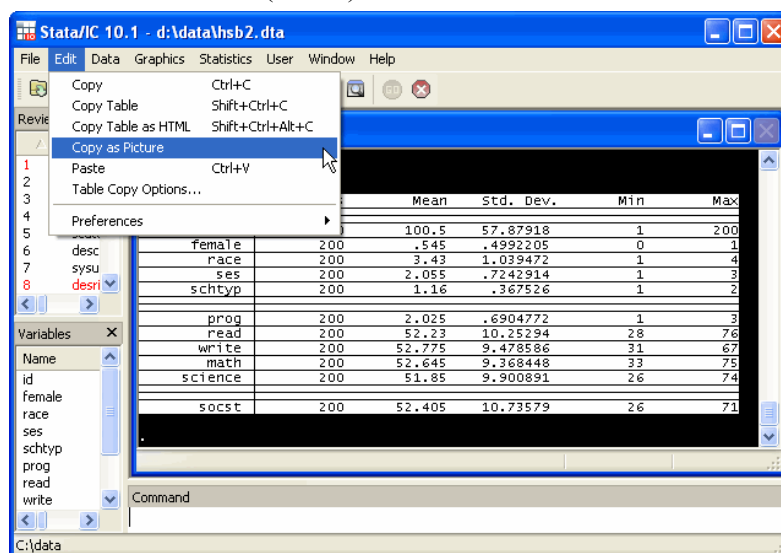


Рис. 2 Вікно програми Stata

В Stata дозволено використовувати в командному рядку умови, наприклад, розраховувати сумарні статистики не у всій аналізованій групі даних, а за певним піднабором даних. Повністю представлені різні методики регресійного аналізу.

В цілому Stata орієнтована на користувачів, які володіють деякими знаннями як у галузі статистичного опрацювання даних, так і в програмній реалізації статистичних алгоритмів. Для цієї категорії студентів статистичний пакет представляє потужний, швидкий і компактний інструмент. Разом з тим лише текстові висновки у таблицях, складність доступу до деяких команд робить цю програму недостатньо простою у використанні для студента-гуманітарія [2].

Statgraphics +. Дану систему було розроблено ще для операційної системи MS DOS. У ті часи вона відкрила перед користувачами систему меню, чітку графіку високих роздільних характеристик, великі можливості з експорту графічних зображень у поєднанні з досить повним набором статистичних алгоритмів.

Проте на комп'ютерах, оснащених операційною системою Windows, Statgraphics поступився своїми позиціями в якості "статистичної системи №1 для початківців" пакету Statistica та S-Plus. Разом з тим досі Statgraphics є системою статистичного опрацювання даних, орієнтованою на початківців-користувачів у поєднанні з потужними можливостями візуалізації даних.

Слід зауважити, що структура Statgraphics досить сильно відрізняється від такої в системах Statistica або SPSS. Справа в тому, що процедури у даній програмі згруповані за типами аналізу, а не за особливостями алгоритмів. Так, пункти меню носять наступні назви – "Порівняти", "Проаналізувати зв'язки", "Описати" – що значно полегшує вибір потрібних процедур, наприклад, в

порівнянні з таким пунктом меню, як "Базова статистика". При цьому методики параметричної і непараметричної статистики зазвичай знаходяться в одному пункті меню і можуть бути використані при перегляді опцій даного типу аналізу. Після кожного аналізу йде короткий коментар того, що було отримано, і даються пропозиції щодо використання додаткових методик. Активно використовуються команди контекстного меню.

Якщо дослідник звик працювати з іншими пакетами, де ставляться питання до тих пір, поки поставлене завдання не буде виконано, робота в Statgraphics може здатися дещо незграбною. Однак для студентів, які починають роботу з цієї програми, даний підхід може здатися природним - вибрати тип аналізу, вказати змінні, потім отримати коментар з приводу даних і початкові результати, а після цього вибрати уточнюючі методики аналізу.

Також треба зазначити, що однією з найбільш сильних сторін Statgraphics є його засоби візуалізації даних.

