

XXVII. Нововведения в 11-ой версии SPSS

- [Использование программы SPSS в качестве ядра для современных маркетинговых исследований](#)
- [Конкретные нововведения в SPSS 11](#)

27.1. Использование программы SPSS в качестве ядра для современных маркетинговых исследований

Новая, 11-ая версия SPSS появилась в мае этого года, разумеется в английской локализации. Разработчики пакета сочли, что пользовательский интерфейс в последней версии уже является достаточно совершенным, поэтому подавляющее большинство изменений в новой версии связаны с усовершенствованием или добавлением статистических процедур, которые более полно обеспечивают потребности пользователей в современных методах обработки информации, возникающей в результате маркетинговых исследований, а также исследований в области социологии и психологии.

Основное внимание уделено расширению функциональных возможностей специальных модулей SPSS, таких как SPSS Categories, SPSS Advanced Models и другие. Собственно расширение программного наполнения специальных модулей и делает новую версию программы, которая уже успела стать стандартом в области обработки данных, еще более привлекательной для пользователей с весьма широким диапазоном деловых и научных интересов.

Если раньше данная программа широко использовалась в таких «классических» областях науки и бизнеса, как биология, социология, психология, управление качеством производства, общие маркетинговые исследования и экономическое прогнозирование, то сейчас новую версию можно с успехом применять в таких актуальных специализированных областях, как маркетинг, основанный на использовании баз данных, Data Mining, Data Warehousing и другие. Особенного внимания заслуживает тот факт, что изменения, внесенные в модуль SPSS Regression Models, позволяют использовать SPSS при решении задач управления лояльностью клиентов (CRM). Отметим, что данная тема представляет собой один из наиболее популярных разделов современного практического маркетинга.

Отдельного упоминания заслуживает то факт, что большинство наиболее популярных статистических методов прогнозирования, включенных в модуль SPSS Regression Models, позволяют работать с большим объемом недоступной информации. В математике в таком случае говорят о повышении робастности метода, то есть его устойчивости по отношению к неопределенностям и существенным отклонениям от диапазона параметров, для которого разрабатывался метод. Такое повышение робастности весьма желательно в маркетинговых исследованиях и в социологии, где всегда присутствует большой объем отсутствующих или недостоверных данных. Небесполезно данное улучшение и в области управления качеством, где всегда существует компромисс между подробностью информации о производственном процессе и его усложнением.

Изменения коснулись и техники вычислений. Подобные изменения не сказываются на интерфейсе и прочих видимых функциональных особенностях программы, но однако они затрагивают вычислительное ядро, которое используется в ходе проведения конкретных расчетов. Здесь основное внимание было сосредоточено на повышении эффективности статистических алгоритмов, в некоторых случаях эффективность повысилась до 50 раз.

Эффективность одной из наиболее часто используемых статистических процедур, общей линейной модели (GLM), возросла в 10 раз, что несомненно скажется на общей производительности при выполнении статистических исследований, особенно в области обработки больших массивов экспериментальных данных, которые возникают, например, в решении задач управления качеством, социологии и медицины.

В два раза выросла скорость выполнения самых массовых статистических процедур, таких как расчет дисперсии и вычисление средних. Можно смело сказать, что пользователь, который нуждается только в самых простых статистических методах, заметит именно двукратное повышение эффективности работы программы.

Особенно повышение быстродействия чувствительно в случае, когда речь идет о методах кластерного анализа, широко используемого в маркетинге, социологии, психологии и медицине, которые иногда требовали многочасовых расчетов даже на мощных компьютерах, для чего в предыдущих версиях SPSS был предусмотрен пакетный режим выполнения задач.

Следует отметить, что только одно столь существенное повышение производительности уже может быть основанием для выпуска новой версии программы. Снижение затрат времени, которое обеспечивает новая версия SPSS, позволяет более интенсивно использовать эту программу в практических маркетинговых исследованиях, анализировать большее количество вариантов, обрабатывать более широкие и представительные выборки. В результате издержки, связанные с исследованиями падают, а степень достоверности информации повышается.

Изменения, которые были внесены в изобразительную и презентационную части программы в основном затрагивают гибкость отображения результатов статистической обработки данных и включают несколько более показательных видов графиков. Например, при выводе информации о приближении данных с помощью выбранного метода аппроксимации, на графике приводится информация о том насколько хорошо полученное приближение. Такая дополнительная возможность может оказаться весьма полезной для не слишком опытных пользователей или пользователей не имеющих и не нуждающихся в глубокой математической подготовке. В целом изменения, которым подверглась графическая и презентационная часть программы направлены на упрощение работы и облегчение интерпретации результатов вычислений неподготовленными пользователями.

Рассматривая изменения, внесенные в техническую часть программы, необходимо упомянуть, что новая версия SPSS способна конвертировать базовые и переносимые файлы программы SAS (www.sas.com), своего наиболее мощного конкурента в области статистической обработки данных. Очень многие массивы общедоступной информации, имеющие отношения к маркетинговым исследованиям и социальной статистике, например данные по исследованию уровня жизни США и других стран (в том числе и России — знаменитый RLMS, www.unc.edu), проводимые американскими исследователями имеют формат переносимых файлов SAS.