S-plus. S-plus є, напевно, одним з найрозвиненіших пакетів статистичного опрацювання даних. Основою S-plus є мова з такою ж назвою, розроблена більше двадцяти років тому в лабораторіях АТТ. Цей пакет спеціально призначений для аналізу та дослідницької роботи в галузі статистики. Сучасні версії використання S-plus базуються на ядрі мови і графічному пакеті Axiom, в якому також реалізовано графічний інтерфейс користувача. Основні можливості S-plus представлено в таблиці 2. У пакеті разом зі стандартними методами аналізу включено нечіткий кластерний аналіз і ряд інших додаткових можливостей. В S-plus передбачено підключення додаткових модулів вже в скомпільованому вигляді, тому розширення системи не викликає труднощів. Доступні додаткові модулі статистики, нейромережевого моделювання та ряду інших типів аналізу. Звичайно, робота з додатковими модулями не настільки зручна, як з вбудованими, проте можливість їх безкоштовного отримання в мережі Інтернет є дуже привабливою.

Графіка завжди була однією з сильних сторін S-plus, студенту надається широкий вибір різних високоякісних діаграм і дозволяється достатньо легко маніпулювати ними.

Але існує ряд недоліків. Виведення даних в S-plus не завжди зручне для інтерпретації. Так, наприклад, за командою кореляційного аналізу видаються стовпці значень коефіцієнтів з точністю до восьмого знака після коми, однак без розрахунку вірогідності коефіцієнтів чи допоміжних статистик.

Використання командного рядка (а значить і повністю всіх призначень S-plus) вимагає вивчення мови. Треба зауважити, що використання командного рядка дещо складніше, ніж в пакетах SAS, SPSS або Stata.

S-plus є системою, розрахованою, в основному, на викладачів у галузі статистичного опрацювання даних, науковців-дослідників у галузі статистики та навчання студентів-статистиків. Тому ця система є дуже поширеною [4].

Minitab – один з перших статистичних пакетів опрацювання даних для персонального комп'ютера. Нині Minitab явно втрачає свої лідируючі позиції. Незважаючи на досить великий набір статистичних процедур в ньому відсутні деякі з них, які вважаються стандартом і необхідні для аналізу педагогічних даних. Так, наприклад, відсутність непараметричних коефіцієнтів кореляції взагалі важко пояснити. Графіка в цьому пакеті досить розвинена, але використання великої кількості команд для побудови графіків ніяк не можна назвати інтуїтивним. Взагалі користувацький інтерфейс Minitab є однією з найслабкіших його сторін. Minitab має свою командну мову та є ряд макросів для розширення можливостей використання статистичного пакета. Разом з тим серед цих макросів немає тих, які кардинально б розширювали функціональність пакета, а їх число явно поступається такому в пакетах Stata або SAS.

Великою перевагою статистичного пакету є можливість отримати безкоштовну повну версію, що обмежена 30 днями роботи. Жодний з інших пакетів (включаючи Statgraphics) не пропонується для тестування в повному обсязі.

В цілому можна зазначити, що Minitab розрахований на початківців і може ними успішно використовуватися, проте він явно поступається за своїми характеристиками S-Plus [3].

SAS. Один з найперших і найбільш часто використовуваних пакетів опрацювання статистичних даних. SAS було розроблено для великих ЕОМ, але і нині його реалізовано для різних комп'ютерних платформ. SAS має програмно-модульну структуру, а це означає, що існують спеціалізовані модулі опрацювання даних (статистика – STAT, підтримка прийняття рішень – OR, графіка тощо), а всередині модуля є програми для цього опрацювання. SAS практично не розроблявся окремо для середовища Windows і тому версії для цієї операційної системи виглядають так само, як і для інших ОС.

Графічний інтерфейс користувача поставляється в окремому модулі – SAS / ASSIST – і не надає доступу не тільки до всіх опцій пакету, але навіть до їх більшої частини. Справа в тому, що цей модуль є оболонкою не для всієї системи, а лише для блоку загального аналізу даних, і тому повнота охоплення процедур обмежена. Слід зазначити, що на ринку існує і розповсюджується

безкоштовно оболонка для SAS – Overstat, за допомогою якої перетворюється ця програма в систему, більш звичну для користувача Windows, з повністю налагоджуваним меню і допоміжними вікнами. Дану програму розраховано на роботу з ранніми версіями SAS (6.03-6.04), проте може використовуватися і для нових версій.

Сильною стороною SAS є запрограмовані в ньому процедури опрацювання даних та повнота представлених процедур. Графіка також досить вражаюча, проте не може зрівнятися з генерованими графічними даними за допомогою S-Plus або Statistica.

Ще однією сильною стороною SAS є його поширюваність. Пакет включає командну мову, мову роботи з матрицями (IML) і підтримку макросів. Тож не дивно, що на ринку можна знайти достатньо велику кількість готових підпрограм і макросів для розв'язування різних статистичних задач.

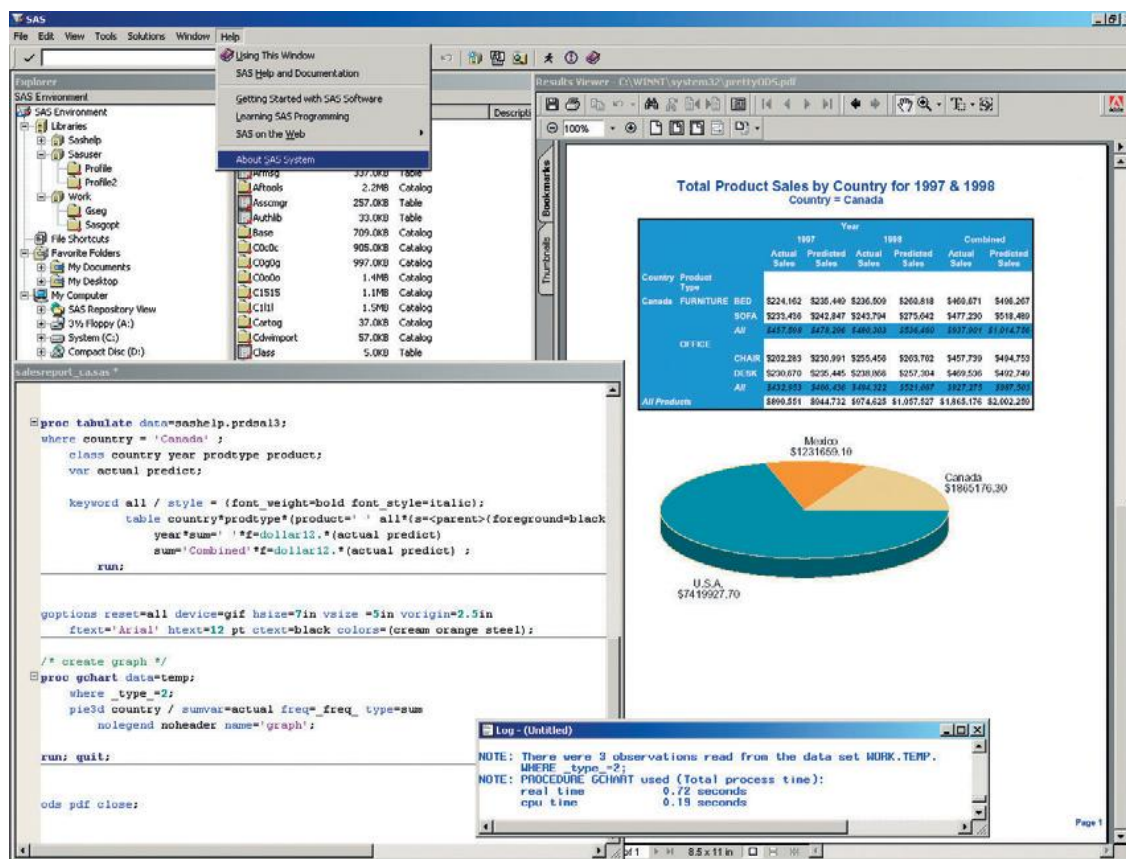


Рис. 3 SAS Analytics Pro забезпечує аналіз даних, звітність і графічне подання результатів в одному пакеті

Кожна процедура SAS має безліч опцій, за допомогою яких виконуються додаткові тести і специфікуються додаткові моделі. Природно, повне і гнучке використання всіх цих можливостей вимагає знання командної мови або створення детальних меню в Overstat (Рис. 3).