Кроме того, следуя тенденции к превращению SPSS в мощное средство для проведения маркетинговых исследований и анализа разнородной информации, в 11-ой версии существенно расширено удобства доступа к различным форматам баз данных. В список поддерживаемых форматов теперь входят Sybase 11 и 12; Infomix 7.3+, 9.14; Infomix 2000 (9.20); UDB (DB2 6.1 и 7.1); SQL Server 2000; Oracle 8.06; Oracle! Releases 2 and 3 (8.1.6, 8.1.7). Улучшена связь с Microsoft Data Access pack. Более мощным стал язык запросов, появилась возможность на уровне запроса формировать и имена переменных и метки, что облегчает интерпретацию результатов и повышает их наглядность. Повысилась гибкость и функциональные возможности мобильных таблиц — это изменение затрагивает модуль SPSS Tables.

Если подытожить все вышесказанное, то можно сделать вывод, что переход конкретного пользователя на новую версию SPSS оправдан в том случае, когда этот пользователь нуждается в расширении своего арсенала методов статистической обработки данных, реализации наиболее современных тенденций в маркетинговых исследованиях, испытывает проблемы с полнотой информации, на основании которой необходимо сделать надежные и достоверные выводы, а также сталкивается с жесткими временными ограничениями, накладываемыми на процесс проведения исследований. Кроме того, работа с 11-ой версией SPSS снимает часть проблем, с которыми сталкиваются пользователи на начальных этапах освоения сложных методов статистического анализа, применяемого в широком спектре областей деятельности.

Переход к новой версии желателен, если пользователь собирается более широко применять современные методы анализа данных в ходе своей деятельности и уже имеет опыт использования аппарата математической статистики.

И, наконец, переход к новой версии не столь необходим, если в ходе деятельности пользователя возникает потребность в применении только самых простых статистических методах, а интерпретация результатов не вызывает особенных проблем. Кроме того, необходимо отметить, что предыдущая версия SPSS 10.1 имеет русскую локализацию, а выход локализованной 11-ой версии может стать весьма отдаленной перспективой.

27.2. Конкретные нововведения в SPSS 11

Приведем краткий список конкретных изменений, внесенных в вычислительную часть программы.

Следуя намеченному выше плану, начнем с нововведений в специализированные модули программы.

- В модуле Advanced Models добавлена новая статистическая процедура Linear mixed models (смешанные линейные модели), также известная как Hierarchical Linear Models (иерархические линейные модели), которая используется для получения наиболее точной прогнозирующей модели при работе с вложенной структурой данных. Гибкость этой модели позволяет существенно расширить применимость методов прогноза в решении задач экономического анализа и маркетинга: ANOVA фиксированными эффектами Model, Randomized Complete Blocks Design, Split-Plot Design, Purely Random Effects Model, Random Coefficient Model, Multilevel Analysis, модель безусловного линейного роста, модель линейного роста с a person-level covariate, Repeated Measures Analysis and Repeated Measures Analysis with time-dependent covariate. Кроме того, данная модель позволяет анализировать данные с повторными измерениями, включая неполные повторные измерения, когда объем информации изменяется от объекта к объекту.
- Внедрена новая процедура Descriptive ratio statistics (Дескриптивные статистики отношений)
- Использование медианы в качестве статистики при агрегировании данных, отсутствие данной возможности вызывало удивление у опытных пользователей SPSS.
- Метод многомерной логистической регрессии теперь можно использовать для построения психологического профиля клиента при проведении маркетинговых исследований.
- Улучшена масштабируемость и повышена производительность методов многомерной логистической регрессии (Multinomial Logistic Regression), иерархического Кластерного анализа (Hierarchical Cluster Analysis). Скорость выполнения иерархического кластерного анализа возросла от 5 до 50 раз в зависимости от типа вычислений. А скорость выполнения мультиномиальной логистической регрессии и Общих Линейных Моделей (General Linear Models) увеличилась в 10 раз
- Процедуры MLR, PLUM, GLM и MIXED позволяют сохранять предсказываемые значения даже при наличии пропущенных значений. Данные процедуры также можно использовать для определения целевых групп клиентов в маркетинговых исследованиях.
- Усовершенствована процедура категориальной регрессии (CATREG): возможность преобразования переменных, возможность обработки пропущенных значений, улучшение вывода таблиц и другие
- Усовершенствована процедура преобразования данных с использованием монотонных и немонотонных сплайнов с назначаемой пользователем степенью и числом узлов

В технической и презентационной части программы произошли следующие изменения:

- Для визуализации качества построенного приближения данных используется новый вид графиков

- Для получения большего объема данных о корреляционных связях введена новая форма выходных таблиц
- В выводимой информации появилась возможность использования длинных строк
- Расширились возможности задания вариантов обработки пропущенных значений
- Появились дополнительные возможности сохранения файлов данных
- В OLAP-кубах появилась статистика процентного изменения.
- Повышена защита данных путем шифрования при обмене между клиентом SPSS для Windows и сервером SPSS Server.
- Усовершенствован конструктор чтения баз данных (Database Wizard).
- Возможность делать выборку при импорте данных через ODBC.
- Возможность доступа к данным только для чтения.
- Не требуется login при доступе к данным MS Access.
- Улучшена графика и возможности работы с таблицами в SPSS Viewer