В цілому слід зазначити, що SAS є найбільш гнучким і розвиненим пакетом опрацювання даних, і може бути використаний і студентами-гуманітаріями, якщо вони скористаються оболонкою Overstat або SAS / ASSIST.

SPSS. SPSS є, разом з SAS, одним з найвідоміших пакетів статистичного аналізу даних. Однак, на відміну від SAS, компанія, що виробляє SPSS, завжди була значною мірою орієнтована на непрофесіоналів, і тому вже з ранніх версій у пакеті було розроблено розгалужену систему меню. При цьому система меню була лише оболонкою (front-end) для ядра пакету, керованого командною мовою. Використання меню призводило до автоматичного формування команд для SPSS, які потім необхідно було передати командному процесору (той же принцип використовується і в Overstat). В ході еволюції SPSS принцип взаємофункціонування оболонки з ядром залишався тим самим, однак ставав все більш і більш "прихованим" від користувача.

Тривалий час ядро SPSS залишалося без змін, однак у версії 7.5 один з основних компонентів – загальну лінійну модель (GLM) було переписано.

Крім того, фірма-виробник відмовилася від розвитку SPSS на платформах, відмінних від Windows, що призвело до можливості створити пакет, який має звичну для користувачів Windows структуру, скорочуючи, таким чином, час на навчання користувачів.

Нині в SPSS включено велику кількість статистичних процедур, можливості маніпуляції даними та створення графіків. Опрацювання статистичних алгоритмів проводиться надзвичайно ретельно і дозволяє добре контролювати процес опрацювання даних. Більшість команд доступно з меню та спеціальних вікон, що вигідно відрізняє SPSS від оболонок SAS.

Накопичення результатів дослідження в даному пакеті ведеться в редакторі SPSS Data Editor, дані зберігаються в форматі цього редактора (*.sav). Кожному показнику в таблиці відповідає певний рядок, а кожній змінній - певний стовбець (Рис. 4).

The screenshot shows the SPSS Data Editor interface. The title bar reads '1 - SPSS Data Editor'. The menu bar includes 'File', 'Edit', 'View', 'Data', 'Transform', 'Statistics', 'Graphs', 'Utilities', 'Window', and 'Help'. Below the menu bar is a toolbar with various icons. The main window displays a data table with the following content:

	age	social	inform	come
1	<18	Учащийся	Имеет правильную инфо	Буду участвовать
2	<18	Учащийся	Не имеет правильной ин	Не буду участвовать
3	<18	Учащийся	Имеет правильную инфо	Не буду участвовать
4	<18	Учащийся	Не имеет правильной ин	Не определился
5	<18	Учащийся	Имеет правильную инфо	Не буду участвовать
6	<18	Учащийся	Имеет правильную инфо	Не буду участвовать
7	<18	Учащийся	Имеет правильную инфо	Не буду участвовать
8	<18	Учащийся	Имеет правильную инфо	Не буду участвовать
9	<18	Учащийся	Не имеет правильной ин	Не буду участвовать
10	<18	Учащийся	Имеет правильную инфо	Не определился
11	<18	Учащийся	Имеет правильную инфо	Не буду участвовать
12	18-24	Учащийся	Не имеет правильной ин	Не определился

At the bottom of the window, the status bar indicates 'SPSS Processor is ready'.

Рис. 4 Подання результатів опитування в пакеті SPSS Data Editor

Разом з тим, SPSS поступається ряду інших статистичних систем за кількома параметрами. По-перше, політика фірми спрямована на те, що багато додаткових модулів існують у вигляді окремих програмних продуктів, що інтегруються в систему завдяки старанням користувача (і не можуть викликатися в прямому вигляді з командного процесора). По-друге, різні модулі можуть давати результати в несумісних форматах.

Система підказок в SPSS досить рудиментарна, і відрізняється в гірший бік від підказки у Statistica, а Statistical Coach не йде ні в яке порівняння зі Statistical Advisor в Statgraphics та S-Plus [3].

NCSS. Пакет NCSS є мало поширеним, однак це компенсується повнотою охоплення статистичних процедур усередині одного пакету (без додаткових модулів). За кількістю пропонованих процедур NCSS нагадує S-plus: за його допомогою виконують велику кількість процедур кластерного аналізу, детальну описову статистику, побудову графіків та багато інших статистичних методик. Зовнішній вигляд програми також нагадує S-plus. Перевагою пакету також є те, що всі його опції доступні з спадаючих меню, а супроводжуючий пакет – система підказок містить набір покрокових інструкцій з прикладами, що дозволяє швидко опанувати його основними операціями. Результати, що генеруються за програмою, автоматично зберігаються в ttf-файл, який потім легко прочитати і редагувати за допомогою будь-якого сучасного текстового редактора. Самі результати організовані так, що їх легко переглядати та аналізувати. Аналогічно пакету SAS, за допомогою даної системи за замовчуванням розраховують велику кількість статистик, що дозволяє охоплювати їх всіх одним поглядом і аналізувати.

Недоліки статистичного пакету: генеровані графіки не можна редагувати. Однак за іншими характеристиками NCSS є досить привабливою системою для студентів гуманітаріїв.

SYSTAT. Даний пакет позиціонується як повномасштабний статистичний пакет опрацювання даних для дослідників. Структура SYSTAT дуже схожа на структуру всіх інших пакетів, що базуються на ядрі командного інтерпретатора з оболонкою у вигляді меню і панелі кнопок. Разом з тим, за цілим рядом параметрів SYSTAT дійсно є дуже зручною для дослідників системою. По-перше, цей пакет містить досить широкий набір статистичних процедур в рамках одного модуля,

досить компактного і швидкого. Крім того, в SYSTAT підтримується спеціальний набір процедур для статистичного розпізнавання сигналу. Демо-версія, працює протягом 60 днів і обмежена тільки об'ємом вибірки (менше 250 спостережень). Недоліків у цього пакету небагато, до них можна віднести відсутність модулів розширення, не дуже інтуїтивні допоміжні вікна вибору опцій аналізу і відсутність форматуваних таблиць в стилі SPSS. Не дуже добре організований і модуль непараметричної статистики. Так, більшість тестів для пов'язаних сукупностей винесені в окремі пункти меню, а для незалежних - взагалі відсутні. Так само не цілком інтуїтивна і організація результатів параметричного дисперсійного аналізу [4].

В цілому система SYSTAT є потужним і зручним інструментом для студентів-гуманітаріїв.

В таблиці 2 наведено основні характеристики та компоненти названих вище статистичних пакетів опрацювання даних.

Таблиця 2

	Статистичні пакети опрацювання даних								
	SPSS	SAS	Statistica	NCSS	S-Plus	STATA	Systat	Stat-graphics+	Mini-tab
Система підказки	+	+	+	+	+	+	+	+	+
Наявність "помічника"	+	-	+	-	-	-	-	+\$	-
Описова статистика	+	+\$	+	+	+	+	+	+	+
Дисперсійний аналіз/порівняння груп	+	+\$	+	+	+	+	+	+	+
Кореляційний аналіз	+	+\$	+	+	+	+	+	+	+
Нелінійна регресія	+	+\$*	+	+	+	+\$	+	+	+
Кластерний аналіз	+	+\$	+	+	+	-	+	+	+
Факторний аналіз	+	+\$	+	+	+	+\$	+	+	+
Непараметрична статистика	+	+\$	+	+	+**	+	+	+*	+^
Аналіз часових рядів	+	+\$	+	+	+	+\$	+	+	+
Дизайн експерименту і контроль якості	-	-	+	+	+	-	+	+	+
Кластерний аналіз на основі нечітких множин	-	-	-	+	+	-	-	-	-
Побудова характеристичних кривих	-	-	-	+	-	+**	+	-	-

+ - компонент наявний;

-- - компонент відсутній;

* - наявний у вигляді окремих опцій аналізу;

** - є лише окремі компоненти;

\$ - велика частина опцій доступна лише через командну мову;

\$ \$ - повний коментар, що пояснює результати аналізу та можливості додаткового опрацювання;

^ - відсутня низка стандартних опцій (рангові коефіцієнти кореляцій, моделювання за Коксем).

Порівнюючи різні статистичні пакети опрацювання даних, можна зробити висновок, що практично всі вони мають набір стандартних процедур. Алгоритми, що використовуються в пакетах, здебільшого, стандартні і відмінностей при використанні тієї чи іншої програми немає. Тому при виборі статистичного пакета опрацювання даних необхідно звернути увагу на інтерфейс, повноту охоплення сучасних статистичних методів, програмованість, наявність додаткових модулів розширення і можливість використання отриманих результатів (графіків, таблиць, тощо) в інших програмах. Не всі перераховані вище вимоги можуть виконуватися одночасно. Так, наприклад, програмованість і поширюваність погано поєднуються з повнотою користувацького інтерфейсу. На одному рівні знаходяться в цьому відношенні пакети STATA і SAS - керовані переважно з командного рядка, але при цьому з великою кількістю підключених і використовуваних додаткових модулів. На іншому - Statgraphics +, NCSS і Statistica, які мають надзвичайно привабливий інтерфейс, повний і зручний для студентів-гуманітаріїв, однак при майже повній відсутності додаткових (безкоштовних) модулів і підпрограм.

Ті, хто в своїй роботі стикається з необхідністю застосовувати незвичайні статистичні підходи чи хоче скористатися ідеями, створеними і реалізованими іншими людьми, повинен звернути увагу на пакети, керовані з командного рядка. Обраний пакет повинен містити велику кількість готових підпрограм, процедур та мову маніпулювання матрицями, що надзвичайно зручно для написання статистичних програм.

Наявність оболонок для SAS (власний ASSISST і безкоштовний OverStat) дозволяє розв'язувати складні завдання опрацювання даних і аналізувати їх досить швидко. Stata дещо відстає від SAS щодо зручності програмування, однак це компенсується наявністю великої кількості додаткових модулів і передплатним листом, в якому користувачі наводять свої програми, розраховані на аналіз різних статистичних моделей.

Природно, ці два статистичні пакети навряд чи можуть бути рекомендовані для студентів-гуманітаріїв для опрацювання даних, які не збираються часто вдаватися до нестандартних статистичних процедур. Це інструменти осіб, які часто займаються аналізом великих масивів даних, що використовують різні статистичні підходи і бажають мати контроль над процесом опрацювання результатів.

Проміжну позицію займає SPSS, який одночасно є і статистичним пакетом з потужною мовою програмування і в ньому розроблено досить «дружній» до користувача інтерфейс. Взагалі ряд послуг, передбачених в SPSS, особливо в галузі факторного аналізу, є найбільш широким серед всіх описаних пакетів. Наприклад, в Statistici не пропонуються методики кутового обертання факторної структури (за винятком т.з. ієрархічного факторного аналізу), в SAS пропонуються тільки обертання за методом Promax, а в SPSS можна вибирати Promax або Oblimin.

Мова SPSS досить проста і дозволяє автоматизувати часто повторювані завдання, проте не завжди використання вбудованих команд приводить до очікуваного результату - наприклад, кореляційна матриця, створена при використанні підпрограми розрахунку коефіцієнтів Спирмена, не розпізнається підпрограмою факторного аналізу.

В цілому SPSS може бути рекомендовано науковцям, які хочуть мати систему з простим, інтуїтивним інтерфейсом, відносно розвиненою графікою і періодично використовувати мову програмування для автоматизації складніших завдань.

При навчанні студентів-гуманітаріїв слід звернути свою увагу на пакети Statistica або S-Plus. «Доброзичливий» інтерфейс, розвинена система добору і повнота поданих статистичних процедур дозволяють рекомендувати цю систему студентам, що часто використовують у своїй роботі статистичні методи аналізу даних.

Література

1. Гончаренко С.У. Методика навчання і наукових досліджень у вищій школі. – Київ: Вища школа, 2003. – 323 с.
2. Боровиков В.П. Statistica: Искусство анализа данных на компьютере / Владимир Боровиков. – СПб. [и др.]: Питер, 2001. – 650 с.: ил., табл. + 1 CD-ROM. – (Для профессионалов).
3. Бююль А. SPSS: искусство обработки информации: [пер. с нем] / Ахим Бююль, Петер Цёфель. – М.: DiaSoft, 2005. – 602 с. : ил., табл.

4. Шишкіна М.П. Вимоги до засобів інформатизації науково-педагогічного дослідження.
<http://www.nbu.gov.ua/e-journals/ITZN/em9/content/09smprrps.htm> 08.03.2010.